

Ruimtelijke ontwikkelingen woningbouw Nederland 1980 - 1995

*Een historisch-kwantitatieve analyse van de ruimtelijke
ontwikkelingen in de woningbouw 1980 – 1995
ter ondersteuning van de Omgevings-effectrapportage
Vijfde Nota Ruimtelijke Ordening*

februari, 2000

Faculteit der Economische Wetenschappen, Vrije Universiteit Amsterdam

drs Alfred Wagtendonk
Prof. dr Piet Rietveld

m.m.v.

dr Roland Goetgeluk (RIVM)
drs Maarten Hilferink (YuseSGO/ObjectVision)
drs Mathijs Ransijn (Provincie Noord Brabant)

vrije Universiteit

amsterdam



Samenvatting

Aanleiding

Dit onderzoek is aan de Faculteit der Economische Wetenschappen van de Vrije Universiteit uitgevoerd in de periode januari – mei 1999, als deel van een groter beleidsondersteunend onderzoek voor de Vijfde Nota Ruimtelijke Ordening. Dit onderzoek is verricht door RIVM en TNO Inro (Goetgeluk et al. 2000). Doel van dit onderzoek is om de toename in de woningvoorraad en werkgelegenheid tussen 1995 en 2020 voor drie Lange Termijn scenario's van het CPB, die geleverd worden op het ruimtelijke schaalniveau van COROP-gebieden en landsdelen, ruimtelijk te desaggregeren op het ruimtelijke schaalniveau van 500 bij 500 meter. Deze ruimtelijke allocatie vindt plaats op basis van trendanalyses en het doortrekken van bestaand beleid. Gebruikt worden de respectievelijk de modellen Ruimtescanner (RIVM) en Opera (TNO Inro). De VU is ingeschakeld om een trendanalyse te maken van de woningbouw opdat de resultaten (modelschatting) geïmplementeerd worden in de Ruimtescanner.

Doel en onderzoeksvragen van de VU-studie

Het doel van de op de VU uitgevoerde trendanalyse is:
het verkrijgen van een integraal beeld van de feitelijke ruimtelijke woningbouwontwikkelingen in de afgelopen 15 jaar, welke de basis vormen voor het vooruit berekenen van de autonome woningbouwontwikkelingen tot 2020.

Deze doelstelling is vertaald in de volgende vier onderzoeksvragen:

- Op welke wijze kunnen deze ruimtelijke ontwikkelingen op een analytische wijze worden verklaard?
- Welke schattingsmethodiek sluit zo optimaal mogelijk aan op de ruimtelijke allocatieprocedure van de Ruimtescanner?
- Welke software is geschikt om de ruimtelijke analyses met zo min mogelijke conversiestappen van gegevens te analyseren?
- Wat zijn de ruimtelijke ontwikkelingen met betrekking tot de uitbreiding van het wonen tussen 1980 en 1995?

Model en gegevens

Het VU-rapport beschrijft de ruimtelijke ontwikkelingen in de Nederlandse woningbouw van 1980 tot en met 1995 aan de hand van een ruimtelijk statistische analyse met behulp van een logistische regressiemodel (kansenmodel) en geografische informatie systemen. Het is een vrij technische studie. De nadruk ligt op de gemaakte keuzes met betrekking tot de verklarende variabelen, de gegevensmanipulatie, de keuze van het model en dergelijke (H2). Met dit inzicht kunnen de resultaten (H3) op hun waarde worden beoordeeld.

In het VU-onderzoek worden twee hoofdtypen woningen onderscheiden, eengezins- en meergezinswoningen. Specifiek is gekeken naar de ruimtelijke relaties tussen de groei van de woningvoorraad per woningtype (op het voor woningbouw beschikbare areaal) en de ruimtelijke factoren die deze groei beïnvloeden. De analyses zijn uitgevoerd op een schaalniveau van gridcellen van 500 x 500 meter.

De te verklaren variabele is de kans op de toename van de woningdichtheid naar woningtype tussen 1980 en 1995. De verklarende variabelen zijn geselecteerd op basis van bestaande wetenschappelijke kennis op het terrein van het ruimtelijke keuzegedrag van woonconsumenten. Er is geen rekening gehouden met andere actoren in het spel, zoals gemeenten, grondeigenaren en institutionele investeerders. Statistisch is getoetst in welke mate deze verklarende variabelen de toename van de dichtheid goed beschrijven. Dit heeft plaatsgevonden op basis van ten eerste een correlatie- en ten tweede een logistisch regressiemodel. Daar de ruimtelijke spreiding in de toename ongelijk kan zijn, zijn de analyses apart gemaakt voor de Randstad, de Intermediare zone en de Periferie. Deze indeling is eerder gebruikt in onderzoek. Uit de analyses blijkt dat de heterogeniteit tussen beide laatste zones gering is en zijn beide samengevoegd.

Er is gebruik gemaakt van een logistisch regressie-model. De modelschattingen $f(x)$ kunnen rechtstreeks worden geïmplementeerd in de Ruimtescanner. De parameters (beta's die de invloed van de verklarende variabelen kwantitief weergeven) van dit model zijn niet zo eenvoudig te interpreteren als in een gewone multiple regressiemodel. De parameters worden namelijk ingevuld in een exponentiële kansfunctie $P = e^{f(x)} / (1 + e^{f(x)})$, waarbij $f(x)$ de regressievergelijking is, P de kans is en e staat voor een e-machtfunctie ($e^{-2,7}$). Indien de regressievergelijking $f(x)$ de waarde nul heeft wordt de kans op een toename ($e^0 / (1 + e^0)$), hetgeen identiek is ($1/1+1$) ofwel 0,50. Indien de functie de waarde 1 heeft dan wordt de kans 0,73 ($(e^1 / (1 + e^1))$ ofwel $(2,7 / (1 + 2,7))$). Is de waarde van de functie -1 dan wordt de kans 0,27. Bovenstaande drie voorbeelden maken inzichtelijk hoe de beta's moeten worden behandeld.

De resultaten

Uit de regressie-analyses komt een aantal significant bijdragende variabelen naar voren, waarmee de ruimtelijke ontwikkelingen in de woningbouw in de periode van 1980 tot en met 1995 goed statistisch te verklaren zijn. Dit wil zeggen dat de berekende verdeling van kansen op toename van woningen visueel en statistisch vergelijkbaar zijn met de verdeling van de werkelijke toename van deze woningtypen tussen 1980 en 1995. Bij deze vergelijking blijkt dat de kansverdeling en het model in het algemeen goed met elkaar overeen komen. Met name de bestaande woningbouw, de ligging in groeikernen (bouwlocaties), de nabijheid van werkgelegenheid en de afstanden tot NS-stations en -in iets mindere mate- op- en afritten van snelwegen, dragen in positieve wijze aan de verklaring van de groei in de woningvoorraad bij. Ook de nabijheid van bos, natuur en recreatiegebieden levert een kleine, maar significante, bijdrage aan de verklaring van de toename van de dichtheid in de woningvoorraad. Bovenstaande tabel toont de uitkomsten, waarbij de t-waarden aangeven of een beta significant afwijkt van nul.

Randstad				
	Eengezins		Meergezins	
	B	t-waarde	B	t-waarde
Startwaarde/intercept	-7,02	-389,59	-20,31	-255,65
Kennisdiensten	Nvt	Nvt	1,00	160,00
Recreatief bosgebied	-0,02	-27,46	0,06	36,386
Wegen	-0,34	-94,70	-0,36	-69,85
Rail	-0,58	-91,31	Nvt	Nvt
Stations	-0,14	-103,47	-0,13	-104,60
Op- en afritten	0,04	33,04	0,12	61,27
Groeikern	1,13	264,04	0,74	100,819
Eengezinswoningen	0,73	548,00	0,54	242,253
Meergezinswoningen	0,06	75,67	0,39	331,891
Overig Nederland				
	Eengezins		Meergezins	
	B	t-waarde	B	t-waarde
Startwaarde/intercept	-12,06	-462,24	-15,02	-329,83
Industrie	0,34	121,35	Nvt	Nvt
Kennisdiensten	Nvt	Nvt	0,34	79,31
Recreatief bosgebied	Nvt	Nvt	Nvt	Nvt
Natte natuur	0,08	75,27	0,11	39,59
Wegen	-0,11	-41,34	-0,15	-27,00
Rail	-0,37	-85,23	-0,30	-35,77
Stations	0,12	-115,32	-0,18	-109,74
Op- en afritten	0,05	57,14	Nvt	Nvt
Groeikern	1,62	364,66	1,08	121,58
Eengezinswoningen	0,83	933,53	0,88	300,76
Meergezinswoningen	0,10	126,37	0,50	328,66

Aanbevelingen voor verder onderzoek

- Een aantal van de beslissingen aangaande de te volgen methodiek zijn genomen op grond van de beschikbaarheid van de op dat moment gewenste data. Zo kon er bijvoorbeeld niet met de gewenste woonmilieu's gewerkt worden, maar moest met vijf woontypes gewerkt worden die vervolgens weer tot twee verschillende woontypen samengevoegd moesten worden. Aangezien gewerkt wordt op een schaalniveau van 500 metercellen was ook de homogene verdeling van aantallen arbeidsplaatsen per postcodegebied naar 500 meter cellen, geen ideale oplossing. In verband met het te volgen tijdsplan kon echter niet gewacht worden op nieuwe data. Indien de gewenste data alsnog ter beschikking komt kunnen de regressie-analyses verder verbeterd worden.
- Deze trendstudie is valide voor het onderzoek van RIVM en TNO Inro waarin de aanname centraal staat dat de bestaande verhouding tussen enerzijds de regulerende overheid (mogelijkheden en beperkingen) als publiek orgaan en anderzijds de woonconsumenten, en indirect de overheid als privaat orgaan en aanbieders van woonruimte en woondiensten, eenvoudig kon worden doorgetrokken. Voor de ontwikkeling van alternatieve ruimtelijke scenario's voldoet een dergelijke trendanalyse op basis van ruimtelijke informatie niet. In het laatste geval dient een modelspecificatie te worden gebruikt op basis van een gedragsmatige onderbouwing van het besluitvormingsgedrag van de relevante actoren in het spel

- De verklarende variabele grondprijzen is niet in de analyse meegenomen omdat hier geen geschikte data van voorhanden was. Vanwege het verwachte verband tussen grondprijzen en de verschillende woningtypen wordt aanbevolen deze variabele (indien beschikbaar) op te nemen in de trendanalyses. In deze studie hoefde de grondprijs niet te worden opgenomen vanwege de aannames in de RIVM en TNO Inro studie (Goetgeluk et al. 2000).
- Behalve de beschikbaarheid van de data leverde de omvang van de in S-Plus (statistisch pakket) in te voeren data-set problemen op met de rekenkracht van de gebruikte computers. Geprobeerd is te werken met samples van de dataset (25.000 gridcellen in plaats van 140.000), dit leverde echter afwijkende resultaten op ten opzichte van de volledige dataset. Om sneller te kunnen werken wordt daarom het gebruik van krachtiger computers met veel opslagruimte aangeraden.
- In de huidige opzet is gewerkt met één tijdperiode van 1980 tot en met 1995. Indien meer tijd (en meer data) beschikbaar is, kunnen de analyses per tijdperiode (bijvoorbeeld perioden van vijf jaar) verricht worden, zodat ook ontwikkelingen in de tijd van variabelen en coëfficiënten kunnen worden waargenomen. Deze trends kunnen vervolgens weer naar de toekomst worden doorgetrokken.
- Uitgegaan is van een maximaal aantal woningen per woningtype per cel in 1995. Dit aantal hoeft niet de hoogste dichtheid van woningen per cel te vertegenwoordigen. Wellicht kan beter met de hoogst mogelijke dichtheid gewerkt worden. Deze kan gevonden worden door eerst het niet voor wonen geschikte areaal van alle gridcellen af te trekken en het aantal woningen dan door de resterende celoppervlakte te delen. Vervolgens kan uitgegaan worden van een maximaal aantal woningen per voor wonen geschikte hectare.
- De ruimtelijke schaal waarop claims worden gebruikt bepaalt in hoge mate de uitkomsten. Uit de analyses kwamen duidelijke regionale verschillen naar voren tussen de Randstad en overig Nederland. Indien het onderzoek wordt uitgebreid kunnen meer regionale indelingen geanalyseerd worden, zoals een analyse per woningmarktgebied. Deze worden omschreven als het gebied waar de meeste mensen verhuizen omwille van woonredenen zonder dat hierbij de bereikbaarheid van familie, vrienden en het werk in geding komt.

Inhoud

SAMENVATTING	2
AANLEIDING	2
MODEL EN GEGEVENS	2
TOEPASSING VAN DE RESULTATEN IN DE RUIMTESCANNER	ERROR! BOOKMARK NOT DEFINED.
DE RESULTATEN	3
AANBEVELINGEN VOOR VERDER ONDERZOEK	4
INHOUD	7
VOORWOORD	9
INLEIDING	11
INLEIDING	11
HET KADER VAN DIT ONDERZOEK	11
DOELSTELLING VAN HET VU-DEELPROJECT.....	12
OPZET RAPPORTAGE.....	14
MODELLEN & DATA	17
INLEIDING	17
GLOBALE WERKING MODEL RUIMTESCANNER	17
SCHATTINGSMETHODIEK.....	18
<i>Theoretische en pragmatische onderbouwing voor een ruimtelijk analytisch model om de attractiviteitskaarten te bepalen</i>	18
<i>De achtergrond van de gebruikte schattingsmethodiek</i>	20
<i>Maximum Likelihood schattingsprocedure</i>	22
<i>Logistische regressiemodel en correlatie-analyse</i>	23
<i>Conclusie</i>	25
DE VARIABELEN	25
<i>De te verklaren variabele (Y)</i>	25
<i>De verklarende variabelen (X)</i>	28
<i>Conclusie</i>	33
GEBRUIKTE STATISTISCHE SOFTWARE	33
TRENDANALYSE	35
INLEIDING	35
CORRELATIE-ANALYSE	35
<i>Resultaten</i>	35
<i>Conclusies ten aanzien van keuze variabelen</i>	38
REGRESSIE-ANALYSE	39
<i>Logistische regressies Randstad en eengezinswoningen</i>	39
<i>Logistische regressies Randstad en meergezinswoningen</i>	41
<i>Logistische regressies intermediar/perifeer Randstad en eengezinswoningen</i>	41
<i>Logistische regressies intermediar/perifeer Randstad en meergezinswoningen</i>	41
VISUELE TOETSING REGRESSIE-RESULTATEN	42
CONCLUSIES	45
INLEIDING	45
EEN TERUGBLIK	45
AANBEVELINGEN VOOR ONDERZOEK	45
LITERATUUR	47

BIJLAGE 1 ML- PARAMETER SCHATTINGEN VAN DE FACTOREN VOOR WONINGBOUWGROEL.....	49
BIJLAGE 2 SUMMARY STATISTICS 1980.....	51
BIJLAGE 3 ONDERLINGE CORRELATIES 'VERKLARENDE' VARIABELEN.....	52
BIJLAGE 4 STEPWISE REGRESSION ANALYSIS.....	55

Voorwoord

Graag willen wij onze dank betuigen aan Mathijs Ransijn (tegenwoordig provincie Noord-Brabant) en Maarten Hilferink (YUSE GSO Object Vision) die waardevolle bijdragen hebben geleverd in de beginfase van dit onderzoek. Daarnaast gaat onze speciale dank uit naar Roland Goetgeluk (RIVM) die ons met onverminderd enthousiasme heeft geholpen bij de voorbereidingen voor de wetenschappelijke onderbouwing, statistische analyses, commentaar, en de eindredactie van dit rapport.

Faculteit der Economische Wetenschappen, Vrije Universiteit Amsterdam
Amsterdam, dinsdag 8 februari 2000

Drs. A. Wagtendonk

Prof. dr. P. Rietveld

Inleiding

Inleiding

Dit rapport beschrijft de werkzaamheden die zijn uitgevoerd door de Vrije Universiteit Amsterdam (VU) als onderdeel van een groter project dat wordt uitgevoerd in het kader van de beleidsvoorbereiding van de PKB2 Vijfde Nota Ruimtelijke Ordening (VIJNO). In het VIJNO-project worden de verwachte economische en demografische ontwikkelingen en het vastgestelde ruimtelijke beleid in Nederland vertaald naar een ruimtelijk beeld van het ruimtebeslag en de distributie van de verschillende gebruiksfuncties tot 2020. Dit onderzoek is uitgevoerd door RIVM en TNO Inro (Goetgeluk et al. 2000). Vervolgens worden de mogelijke gevolgen van deze ruimtelijke ontwikkelingen voor de kwaliteit van de leefomgeving ingeschat door RIVM en partners.

Het VIJNO-project wordt gecoördineerd door het Rijksinstituut voor Volksgezondheid en Milieu (RIVM) en TNO Inro. Bij de algemene omschrijving van het project is gebruik gemaakt van het projectvoorstel zoals door het RIVM en TNO Inro is opgesteld (zie Goetgeluk et al. 2000). In dit hoofdstuk behandelen we eerst de probleemstelling en doelstellingen van het VIJNO-project. Vervolgens zoomen we in op de probleem- en doelstellingen van het projectonderdeel van de VU.

Het kader van dit onderzoek

In de regeringsverklaring wordt de PKB Vijfde Nota Ruimtelijke Ordening (VIJNO) aangekondigd. De VIJNO schetst de ruimtelijke hoofdstructuur van Nederland voor de komende decennia. De ruimtelijke hoofdstructuur is een middel om zorg te dragen voor de kwaliteit van de leefomgeving. Deze kwaliteit is een duurzame specifieke mix van economische, ecologische en sociale factoren, die van belang zijn voor het gezamenlijk functioneren van huishoudens, bedrijven en instellingen. De hoofdstructuur zal daarom ruimte bieden en grenzen stellen aan individuele keuzen van huishoudens, bedrijven en instellingen.

In 1997 zijn door de planbureaus verkenningen uitgebracht waaruit blijkt, dat de groei van de bevolking, economie en mobiliteit de kwaliteit van de leefomgeving onder druk zet. Voor de VIJNO is daarom de vraag relevant hoe via de ruimtelijke hoofdstructuur kan worden bijgedragen aan de kwaliteit van de leefomgeving.

De RPD heeft de planbureau's, waaronder het RIVM gevraagd, om een bijdrage te leveren aan de wetenschappelijke onderbouwing van de VIJNO. De kennisvraag aan de planbureaus kan worden verdeeld in vier vragen:

- Welke ruimtelijke ontwikkelingen zijn tussen 1995 en 2020 te verwachten op grond van verwachte economische en demografische ontwikkelingen en het vastgestelde ruimtelijk beleid (de autonome ontwikkelingen)?

² Planologische KernBeslissing

- Hoe verhouden deze ruimtelijke ontwikkelingen zich tot de doelstellingen van het vastgestelde ruimtelijke beleid en de visie Stedenland-plus?
- Wat zijn de gevolgen van deze ruimtelijke ontwikkeling voor de kwaliteit van de leefomgeving (milieu, natuur, landschap, bereikbaarheid, leefbaarheid en economische vitaliteit)?
- Wat zijn tegen deze achtergrond de effecten van beleidsopties voor de ruimtelijke inrichting om de gesignaleerde knelpunten te kunnen opheffen en meer in het bijzonder hoe kan via de ruimtelijke inrichting zowel de ruimtelijk-economische dimensie als de sociale en ecologische dimensie worden geoptimaliseerd?

De doelstelling van het VIJNO-project 'Autonome ruimtelijke ontwikkelingen 2020' is het simuleren van de ruimtelijke verdeling van bevolking en werkgelegenheid en het bepalen van de neerslag hiervan op de bebouwde omgeving in de periode 1995-2020, op basis van een empirische analyse van de historische ruimtelijke ontwikkeling, rekening houdend met CPB lange termijn scenario's (LT'97) en het vastgestelde ruimtelijke beleid. Met deze kennis kan de 'distance to target' worden bepaald ten opzichte van actuele beleidsdoelen en de ruimtelijke visie Stedenland-plus.

Doelstelling van het VU-deelproject

De onderzoeksoopdracht van het RIVM aan de VU heeft hoofdzakelijk betrekking op de eerst genoemde kennisvraag: welke ruimtelijke ontwikkelingen op het terrein van het wonen zijn tussen 1995 en 2020 te verwachten op grond van verwachte economische en demografische ontwikkelingen en het vastgestelde ruimtelijk beleid? Beantwoording van deze vraag is een noodzakelijke eerste fase in de beantwoording van de overige vragen. De vraag spitst zich specifiek toe op de analyse van ruimtelijke ontwikkelingen in de woningbouw in de periode 1980 – 1995. Het onderzoek naar de ruimtelijke ontwikkelingen in de sector arbeid is in een afzonderlijk onderzoek uitgevoerd door TNO INRO (Goetgeluk et al. 2000).

Het doel van de op de VU uitgevoerde trendanalyse is:

het verkrijgen van een integraal beeld van de feitelijke ruimtelijke woningbouwontwikkelingen in de afgelopen 15 jaar, welke de basis vormen voor het vooruit berekenen van de autonome woningbouwontwikkelingen tot 2020.

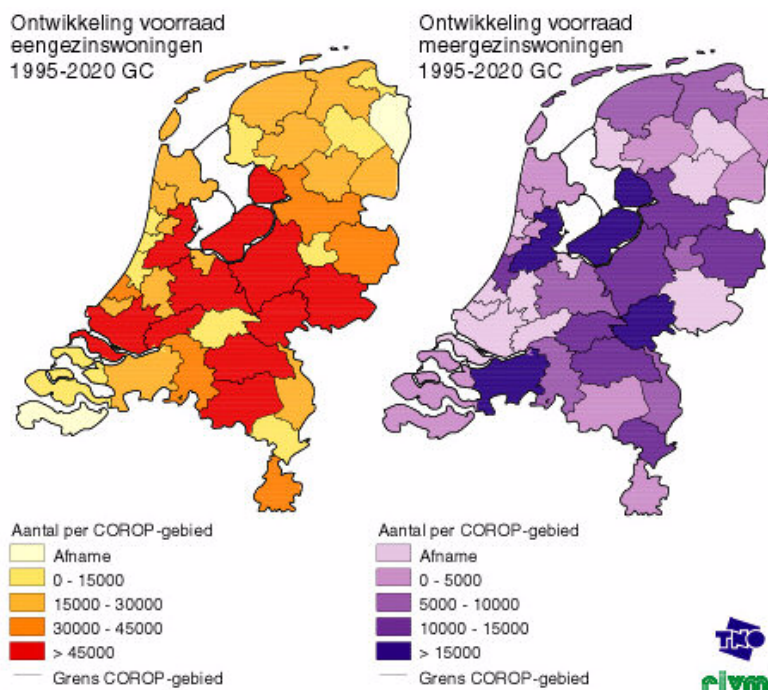
Het integrale beeld betreft de kwantitatieve relaties (in de vorm van rekenregels) tussen de groei van de dichtheid in de woningvoorraad in de periode 1980 – 1995 en de ruimtelijke factoren die deze groei beïnvloeden. De resulterende rekenregels dienen direct ingevoerd te kunnen worden in het allocatie-model (Model Ruimtescanner) dat door het RIVM gebruikt wordt om de trend uit het verleden te vertalen naar toekomstige ontwikkelingen.

Deze doelstelling is vertaald in de volgende vier onderzoeksvragen:

- Op welke wijze kunnen deze ruimtelijke ontwikkelingen op een analytische wijze worden verklaard?
- Welke schattingsmethodiek sluit zo optimaal mogelijk aan op de ruimtelijke allocatieprocedure van de Ruimtescanner?

- Welke software is geschikt om de ruimtelijke analyses met zo min mogelijke conversiestappen van gegevens te analyseren?
- Wat zijn de ruimtelijke ontwikkelingen met betrekking tot de uitbreiding van het wonen tussen 1980 en 1995?

De verkregen regressieresultaten ($f(x)$) kunnen rechtstreeks worden geïmplementeerd in het model Ruimtescanner dat gebruikt wordt om de toekomstige verwachte uitbreiding voor de woningvoorraad. Deze uitbreiding wordt geleverd op het ruimtelijke schaalniveau van COROP-gebieden (kaart 1.1).

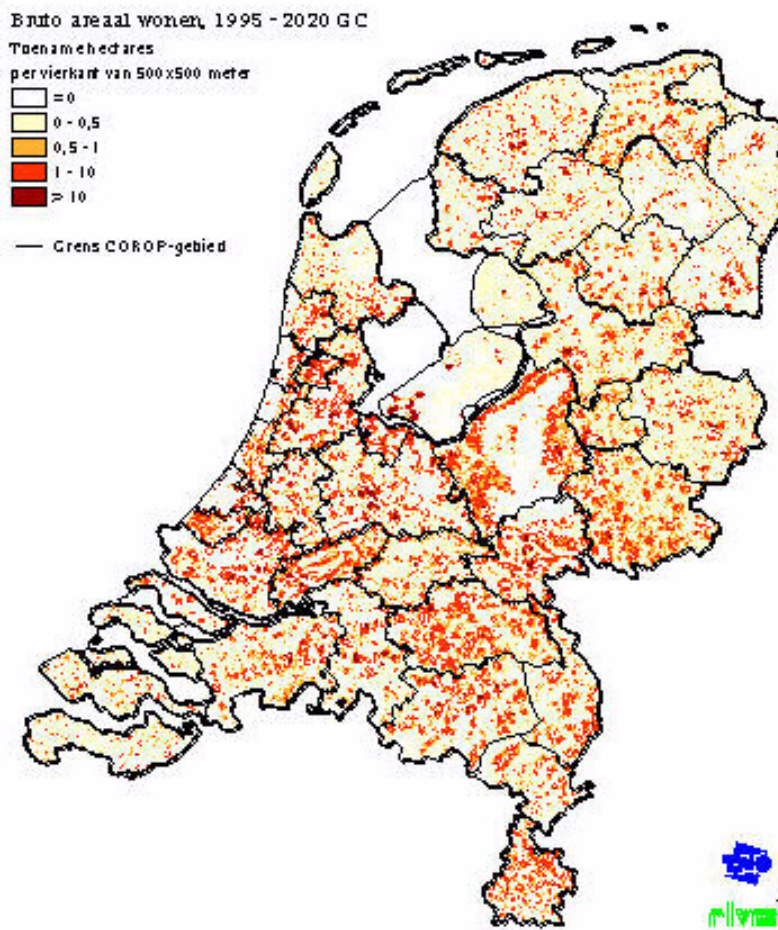


Kaart 1.1 Voorbeelden van de uitbreidingsvraag per COROP-gebied die op een lager ruimtelijk schaalniveau moeten worden gealloceerd (bron: RIVM/TNO)

Bovenstaande claims moeten ruimtelijk gealloceerd worden op een lager ruimtelijk schaalniveau (kaart 1.2). De vraag is hoe dat moet. Voor dat doel wordt de Ruimtescanner gebruikt. De allocatieprocedure van dit model is onder andere ontwikkeld door de faculteit der Economische Wetenschappen van de Vrije Universiteit.

Het doel van de Ruimtescanner is om deze uitbreiding ruimtelijk te alloceren op het lage ruimtelijke schaalniveau van 500 bij 500 meter (gridcel). De Ruimtescanner bezit een ruimtelijk allocatiemodule op basis van een logistisch regressiemodel. In de invulling van de parameters van dit model is afhankelijk van het onderzoeksdoel. In dit onderzoek moeten de parameters gebaseerd zijn op een trendanalyse van de toename van de woningvoorraad gegeven het beschikbare areaal. Kenmerkend aan de Ruimtescanner is dat de ruimtelijke claims op een hoogschaalniveau op een lager ruimtelijk schaalniveau zodanig worden

gealloceerd totdat de claim 'op' is. Met andere woorden: verondersteld wordt dat de claim 'waar' is (kaart 1.1).



Kaart 1.2 Het areaal wonen in 1995 (bron: RIVM/TNO)

Opzet rapportage

Dit rapport legt sterk de nadruk op de gebruikte methodiek hoewel natuurlijk de resultaten van de analyses relevant zijn voor het onderzoek van RIVM en TNO Inro. Toch geldt dat de uitkomsten pas relevantie hebben in het licht van de gemaakte keuzes ten aanzien van de variabelen, het regressiemodel en de schattingstechniek. Het zijn immers deze keuzes die noodzakelijkerwijze leiden tot deze uitkomsten.

Hoofdstuk 2 gaat in op de gemaakte keuzes en gaat daarom expliciet in op de beantwoording van de onderzoeksvragen 1 tot en met 3. Het gaat in op de selectie en bewerkingen van de gegevens, de argumentatie voor het gebruik van het logistische regressiemodel in het licht van de toepassing met de Ruimtescanner. In hoofdstuk 3 worden de resultaten van de analyse gepresenteerd en wordt de laatste onderzoeksvraag beantwoord. Hoofdstuk 4 is een concluderend hoofdstuk. Daarnaast kent dit rapport een groot aantal bijlagen die gebruikt

kunnen worden als naslagwerk en omwille van de leesbaarheid van het rapport niet in de hoofdtekst zijn opgenomen.

Modellen & Data

Inleiding

In dit hoofdstuk gaan we in op de gehanteerde schattingsmethodiek voor de allocatie van nieuwe woningen, de gebruikte data en de op de VU gevolgde werkwijze om de modelparameters te bepalen. Er wordt kort ingegaan op de Ruimtescanner. De reden is dat de gevolgde schattingsmethodiek met behulp van ruimtelijke gegevens op het ruimtelijke schaalniveau van 500 bij 500 meter en het gebruikte logistische regressiemodel identiek aan de allocatiemodule van de Ruimtescanner. De keuzes die in dit hoofdstuk worden beschreven bepalen dus hoe de ruimtelijke ontwikkeling van de woningvoorraad 1980-1995 zal worden geïnterpreteerd en hebben dus gevolgens voor het toekomstbeeld voor 2020.

Globale werking Model Ruimtescanner

Het mede door de VU ontwikkelde model Ruimtescanner (Schotten et al., 1997) is in dit project ingezet om de mogelijke verdeling van het ruimtegebruik van wonen en werken in Nederland tot 2020 in beeld te brengen.

Een belangrijke doelstelling bij de ontwikkeling van het model Ruimtescanner was de mogelijkheid om op eenvoudige en snelle wijze ruimtelijke scenario's te kunnen vertalen naar kaarten met mogelijk toekomstig ruimtegebruik. Het model is in principe ook geschikt om de feitelijke ruimtelijke ontwikkelingen uit het verleden door te trekken naar de toekomst en daarmee een referentie-beeld te creëren voor alternatieve scenario's. Dit onderzoek staat daarom ten dienste om een wetenschappelijke onderbouwing te geven van de toekomstige ruimtelijke spreiding van het wonen ten behoeve van het onderzoek van RIVM en TNO Inro. Uitgangspunten bij de simulaties van het toekomstig grondgebruik zijn het huidige grondgebruik en de bekende ruimteclaims voor een bepaald zichtjaar (zie hoofdstuk 1). De claims worden aan een bepaalde locatie toegekend door de attractiviteit (aantrekkelijkheid) van een bepaalde grondgebruiksklasse te operationaliseren aan de hand van rekenregels waarin beleid (ruimtelijke ordening, natuur, etc.), fysische en economische eigenschappen en de ruimtelijke relaties tussen verschillende soorten grondgebruik zijn opgenomen. De attractiviteit van een bepaalde grondgebruiksklasse wordt weergegeven in zogenaamde 'attractiviteitskaarten'. Deze kaarten zijn landsdekkend. In dit project zijn de attractiviteitskaarten samengesteld aan de hand van de rekenregels die volgen uit de trendanalyse wonen 1980-1995.

De uiteindelijke toedeling van de verschillende grondgebruiksklassen vindt plaats op grond van een allocatie-algoritme waarin de kans op toedeling niet alleen afhangt van de attractiviteit van een bepaalde cel maar ook van de grootte van de ruimtevrage per grondgebruiksklasse. Hierop komen later in dit hoofdstuk terug. Resultaat zijn kaarten waarop het grondgebruik in een 'zichtjaar' is weergegeven op een schaalniveau van 500 bij 500 meter.

Schattingsmethodiek

In deze paragraaf wordt eerst ingegaan op de keuze van het soort analysemodel. Vervolgens wordt de aard van de schattingsmethodiek behandeld, die aansluit op de methodiek van de Ruimtescanner. Daarna wordt kort ingegaan hoe de invloed van de verklarende factoren op de te verklaren variabele, de ruimtelijke spreiding van de toename van de woningdichtheid 1980-1995, kan worden geschat. Deze twee paragrafen zijn technisch van aard. Tenslotte wordt nogmaals minder technisch ingegaan op de aard van de logistische regressiemodel, dat een variant is van het multiple regressiemodel, de interpretatie van de parameters, de waarde die aan deze parameters vanuit een statisch optiek moet worden gegeven, de wijze waarop de beste regressievergelijking kan worden gevonden en hoe de set van verklarende variabelen kan worden gereduceerd alvorens met de regressieanalyse te starten.

Theoretische en pragmatische onderbouwing voor een ruimtelijk analytisch model om de attractiviteitskaarten te bepalen

Om de ruimtelijke attractiviteiten van wonen te schatten zijn in de voorbereidende fase twee mogelijke oplossingen aangedragen: een gedragsmatig onderbouwde methode op basis van woningmarktonderzoeken (zie bijvoorbeeld Woonverkenningen 2030, DGVH, 1997 of Goetgeluk, 1997), of een ruimtelijk statistische benadering waarin wordt geprobeerd de toename van de woningdichtheid 1980-1995 zo goed mogelijk verklarend te reproduceren. In overleg met de begeleidingsgroep van dit onderzoek en de Rijksplanologische Dienst, is gekozen voor de tweede methode van ruimtelijk statistische trendanalyse. Daar er geen analyse van woonmilieus mogelijk is geweest vanwege het ontbreken van geschikte data, is besloten de analyses uit te voeren op grond van de verschillende woningtypen.

Inhoudelijk is bovengenoemde pragmatische keuze voor het het onderzoek van RIVM EN TNO Inro ook goed verdedigbaar. Indien autonome ontwikkelingen worden opgevat als het doortrekken van de huidige ontwikkelingen dan is de onderliggende assumptie dat voorkeuren en restricties (beleid) constant blijven in de tijd.

Dit betekent dat een vooruitberekening van de toekomstige spreiding van wonen met behulp van de Ruimtescanner niet zozeer hoeven gebaseerd te zijn op de huidige samenstelling van de attractiviteitskaarten (beleid, consumentenvoorkeuren, fysieke geschiktheden), maar op een attractiviteitskaart die louter en alleen de spreiding van één grondgebruikstype alsmede de samenhang tussen de spreiding van verschillende grondgebruikstypen weergeeft.

De onderliggende verklaring voor deze aanwezige statistische samenhangen hoeven we dan niet gedragsmatig te onderzoeken, maar kunnen we kwalitatief natuurlijk wel duiden op grond van bestaand onderzoek. Hierbij kunnen we denken aan de spreiding van wonen in relatie tot respectievelijk de infrastructuur en werklocaties.

Een methode die hiervoor een uitkomst kan brengen is de ruimtelijke autocorrelatie. In plaats van een analyse van de ruimtelijke autocorrelatie voor 1995 – dus de reproductie van de kaart 1995- is het zinvoller om de ontwikkeling in de tijd op deze manier te modelleren. Afhankelijk van de beschikbaarheid van de gegevens kan voor een aantal intervallen (bijvoorbeeld 1980-1995) de ruimtelijke spreiding van wonen en de samenhang van die spreiding met andere

grondgebruiksklassen worden gemodelleerd. Op deze wijze kan rekening worden gehouden met de dynamiek in de tijd.

De vraag is nu welke periode gekozen moet worden om die dynamiek te schatten. Daarvoor gelden pragmatische en inhoudelijke argumenten. Een fors deel van de woningvoorraad in 1995 staat daar al heel lang. Willen we een verklaring geven van de veranderingen in de woningvoorraad in het licht van de doelstelling van het RIVM en TNO Inro onderzoek, dan is het gewenst de analyse te richten op de nieuwbouw in een recente periode (bijvoorbeeld 1980-1995). De te verklaren variabele is dan niet langer de totale oppervlakte per cel besteed aan wonen op een bepaald moment, maar de groei in de oppervlakte per cel besteedt aan wonen in de afgelopen periode. Deze dient dan te worden verklaard op basis van de kenmerken van de cel aan het begin van de periode.

De keuze voor de periode 1980-1995 valt goed te verdedigen in het licht van de onderzoeksopdracht van RIVM en TNO Inro: het doortrekken van het huidige compacte stadbeleid aangevuld met autonome ontwikkelingen in de ruimtelijke spreiding van wonen van 1995 tot 2020. Met autonomie doelen we dan op de vaststaande verhoudingen tussen burger en overheid. Deze periode is redelijk turbulent geweest. De woningmarkt kende een periode van stagnatie (instorting woningmarkt), herstel en grote bloei. Ook is het een periode waarin beleidsmatig de ruimtelijke ordening van het wonen veranderde van gebundelde deconcentratie naar het compacte stad idee. Dit concept wordt in het onderzoek van RIVM en TNO Inro doorgetrokken tot 2020. Ook is het de periode waarin de marktwerking op de woningmarkt meer en meer gestalte krijgt en verwoord is in de Nota Heerma (1998) die binnenkort wordt geëvalueerd. Deze periode kenmerkt zich daarom ook door een toenemende vrijheid van de woonconsument met name in die gebieden waar de druk op de woningmarkt beperkt is (zie o.a. Goetgeluk 1997) en relatief constante consumentenpreferenties (Bootsma 1998). Bootsma toont aan dat van een reurbanisatieproces op basis van preferenties geen enkele sprake is. Het compacte stadsbeleid is dus woonconsumentenvreemd.

Gegeven de noodzaak de dynamiek in de tijd te meten, uitgedrukt in de toename van de woningdichtheid, kan met de ruimtelijke autocorrelatie wordt aangegeven in welke mate er samenhang bestaat tussen de omvang van het grondgebruik in een zone en in omliggende zones. In de literatuur zijn coëfficiënten ontwikkeld om het niveau van ruimtelijke autocorrelatie te meten. Een bekend voorbeeld is de Moran coëfficiënt. Deze coëfficiënt heeft vooral een beschrijvende betekenis. Interessanter is het om na te gaan of er in de loop van de tijd veranderingen optreden in de onderlinge verhoudingen tussen soorten grondgebruik. Daarnaast is het mogelijk om ruimtelijke autocorrelatie een plaats te geven in een verklarende analyse. Het gaat er dan bijvoorbeeld om dat de omvang van een bepaald type grondgebruik in een zone niet alleen afhangt van de kenmerken van de zone zelf, maar ook van de vraag in hoeverre dit grondgebruik in omliggende zones voorkomt. Dit kan bijvoorbeeld duiden op agglomeratie-voordelen. Ook is het mogelijk om rekening te houden met positieve (of negatieve) ruimtelijke correlaties ten aanzien van andere soorten grondgebruik. De eenvoudigste manier om dit te doen is door de verklarende variabelen (X) voor de toename van de dichtheid van het wonen (Y) de score op de verklarende variabelen van de omliggende gridcellen mee te nemen. Inhoudelijk klopt dit natuurlijk omdat de attractiviteit van een bouwlocatie ook door zijn omgeving wordt bepaald.

Dit is reeds mogelijk in de huidige versie van de Ruimtescanner. De attractiviteit van een lokatie voor wonen hangt van een gedesaggreerd beperkt ruimtelijk interactiemodel

(logistisch) van de omvang van woonactiviteiten in de omgeving. De wijze waarop dit modelmatig kan worden bepaald wordt in de volgende paragraaf behandeld.

De achtergrond van de gebruikte schattingsmethodiek

Bij de schattingstechniek wordt aangesloten bij het grondgebruiksmodel Ruimtescanner (Schotten et al., 1997) zoals dat in paragraaf 2.3 beschreven is. Het model analyseert de allocatie van grond tussen verschillende grondgebruikscategorieën $j=1, \dots, J$ in cellen $c=1, \dots, C$. De attractiviteit van cel c voor grondgebruikscategorie j wordt gerepresenteerd door s_{cj} . Hierbij is s_{cj} een functie van een aantal factoren $k=1, \dots, K$ die tezamen bepalen hoe de attractiviteit is van cel c voor grondgebruik j . Uitgaande van een lineair model gebaseerd op kenmerken x_{ck} komen we uit op:

$$s_{cj} = \beta_{1j}x_{c1} + \dots + \beta_{Kj}x_{cK}, \quad \text{voor alle } c \text{ en } j$$

In s_{cj} zijn alle relevante geschiktheidsbepalende factoren meegenomen op één na: de prijs. De prijs van grond in cel c wordt gerepresenteerd als p_c . Het netto nut van grondgebruik j in cel c is het verschil tussen de geschiktheid en de prijs: $s_{cj} - \alpha_j p_c$:

$$v_{cj} = \beta_{1j}x_{c1} + \dots + \beta_{Kj}x_{cK} - \alpha_j p_c, \quad \text{voor alle } c \text{ en } j$$

waarbij α_j de omreken factor weergeeft tussen geschiktheid voor gebruik j en geld. Hieronder gaan we dieper in op de schatting van de parameters α_j en β_{cj} .

We beschouwen nu een actor (bijvoorbeeld 'gezinnen') die de grond in alle cellen vergelijkt naar hun aantrekkelijkheid (netto nut v_{cj}) als woonlocatie. Deze aantrekkelijkheid hangt steeds af van het verschil tussen geschiktheid en prijs $s_{cj} - \alpha_j p_c$. Hoe hoger deze waarde is hoe hoger de kans dat gezinnen hun activiteit j (wonen) in cel c zullen lokaliseren. Volgens het standaard logit model is de kans q_{cj} dat gezinnen bij activiteit j kiezen voor cel c gelijk aan:

$$q_{cj} = \exp(s_{cj} - \alpha_j p_c) / [\exp(s_{1j} - \alpha_j p_1) + \dots + \exp(s_{cj} - \alpha_j p_c)] \quad \text{voor alle } c \text{ en } j$$

Deze vergelijking heeft nog een tekortkoming; er wordt namelijk impliciet vanuit gegaan dat alle cellen even groot zijn. In werkelijkheid is dat niet zo: sommige cellen zijn kleiner (bijvoorbeeld als ze aan de grens liggen) en daarnaast is het mogelijk dat sommige zones door beleidsmaatregelen (zonerings) slechts voor een deel of helemaal niet in aanmerking komen voor grondgebruikstype j . Er geldt ook een ander aspect. De totale ruimte in een gridcel is ook niet geheel beschikbaar omdat de bestaande bebouwde ruimte niet eenvoudig kan worden ontkend. De ruimte per cel die in de analyse wordt betrokken is dan immers gelijk aan de totale ruimte L_c minus de ruimte R_c die aan het begin van de periode al aan wonen wordt besteed. Overigens kan R_c/L_c wel een rol spelen als verklarende variabele van de verandering van het grondgebruik in een cel. De mate waarin het niet-woon gedeelte van een cel zal worden ingevuld voor wonen kan immers afhangen van de omvang van het woongedeelte in een cel in de uitgangssituatie. Zij L_c de omvang van cel c . Dan kan als volgt expliciet rekening worden gehouden met verschillen in de omvang van cellen:

$$q_{cj} = L_c \cdot \exp(s_{cj} - \alpha_j p_c) / [L_1 \cdot \exp(s_{1j} - \alpha_j p_1) + \dots + L_c \cdot \exp(s_{cj} - \alpha_j p_c)] \quad \text{voor alle } c \text{ en } j$$

Duidelijk is dat beide vorige formules identiek zijn als de omvang van de grond in alle cellen even groot is.

Neem nu aan dat voor de diverse grondgebruikstypen de totale vraag gegeven is (weergegeven als Q_j^D). Dit houdt in dat de vraag exogeen is en op geaggregeerd niveau inelastisch is. We gaan er hier vanuit dat de geaggregeerde vraag gelijk is aan de het totale aanbod van grond in het land. De verwachte hoeveelheid grond die gevraagd wordt voor type j in cel c , Q_{cj} , is:

$$Q_{cj} = \{ [L_c \cdot \exp(s_{cj} - \alpha_j p_c) / [L_1 \cdot \exp(s_{1j} - \alpha_j p_1) + \dots + L_c \cdot \exp(s_{cj} - \alpha_j p_c)] \} \cdot Q_j^D \quad \text{voor alle } c \text{ en } j$$

Hiermee is duidelijk dat de vraag naar grond in cel c voor gebruiksdoel j een dalende functie is van de grondprijs in de eigen cel (via de teller) en een stijgende functie van de prijs in alle andere cellen (via de noemer). Het logit model stelt ons in staat om de parameters van de laatste vergelijking te schatten.

Duidelijk is dat de te schatten gewichten voor de waardering van attractiviteit voor de activiteit wonen in het verleden een combinatie is van zowel vraagelementen, als van overheidsbeleid. Het zou overigens interessant zijn om na te gaan in hoeverre schatting van de parameters β_{kj} in de loop van de tijd leidt tot systematische trends die eventueel zouden kunnen worden doorgetrokken bij het gebruik van de Ruimtescanner voor de schatting van toekomstig grondgebruik.

In bovenstaande is aangegeven dat de prijs een belangrijke factor is om concurrentie tussen grondgebruikstypen te simuleren zoals de concurrentie tussen wonen, werken en infrastructuur te simuleren. In het onderzoek van RIVM en TNO Inro is dit evenwel niet van belang. Aangenomen wordt namelijk dat infrastructuur, natuur (EHS) en de woonlocaties in Bestaande Stedelijk Gebied (BSG) en de VINEX-uitleg sowieso worden gerealiseerd. In het onderzoek wordt ook aangenomen dat op de werklocaties, zoals aangegeven op de plankaart Nederland in Plannen (RPD) alsmede huidige werklocaties, geen woningen mogen komen. De resterende woningclaim per COROP (de totale woningclaim per COROP minus het a priori gealloceerde deel van deze claim per COROP) zal ruimtelijk verspreid worden conform de hier gepresenteerde trendanalyse. Deze restclaim wordt gealloceerd in enerzijds Overige Bouw Locaties (bestemmingsplannen en dergelijke buiten de BSG en VINEX-Uitleg) en anderzijds het landelijke gebied (uitgezonderd BSG). Daarbij wordt aangenomen dat de prijs die de woningbouw kan betalen voor een hectare grond a priori hoger is dan voor andere grondgebruikstypen. Een probleem zou slechts kunnen opdoemen voor locaties die voor werkgelegenheid en voor wonen even toegankelijk zijn. Aangenomen wordt dat er voldoende werklocaties worden gevonden in woongebieden. Hier denke men aan de detailhandel en de zakelijke dienstverlening in de BSG.

Resumerend betekent dit dat de grondmarkt buiten spel staat. De attractiviteit is dus geen resultante van de prijs-kwaliteitsverhouding, maar louter en alleen van de kwaliteit. Deze vergaande veronderstelling kan slechts gelden voor het specifieke scenario in het onderzoek

van RIVM en TNO Inro. Voor verdere details verwijzen wij naar dit rapport (Goetgeluk et al. 2000).

Maximum Likelihood schattingsprocedure

De vraag is nu hoe de coëfficiënten β_{kj} uit vergelijking (2) kunnen worden geschat. In het kader van de huidige studie beperken we ons tot de coëfficiënten die betrekking hebben op het grondgebruik voor het wonen. Op basis van de waargenomen veranderingen in het grondgebruik gedurende de afgelopen periode kunnen de betreffende coëfficiënten als volgt worden geschat.

Het waargenomen aantal gebouwde woningen W_c in grid cel c kunnen we beschouwen als de uitkomst van een kansproces. Dit kansproces hangt af van de scores van de verklarende variabelen x_{c1}, \dots, x_{cK} en de nog onbekende bijpassende coëfficiënten β_{kj} . Gegeven het type woning kunnen er maximaal M woningen per ha gebouwd worden. Neem aan dat er in een cel L_c ha beschikbaar is voor woningen dan zouden er in het uiterste geval in cel c $N_c = M \cdot L_c$ woningen kunnen worden gebouwd. Het werkelijke aantal gebouwde woningen W_c van het betreffende type is doorgaans uiteraard kleiner dan N_c . We beschouwen nu de kans op de realisatie van W_c woningen als de uitkomst van een binomiaal kansproces met N_c trekkingen met een kans P_c op realisatie van de woonfunctie. Deze kans P_c is een functie van de geschiktheid v_{cj} van cel c voor het wonen, die zoals aangegeven in (2) ondermeer afhangt van de te schatten coëfficiënten β_{kj} . Voor de functionele vorm maken we daarbij gebruik van de basisvorm van het logit model:

$$P_c = e^{(v_{cj})} / (1 + e^{(v_{cj})})$$

Duidelijk is dat bij een zeer hoge geschiktheid v_{cj} de kans op realisatie bijna gelijk zal worden aan 1, wat volgens het kansmodel zou leiden tot een zeer hoog aandeel van de beschikbare grond in cel c die de bestemming wonen krijgt. De kans op een realisatie van W_c woningen in cel c is nu volgens het binomiale model:

$$L_c(W_c, v_{cj}) = [W_c! (N_c - W_c)! / N_c!] \cdot [P_c^{W_c}] \cdot [1 - P_c]^{[N_c - W_c]}$$

De tweede en de derde factor zijn afhankelijk van de te schatten coëfficiënten. Dat geldt niet voor de eerste factor: deze hangt niet af van de te schatten coëfficiënten. Deze factor geeft het aantal manieren aan waarop W_c woningen op N_c percelen kunnen worden gebouwd. De waargenomen realisaties per cel stellen ons nu in staat om met de behulp van de maximum likelihood methode de waarden van de coëfficiënten β_{kj} te schatten. Voor details verwijzen we naar bijlage 1.

Logistische regressiemodel en correlatie-analyse

Uit de vorige paragraaf en uit bijlage 1 blijkt dat gebruik wordt gemaakt van een regressiemodel die de kans (P) weergeeft op een gebeurtenis. In dit onderzoek is die gebeurtenis de toename van de woningdichtheid 1980-1995 in een bepaalde gridcel als gevolg van de invloed van de verklarende variabelen (X). Aangetoond kan worden dat een multiple regressiemodel voor het doel van dit onderzoek niet geschikt is (Wrigley 1985 pp. 21-62) en dat een logistische regressiemodel het alternatief is (Wrigley 1985, pp. 28-29).

Een logistische regressiemodel wordt gebruikt om de relatie tussen het studieonderwerp en meerdere verklarende variabelen te onderzoeken. Met name als de meetschalen (bijvoorbeeld nominaal versus interval of ratio) van de verklarende en de te verklaren variabelen verschillend zijn, is het logistische regressiemodel goed toepasbaar. Ook sluit de logistische regressie goed aan bij de manier waarop Ruimtescanner grondgebruik toedeelt, namelijk in termen van kans op voorkomen tussen 0 en 1. De relatie in een logistische regressie-analyse wordt als volgt uitgedrukt:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + e$$

Hierin is Y de bestudeerde (afhankelijke) variabele en X de verklarende (onafhankelijke) variabele. β_0 staat voor de intercept en β_n stelt de berekende parameter voor van de variabele X_n . De intercept staat voor de gemiddelde waarde (geometrisch gemiddelde) van Y indien geen rekening wordt gehouden met de verklarende variabelen. De beta's geven de richting (negatief en positief) en de sterkte van het verband tussen X en Y aan. De letter e staat voor de random verdeelde fout in deze relatie.

De Y in bovenstaande vergelijking is identiek aan de V_{cj} in de vorige paragraaf. In de vorige paragraaf hebben we reeds gezien dat de V_{cj} wordt ingepast in de vergelijking om P_c te bepalen. De parameters (beta's die de invloed van de verklarende variabelen kwantitief weergeven) van dit model zijn niet zo eenvoudig te interpreteren als in een gewone multiple regressiemodel. De parameters worden namelijk ingevuld in een exponentiële kansfunctie $P = e^{f(x)} / (1 + e^{f(x)})$, waarbij $f(x)$ de regressievergelijking is, P de kans is en e staat voor een e-machtfunctie ($e \sim 2,7$). Indien de regressievergelijking $f(x)$ de waarde nul heeft wordt de kans op een toename ($e^0 / (1 + e^0)$), hetgeen identiek is ($1 / (1 + 1)$) ofwel 0,50. Indien de functie de waarde 1 heeft dan wordt de kans 0,73 ($(e^1 / (1 + e^1))$ ofwel ($2,7 / (1 + 2,7)$)). Is de waarde van de functie -1 dan wordt de kans 0,27. Bovenstaande drie voorbeelden maken inzichtelijk hoe de beta's moeten worden behandeld.

De vraag die nu gesteld wordt, is welke variabelen een zinvolle bijdrage leveren aan de verklaring in de toename van het aantal woningen tussen 1980-1995. Bij de uitvoering van de regressie-analyse worden de parameters (β s) van de verschillende variabelen berekend. De parameters bepalen het relatieve gewicht per x-variabele in de distributie van de te verklaren y-variabele. In het algemeen kan gesteld worden dat in een efficiënt ruimtelijk model met een zo klein mogelijk aantal bepalende variabelen toch een voldoende precieze voorspelling kan worden gedaan (Yue-Hong Chou, 1997).

De volgende belangrijke stap is daarom te bepalen welke van de variabelen in het regressie-model een significante bijdrage hebben aan de verklaring van de y-variabele. Onder deze

bijdragereg wordt verstaan een bijdrage die niet afwijkt van nul. Dit wordt gedaan aan de hand van significantie testen van de berekende parameters (β s). De gebruikte significantie-test is de t-test van statistische significantie. De t-waarde is als volgt gedefinieerd:

$$t = \beta / S\beta$$

De $S\beta$ is de standaard fout ofwel de spreiding (standardevatie) van waarden om de beta, die met de ML is geschat. Met andere woorden: een beta met een zeer geringe spreiding een 'betere' beta dan met een grote spreiding. In de gebruikte logit-regressie in S-plus wordt de t-waarde automatisch berekend, die vervolgens vergeleken wordt met de waarde in de t-tabel met een significantie-niveau van 0,05 en een oneindig aantal vrijheidsgraden. Zolang de bepaalde t-waarde hoger is dan de t-waarde uit de tabel kan gesteld worden dat de betreffende beta afwijkt van nul.

Wat betekent zo'n significantieniveau van 0,05? We stellen als onderzoeker dat bij een oneindig aantal schattingen we vijf van de honderd keer onterecht beweren dat de beta van nul afwijkt. Deze fout wordt een fout van eerste orde ofwel een alpha-fout genoemd. Deze fout kan namelijk exact worden bepaald op basis van de kanstheorie. Indien we geen relatie verwachten op tussen Y en X zal de waarde van de beta van de X eenvoudig het resultaat zijn van toevaltreffers. Indien de waarde van de beta echter groter is dan de toevalstrefferwaarde dan toetsen we statistisch binnen welke marge deze afwijking 'echt' afwijkt van nul. En dat is de fout op de eerste orde ofwel de alpha-fout.

Hieruit volgt tevens op welke wijze bepaald kan worden welke verklarende variabelen nu daadwerkelijk invloed hebben op de toename van de dichtheid van het wonen (Y). Indien we deze variabelen uit de regressievergelijking halen, kunnen we toetsen of de geschatte toename van de dichtheid van het wonen significant afwijkt van de waargenomen toename van de dichtheid. De toets is nu of het verschil tussen schatting en waarneming significant afwijkt van nul. Indien dit niet zo is dan betekent dit dat de gereduceerde regressievergelijking uitstekend in staat is om de werkelijkheid te benaderen. Dit is ook het doel van de analyse: een zo simpel mogelijke regressievergelijking.

Uit bijlage 4 blijkt dat dit reductieproces niet in één stap kan: het weglaten van één variabele kan de invloed van de andere variabelen doen veranderen. Dit is mogelijk als er samenhang bestaat tussen de verklarende variabelen. Hypothetisch kunnen de variabelen geheel met elkaar samenhangen, zowel negatief als positief, geheel niet en een beetje. Het is evident dat variabelen die geheel met elkaar overeenkomen niet beiden in de regressievergelijking hoeven voor te komen. De vraag is hoe die samenhang kan worden achterhaald om tot een reductie van de variabelenset te komen. Het antwoord is een correlatie analyse.

Met de correlatie analyse wordt de grootte van het onderlinge verband tussen verschillende variabelen gemeten. Dit wordt uitgedrukt in de correlatie-coëfficiënt r, die een waarde tussen -1 en +1 kan aannemen. Deze correlatiecoëfficiënt geeft grofweg de hoeveelheid spreiding aan rondom de lineaire 'least square' vergelijking. Als alle x, y punten van twee tegen elkaar uitgezette variabelen precies op één lijn liggen dan is r gelijk aan 1 of -1, afhankelijk of de relatie positief of negatief is. Overigens wil een r waarde van nul nog niet zeggen dat er geen relatie is tussen twee variabelen. De mogelijkheid bestaat dat er een niet-lineair verband tussen de twee variabelen bestaat. Dit is het best te controleren door een 'scatterdiagram' (puntenwolfdiagram) van de variabelen te maken, waarmee ook ver weg liggende waarden goed geïdentificeerd kunnen worden. Verder kunnen de r-waarden tussen verschillende

variabelen alleen maar globaal vergeleken worden omdat de correlatie met weinig voorkomende variabelen.

Conclusie

Het is gegeven de aannames van het onderzoek van RIVM en TNO Inrois het mogelijk om te kiezen voor een ruimtelijk-analytisch verhaal in plaats van een gedragsmatig onderbouwing, waarin woonconsumentenpreferenties worden afgezet tegen door onder andere door het beleid gecreëerde mogelijkheden en restricties. Een logistisch regressiemodel kan worden geschat. Dit regressiemodel blijkt bovendien geheel overeen te komen met het ruimtelijk allocatie algoritme van de Ruimtescanner. Dit laatste model wordt gebruikt om de toename van de woningvoorraad 1995-2020 op COROP-niveau op het ruimtelijke schaalniveau van 500 bij 500 meter te alloceren. De volgende vragen die nu beantwoord moeten worden is wat de verklaren en wat de verklarende variabelen zijn in de analyse.

De variabelen

In deze paragraaf wordt eerst ingegaan op de keuze en de bewerkingen van de te verklaren variabele. Vervolgens wordt ingegaan op de selectie en de bewerkingen van de verklarende variabelen.

De te verklaren variabele (Y)

Voor de jaren 1980 en 1995 zijn landsdekkende bestanden (PAP) gebruikt met voor iedere gridcel van 500 x 500 meter het aantal woningen van een bepaald type. De volgende woningtypen zijn onderscheiden:

- Woningtype 1 = ééngezins, vrijstaand/2-kap/boerderij
- Woningtype 2 = ééngezins, rijtjes
- Woningtype 3 = meergezins, etage/beneden/bovenwoning
- Woningtype 4 = meergezins, flat
- Woningtype 5 = BAR (onzelfstandige woonruimten) en overig

Deze woningtypes zijn evenwel samengevoegd tot een- en meergezinswoningen omdat in het onderzoek van RIVM en TNO Inro de toename van de woningvoorraad 1995-2020 per COROP-gebied slechts bestaat uit deze twee klassen. Opgemerkt zij dat hierdoor a priori heterogeniteit ontstaat. Intuïtief is immers duidelijk dat vrijstaande woningen en tweekappers vaak op andere locaties dan rijtjeswoningen worden gevonden. Idem dito geldt dit voor de ruimtelijke spreiding van flatwoningen en boven- en benedenwoningen.

Tevens hebben we ten behoeve van de regressie-analyses een kaart met het aantal hectares woongebied per gridcel in 1981 gebruikt (Bron: CBS, Bodemstatistiek 1981). Anders gesteld, er kan dus bepaald worden wat de woondichtheden naar type per 500 bij 500 meter gridcel zijn.

De volgende bewerkingen zijn in S-plus gemaakt. In S-plus zijn derhalve twee kolommen toegevoegd, dywt12 en dywt345 waarin de toename van respectievelijk het aantal ééngezins en het aantal meergezinswoningen in de periode 1980-1995 is weergegeven.

dywt12 = d80.95wt1 + d80.95wt2 # toename aantal ééngezinswoningen 1980-1995
 dywt345 = d80.95wt3 + d80.95wt4 + d80.95wt5 # toename aantal meergezinswoningen 1980-1995

De volgende stap is de maximale toename van het aantal woningen per cel3, te relateren aan de beschikbare ruimte voor woningbouw. Het voor woningbouw beschikbare areaal wordt berekend door de maximale celoppervlakte (25 hectare) te verminderen met het bestaande woningareaal en het waterareaal die beiden als onveranderlijk worden beschouwd. Hier aan voorafgaand zijn eerst zowel het woonareaal 1981 en het waterareaal 1981 (CBS, 1981) als kolommen aan de hoofdtabel van S-plus toegevoegd.

wywt12.h = (1280 / 25) * (25.001-WOONAREA81-WATERAREA8)
 # maximaal aantal ééngezinswoningen per hectare beschikbaar woonareaal

wywt345.h = (3846 / 25) * (25.001-WOONAREA81-WATERAREA8)
 # maximaal aantal meergezinswoningen per hectare beschikbaar woonareaal

In de berekening van het voor woningbouw beschikbare woonareaal is 25.001 hectaren genomen in plaats van 25 hectaren⁴. De reden hiervoor is dat als woonarea81 + waterarea8 samen 25 hectare is, het voor de woningbouw beschikbare areaal dus nul is. Deze uitkomst is ongewenst omdat in de volgende stap de woningtoename door het beschikbare areaal (25 min woonoppervlak 1981 min wateroppervlak 1981) gedeeld wordt en delen door nul een 'missing value' oplevert. Deze kleine afwijking wordt acceptabel geacht omdat de berekening van het voor woningbouw beschikbare areaal al een grove benadering van de werkelijkheid is. Men zou er bijvoorbeeld ook van uit kunnen gaan, dat het areaal dat in beslag wordt genomen door infrastructuur zoals spoor- en snelwegen niet beschikbaar komt voor woningbouw.

Er zijn evenwel enkele extra condities. De eerste conditie waar de y-variabele aan moet voldoen is dat het verschil van het aantal ééngezins- of meergezinswoningen tussen 1980 en 1995 positief moet zijn (alleen toenames), anders wordt de toename op nul gesteld. Dit wordt bereikt met een ifelse-constructie:

tusywt12.h = # tussenresultaat conditie1 (ééngezinswoningen)
 ifelse(dywt12>0,(dywt12/wywt12.h),0)

tusywt345.h = # tussenresultaat conditie1 (meergezinswoningen)
 ifelse(dywt345>0,(dywt345/wywt345.h),0)

We modelleren dus bewust de kans dat de beschikbare percelen bebouwd worden en dat het aantal woningen dus toeneemt. De kans dat woningen gesloopt worden en er dus afname

³ Maximum aantal ééngezinswoningen per cel in 1995: 1280 woningen.
 Maximum aantal meergezinswoningen per cel in 1995: 3846 woningen.

⁴ Een betere manier zou geweest zijn deze cellen eruit te filteren, waarmee echter wel de mogelijkheid tot het berekenen van de kans op de vergroting van de dichtheid van het aantal woningen wordt verwijderd.

plaatsvindt is het resultaat van andere, buiten beschouwing gelaten processen. Hiervoor kan als argument gelden dat in het onderzoek van RIVM en TNO Inro het huidige beleid wordt doorgetrokken waarin de toename van de dichtheid eerder gerealiseerd zal worden dan een afname.

De tweede conditie voor de y-variabele is dat de maximale toename van het aantal nieuwe woningen wordt begrensd op het maximale aantal woningen dat in 1995 in een cel is waargenomen⁵. Is de toename groter dan dit aantal, dan is kennelijk sprake van verdichting en wordt het resultaat op 1 gesteld. Dit wordt bereikt met de volgende ifelse-constructie:

ywt12.h = ifelse(tusywt12.h>1,1,tusywt12.h) # eindresultaat conditie 2 (ééngesins)

ywt345.h = ifelse(tusywt345.h>1,1,tusywt345.h) # eindresultaat conditie 2 (meergezins)

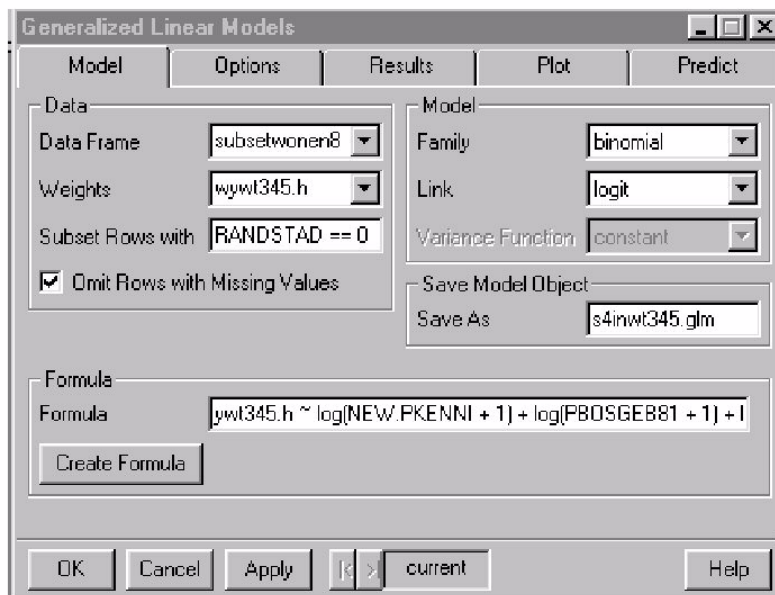
Omdat in de analyse het belang van gebieden met veel beschikbare woonruimte groter is dan bijna volgebouwde gebieden, wordt in de regressie-berekeningen aan de eerst genoemde gebieden een groter gewicht toegekend dan aan de reeds dichtbebouwde gebieden. Ter illustratie (tabel 2.1) een simpel rekenvoorbeeld met twee willekeurige celwaarden voor meergezinswoningen (met een maximum van 3846 woningen per cel in 1995):

Tabel 2.1 Een rekenvoorbeeld voor de bewerkingen van de te verklarenvariabele

A: dywt345 Toename aantal woningen 1980-1995 (Wc)	B: 25-wonen-water Beschikbaar woonareaal 1980 (Lc)	C: wywt345.h maximaal aantal woningen 1995 op beschikbare woonareaal (Nc)	D A/C
107	1	154	0,96
290	14	2154	0,13

D (A/C) is de te verklaren variabele en een geschikte gewichtsmaat is C (wywt345.h). Bovenstaande weging kan intuïtief duidelijk worden als men beseft dat de informatie dat 69% van het beschikbare areaal voor wonen wordt gebruikt, een grotere betekenis heeft wanneer het gaat om een areaal voor 100 woningen vergeleken met een areaal voor 1000 woningen. In het eerste geval is immers meer kans dat de score van 69% een toevalstreffer is dan in het tweede geval. Het gewicht wordt eenvoudigweg als kolomnaam in het invulscherm van de lineaire regressie-analyse ingevuld (zie figuur 2.1)

⁵ Hiervoor had overigens ook uitgegaan kunnen worden van de hoogste dichtheid woningen per cel, dus ten opzichte van het voor wonen beschikbare areaal (zonder water).



Figuur 2.1: invulscherm lineaire regressies S-Plus

De verklarende variabelen (X)

Hieronder volgt een lijst van variabelen waarvan op basis van wooncosumentenonderzoek en beschikbare gegevens verwacht mag worden dat ze van invloed zijn op de groei van de woningvoorraad. De in het model op te nemen verklarende variabelen voor de kans dat een cel voor het grondgebruikstype wonen wordt gebruikt betreffen:

- De mate waarin in omliggende cellen de functie wonen voorkomt. Het Nederlandse ruimtelijk beleid wordt immers mede geleid door de wens om tot een zekere ruimtelijke bundeling te komen van woonactiviteiten. Het ligt voor de hand dat de directe woonomgeving hier zwaarder telt dan de verder afgelegen woonomgeving. Verder spelen persoonlijke voorkeuren een belangrijke rol, mensen wonen graag in de buurt van andere mensen (voorzieningen, veiligheid, etc.). Concreet betekent dit dat de directe omgeving van belang zal zijn.
- De mate waarin in de (ruimere) omgeving van een cel werkgelegenheids-concentraties te vinden zijn. Een fors deel van de Nederlandse bevolking participeert immers op de arbeidsmarkt en zal dus bij de woonlocatiekeuze mede op de ligging ten opzichte van snelwegen letten. Zo werkt meer dan 50 % van de mensen op een reisafstand van minder dan 15 minuten (Ham, 1999). Nu ligt het voor de hand de ruimere omgeving van invloed zal zijn op de attractiviteit van een gridcel.

- De mate waarin de omgeving van een cel landschappelijk aantrekkelijk is (bijvoorbeeld natuurlijk grondgebruik). Ook nu ligt het voor de hand dat de ruimere omgeving een rol speelt.
- De ontsluiting van een cel in termen van de ligging ten opzichte van een oprit van een snelweg. Hier is de afstand de cel ten opzichte van de snelweg het meest voor de hand liggend.
- De ontsluiting van een cel in termen van de ligging ten opzichte van een NS station. Maatstaf: afstand tot het station.

In het hiernavolgende wordt dieper ingegaan per type variabele. Bijlage 2 toont statistieken.

Werkgelegenheid

De voor dit onderzoek beschikbaar gestelde gegevens van TNO Inro over arbeidsplaatsen hebben betrekking op het jaar 1983; per gemeente met een doorvertaling naar viercijferige postcodegebieden: onderscheiden naar 44 economische sectoren, van waaruit optellingen zijn gemaakt naar een negental bredere economische sectoren (afkomstig uit het OPERA-model van TNO Inro, Louter, 1997): landbouw, industrie, bouwnijverheid, transport, detailhandel, groothandel, zakelijke dienstverlening, overige dienstverlening, en overige sectoren (haven/luchthaven).

Een postcode gebied beslaat over het algemeen meerdere gridcellen. Na het verrasteren van de postcode-gebieden (polygoon) naar 500 x 500 meter gridcellen is het aantal arbeidsplaatsen per postcode-gebied gedeeld door het aantal gridcellen per postcode-gebied. De resultaten van de deling zijn naar beneden afgerond. Consequentie van deze procedure is dus dat het aantal arbeidsplaatsen binnen de postcode-gebieden homogeen over de gridcellen zijn verdeeld, terwijl dit in werkelijkheid meestal een heterogene verdeling zal zijn. Op het moment van dit onderzoek was er echter nog geen alternatief beschikbaar.

Infrastructuur

Gebruik is gemaakt van naar 500 meter cellen verrasterde vektorkaarten met respectievelijk de autosnelwegen van 1983 en de spoorwegen van 1980 (Goetgeluk et al. 2000). De wegen zijn in Arc Info via een automatische vektor-raster operatie verrasterd. Dat wil zeggen dat aan cellen de waarde één of nul is toegekend afhankelijk van het wel of niet voorkomen van de gedigitaliseerde vektor van de snel- of spoorweg in de betreffende cel. Het resultaat van deze procedure is een Boolean kaart (wél of géén voorkomen van de beschouwde grondgebruiksklasse).

Tevens hebben we gebruik gemaakt van naar 500 meter cellen verrasterde puntkaarten (Goetgeluk et al. 2000) met respectievelijk de op- en afritten van snelwegen van 1983 en de NS-stations van 1980. Ook hier is bij aanwezigheid van een op- en afrit of een NS-station de hele cel als zodanig geassocieerd. Het resultaat van deze procedure is eveneens een Boolean kaart (wél of géén voorkomen van de beschouwde grondgebruiksklasse).

Natuur

Onderscheiden zijn respectievelijk droge- en natte natuurterreinen en bosgebieden in 1981. Dit betreffen gridkaarten waarin per 500 meter cel de oppervlakte van deze gebieden is aangegeven (CBS, 1997).

Water

Gebruik is gemaakt van een kaart met overig water breder dan 6 meter. Hieronder vallen rivieren, kanalen, grachten, vaarten en sloten breder dan 6 meter, meren en plassen, slikken, havens en water in natte natuurlijke terreinen (CBS, 1997).

Recreatie

We hebben geen data gebruikt voor binnenstedelijk groen vanwege het bevolkingsvolgende karakter. Derhalve is alleen gebruik gemaakt van gridkaarten waarin per 500 meter cel de oppervlakte met bosrecreatie-gebieden is aangegeven (CBS, 1997). Hiertoe worden gerekend bosterreinen die door hun ligging te midden, of aan de rand, van de bebouwing nagenoeg dezelfde functie hebben als de meer gecultiveerde en veelal zorgvuldiger onderhouden parken. De bossen zijn voor een ieder vrij toegankelijk. Houtopbrengst is een nevenfunctie.

Regio-indelingen

Omdat we regionale verschillen verwachten in de ruimtelijke relaties tussen de groei van de voorraad woningen en de ruimtelijke factoren die deze groei beïnvloeden, hebben we kaarten gemaakt waarin drie regio's worden onderscheiden, de Randstad, de Intermediaire zone en de Perifere zone. De kaarten zijn gebaseerd op een naar 500 meter cellen verrasterde kaart van de 40 COROP-gebieden in Nederland (Vocht, 1992). De gebruikte regio-indeling is weergegeven in kaart 2.1

Kaart 2.1 Randstad, Intermediaire zone en de Periferie

Binnen de hierboven genoemde hoofdgroepen van variabelen heeft een verdere uitsplitsing plaatsgevonden. De variabele werkgelegenheid bijvoorbeeld is gebaseerd op een bestand met aantallen arbeidsplaatsen per sector per gridcel van 500 x 500 meter. Voor dit onderzoek is een indeling gebruikt in negen verschillende arbeidssectoren die elk hun eigen specifieke verband (kunnen) hebben met de toename van woningbouw. Aan de hand van een correlatie-analyse (zie hoofdstuk 3) hebben we beoordeeld welke van deze negen subvariabelen het sterkste verband heeft met de toename van woningbouw. Tevens hebben we gekeken naar de onderlinge correlatie van de subvariabelen.

In paragraaf 2.3.1 is aangegeven dat de attractiviteit van een gridcel voor wonen afhankelijk is van de omgeving. De vraag is wat onder omgeving moet worden verstaan. In de volgende paragraaf wordt hierop ingegaan.

De omgeving geoperationaliseerd

Het verband tussen nieuwe woningbouw en de aanwezigheid van bijvoorbeeld het aantal arbeidsplaatsen in de omgeving zou op het niveau van individuele cellen bepaald kunnen worden. De kans is echter groot dat de op deze wijze gevonden verbanden erg klein zijn omdat hierin geen rekening is gehouden met de mobiliteit van mensen en de bereidheid een bepaalde afstand (buiten de eigen cel) af te leggen tot bijvoorbeeld werk of recreatiemogelijkheden. Zo is intuïtief duidelijk dat een stadspark op 500 meter meer invloed heeft dan op 20 kilometer.

Een maat om de attractiviteit te laten samenhangen met de afstand is de potentiaalmaat. De potentiaal is dus een maat die de bereikbaarheid aangeeft, oftewel het potentiaal aan te bereiken banen, bos en dergelijke. Hierin wordt per cel aangegeven wat de potentiaalwaarde is ten opzichte van een bepaalde factor zoals het aantal arbeidsplaatsen binnen een bepaalde, qua reistijd acceptabele, afstand. Cellen met het aantal arbeidsplaatsen op een grotere afstand krijgen hierbij een lager gewicht mee dan cellen met arbeidsplaatsen op kortere afstand, die aantrekkelijker worden geacht. De wijze waarop de afstand invloed heeft op de aantrekkelijkheid van de arbeidsplaatsen in een bepaalde cel wordt vastgelegd in een bepaalde afstandsvervalfunctie. In dit onderzoek is gebruik gemaakt van reeds bekende uit de literatuur afkomstige afstandsvervalfuncties (Hilbers & Verroen, 1993). De afstanden worden in hemelsbrede richting gemeten (Euclidian distance). Hierbij is geen rekening gehouden met werkelijke reistijden, die afhankelijk zijn van het wegennetwerk, openbaar vervoer en verkeersdrukte. Hiervoor zijn gegevens op celbasis nodig die niet direct beschikbaar zijn. Verder wordt er van uitgegaan dat de bij de woonkeuze de afstanden tot werk, recreatie etc. ook op een globale manier bepaald worden. De potentiaal per cel wordt tenslotte bepaald door de door de afstandsvervalfunctie gecorrigeerde celwaarden in het omringende invloedsgebied bij elkaar op te tellen. Deze operaties zijn verricht met 'focal sum' functies in de gridmodule van Arc Info (zie hieronder).

Een belangrijke keuze is dus de afstandsvervalfunctie omdat deze immers de attractiviteit beïnvloedt. Er wordt gebruik gemaakt van de gevalideerde afstandsvervalfunctie van TNO Inro (Hilbers & Verroen, 1993). Dit is ook weer een exponentiële functie die vergelijkbaar is met die van de Ruimtescanner:

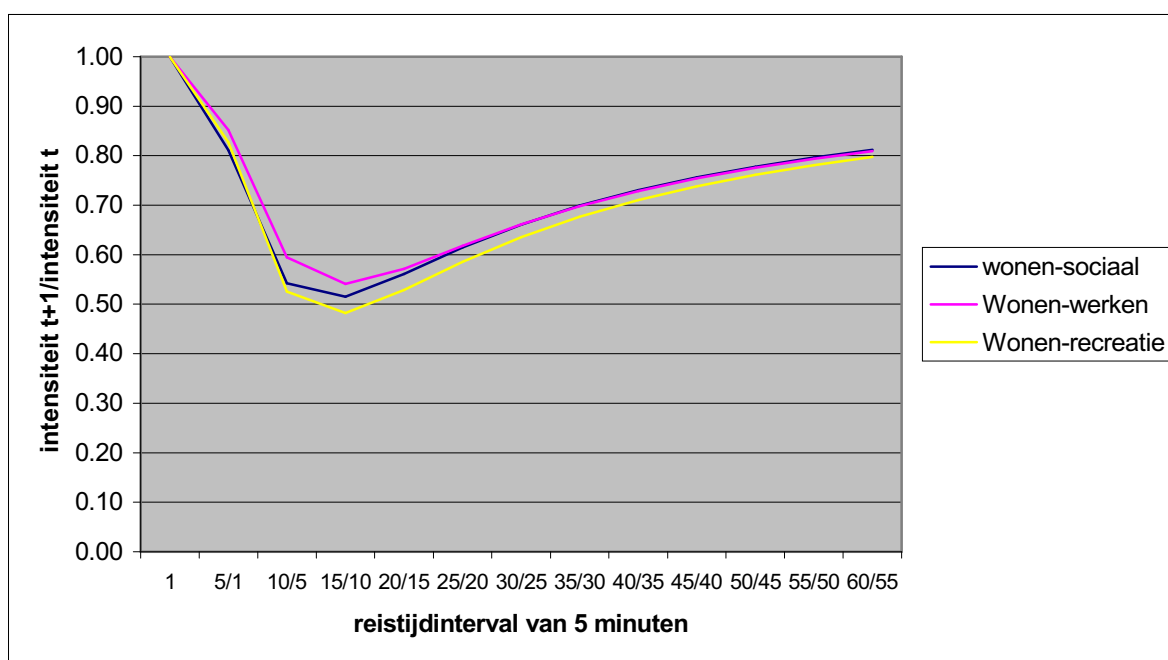
$$P(x)=1/(1+e^{-(a+b*\ln x)})$$

Hierbij staat de X voor minuten reistijd, de a de intercept en de b staat voor de beta.

De geselecteerde vervalfuncties zijn:

- Wonen-werken totaal (max 60 kilometer)
- Wonen-recreatie totaal (max. 15 kilometer)
- Wonen-sociaal totaal (max. 0,5 kilometer)

Voor wonen-wonen wordt geen vervalfunctie gebruikt, maar wordt direct de potentiaal berekend (zie hieronder). Figuur 3.2 toont voor bovengenoemde functies het afstandsverval door het verschil te nemen tussen twee reistijden. Hieruit wordt duidelijk dat na een aantal minuten reistijd (afstand gedeeld door een gemiddelde snelheid) mensen indifferent worden. Mensen zoeken dus een woning dicht bij hun baan en vice versa (van Ham, 1999). Echter,



zodra een drempelwaarde wordt overschreden, maakt het niet meer zo veel uit of het om 15 of om 30 minuten gaat.

Figuur 3.3 Afstandsvervalcurves voor drie vormen van ruimtelijke interactie

Op welke wijze worden nu de potentiaalkaarten gemaakt? De potentiaalkaarten worden in Arc Info gemaakt met de grid-functie 'focalsum'. Hiermee wordt voor elke cel in het grid de potentiaal berekend volgens een maximale invloedsafstand en een afstandsverval-functie die worden opgegeven in een zogenaamde 'Kernel'-ascii bestand, zoals hieronder wordt geïllustreerd:

Potentiaalgrid = focalsum(sourcegrid, irregular, kernelfile)

Op deze wijze worden voor alle negen sectoren werk, de vier verschillende vormen van recreatie en de twee onderscheiden woontypen aparte potentiaal grids gemaakt. Voor de

potentiaalkaart wonen wordt zoals gezegd geen afstandsverval-functie gebruikt. Hier wordt een 'focal sum' met een 'focalpoint' van 0 berekend waarin per cel de waarden van de acht omringende cellen (dus maximaal 0,5 kilometer) bij elkaar worden opgeteld en toegekend aan de desbetreffende cel. De potentiaalwaarden van de verschillende gridcellen vertegenwoordigen hiermee de waarden van de verklarende variabelen die vervolgens als input voor de statistische analyses worden gebruikt.

Van de kaarten met de op-en afritten van snelwegen en de NS-stations, zijn bereikbaarheidskaarten gemaakt. Deze kaarten geven per cel aan wat de afstand is tot de dichtst bijzijnde op-en afrit of NS-station. Om deze kaarten te maken zijn de 'boolean'-kaarten van de op- en afritten en de NS-stations eerst omgezet in kaarten waarbij de nul-waarde van alle cellen waarin geen op- en afrit of NS-station is gelegen is omgezet naar NODATA. Van de resulterende kaarten kunnen vervolgens de gewenste bereikbaarheids kaarten worden afgeleid door de 'Euclidean distance'⁶ van alle gridcellen tot de met data gevulde cellen te berekenen. We hebben met opzet gewerkt met de afstand tot de dichtst bijzijnde op-en afrit of NS-station en niet met de potentiaal (de som van het aantal op-en afritten of NS-stations binnen een bepaalde afstand). We gaan ervan uit dat de afstand tot de dichtst bijzijnde op- en afrit of NS-station het meest bepalend is, daar het over het algemeen gunstiger is één op- en afrit of NS-station vlak in de buurt te hebben (waarvandaan de andere knooppunten makkelijk bereikt kunnen worden) dan meerdere op- en afritten of NS-stations op grotere afstand.

Conclusie

De te verklaren en de verklarende variabelen zijn grotendeels beschikbaar voor het doel van dit onderzoek. In de bewerkingen zitten aannames die binnen het kader van het onderzoek van RIVM EN TNO Intro verdedigbaar zijn, maar in een andere invulling van het onderzoek te kort zouden schieten omdat deze aannames 'te star' zijn. Het moge duidelijk zijn dat de bewerkingen en de analyse arbeidsintensief zijn en vragen om een koppeling tussen enerzijds een geografisch informatie systeem en anderzijds een krachtig statistisch pakket. De vraag hoe deze koppeling kan worden gemaakt, is dan ook een expliciete onderzoeksvraag geweest. In onderstaande paragraaf wordt daar op ingegaan.

Gebruikte statistische software

Alle kaarten zijn door het RIVM geleverd op Arc Info gridformaat. Deze kaarten zijn op de VU omgezet naar Arc View formaat. Omdat Arc View gekoppeld kan worden aan het statistische pakket S-plus, is het in principe mogelijk een aantal ruimtelijk statische analyses uit te voeren. De functionaliteit van de S-Plus extensie van Arc View, was echter niet toereikend om de uiteindelijk gewenste statistische analyses uit te voeren. Daarom was het

⁶ Dit is de hemelsbrede afstand. In een grid wordt deze gemeten vanaf het middelpunt van de cel tot het middelpunt van de doelcel. Horizontaal en vertikaal betreft dit dus het aantal tussenliggende cellen maal de celgrootte. Diagonaal wordt de afstand tussen de middelpunten volgens de functie van Phytagoras uitgerekend.

noodzakelijk alle kaarten stuk voor stuk te koppelen aan een basisgrid⁷ waarin elke gridcel een uniek nummer krijgt. Doordat alle kaarten op deze wijze gerefereerd zijn aan dezelfde sleutelkolom met unieke gridcelwaarden, kunnen de kaarten in tabelvorm met elkaar vergeleken worden. Om de kaarten op de juiste wijze te koppelen is het van belang dat ook de zogenaamde VAT (value attribute tables) van de verschillende gridkaarten aan elkaar gekoppeld worden. De koppeling van de verschillende gridkaarten en het basisgrid is uitgevoerd met de functie 'Combine' die in Arc View (met Spatial analyst extension) beschikbaar is met de extensie 'Grid Conversion Tools'. Maximaal twintig grids kunnen tegelijk worden gecombineerd waarna een nieuw grid met unieke celwaarden wordt gemaakt. De resulterende VAT van dit grid is in zijn geheel geconverteerd naar een D-base tabel (DBF). Een DBF-bestand is eenvoudig in te lezen in S-Plus, door bij de data-import functie aan te geven welk DBF-bestand moet worden ingelezen. In dit scherm is een filter-tabblad aanwezig waarin kan worden opgegeven welke selectie in de records moet worden gemaakt en of er een steekproef moet worden genomen.. Om het aantal cellen in de berekening tot het minimum te beperken zijn bijvoorbeeld alle cellen die volledig uit water bestaan weggefilterd (WHERE WATER < 25: selecteer alle cellen met een water oppervlakte kleiner dan 25 hectare). De in S-plus geïmporteerde tabel is nu in principe geschikt om de gewenste correlatie en regressie-analyses uit te voeren. Tevens kunnen er berekeningen worden uitgevoerd tussen de kolommen onderling en kunnen de resultaten in nieuwe kolommen weggeschreven worden.

⁷ Als basisgrid is één van de door het RIVM geleverde grids genomen (droge natuurterreinen 1981, CBS 1997) waarin alle gridcellen in de Noordzee, de Zeeuwse wateren, het IJsselmeer en de Randmeren de waarde 'NODATA' hebben. Dit betreft derhalve een grid met 139620 duizend uniek genummerde cellen. Uitgangspunt hierbij is dat genoemde watergebieden niet relevant zijn als potentieel woongebied (indien ontwikkelingen uit het verleden worden doorgetrokken, lokaties in het water als IJburg worden hier dus van uitgezonderd). Het grid is zo klein mogelijk genomen om de grote hoeveelheid benodigde rekentijd van de regressie-analyses in S-plus zo veel mogelijk te beperken.

Trendanalyse

Inleiding

In de trendanalyse wordt geprobeerd zo goed mogelijk de ruimtelijke verdeling in de toename van de woningdichtheid 1980–1995) te verklaren. In dit hoofdstuk worden de in hoofdstuk 2 geselecteerde variabelen verder onderzocht op hun bijdrage aan de verklaring van de te verklaren variabele. Dit gebeurt eerst op een globale wijze door middel van correlatie-analyse, waarin de sterkte van het verband tussen elke afzonderlijke x-variabele met de y-variabele en de sterkte van de relaties tussen de verklarende variabelen worden bepaald. Als er op grond hiervan redenen zijn om te geloven dat bepaalde variabelen geen nuttige bijdrage leveren in het regressie-model, worden deze verwijderd. Ook kan het zijn dat uit de onderlinge correlatie blijkt dat sommige x-variabelen sterk aan elkaar gecorreleerd zijn waardoor meer dan één van deze variabelen niet nodig of niet nuttig zijn. Na deze voorselectie van variabelen kan de eigenlijke regressie-analyse beginnen, waarin niet alleen de sterkte van de relaties bepaald wordt, maar ook de richting van het verband tussen een set verklarende variabelen en de te verklaren variabele.

Correlatie-analyse

Resultaten

In tabel 3.1 zijn de correlaties weergegeven tussen de toename van het aantal woningen tussen 1980 en 1995 en alle 22 geselecteerde variabelen waarvan aangenomen wordt dat ze een meer of minder grote invloed hebben op de lokatie van nieuwbouwwoningen (bijlage 3). Opgemerkt zij dat in de gepresenteerde tabel de woningtypen enerzijds of een- en meergezinswoningen zijn (deze classificatie wordt in de Ruimtescanner gebruikt) en anderzijds een verdergaande uitsplitsing naar type. Deze uitsplitsing is toch gebruikt om aan te tonen dat de noodzakelijke dichotomie naar woningtype leidt tot heterogeniteit in de schattingen moet leiden.

De r-waarden zijn over het algemeen erg laag. Ze laten wel een logisch en consistent patroon zien en worden daarom bruikbaar geacht. De toename van de voorraad van ééngezinswoningen, typen 1 en 2, correleert naar verwachting redelijk goed met de potentiaal van de ééngezinswoningen en rondt slecht met de potentiaal van meergezinswoningen. Maar ook de meergezins flatwoningen (type 4) correleren goed met de potentiaal van de ééngezinswoningen, wat kan wijzen op een verschuiving in de laatste decennia in de lokalisering van deze woningen. Nog opvallender echter is de negatieve r-waarde van de woninggroei van type 3 (meergezins, etage-, beneden-, bovenwoningen), met de potentiaal van meergezinswoningen. Hieruit kan geconcludeerd worden dat dit soort woningen de laatste 15 jaar kennelijk in mindere mate in de buurt van de bestaande lokaties voor deze typen woningen zijn gebouwd (de binnensteden).

Tabel 3.1: Correlaties tussen de verklarende en de te verklaren variabele

De correlatie tussen nieuwbouwwoningen en de potentiaalwaarden voor bosgebieden is over het algemeen negatief en klein. Naar verwachting is de positieve correlatie met de vrijstaande ééngezinswoningen van woningtype 1. Dit type woningen is in de afgelopen 15 jaar kennelijk meer in de omgeving van de bosrijke gebieden gebouwd of bos is ontstaan bij deze locaties. Alle nieuwe woningbouw scoort positief met de potentiaal voor bosrecreatie. Hieruit kan geconcludeerd worden dat er een tendens is nieuwe woningbouw dicht bij bosgebieden met bosrecreatie te situeren.

De r-waarden voor de potentiaal van droge en natte natuurterreinen zijn zo laag ($< 0,01$, behalve voor woningtype 1) dat geconcludeerd kan worden dat er geen duidelijk verband bestaat met nieuwe woningbouw. Een kleine positieve correlatie is aanwezig tussen gebieden in de buurt van water en nieuwe woningbouw. Alleen woningen van type 1 scoren negatief in dat opzicht, wat correspondeert met de ligging van deze woningbouwlocaties in bosgebieden die over het algemeen hoger en verder weg gelegen zijn van water.

Alle nieuwe woningbouw correleert duidelijk positief met de potentiaal van het aantal arbeidsplaatsen per sector, dat wil zeggen indien deze dicht in de buurt zijn gelegen van gebieden met veel arbeidsplaatsen. Dit geldt met name voor woningtypen 2 en 4 (eengezins rijtjeswoningen en meergezins flatwoningen) die voor bijna alle arbeidssectoren r-waarden boven de 0,1 hebben. De verschillen tussen de arbeidssectoren onderling zijn klein te noemen.

Woningtypen 2 tot en met 5 correleren het sterkst met de detailhandel en de kwartair- en kennisdiensten. Woningtype 1 correleert het best met de sector industrie, bouw en landbouw, hetgeen naar verwachting is. Een uitzondering op het algemene patroon wordt gevormd door de sector landbouw, die uitgezonderd voor woningtype 1, slechter correleert dan de andere sectoren. Arbeidsplaatsen in de landbouw zijn juist buiten de stedelijke woongebieden gelegen en zijn meer gespreid over het land, met uitzondering van het Westlandgebied waar een hoge concentratie arbeidsplaatsen is te vinden in de kassenbouw.

De Nederlandse snelwegen zijn als een boolean-kaart (waarde 1 of 0) als variabele opgenomen. Verwacht werd dat deze gebieden door met name geluidsoverlast onaantrekkelijk zouden zijn voor nieuwe woningbouw. Er is echter een positieve correlatie gevonden tussen de directe omgeving van snelwegen en nieuwe woningbouw. Met name de vrijstaande eengezinswoningen van type 1 ($> 0,05$) gevolgd door de rijtjeswoningen (woningtype 2) correleren, weliswaar niet sterk maar, positief met de snelwegen. Een mogelijke verklaring is dat relatief veel nieuwe woonwijken vlakbij snelwegen zijn gebouwd, waarbij de geluidsoverlast al dan niet is afgeschermd met geluidsisolerende voorzieningen. Voor de aanwezigheid van spoorwegen geldt juist dat deze het hoogst gecorreleerd zijn met de meergezins woningtypen 3 en 4, met name de flatwoningen van type 4 scoren relatief hoog ($> 0,05$).

De nabijheid van een op- en afrit van een snelweg correleert positief met nieuwe woningbouw. Deze correlatie is het sterkst voor woningtype 1, 2 en 4. Hetzelfde geldt voor de afstand tot het dichtst bij gelegen station. Woningtype 2, de eengezins-rijtjeshuizen correleert zowel bij de op- en afritten als de stations het sterkst.

Tenslotte valt een hoge positieve correlatie waar te nemen tussen de gebieden die in de jaren zeventig als groeikern zijn aangewezen en de nieuwbouw van eengezins rijtjeshuizen van type 2 ($> 0,1$). De correlatie met de overige woningbouw typen is ook positief maar beduidend lager ($< 0,05$). De als groeisteden aangewezen gebieden laten geen opvallend positieve correlaties zien voor één van de woningtypen.

In bijlage 3 is een tabel opgenomen met de onderlinge correlaties tussen de 'verklarende' variabelen. Wat betreft de onderlinge correlaties is een aantal verbanden interessant. De potentiaalkaart bosgebieden correleert sterk met de potentiaalkaart droge natuurterreinen. Voor de regressie-analyse betekent dit dat één van beide variabelen gekozen moet. Ook bosgebieden en bosrecreatie correleren goed maar niet zo sterk dat we op voorhand kunnen besluiten één van beide variabelen te laten vallen.

De onderlinge correlaties van de verschillende arbeidssectoren zijn zeer hoog te noemen. De r-waarden liggen in de meeste gevallen boven de 0,9. Alleen industrie (ca. tussen 0,8 en 0,9) en landbouw (ca. tussen 0,5 en 0,7) hebben enigszins lagere correlaties met de andere sectoren, hetgeen consistent is. Geconcludeerd kan worden dat de onderlinge correlatie van de arbeidssectoren over het algemeen zo sterk is dat in de regressie-analyses gekozen moet worden voor één van de variabelen, omdat meerdere variabelen niet bijdragen aan de verklaring van de y-variabele. Wel kan eventueel industrie of landbouw als aparte variabele opgenomen worden.

Conclusies ten aanzien van keuze variabelen

Tabel 3.1 geeft een overzicht van de geselecteerde variabelen. De potentiaalkaarten van bosgebieden en bosrecreatie tonen in de correlatie-analyse een licht maar specifiek verband aan met de toename van de woningvoorraad en worden daarom in de regressie-analyse opgenomen. Ondanks de zeer lage correlatie van de toename in de woning-voorraad met de potentiaal van droge en natte natuurterreinen hebben we besloten ook deze twee variabelen voorlopig te handhaven. Ook de potentiaalkaart met watergebieden laat in het algemeen een licht positieve correlatie met nieuwe woningbouw zien en is daarom voorlopig als variabele opgenomen voor de regressie-analyse.

Tabel 3.1 De geselecteerde variabelen voor de regressie-analyse op basis van de correlatie-analyse

Variabelen
Potentiaal arbeidsplaatsen landbouw 1983
Potentiaal arbeidsplaatsen industrie 1983
Potentiaal arbeidsplaatsen transportsector 1983
Potentiaal arbeidsplaatsen kennissector 1983
Potentiaal bosgebieden 1981
Potentiaal bosrecreatie 1981
Potentiaal droge natuurterreinen 1981
Potentiaal natte natuurterreinen 1981
Potentiaal watergebieden 1981
Aanwezigheid autosnelwegen 1983 in gridcel (ja/nee)
Aanwezigheid spoorwegen 1983 in gridcel (ja/nee)
Afstand tot dichtst bij gelegen op- en afrit 1983
Afstand tot dichtst bij gelegen NS-station 1980
Ligging in groeikern in 1980(?) (ja/nee)
Ligging in groeistad in 1980(?) (ja/nee)
Potentiaal woongebied eengezinswoningen 1980
Potentiaal woongebied meergezinswoningen 1980

De potentiaal van het aantal arbeidsplaatsen correleert duidelijk positief met nieuwe woningbouw, de verschillen per sector zijn echter klein te noemen. Uitzonderingen hierop vormen de sectoren landbouw en industrie die slechter correleren met woningtypen 1 t/m 4, maar beter met woningtype 1. Deze twee sectoren zijn vanwege de afwijkende correlatie als aparte variabelen geselecteerd. Van de overige sterker correlerende arbeidssectoren zijn de kennis- en de transportsector gekozen, omdat aangenomen wordt dat de kennissector meer in de stedelijke gebieden voorkomt en de transportsector meer in goed bereikbare gebieden of in de buurt van knooppunten van infrastructuur. Met de keuze van deze vier werkgelegenheidsvariabelen is zowel een goede ruimtelijke spreiding tussen de verschillende sectoren gewaarborgd als een goede spreiding van potentiële werknemers c.q. bewonersgroepen, zodat de mogelijke verschillen zo groot mogelijk zijn en de overlap tussen de variabelen dus zo klein mogelijk.

Alle infrastructuurvariabelen laten meer of minder duidelijke correlaties zien met de toename van de woningvoorraad. Aangezien deze variabelen geografisch gezien op verschillende lokaties voorkomen is het gevaar voor overlap in dit geval niet van belang en hebben we ze dus alle vier als variabele voor de regressie-analyse opgenomen. Tenslotte hebben we de groeikernen als variabele opgenomen daar deze zeer duidelijk positief correleren met één van de woningtypen (type 2). Deze variabele kan in de toekomstige simulaties dienen als proxy voor nieuw aan te wijzen groeigebieden, bijvoorbeeld rond de transportcorridors. Ook al laten de als groeisteden aangewezen gebieden geen opvallende correlaties zien met één van de woningtypen, hebben we toch besloten deze variabele in de eerste regressie-analyses op te nemen.

Regressie-analyse

Onderstaande tabel 3.2 toont zijn voor respectievelijk de Randstad en de Rest van Nederland en voor een- en meergezinswoningen de resultaten van de regressieanalyses beschreven. In bijlage 4 is de totale uitvoer van de step-wise-regression opgenomen opdat duidelijk wordt welke variabelen waarom (in statistische zin) niet worden opgenomen in het definitieve resultaat. In de afzonderlijke paragrafen wordt dieper ingegaan op de interpretatie van tabel 3.2

Logistische regressies Randstad en eengezinswoningen

Na een eerste regressie met de startset van 16 variabelen⁸ per regio, zijn daarna regressies uitgevoerd met kleinere groepen variabelen, waarbij in eerste instantie is uitgegaan van de sets variabelen die zijn overgebleven uit de step-wise-regression.

Geconcludeerd kan worden dat de factor arbeid als verklarende factor voor de groei van de voorraad eengezinswoningen niet van significant belang is. De belangrijkste reden hiervoor is waarschijnlijk dat ongeacht de lokatie in de Randstad de bereikbaarheid van gebieden met voldoende werkgelegenheid altijd voldoende is (Van am 1999).

⁸ Alleen de variabele groeisteden is verwijderd, omdat niet duidelijk is of de opgetreden groei het resultaat was van beleid of van autonome ontwikkelingen en er bovendien geen equivalent voor deze variabele naar de toekomst is.

Voor de (zeer kleine) negatieve waarde van de potentiaal bosrecreatie kan geen directe verklaring gegeven worden, mogelijk houdt dit verband met de gekozen afstandsverval functie (maximale invloedsafstand 15 kilometer).

Tabel 3.2 *Beta 's en t-waarden van de regressievergelijkingen voor de groei van het aantal een- en meergezinswoningen in de Randstad en Overig Nederland*

	Randstad		Overig Nederland	
	Eengezins	Meergezins	Eengezins	Meergezins
	B	t-waarde	B	t-waarde
Startwaarde/intercept	-7,02	-389,59	-12,06	-462,24
Kennisdiensten	Nvt	Nvt	0,34	121,35
Recreatief bosgebied	-0,02	-27,46	Nvt	Nvt
Wegen	-0,34	-94,70	Nvt	Nvt
Rail	-0,58	-91,31	0,08	75,27
Stations	-0,14	-103,47	-0,11	-41,34
Op- en afritten	0,04	33,04	-0,37	-85,23
Groeikern	1,13	264,04	0,12	-115,32
Eengezinswoningen	0,73	548,00	0,05	57,14
Meergezinswoningen	0,06	75,67	1,62	364,66
			0,83	933,53
			0,10	126,37
			0,50	328,66

Niet onverwacht zijn de negatieve waarden voor wegen en spoorwegen. De (kleine) positieve waarde voor de afstand tot op- afritten van snelwegen, betekent juist een kleinere kans op woningbouw vlak bij op- en afritten. Dit kan betekenen dat het belang van een dicht bij gelegen op- en afrit voor de woningbouw vanaf 1980 relatief gering is. Dit in tegenstelling tot de groei van bedrijfsterreinen en kantoorlokaties die juist het hoogst is vlakbij op- en afritten (Wagtendonk, 1999). Voor NS-stations geldt dat de woningbouw sneller groeit dicht bij de stations dan op grotere afstand. Bij bedrijfsterreinen is dit verband weer precies andersom. Verder wordt uit de β -waarden duidelijk dat de ligging in een groeikern een belangrijke positieve factor is en ook de potentiaalwaarde van de bestaande ééngezinswoningen zwaar meetelt. Een hoge positieve β -waarde wordt tenslotte gevonden voor de potentiaal van de bestaande voorraad ééngezinswoningen.

Logistische regressies Randstad en meergezinswoningen

De hoogste (positieve) β -waarde wordt gevonden voor de arbeidspotential kennis. Dit is geen onverwacht resultaat, daar zowel de meergezinswoningen als de kenniscentra zich hoofdzakelijk in de grote steden concentreren. De bosgebieden laten een licht positief verband zien met de toename van meergezinswoningen, mogelijk omdat het aanwezige bos in de Randstad dicht bij de grote steden (met meergezinswoningen) gelegen is (bijvoorbeeld de Utrechtse heuvelrug). In vergelijking met de regressieresultaten van de ééngezinswoningen in de Randstad valt voor de meergezinswoningen de (licht) positieve β -waarde op voor de aanwezigheid van spoorwegen (op korte afstand). Dit betreft waarschijnlijk de bouw van gestapelde meergezinswoningen in de grote steden vlakbij het spoor. De negatieve waarde voor de autosnelwegen is niet verbazingwekkend, daar deze over het algemeen verder buiten de stad gelegen zijn. Het verband tussen de nabijheid van op-en afritten van snelwegen en NS-stations (respectievelijk negatief en positief) is vergelijkbaar met het verband bij de ééngezinswoningen. De ligging in een groeikern is gezien de lagere β -waarde van aanmerkelijk minder belang voor meergezinswoningen in de Randstad dan ééngezinswoningen, wat te verklaren is door het relatief geringe aandeel van meergezinswoningen in de groeikernen. Tenslotte valt op dat de potential van ééngezinswoningen leidt tot een hogere β -waarde dan de potential met meergezinswoningen, wat gedeeltelijk een gevolg is van het kleinere aandeel meergezinswoningen⁹.

Logistische regressies intermediair/perifeer Randstad en eengezinswoningen

Buiten de Randstad levert van de onderzochte variabelen de arbeidspotential de hoogste β -waarde op voor de categorie ééngezinswoningen. In vergelijking met de Randstad valt op dat nu ook een licht positief verband wordt gevonden voor de potential van natte natuur. De aanwezigheid van een waterrijke omgeving is buiten de Randstad kennelijk van groter belang. Omdat natte natuur en de bosgebieden die ook een licht positief verband laten zien, onderling niet correleren, zijn beide variabelen in dit geval opgenomen. Zowel vlakbij gelegen autowegen als spoorwegen laten buiten de Randstad een duidelijk negatief verband zien met de toename van de voorraad ééngezinswoningen. Ook de nabijheid van op- en afritten en NS-stations laten een normaal verband zien (respectievelijk negatief en positief). Tenslotte worden hoge β -waarden gevonden voor de ligging in groeikernen en voor gebieden in de nabijheid van bestaande concentraties ééngezinswoningen.

Logistische regressies intermediair/perifeer Randstad en meergezinswoningen

Voor meergezinswoningen buiten de Randstad geldt net als voor meergezinswoningen in de Randstad dat de arbeidspotential kennis de hoogste β -waarde oplevert. Net als bij de ééngezinswoningen levert nu ook de potential van natte natuurgebieden een duidelijk positief verband op. De bosgebieden doen echter weinig en zijn daarom weggelaten. Ook hier scoren

⁹ in 1980 ca. 1,7 miljoen meergezinswoningen t.o.v. 4,4 miljoen ééngezinswoningen (zie summary statistics bijlage 2).

zowel autosnel-wegen als spoorwegen negatief. In dit geval wordt geen significant verband gevonden met de nabijheid van op- en afritten. Wel wordt wederom een positief verband gevonden met de nabijheid van NS-stations. Ook de ligging in een groeikern levert weer een sterk positief verband op, maar minder sterk dan in het geval van de ééngezinswoningen. Net als in de Randstad weegt de potentiaal van ééngezinswoningen zwaarder dan de potentiaal van meergezinswoningen.

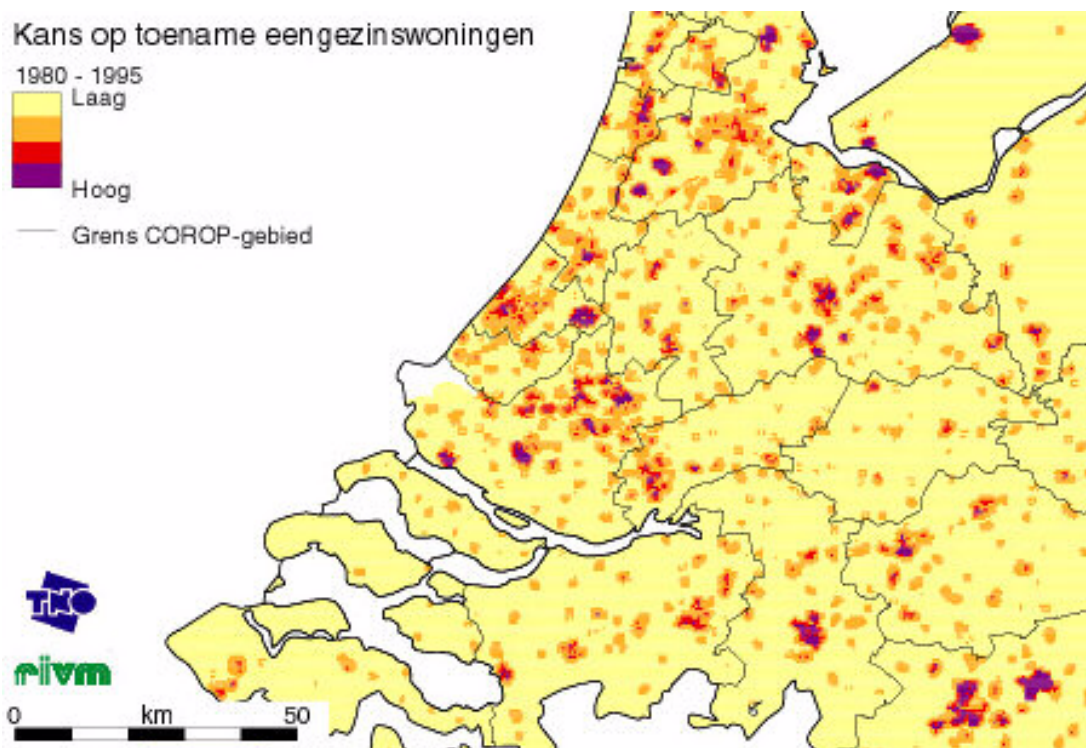
Visuele toetsing regressie-resultaten

De regressie-resultaten kunnen gevisualiseerd worden door de regressie-vergelijkingen met bijbehorende β -waarden om te zetten in kansenkaarten. De reden is om de abstracte resultaten te toetsen op hun plausibiliteit.

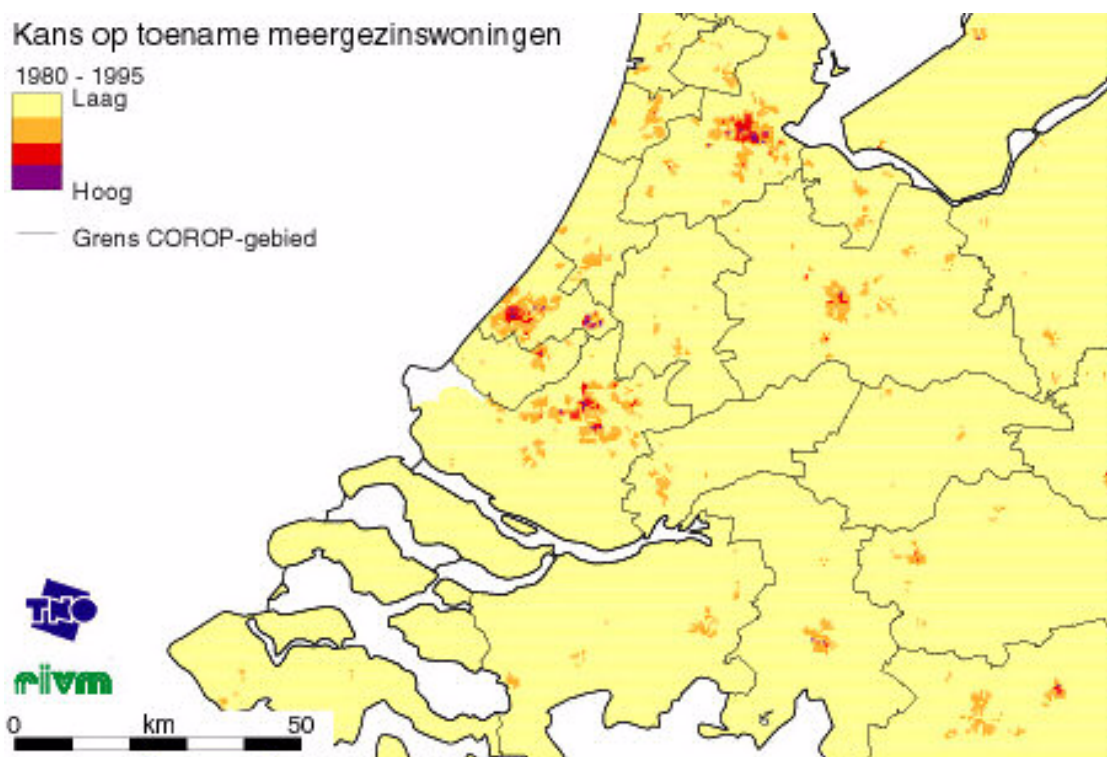
Hiertoe wordt per woningtype de y-waarde op onderstaande wijze uit de regressie-vergelijking berekend. Hierbij worden de regio's Randstad en overig Nederland apart berekend. De berekende kans is de kans op de maximale toename in een cel van respectievelijk 1280 ééngezinswoningen of 3864 meergezinswoningen. In tabel 3.3 zijn de beide regressievergelijkingen zoals deze in praktijk worden gebruikt weergegeven. Deze vergelijking wordt ook in de Ruimtescanner gebruikt.

Tabel 3.3 De regressievergelijkingen voor eengezins- en meergezinswoningen zoals geïmplementeerd in de Ruimtescanner

Kansfuncties	$P(y) = \exp(y) / (1 + \exp(y))$
Eengezinswoningen	$y = \text{ifelse}(\text{RANDSTAD} == 0, (-11.81059046 + (0.28298343 * \log(\text{P.INDUS83} + 1)) + (0.03520182 * \log(\text{P.BOSGEB81} + 1)) + (0.07236123 * \log(\text{P.NATNTR81} + 1)) + (-0.10723159 * \text{WEGEN83}) + (-0.37140986 * \text{RAIL80}) + (0.04993489 * \log(\text{D.OPAF83} + 1)) + (-0.11796565 * \log(\text{D.STAT80} + 1)) + (1.63244323 * \text{GROEIKRN}) + (0.82802218 * \log(\text{P.EENGEZ80} + 1)) + (0.09820695 * \log(\text{P.MEERGEZ80} + 1))),$ $-7.02037990 + (-0.02344555 * \log(\text{P.BOSREC81} + 1)) + (-0.34348735 * \text{WEGEN83}) + (-0.58014368 * \text{RAIL80}) + (-0.13722021 * \log(\text{D.STAT80} + 1)) + (0.04098646 * \log(\text{D.OPAF83} + 1)) + (1.12648636 * \text{GROEIKRN}) + (0.73087098 * \log(\text{P.EENGEZ80} + 1)) + (0.06201094 * \log(\text{P.MEERGEZ80} + 1)))$
Meergezinswoningen	$y = \text{ifelse}(\text{RANDSTAD} == 0, (-11.81059046 + (0.28298343 * \log(\text{P.INDUS83} + 1)) + (0.03520182 * \log(\text{P.BOSGEB81} + 1)) + (0.07236123 * \log(\text{P.NATNTR81} + 1)) + (-0.10723159 * \text{WEGEN83}) + (-0.37140986 * \text{RAIL80}) + (0.04993489 * \log(\text{D.OPAF83} + 1)) + (-0.11796565 * \log(\text{D.STAT80} + 1)) + (1.63244323 * \text{GROEIKRN}) + (0.82802218 * \log(\text{P.EENGEZ80} + 1)) + (0.09820695 * \log(\text{P.MEERGEZ80} + 1))),$ $-7.02037990 + (-0.02344555 * \log(\text{P.BOSREC81} + 1)) + (-0.34348735 * \text{WEGEN83}) + (-0.58014368 * \text{RAIL80}) + (-0.13722021 * \log(\text{D.STAT80} + 1)) + (0.04098646 * \log(\text{D.OPAF83} + 1)) + (1.12648636 * \text{GROEIKRN}) + (0.73087098 * \log(\text{P.EENGEZ80} + 1)) + (0.06201094 * \log(\text{P.MEERGEZ80} + 1)))$



Kaart 3.1 De kans op toename van eengezinswoningen 1980-1995 (bron: RIVM/TNO bewerking)



Kaart 3.2 De kans op toename van meergezinswoningen 1980-1995 (bron: RIVM/TNO bewerking)

Met behulp van beide regressievergelijkingen wordt aan iedere gridcel een kans toegekend aan de toename van de dichtheid. Deze kansen kunnen natuurlijk gevisualiseerd worden in GIS-kaarten. Kaart 3.1 toont de kans op de maximale toename van ééngezinswoningen weergegeven. Kaart 3.2 toont de kans op de maximale toename van meergezinswoningen weergegeven. Tevens is in beide kaarten het bij de kansen horende aantal woningen vermeld.

Op basis van een visuele interpretatie kan gezegd worden dat zowel de kansenverdeling van de ééngezins- als de meergezinswoningen goed overeen lijkt te komen met de verdeling van de werkelijke toename van deze woningtypen tussen 1980 en 1995. Alleen de kans op woningbouwtoenames in de stad Almere lijkt te laag ingeschat in vergelijking met de werkelijkheid. Dit wekt echter geen verbazing, omdat in het geval van Almere sprake is geweest van een bijzondere beleidsinspanning gedurende de beschouwde periode

Conclusies

Inleiding

In dit laatste hoofdstuk gaan we kort in op de belangrijkste conclusies. Eert wordt kort ingegaan op de waarde van de gebruikte methodiek en de resultaten. Vervolgens worden aanbevelingen voor verder onderzoek gepresenteerd.

Een terugblik

De correlatie en regressie-analyses hebben voor de twee verschillende regio's geleid tot een selectie van een aantal verklarende variabelen per onderscheiden woningtype. Ondanks relatief lage r-waarden in de correlatie-analyse kunnen de positieve en negatieve verbanden en de relatieve verschillen uit de regressie-analyses, over het algemeen goed verkaart worden, waaruit geconcludeerd wordt dat de gekozen methode en de gekozen variabelen bruikbaar zijn voor de uitgevoerde historische analyse van ruimtelijke ontwikkelingen in de woningbouw in de periode 1980-1995. Opvallend daarbij is, is dat ook de ruimtelijke factoren als natuur en recreatie enig significant verband laten zien met de toename van de woningvoor-raad. De gekozen schaalfactor met gridcellen van 500 bij 500 meter is daarbij waarschijnlijk van doorslaggevend belang. In een vergelijkbare studie in Belgisch Vlaanderen (IGEAT- ULB, 1999) op gemeenteniveau, werd bijvoorbeeld alleen een significante invloed gevonden van de bereikbaarheid tot werkgelegenheid. Voor variabelen als natuur en recreatie werd in die studie geen significante invloed gevonden. Ook bij een visuele vergelijking van de met het regressiemodel berekende kansen op toename van woningen met de verdeling van de werkelijke toename van deze woningtypen tussen 1980 en 1995, blijken deze over het algemeen goed met elkaar overeen te komen.

Aanbevelingen voor onderzoek

- Een aantal van de beslissingen aangaande de te volgen methodiek zijn genomen op grond van de beschikbaarheid van de op dat moment gewenste data. Zo kon er bijvoorbeeld niet met de gewenste woonmilieu's gewerkt worden, maar moest met vijf woontypes gewerkt worden die vervolgens weer tot twee verschillende woontypen samengevoegd moesten worden. Aangezien gewerkt wordt op een schaalniveau van 500 metercellen was ook de homogene verdeling van aantallen arbeidsplaatsen per postcodegebied naar 500 meter cellen, geen ideale oplossing. In verband met het te volgen tijdsplan kon echter niet gewacht worden op nieuwe data. Indien de gewenste data alsnog ter beschikking komt kunnen de regressie-analyses verder verbeterd worden.
- Deze trendstudie is valide voor het onderzoek van RIVM en TNO Inro waarin de aanname centraal staat dat de bestaande verhouding tussen enerzijds de regulerende overheid (mogelijkheden en beperkingen) als publiek orgaan en anderzijds de woonconsumenten, en indirect de overheid als privaat orgaan en aanbieders van woonruimte en woondiensten,

eenvoudig kon worden doorgetrokken. Voor de ontwikkeling van alternatieve ruimtelijke scenario's voldoet een dergelijke trendanalyse op basis van ruimtelijke informatie niet. In het laatste geval dient een modelspecificatie te worden gebruikt op basis van een gedragsmatige onderbouwing van het besluitvormingsgedrag van de relevante actoren in het spel

- De verklarende variabele grondprijzen is niet in de analyse meegenomen omdat hier geen geschikte data van voorhanden was. Vanwege het verwachte verband tussen grondprijzen en de verschillende woningtypen wordt aanbevolen deze variabele (indien beschikbaar) op te nemen in de trendanalyses. In deze studie hoefde de grondprijs niet te worden opgenomen vanwege de aannames in de RIVM en TNO Inro studie (Goetgeluk et al. 2000).
- Behalve de beschikbaarheid van de data leverde de omvang van de in S-Plus (statistisch pakket) in te voeren data-set problemen op met de rekenkracht van de gebruikte computers. Geprobeerd is te werken met samples van de dataset (25.000 gridcellen in plaats van 140.000), dit leverde echter afwijkende resultaten op ten opzichte van de volledige dataset. Om sneller te kunnen werken wordt daarom het gebruik van krachtiger computers¹⁰ met veel opslagruimte aangeraden.
- In de huidige opzet is gewerkt met één tijdsperiode van 1980 tot en met 1995. Indien meer tijd (en meer data) beschikbaar is, kunnen de analyses per tijdsperiode (bijvoorbeeld perioden van vijf jaar) verricht worden, zodat ook ontwikkelingen in de tijd van variabelen en coëfficiënten kunnen worden waargenomen. Deze trends kunnen vervolgens weer naar de toekomst worden doorgetrokken.
- Uitgegaan is van een maximaal aantal woningen per woningtype per cel in 1995. Dit aantal hoeft niet de hoogste dichtheid van woningen per cel te vertegenwoordigen. Wellicht kan beter met de hoogst mogelijke dichtheid gewerkt worden. Deze kan gevonden worden door eerst het niet voor wonen geschikte areaal van alle gridcellen af te trekken en het aantal woningen dan door de resterende celoppervlakte te delen. Vervolgens kan uitgegaan worden van een maximaal aantal woningen per voor wonen geschikte hectare.
- De ruimtelijke schaal waarop claims worden gebruikt bepaalt in hoge mate de uitkomsten. Uit de analyses kwamen duidelijke regionale verschillen naar voren tussen de Randstad en overig Nederland. Indien het onderzoek wordt uitgebreid kunnen meer regionale indelingen geanalyseerd worden, zoals een analyse per woningmarktgebied. Deze worden omschreven als het gebied waar de meeste mensen verhuizen omwille van woonredenen zonder dat hierbij de bereikbaarheid van familie, vrienden en het werk in geding komt.

Literatuur

- Blalock, H.M.(1979), *Social statistics*. McGraw-Hill book co., Singapore.
- Bootsma, H.G. (1998), *The myth of reurbanization; location dynamics of households in the Netherlands* Thela Thesis/NethurD Publications, Amsterdam
- CBS (1997), *Statistiek van het bodemgebruik 1993*. Centraal Bureau voor de Statistiek, Voorburg/Heerlen.
- Chou Y.H.(1997), *Exploring spatial analysis in geographic information systems* Onword Press, Santa Fe, USA.
- Cramer, J.S. (1991), *The LOGIT model: an introduction for economists*. Edward Arnold, Londen..
- Cramer, J.S. (1986), *Econometric applications of Maximum Likelihood methods*. Cambridge University Press.
- Dubois, O., E. Wolff & M.F. Godart (1999), *A model of new residential housing location in Wallonia*. IGEAT (ULB)-Brussel. Poster gepresenteerd op de 'Euro-conference of Socio-economic analysis and GIS' in Espinho, Portugal, 22 -27 mei 1999.
- Goetgeluk, R.W. (1997), *Bomen over Wonen; woningmarktonderzoek met beslissingsbomen* Koninklijk Nederland Aardrijkskundig Genootschap 235, dissertatie, Utrecht.
- Goetgeluk, R.W., P.J. Louter, J.A.M. Borsboom-Van Beurden, M.A.J. Kuijpers-Linde, J.F.M. van der Waals & K.T. Geurs (2000), *Wonen en werken ruimtelijk verkend, waar wonen en werken we in 2020 volgens een compacte inrichtingsvariant voor de Vijde Nota Ruimtelijke ordening*. RIVM, Biltthoven, RIVM-rapport 711931001.
- Ham, M. van (1997), *Job access in a polynucleated metropolitan region; commuting tolerance and job access in the Netherlands*. Paper for the ENHR-YHR Educational Seminar Istanbul, Turkey, 24-28 May 1999.
- Harts J.J, C. Maat, M. & M.J. Zeijlmans van Emmichhoven (1999), *Meervoudig stedelijk ruimtegebruik: methoden en analyse*. Delft University press, reeks stedelijke en regionale verkenningen No 20.
- Hilbers, H.D. & E.J. Verroen (1993), *Het beoordelen van de bereikbaarheid van lokaties: definiëring, maatsstaven, toepassing en beleidsimplicaties*. Rapport 93/NV/099 TNO Inro, Delft.
- Louter, P.J. (1997), *De economische kaart van Nederland in 2015, beschrijving van een prototype van het OPERA-model: een verklarende shift-share analyse voor werkgelegenheids-groei*. TNO Inro, Delft.
- Ransijn, M. & R. Vreeker (1998), *Historische analyse van ruimtegebruik en ruimtelijk beleid in Nederland*. Vrije Universiteit. Amsterdam.
- Rietveld P.(1998), *Modelontwikkeling van de ruimtescanner*, Vrije Universiteit, Amsterdam.
- Schotten, C.G.J., R.J. van de Velde en H.J. Scholten, W.T. Boersma, M. Hilferink, M. Ransijn, P. Rietveld & R. Zut (1997), *De Ruimtescanner, geïntegreerd ruimtelijk informatiesysteem voor de simulatie van toekomstig ruimtegebruik* RIVM-rapport 711901 002. RIVM, Biltthoven.
- Vocht, A. de (1992), *Indeling van Nederlandse gemeenten*. Faculteit Ruimtelijke wetenschappen, Universiteit Utrecht.

Wagtendonk, A.J. & C.G.J. Schotten (2000), *Ruimtelijke ontwikkelingen van bedrijfsterreinen in Nederland, 1981 – 2005*. Een historische analyse ter ondersteuning van het Model Ruimtescanner, op een schaalniveau van 500 meter cellen. RIVM, Bilthoven, RIVM-rapport 711901028 (in voorbereiding).

Wrigley, N. (1985), *Categorical data analysis for geographers and environmental scientists* Longman, London/New York.

Bijlage 1 ML- parameter schattingen van de factoren voor woningbouwgroei

In grid cel c kunnen voor een bepaald woningtype maximaal $N_c := M_c \cdot L_c$ woningen gebouwd worden, waarbij L_c het aantal mogelijk te bebouwen ha grond in grid cel c is en M_c is het maximum aantal woningen van dat type per ha. (M_c = intensiteit).

We kunnen het waargenomen aantal gebouwde woningen W_c beschouwen als uitkomst van een binominaal kansproces met N_c trekkingen met kans P_c waarbij P_c een functie is van Y_c , waarbij

$$Y_c := X_c \cdot \beta = \sum_k [X_{ck} \cdot \beta_k],$$

met X_c een vector met de exogene kansbepalende factoren is en β de vector van de te bepalen modelparameters.

Voor de relatie tussen P_c en Y_c is gekozen voor de logit functie:

$$P(y) := \exp y / (1 + \exp y)$$

We definiëren:

$$Q(y) := 1 - P(y) = 1 / (1 + \exp y)$$

We kunnen nu $P'(y)$ afleiden:

$$\begin{aligned} P'(y) &= P(y) \cdot Q(y) \quad \text{en} \\ Q'(y) &= - P(y) \cdot Q(y) \end{aligned}$$

De inverse van de P functie (dit is de link functie in S-plus),

$$y(p) = \ln p - \ln (1-p) = \ln [p / (1-p)]$$

Likelihood $L_c(W_c; \beta, X_c)$ van waargenomen aantal nieuwe woningen W_c gegeven parameter vector β en exogene factoren X_c wordt gedefinieerd als:

$$L_c(W_c; \beta, X_c) := P(X_c \cdot \beta)^{W_c} \cdot Q(X_c \cdot \beta)^{(N_c - W_c)} \cdot C(N_c, W_c)$$

Er wordt dus een binominaal keuzeprocess (wel of niet woning bouwen) geschat met kans $P(\beta, X_c)$ en aantal waarnemingen N_c . Het is eenvoudig dit model uit te breiden met meerdere keuzemogelijkheden tot een multinominaal model. De van de modelparameters onafhankelijke term $C(W_c, N_c)$ staat voor het aantal mogelijke manieren waarop men W_c huizen op N_c percelen kan bouwen en is als volgt gedefinieerd:

$$C(N_c, W_c) := N_c! / [W_c! \cdot (N_c - W_c)!]$$

De likelihood van de waargenomen wel/niet woningbouw in alle gridcellen tezamen:

$$L(W; \beta, X) := \prod_c [L_c(W_c; \beta, X_c)]$$

Hieruit volgt dat:

$$\ln L(W; \beta, X) := \sum_c [W_c \cdot \ln P(X_c \cdot \beta) + (N_c - W_c) \cdot \ln Q(X_c \cdot \beta) + \ln C(N_c, W_c)]$$

De Maximum Likelihood schatter voor β wordt gevonden door $\ln L(W; \beta, X)$ te maximeren, wat impliceert dat de partiele afgeleiden naar $\beta_k = \partial \ln L(W; \beta, X) / \partial \beta_k$ aan 0 gelijk gesteld kunnen worden.

$$\begin{aligned} \partial \ln L(W; \beta, X) / \partial \beta_k &= \\ \sum_c [(W_c / P(X_c \cdot \beta)) \cdot P'(X_c \cdot \beta) \cdot X_{ck} + ((N_c - W_c) / Q(X_c \cdot \beta)) \cdot Q'(X_c \cdot \beta) \cdot X_{ck}] &= \\ \sum_c X_{ck} \cdot P'(X_c \cdot \beta) \cdot [W_c / P(X_c \cdot \beta) - (N_c - W_c) / Q(X_c \cdot \beta)] &= \\ \sum_c X_{ck} \cdot [W_c \cdot Q(X_c \cdot \beta) - (N_c - W_c) \cdot P(X_c \cdot \beta)] &= \\ \sum_c X_{ck} \cdot [W_c - N_c \cdot P(X_c \cdot \beta)] & \end{aligned}$$

Hieruit blijkt dat de term $\ln C(N_c, W_c)$ geen rol speelt in het bepalen van de ML-estimator voor β . Dit ondersteunt de constatering dat $X_c \cdot \beta$ als attractiviteitswaarde voor wonen in gridcel c genomen kan worden in het allocatiemodel in de ruimtescanner.

Appendix

De afleiding van $P'(y) = P(y) \cdot Q(y)$ is gebaseerd op:

$$d(a/b) = da \cdot b - db \cdot a / b^2$$

$$P'(y) := [\exp(y) \cdot [1 + \exp(y)] - \exp(y)^2] / [1 + \exp(y)]^2$$

$$= \exp(y) / [1 + \exp(y)]^2 = P(y) \cdot (1 - P(y))$$

$$\partial P_c / \partial \beta_k = X_{ck} \cdot P'(X_c \cdot \beta)$$

$$\begin{aligned} \partial \ln L(W; \beta, X) / \partial \beta_k &= \sum_c [W_c / P_c \cdot \partial P_c / \partial \beta_k - (N_c - W_c) / (1 - P_c) \cdot \partial P_c / \partial \beta_k] \\ &= \sum_c [\partial P_c / \partial \beta_k \cdot [W_c / P_c - (N_c - W_c) / (1 - P_c)]] \\ &= \sum_c [X_{ck} \cdot P_c \cdot (1 - P_c) \cdot [W_c / P_c - (N_c - W_c) / (1 - P_c)]] \\ &= \sum_c [X_{ck} \cdot [(1 - P_c) \cdot W_c - P_c \cdot (N_c - W_c)]] \\ &= \sum_c [X_{ck} \cdot (W_c - P_c \cdot N_c)] \end{aligned}$$

gegeven alle andere β_k , los op $\partial \ln L(b) / \partial \beta_k = 0$

$$\sum_c [X_{ck} \cdot (W_c - P_c \cdot N_c)] = 0$$

merk op dat $\sum_c [X_{ck} \cdot (W_c - P_c \cdot N_c)]$ monotoon dalend is in β_k

$$\partial (\sum_c [X_{ck} \cdot (W_c - P_c \cdot N_c)]) / \partial \beta_k = - \sum_c [X_{ck}^2 \cdot N_c \cdot P_c \cdot (1 - P_c)]$$

Bijlage 2 summary statistics 1980

```

*** Summary Statistics for data in: totaalwonen80 ***

      PSTAT80  RAIL80  PBOSGEB81  PBOSREC81  PDRGNTR81  PNATNTR81  POVOWTR81  PWON80.T34  NEW.PINDUS
Min:         0         0         0         0         0         0         0         0         261
Mean:       7428         0        9385        526        2130        1956        3381         10       17872
Median:     6082         0        5668        182         744        1038        1931         0       14135
Max:       33837         1       67774        5342       31807       19913       21900       3953       74359
Total N:   139620   139620   139620   139620   139620   139620   139620   139620   139620
NA's :      0         0         0         0         0         0         0         0         0
Variance:  28355139         0   119433437   655147   14338154   6699336   12729648   7452   171464571
Std Dev.:   5325         0       10929        809        3787        2588        3568         86       13094
Sum: 1037092996       7104 1310376188   73396893  297328748  273095832  472101500  1388724  2495288174

      NEW.PKENNI  NEW.PLANDB  NEW.PTRANP  POPAF83  WEGEN83  P.EENGEZ80  GROEIKRN  WYWT12.H  YWT12.H
Min:         151         171         38         0         0         0         0         0         0
Mean:       9194        4089        2668        5430         0        189         0        1179         0
Median:     4474        3956        1390        4031         0         22         0        1280         0
Max:       90593       13692       25541       37576         1       6576         1        1280         1
Total N:   139620   139620   139620   139620   139620   139620   139620   139620   139620
NA's :      0         0         0         0         0         0         0         0         0
Variance: 151429790  4299924  11974239  23656653         0   196733         0       61342         0
Std Dev.:  12306        2074        3460        4864         0         444         0        248         0
Sum:1283708401  570901477  372445207  758116341       32019  26441159       3888  164594617       1990

      WYWT345.H  YWT345.H  INTERMEDIA  RANDSTAD
Min:         0         0         0         0
Mean:       3542         0         0         0
Median:     3846         0         0         0
Max:       3846         1         1         1
Total N:   139620   139620   139620   139620
NA's :      0         0         0         0
Variance:  553809         0         0         0
Std Dev.:   744         0         0         0
Sum:  494555389       346       61202       23391
    
```

Bijlage 3 Onderlinge correlaties 'verklarende' variabelen

De naamgeving van de variabelen in de Splus uitgevoerde analyses correspondeert niet geheel met de gehanteerde naamgeving in dit rapport. Hieronder volgt derhalve eerst een verklarend overzicht:

<u>In rapport</u>	<u>In S-Plus</u>
Potentiaal arbeidsplaatsen bouw 1983	NEW.PBOUW83
Potentiaal arbeidsplaatsen detailhandel 1983	NEW.PDTHAN83
Potentiaal arbeidsplaatsen groothandel 1983	NEW.PGRHAN83
Potentiaal arbeidsplaatsen kwartaire sector 1983	NEW.PKWART83
Potentiaal arbeidsplaatsen landbouw 1983	NEW.PLANDB83
Potentiaal arbeidsplaatsen industrie 1983	NEW.PINDUS83
Potentiaal arbeidsplaatsen transportsector 1983	NEW.PTRANP83
Potentiaal arbeidsplaatsen kennissector 1983	NEW.PKENNI83
Potentiaal arbeidsplaatsen overige sectoren 1983	NEW.POVERI83
Potentiaal bosgebieden 1981	PBOSGEB81
Potentiaal bosrecreatie 1981	PBOSREC81
Potentiaal droge natuurterreinen 1981	PDRGNTR81
Potentiaal natte natuurterreinen 1981	PNATNTR81
Potentiaal watergebieden 1981	POVOWTR81
Aanwezigheid autosnelwegen 1983 in gridcel (ja/nee)	WEGEN83
Aanwezigheid spoorwegen 1983 in gridcel (ja/nee)	RAIL80
Afstand tot dichtst bij gelegen op- en afrit 1983	POPAF83
Afstand tot dichtst bij gelegen NS-station 1980	PSTAT80
Ligging in groeikern in 1980(?) (ja/nee)	GROEIKRN
Ligging in groeistad in 1980(?) (ja/nee)	GROEISTD
Potentiaal gebied ééngezinwoningen type 1 - 1980	PWON80.T1
Potentiaal gebied ééngezinwoningen type 2 - 1980	PWON80.T2
Potentiaal gebied meergezinwoningen type 3, 4 en 5 - 1980	PWON80.T345

	PSTAT80	RAIL80	PBOSGEB81	PBOSREC81	PDRGNTR81	PNATNTR81
PSTAT80	1.00000000	-0.180129655	-0.046658048	-0.07517564	0.0298172199	0.089897233
RAIL80	-0.18012966	1.00000000	0.001287485	0.02678678	-0.0044872931	-0.044603979
PBOSGEB81	-0.04665805	0.001287485	1.00000000	0.31576660	0.6747652719	-0.170973546
PBOSREC81	-0.07517564	0.026786784	0.315766603	1.00000000	0.2560098939	-0.073821581
PDRGNTR81	0.02981722	-0.004487293	0.674765272	0.25600989	1.000000000	-0.089780033
PNATNTR81	0.08989723	-0.044603979	-0.170973546	-0.07382158	-0.0897800327	1.000000000
POVOWTR81	-0.18203792	0.030724437	-0.364646499	-0.04250923	-0.1456465581	0.348174226
PWON80.T1	-0.17508775	0.130581244	0.026231033	0.06915530	-0.0123856547	-0.089803069
PWON80.T2	-0.20840853	0.148505959	-0.039776548	0.11849862	0.0089143943	-0.031915623
PWON80.T345	-0.11038963	0.060407690	-0.044414690	0.09493409	0.0011889225	0.001334504
NEW.PBOUW83	-0.31558806	0.099937014	-0.103451340	0.22615328	0.0061919612	0.021410923
NEW.PDTHAN83	-0.33341580	0.097988263	-0.059418663	0.27505092	0.0629348582	0.023749701
NEW.PGRHAN83	-0.28616614	0.086917275	-0.118034753	0.24358639	0.0136143060	0.048569560
NEW.PINDUS83	-0.34115883	0.104960246	0.037695609	0.25595533	0.0175745547	-0.043353839
NEW.PKENNI83	-0.28997020	0.091790390	-0.099017248	0.25965671	0.0538620232	0.042786136
NEW.PKWART83	-0.34142942	0.101918770	-0.033837582	0.29731576	0.0885749020	-0.007737662
NEW.PLANDB83	-0.33053265	0.061687477	-0.174848914	0.07630326	0.0028040665	-0.121176989
NEW.POVERI83	-0.24498826	0.088992148	-0.203495491	0.19086148	0.0144326150	0.027635281
NEW.PTRANP83	-0.27028378	0.096209484	-0.142741974	0.19464208	-0.0096779218	0.039714690
POPAF83	0.43901915	-0.112040216	-0.125609233	-0.18947875	0.0296168574	0.104260120

Ruimtelijke ontwikkelingen woningbouw 1980 - 1995

WEGEN83	-0.09216629	0.074103084	-0.015820721	0.01841470	-0.0102465382	-0.035038479
	PWON80.T1	PWON80.T2	PWON80.T345	NEW.PBOUW83	NEW.PDTHAN83	
PSTAT80	-0.17508775	-0.208408532	-0.110389625	-0.315588063	-0.33341580	
RAIL80	0.13058124	0.148505959	0.060407690	0.099937014	0.09798826	
PBOSGEB81	0.02623103	-0.039776548	-0.044414690	-0.103451340	-0.05941866	
PBOSREC81	0.06915530	0.118498624	0.094934089	0.226153280	0.27505092	
PDRGNTR81	-0.01238565	0.008914394	0.001188922	0.006191961	0.06293486	
PNATNTR81	-0.08980307	-0.031915623	0.001334504	0.021410923	0.02374970	
POVOWTR81	-0.02420118	0.116596176	0.111638133	0.540117922	0.52777563	
PWON80.T1	1.00000000	0.505402281	0.120104633	0.092022021	0.09550383	
PWON80.T2	0.50540228	1.000000000	0.416224330	0.292569685	0.29663279	
PWON80.T345	0.12010463	0.416224330	1.000000000	0.260995695	0.28092706	
NEW.PBOUW83	0.09202202	0.292569685	0.260995695	1.000000000	0.97072217	
NEW.PDTHAN83	0.09550383	0.296632790	0.280927061	0.970722170	1.00000000	
NEW.PGRHAN83	0.06479695	0.268691585	0.258074990	0.970724824	0.98217155	
NEW.PINDUS83	0.14176318	0.291391798	0.227284059	0.916513149	0.90562270	
NEW.PKENNI83	0.06576014	0.282997817	0.316042259	0.941693520	0.97751113	
NEW.PKWART83	0.09740616	0.302539539	0.291530534	0.960684812	0.98260656	
NEW.PLANDB83	0.05899202	0.142985963	0.103490436	0.658723153	0.62600344	
NEW.POVERI83	0.05005160	0.272573001	0.293326433	0.930222190	0.91755596	
NEW.PTRANP83	0.06600112	0.282339025	0.277239264	0.974440208	0.94686976	
POPAF83	-0.14380615	-0.177575758	-0.086715871	-0.319715479	-0.33458065	
WEGEN83	0.18516815	0.133733309	0.021633389	0.072730887	0.07298133	
NEW.PGRHAN83	NEW.PKENNI83	NEW.PKWART83	NEW.PLANDB83	NEW.PTRANP83	NEW.PINDUS83	
PSTAT80	-0.28616614	-0.28997020	-0.341429422	-0.330532653	-0.270283784	-0.34115883
RAIL80	0.08691728	0.09179039	0.101918770	0.061687477	0.096209484	0.10496025
PBOSGEB81	-0.11803475	-0.09901725	-0.033837582	0.174848914	-0.142741974	0.03769561
PBOSREC81	0.24358639	0.25965671	0.297315762	0.076303264	0.194642082	0.25595533
PDRGNTR81	0.01361431	0.05386202	0.088574902	0.002804066	-0.009677922	0.01757455
PNATNTR81	0.04856956	0.04278614	-0.007737662	-0.121176989	0.039714690	-0.04335384
POVOWTR81	0.57238449	0.53301320	0.491851429	0.237286978	0.550355815	0.41927435
PWON80.T1	0.06479695	0.06576014	0.097406158	0.058992017	0.066001122	0.14176318
PWON80.T2	0.26869158	0.28299782	0.302539539	0.142985963	0.282339025	0.29139180
PWON80.T345	0.25807499	0.31604226	0.291530534	0.103490436	0.277239264	0.22728406
NEW.PBOUW83	0.97072482	0.94169352	0.960684812	0.658723153	0.974440208	0.91651315
NEW.PDTHAN83	0.98217155	0.97751113	0.982606560	0.626003438	0.946869756	0.90562270
NEW.PGRHAN83	1.00000000	0.96936189	0.960182811	0.623372737	0.959129812	0.87302211
NEW.PINDUS83	0.87302211	0.83202523	0.875788978	0.689517089	0.868688434	1.00000000
NEW.PKENNI83	0.96936189	1.00000000	0.973406126	0.561135581	0.943312472	0.83202523
NEW.PKWART83	0.96018281	0.97340613	1.000000000	0.636767444	0.935970369	0.87578898
NEW.PLANDB83	0.62337274	0.56113558	0.636767444	1.000000000	0.583141587	0.68951709
NEW.POVERI83	0.92744983	0.93921164	0.911613371	0.536514015	0.963054328	0.81033360
NEW.PTRANP83	0.95912981	0.94331247	0.935970369	0.583141587	1.000000000	0.86868843
POPAF83	-0.29740289	-0.28399655	-0.330862742	-0.332034052	-0.270800951	-0.36993085
WEGEN83	0.06509668	0.06674060	0.075430402	0.046021856	0.066550887	0.06546784
	NEW.POVERI83	WEGEN83	POPAF83	POVOWTR81		
PSTAT80	-0.24498826	-0.092166290	0.43901915	-0.18203792		
RAIL80	0.08899215	0.074103084	0.11204022	0.03072444		
PBOSGEB81	-0.20349549	-0.015820721	0.12560923	-0.36464650		
PBOSREC81	0.19086148	0.018414704	0.18947875	-0.04250923		
PDRGNTR81	0.01443261	-0.010246538	0.02961686	-0.14564656		
PNATNTR81	0.02763528	-0.035038479	0.10426012	0.34817423		
POVOWTR81	0.55030885	0.025181970	0.12652560	1.00000000		
PWON80.T1	0.05005160	0.185168155	0.143806151	-0.02420118		
PWON80.T2	0.27257300	0.133733309	0.17757576	0.11659618		
PWON80.T345	0.29332643	0.021633389	0.08671587	0.11163813		
NEW.PBOUW83	0.93022219	0.072730887	0.31971548	0.54011792		
NEW.PDTHAN83	0.91755596	0.072981329	0.33458065	0.52777563		
NEW.PGRHAN83	0.92744983	0.065096681	0.29740289	0.57238449		
NEW.PINDUS83	0.81033360	0.065467837	0.36993085	0.41927435		
NEW.PKENNI83	0.93921164	0.066740595	0.28399655	0.53301320		

Ruimtelijke ontwikkelingen woningbouw 1980 - 1995

NEW.PKWART83	0.91161337	0.075430402	0.33086274	0.49185143
NEW.PLANDB83	0.53651401	0.046021856	0.33203405	0.23728698
NEW.POVERI83	1.00000000	0.061170199	0.23916178	0.55030885
NEW.PTRANP83	0.96305433	0.066550887	0.27080095	0.55035581
POPAP83	-0.23916178	-0.116174990	1.00000000	-0.12652560
WEGEN83	0.06117020	1.000000000	-0.11617499	0.0251

Bijlage 4 Stepwise regression analysis

De naamgeving van de variabelen in de Splus uitgevoerde analyses correspondeert niet geheel met de gehanteerde naamgeving in dit rapport. Hieronder volgt derhalve eerst een verklarend overzicht:

In rapport

Potentiaal arbeidsplaatsen landbouw 1983
Potentiaal arbeidsplaatsen industrie 1983
Potentiaal arbeidsplaatsen transportsector 1983
Potentiaal arbeidsplaatsen kennissector 1983
Potentiaal bosgebieden 1981
Potentiaal bosrecreatie 1981
Potentiaal droge natuurterreinen 1981
Potentiaal natte natuurterreinen 1981
Potentiaal watergebieden 1981
Aanwezigheid autosnelwegen 1983 in gridcel (ja/nee)
Aanwezigheid spoorwegen 1983 in gridcel (ja/nee)
Afstand tot dichtst bij gelegen op- en afrit 1983
Afstand tot dichtst bij gelegen NS-station 1980
Ligging in groeikern in 1980(?) (ja/nee)
Ligging in groeistad in 1980(?) (ja/nee)
Potentiaal woongebied eengezinswoningen 1980
Potentiaal woongebied meergezinswoningen 1980

In S-Plus

NEW.PLANDB83
NEW.PINDUS83
NEW.PTRANP83
NEW.PKENNI83
PBOSGEB81
PBOSREC81
PDRGNTR81
PNATNTR81
POVOWTR81
WEGEN83
RAIL80
POPAF83
PSTAT80
GROEIKRN
GROEISTD
P.EENGEZ80
PWON80.T345

RANDSTAD – WONINGTYPEN 1/2

Startset (f1rswt12.glm)

*** Generalized Linear Model ***

```
Call: glm(formula = ywt12.h ~ log(NEW.PLANDB83 + 1) + log(NEW.PINDUS83 + 1) + log(NEW.PTRANP83 + 1) + log(NEW.PKENNI83 + 1) + log(PBOSGEB81 + 1) + log(PBOSREC81 + 1) + log(PDRGNTR81 + 1) + log(PNATNTR81 + 1) + log(POVOWTR81 + 1) + WEGEN83 + RAIL80 + log(PSTAT80 + 1) + log(POPAF83 + 1) + GROEIKRN + log(P.EENGEZ80 + 1) + log(PWON80.T345 + 1), family = binomial(link = logit), data = totaalwonen80, weights = wywt12.h, subset = RANDSTAD == 1, na.action = na.omit, control = list(epsilon = 0.001, maxit = 50, trace = F))
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-26.95986	-4.900954	-2.395778	-0.8774457	104.7534

Coefficients:

	Value	Std. Error	t value
(Intercept)	-10.647005926	0.1207794334	-88.1524745
log(NEW.PLANDB83 + 1)	0.091911993	0.0077050542	11.9287926
log(NEW.PINDUS83 + 1)	0.391064937	0.0144493277	27.0645767
log(NEW.PTRANP83 + 1)	-0.490010600	0.0115611844	-42.3841176
log(NEW.PKENNI83 + 1)	0.296271645	0.0096242614	30.7838319
log(PBOSGEB81 + 1)	0.066350313	0.0018149172	36.5583153
log(PBOSREC81 + 1)	-0.045947972	0.0016866239	-27.2425715

Ruimtelijke ontwikkelingen woningbouw 1980 - 1995

log(PDRGNTR81 + 1)	-0.059617419	0.0013258213	-44.9664067
log(PNATNTR81 + 1)	-0.032804946	0.0031092404	-10.5507913
log(POVOWTR81 + 1)	0.003222355	0.0049927090	0.6454121
WEGEN83	-0.330422279	0.0036378354	-90.8293641
RAIL80	-0.557438297	0.0063312214	-88.0459338
log(PSTAT80 + 1)	-0.117433856	0.0013944693	-84.2140153
log(POPAF83 + 1)	0.057579001	0.0013261224	43.4190687
GROEIKRN	1.188813998	0.0044858525	265.0140620
log(P.EENGEZ80 + 1)	0.735797440	0.0013438584	547.5260043
log(PWON80.T345 + 1)	0.062004896	0.0008444312	73.4280063

(Dispersion Parameter for Binomial family taken to be 1)
 Null Deviance: 2228956 on 23390 degrees of freedom
 Residual Deviance: 1479139 on 23374 degrees of freedom
 Number of Fisher Scoring Iterations: 7

Correlation of Coefficients:

(Intercept) log(NEW.PLANDB83 + 1) log(NEW.PINDUS83 + 1)

log(NEW.PLANDB83 + 1)	-0.5951541		
log(NEW.PINDUS83 + 1)	-0.5777843	0.0777516	
log(NEW.PTRANP83 + 1)	0.5798134	-0.2719769	-0.6529897
log(NEW.PKENNI83 + 1)	-0.4716400	0.1823124	-0.0684571
log(PBOSGEB81 + 1)	-0.4036060	0.0797017	0.0090317
log(PBOSREC81 + 1)	0.4722774	-0.0616217	-0.0558637
log(PDRGNTR81 + 1)	-0.2097415	0.2382745	-0.0876547
log(PNATNTR81 + 1)	-0.5124416	0.7582412	0.1263337
log(POVOWTR81 + 1)	0.2909799	-0.3714707	-0.3727512
WEGEN83	-0.0525297	-0.0431318	0.0405199
RAIL80	-0.0712399	0.0041250	0.0106097
log(PSTAT80 + 1)	-0.2204992	-0.0128203	0.1475962
log(POPAF83 + 1)	-0.2543940	0.0215498	0.0966251
GROEIKRN	-0.1131721	0.0666301	0.2148245
log(P.EENGEZ80 + 1)	-0.0641952	0.0151681	-0.0596848
log(PWON80.T345 + 1)	0.0408515	0.0645989	-0.0123417

log(NEW.PTRANP83 + 1) log(NEW.PKENNI83 + 1)

log(NEW.PLANDB83 + 1)		
log(NEW.PINDUS83 + 1)		
log(NEW.PTRANP83 + 1)		
log(NEW.PKENNI83 + 1)	-0.5321495	
log(PBOSGEB81 + 1)	-0.2032521	0.4403123
log(PBOSREC81 + 1)	0.1714008	-0.5829923
log(PDRGNTR81 + 1)	0.0991115	0.1172518
log(PNATNTR81 + 1)	-0.1811174	0.1653350
log(POVOWTR81 + 1)	0.0858935	0.0641715
WEGEN83	-0.0302769	0.0395978
RAIL80	-0.0637086	0.0638209
log(PSTAT80 + 1)	-0.2004684	0.1351071
log(POPAF83 + 1)	-0.1379052	0.1921911
GROEIKRN	-0.2240224	0.0032933
log(P.EENGEZ80 + 1)	-0.0160097	0.0718740
log(PWON80.T345 + 1)	-0.0562107	-0.0516394
log(PBOSGEB81 + 1)		log(PBOSREC81 + 1)

log(NEW.PLANDB83 + 1)		
log(NEW.PINDUS83 + 1)		
log(NEW.PTRANP83 + 1)		
log(NEW.PKENNI83 + 1)		
log(PBOSGEB81 + 1)		
log(PBOSREC81 + 1)	-0.5121452	
log(PDRGNTR81 + 1)	-0.3272928	-0.3388212
log(PNATNTR81 + 1)	-0.1255950	0.0181013
log(POVOWTR81 + 1)	0.3160088	-0.0892133

Ruimtelijke ontwikkelingen woningbouw 1980 - 1995

WEGEN83	0.0479773	-0.0088290	
RAIL80	0.0119284	0.0117564	
log (PSTAT80 + 1)	0.1299628	-0.0433344	
log (POPAF83 + 1)	0.1088986	-0.0169986	
GROEIKRN	0.0703747	0.0486445	
log (P.EENGEZ80 + 1)	0.0459116	-0.0611096	
log (PWON80.T345 + 1)	0.0517990	0.0333156	
	log (PDRGNTR81 + 1)	log (PNATNTR81 + 1)	
log (NEW.PLANDB83 + 1)			
log (NEW.PINDUS83 + 1)			
log (NEW.PTRANP83 + 1)			
log (NEW.PKENNI83 + 1)			
log (PBOSGEB81 + 1)			
log (PBOSREC81 + 1)			
log (PDRGNTR81 + 1)			
log (PNATNTR81 + 1)	0.3642906		
log (POVOWTR81 + 1)	-0.1411356	-0.6701034	
WEGEN83	-0.0425761	-0.0133497	
RAIL80	-0.0179272	0.0016512	
log (PSTAT80 + 1)	-0.0869085	-0.0499180	
log (POPAF83 + 1)	-0.0538721	0.0538566	
GROEIKRN	-0.1773868	0.0394180	
log (P.EENGEZ80 + 1)	-0.0327568	-0.0244223	
log (PWON80.T345 + 1)	-0.0432080	0.0099689	
	log (POVOWTR81 + 1)	WEGEN83	RAIL80
log (NEW.PLANDB83 + 1)			
log (NEW.PINDUS83 + 1)			
log (NEW.PTRANP83 + 1)			
log (NEW.PKENNI83 + 1)			
log (PBOSGEB81 + 1)			
log (PBOSREC81 + 1)			
log (PDRGNTR81 + 1)			
log (PNATNTR81 + 1)			
log (POVOWTR81 + 1)			
WEGEN83	-0.0099725		
RAIL80	0.0227796	0.0059808	
log (PSTAT80 + 1)	0.0337196	0.0120005	0.4757131
log (POPAF83 + 1)	-0.0494079	0.2407504	0.0233592
GROEIKRN	-0.0775361	-0.0011271	0.0477016
log (P.EENGEZ80 + 1)	0.0420766	-0.1067371	-0.0201212
log (PWON80.T345 + 1)	0.0690612	0.0400453	0.0041220
	log (PSTAT80 + 1)	log (POPAF83 + 1)	GROEIKRN
log (NEW.PLANDB83 + 1)			
log (NEW.PINDUS83 + 1)			
log (NEW.PTRANP83 + 1)			
log (NEW.PKENNI83 + 1)			
log (PBOSGEB81 + 1)			
log (PBOSREC81 + 1)			
log (PDRGNTR81 + 1)			
log (PNATNTR81 + 1)			
log (POVOWTR81 + 1)			
WEGEN83			
RAIL80			
log (PSTAT80 + 1)			
log (POPAF83 + 1)	-0.0617814		
GROEIKRN	0.1320861	-0.0204156	
log (P.EENGEZ80 + 1)	0.1048627	0.0539165	0.0085503
log (PWON80.T345 + 1)	0.1375390	0.0296194	0.1130054
	log (P.EENGEZ80 + 1)		
log (NEW.PLANDB83 + 1)			
log (NEW.PINDUS83 + 1)			
log (NEW.PTRANP83 + 1)			
log (NEW.PKENNI83 + 1)			

Ruimtelijke ontwikkelingen woningbouw 1980 - 1995

```

log(PBOSGEB81 + 1)
log(PBOSREC81 + 1)
log(PDRGNTR81 + 1)
log(PNATNTR81 + 1)
log(POVOWTR81 + 1)
  WEGEN83
  RAIL80
log(PSTAT80 + 1)
log(POPAF83 + 1)
  GROEIKRN
log(P.EENGEZ80 + 1)
log(PWON80.T345 + 1) -0.4107154

```

Analysis of Deviance Table
 Binomial model
 Response: ywt12.h

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid. Dev
NULL			23390	2228956
log(NEW.PLANDB83 + 1)	1	9937.6	23389	2219019
log(NEW.PINDUS83 + 1)	1	45819.0	23388	2173200
log(NEW.PTRANP83 + 1)	1	86.7	23387	2173113
log(NEW.PKENNI83 + 1)	1	7866.4	23386	2165246
log(PBOSGEB81 + 1)	1	3021.4	23385	2162225
log(PBOSREC81 + 1)	1	330.3	23384	2161895
log(PDRGNTR81 + 1)	1	665.6	23383	2161229
log(PNATNTR81 + 1)	1	1364.5	23382	2159865
log(POVOWTR81 + 1)	1	1818.3	23381	2158046
WEGEN83	1	1378.6	23380	2156668
RAIL80	1	607.3	23379	2156061
log(PSTAT80 + 1)	1	71786.7	23378	2084274
log(POPAF83 + 1)	1	135.9	23377	2084138
GROEIKRN	1	37341.8	23376	2046796
log(P.EENGEZ80 + 1)	1	562392.8	23375	1484403
log(PWON80.T345 + 1)	1	5264.4	23374	1479139

Set 2 (s1rswt12.glm)

*** Generalized Linear Model ***

```

Call: glm(formula = YWT12.H ~ log(NEW.PINDUS + 1) + log(NEW.PTRANP + 1)
+log(NEW.PKENNI + 1) + log(PBOSGEB81 + 1) + log(PBOSREC81 + 1) +
log(PDRGNTR81 + 1) + WEGEN83 + RAIL80 + log(PSTAT80 + 1) + log(POPAF83 + 1)
+ GROEIKRN +log(P.EENGEZ80 + 1) + log(PWON80.T34 + 1), family =
binomial(link = logit), data = totaalwonen80, weights = WYWT12.H, subset =
RANDSTAD == 1, na.action = na.omit, control = list(epsilon = 0.001, maxit =
50,
trace = F))

```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-26.87188	-4.894456	-2.407648	-0.8873056	105.4756

Coefficients:

	Value	Std. Error	t value
(Intercept)	-10.07753101	0.0953473093	-105.69287
log(NEW.PINDUS + 1)	0.31783310	0.0129886412	24.47008
log(NEW.PTRANP + 1)	-0.43252638	0.0109926721	-39.34679
log(NEW.PKENNI + 1)	0.32222615	0.0091054082	35.38844
log(PBOSGEB81 + 1)	0.05702387	0.0016623804	34.30254
log(PBOSREC81 + 1)	-0.04191147	0.0016602198	-25.24453
log(PDRGNTR81 + 1)	-0.04827763	0.0012168035	-39.67578
WEGEN83	-0.32605960	0.0036326510	-89.75803

Ruimtelijke ontwikkelingen woningbouw 1980 - 1995

```

      RAIL80 -0.55460860 0.0063297581 -87.61924
log(PSTAT80 + 1) -0.12009716 0.0013880437 -86.52261
log(POPAF83 + 1) 0.05926902 0.0013258679 44.70205
      GROEIKRN 1.17641729 0.0044505343 264.33170
log(P.EENGEZ80 + 1) 0.73382925 0.0013396709 547.76830
log(PWON80.T34 + 1) 0.06109098 0.0008381663 72.88647

```

(Dispersion Parameter for Binomial family taken to be 1)

Null Deviance: 2228956 on 23390 degrees of freedom
Residual Deviance: 1480501 on 23377 degrees of freedom
Number of Fisher Scoring Iterations: 7

Analysis of Deviance Table

Binomial model

Response: YWT12.H

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid. Dev
NULL			23390	2228956
log(NEW.PINDUS + 1)	1	55718.3	23389	2173238
log(NEW.PTRANP + 1)	1	113.5	23388	2173124
log(NEW.PKENNI + 1)	1	7874.3	23387	2165250
log(PBOSGEB81 + 1)	1	2962.6	23386	2162287
log(PBOSREC81 + 1)	1	368.8	23385	2161919
log(PDRGNTR81 + 1)	1	663.7	23384	2161255
WEGEN83	1	1383.3	23383	2159872
RAIL80	1	720.4	23382	2159151
log(PSTAT80 + 1)	1	72733.0	23381	2086418
log(POPAF83 + 1)	1	177.7	23380	2086241
GROEIKRN	1	35244.1	23379	2050996
log(P.EENGEZ80 + 1)	1	565311.7	23378	1485685
log(PWON80.T34 + 1)	1	5183.3	23377	1480501

Set 3 (s2arswt12.glm)

*** Generalized Linear Model ***

Call: glm(formula = YWT12.H ~ log(NEW.PTRANP + 1) + log(NEW.PKENNI + 1) + log(PBOSGEB81 + 1) + log(PBOSREC81 + 1) + WEGEN83 + RAIL80 + log(PSTAT80 + 1) + log(POPAF83 + 1) + GROEIKRN + log(P.EENGEZ80 + 1) + log(PWON80.T34 + 1), family = binomial(link = logit), data = totaalwonen80, weights = WYWT12.H, subset = RANDSTAD == 1, na.action = na.omit, control = list(epsilon = 0.001, maxit = 50, trace = F))

Deviance Residuals:

Min	1Q	Median	3Q	Max
-28.395	-4.897149	-2.407148	-0.8965933	105.708

Coefficients:

	Value	Std. Error	t value
(Intercept)	-8.94034141	0.0651313583	-137.26631
log(NEW.PTRANP + 1)	-0.19942170	0.0075589551	-26.38218
log(NEW.PKENNI + 1)	0.34582376	0.0090093040	38.38518
log(PBOSGEB81 + 1)	0.02938814	0.0015207333	19.32498
log(PBOSREC81 + 1)	-0.06263616	0.0014968302	-41.84587
WEGEN83	-0.33490864	0.0036301511	-92.25749
RAIL80	-0.56475299	0.0063474557	-88.97313
log(PSTAT80 + 1)	-0.12881960	0.0013543047	-95.11863
log(POPAF83 + 1)	0.05239254	0.0013013703	40.25952
GROEIKRN	1.12729284	0.0042988079	262.23382
log(P.EENGEZ80 + 1)	0.73427807	0.0013384304	548.61132
log(PWON80.T34 + 1)	0.05868573	0.0008365068	70.15571

(Dispersion Parameter for Binomial family taken to be 1)

Null Deviance: 2228956 on 23390 degrees of freedom

Ruimtelijke ontwikkelingen woningbouw 1980 - 1995

Residual Deviance: 1482513 on 23379 degrees of freedom
 Number of Fisher Scoring Iterations: 7
 Analysis of Deviance Table
 Binomial model
 Response: YWT12.H

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid. Dev
NULL			23390	2228956
log(NEW.PTRANP + 1)	1	49345.6	23389	2179611
log(NEW.PKENNI + 1)	1	10980.4	23388	2168630
log(PBOSGEB81 + 1)	1	2954.9	23387	2165675
log(PBOSREC81 + 1)	1	722.8	23386	2164952
WEGEN83	1	1407.0	23385	2163545
RAIL80	1	897.5	23384	2162648
log(PSTAT80 + 1)	1	74332.4	23383	2088316
log(POPAF83 + 1)	1	143.7	23382	2088172
GROEIKRN	1	35298.6	23381	2052873
log(P.EENGEZ80 + 1)	1	565556.1	23380	1487317
log(PWON80.T34 + 1)	1	4803.7	23379	1482513

Set 4 (s2rswt12.glm)

Call: glm(formula = YWT12.H ~ log(NEW.PKENNI + 1) + log(PBOSGEB81 + 1) +
 WEGEN83 + RAIL80 + log(PSTAT80 + 1) + log(POPAF83 + 1) + GROEIKRN +
 log(P.EENGEZ80 + 1) + log(PWON80.T34 + 1), family = binomial(link = logit),
 data = totaalwonen80, weights = WYWT12.H, subset = RANDSTAD == 1, na.action
 = na.omit, control = list(epsilon = 0.001, maxit = 50, trace = F))

Deviance Residuals:

Min	1Q	Median	3Q	Max
-29.27951	-4.909617	-2.40446	-0.9140801	105.3229

Coefficients:

	Value	Std. Error	t value
(Intercept)	-7.164480688	0.0474477784	-150.997179
log(NEW.PKENNI + 1)	0.005097684	0.0036684948	1.389585
log(PBOSGEB81 + 1)	-0.013983713	0.0010458694	-13.370420
WEGEN83	-0.340257525	0.0036302728	-93.727811
RAIL80	-0.570871355	0.0063416073	-90.019979
log(PSTAT80 + 1)	-0.134439299	0.0013376876	-100.501265
log(POPAF83 + 1)	0.047661736	0.0012863945	37.050638
GROEIKRN	1.119794243	0.0042797994	261.646429
log(P.EENGEZ80 + 1)	0.728845302	0.0013313489	547.448762
log(PWON80.T34 + 1)	0.059091991	0.0008349294	70.774835

(Dispersion Parameter for Binomial family taken to be 1)

Null Deviance: 2228956 on 23390 degrees of freedom
 Residual Deviance: 1484542 on 23381 degrees of freedom
 Number of Fisher Scoring Iterations: 7
 Analysis of Deviance Table
 Binomial model
 Response: YWT12.H

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid. Dev
NULL			23390	2228956
log(NEW.PKENNI + 1)	1	60065.3	23389	2168891
log(PBOSGEB81 + 1)	1	2610.5	23388	2166280
WEGEN83	1	1521.8	23387	2164759
RAIL80	1	992.0	23386	2163767
log(PSTAT80 + 1)	1	73266.9	23385	2090500
log(POPAF83 + 1)	1	96.4	23384	2090403
GROEIKRN	1	36077.1	23383	2054326

```
log(P.EENGEZ80 + 1) 1 564895.4 23382 1489431
log(PWON80.T34 + 1) 1 4888.8 23381 1484542
```

Set 5 (s3rswt12.glm)

*** Generalized Linear Model ***

```
Call: glm(formula = YWT12.H ~ log(NEW.PINDUS + 1) + log(PBOSGEB81 + 1) +
WEGEN83 + RAIL80 + log(PSTAT80 + 1) + log(POPAF83 + 1) + GROEIKRN +
log(P.EENGEZ80 + 1) + log(PWON80.T34 + 1), family = binomial(link = logit),
data = totaalwonen80, weights = WYWT12.H, subset = RANDSTAD == 1, na.action
= na.omit, control = list(epsilon = 0.001, maxit = 50, trace = F))
```

Deviance Residuals:

```
      Min       1Q   Median       3Q      Max
-29.29999 -4.907175 -2.404934 -0.9143482 105.3396
```

Coefficients:

```
              Value Std. Error t value
(Intercept) -7.0980643063 0.0638750773 -111.12415988
log(NEW.PINDUS + 1) -0.0005035823 0.0053281285 -0.09451392
log(PBOSGEB81 + 1) -0.0143598132 0.0010507863 -13.66577906
      WEGEN83 -0.3406331519 0.0036288629 -93.86773738
      RAIL80 -0.5713891249 0.0063345030 -90.20267644
log(PSTAT80 + 1) -0.1347005877 0.0013361191 -100.81480296
log(POPAF83 + 1) 0.0470976196 0.0012643343 37.25092324
      GROEIKRN 1.1203368159 0.0042655300 262.64891473
log(P.EENGEZ80 + 1) 0.7289143072 0.0013343948 546.25084757
log(PWON80.T34 + 1) 0.0593350900 0.0008291675 71.55983722
```

(Dispersion Parameter for Binomial family taken to be 1)

Null Deviance: 2228956 on 23390 degrees of freedom

Residual Deviance: 1484544 on 23381 degrees of freedom

Number of Fisher Scoring Iterations: 7

Analysis of Deviance Table

Binomial model

Response: YWT12.H

Terms added sequentially (first to last)

```
              Df Deviance Resid. Df Resid. Dev
NULL                23390 2228956
log(NEW.PINDUS + 1) 1 55718.3 23389 2173238
log(PBOSGEB81 + 1) 1 3286.7 23388 2169951
      WEGEN83 1 1668.7 23387 2168283
      RAIL80 1 644.7 23386 2167638
log(PSTAT80 + 1) 1 76301.1 23385 2091337
log(POPAF83 + 1) 1 308.9 23384 2091028
      GROEIKRN 1 39564.9 23383 2051463
log(P.EENGEZ80 + 1) 1 561925.5 23382 1489537
log(PWON80.T34 + 1) 1 4993.6 23381 1484544
```

Set 6 (s4rswt12.glm)

*** Generalized Linear Model ***

```
Call: glm(formula = YWT12.H ~ log(NEW.PTRANP + 1) + log(PBOSGEB81 + 1) +
WEGEN83 + RAIL80 + log(PSTAT80 + 1) + log(POPAF83 + 1) + GROEIKRN +
log(P.EENGEZ80 + 1) + log(PWON80.T34 + 1), family = binomial(link = logit),
data = totaalwonen80, weights = WYWT12.H, subset = RANDSTAD == 1, na.action
= na.omit, control = list(epsilon = 0.001, maxit = 50, trace = F))
```

Deviance Residuals:

```
      Min       1Q   Median       3Q      Max
-29.35973 -4.908076 -2.412299 -0.915082 105.4174
```

Coefficients:

	Value	Std. Error	t value
(Intercept)	-6.78394983	0.0430855754	-157.452924
log(NEW.PTRANP + 1)	-0.03153504	0.0038359198	-8.220985
log(PBOSGEB81 + 1)	-0.01667004	0.0010505172	-15.868415
WEGEN83	-0.34217516	0.0036278592	-94.318754
RAIL80	-0.57198733	0.0063373457	-90.256608
log(PSTAT80 + 1)	-0.13525064	0.0013264652	-101.963204
log(POPAF83 + 1)	0.04494428	0.0012510325	35.925745
GROEIKRN	1.12455734	0.0042959308	261.772685
log(P.EENGEZ80 + 1)	0.72944769	0.0013325306	547.415341
log(PWON80.T34 + 1)	0.06060816	0.0008335763	72.708595

(Dispersion Parameter for Binomial family taken to be 1)

Null Deviance: 2228956 on 23390 degrees of freedom

Residual Deviance: 1484477 on 23381 degrees of freedom

Number of Fisher Scoring Iterations: 7

Set 7 (s5rswt12.glm)

*** Generalized Linear Model ***

Call: glm(formula = YWT12.H ~ log(NEW.PINDUS + 1) + log(PBOSGEB81 + 1) + log(PNATNTR81 + 1) + log(POVOWTR81 + 1) + WEGEN83 + RAIL80 + log(PSTAT80 + 1) + log(POPAF83 + 1) + GROEIKRN + log(P.EENGEZ80 + 1) + log(PWON80.T34 + 1), family = binomial(link = logit), data = totaalwonen80, weights = WYWT12.H, subset = RANDSTAD == 1, na.action = na.omit, control = list(epsilon = 0.001, maxit = 50, trace = F))

Deviance Residuals:

Min	1Q	Median	3Q	Max
-29.27766	-4.897611	-2.406516	-0.9151033	105.0123

Coefficients:

	Value	Std. Error	t value
(Intercept)	-6.95960041	0.0668761160	-104.067055
log(NEW.PINDUS + 1)	0.01275761	0.0070187402	1.817649
log(PBOSGEB81 + 1)	-0.01929367	0.0011666531	-16.537624
log(PNATNTR81 + 1)	-0.01418026	0.0018346576	-7.729102
log(POVOWTR81 + 1)	-0.01697887	0.0045283986	-3.749420
WEGEN83	-0.34104245	0.0036316800	-93.907626
RAIL80	-0.57197915	0.0063348458	-90.290934
log(PSTAT80 + 1)	-0.13401429	0.0013385042	-100.122427
log(POPAF83 + 1)	0.04748449	0.0012685707	37.431492
GROEIKRN	1.11978242	0.0042784906	261.723706
log(P.EENGEZ80 + 1)	0.72914358	0.0013356665	545.902418
log(PWON80.T34 + 1)	0.05834496	0.0008356958	69.816024

(Dispersion Parameter for Binomial family taken to be 1)

Null Deviance: 2228956 on 23390 degrees of freedom

Residual Deviance: 1484342 on 23379 degrees of freedom

Number of Fisher Scoring Iterations: 7

Set 8 (s6rswt12.glm)

*** Generalized Linear Model ***

Call: glm(formula = YWT12.H ~ log(PBOSGEB81 + 1) + log(PNATNTR81 + 1) + log(POVOWTR81 + 1) + WEGEN83 + RAIL80 + log(PSTAT80 + 1) + log(POPAF83 + 1) + GROEIKRN + log(P.EENGEZ80 + 1) + log(PWON80.T34 + 1), family = binomial(link = logit), data = totaalwonen80, weights = WYWT12.H, subset =

Ruimtelijke ontwikkelingen woningbouw 1980 - 1995

```
RANDSTAD == 1, na.action = na.omit, control = list(epsilon = 0.001, maxit = 50, trace = F))
```

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-29.32992	-4.898792	-2.406978	-0.9144387	105.0415

Coefficients:

	Value	Std. Error	t value
(Intercept)	-6.85752106	0.036283612	-188.997753
log(PBOSGEB81 + 1)	-0.01916343	0.001163512	-16.470334
log(PNATNTR81 + 1)	-0.01572810	0.001623639	-9.686945
log(POVOWTR81 + 1)	-0.01162891	0.003439022	-3.381460
WEGEN83	-0.34152371	0.003622290	-94.283918
RAIL80	-0.57205245	0.006335930	-90.287053
log(PSTAT80 + 1)	-0.13421356	0.001333605	-100.639647
log(POPAF83 + 1)	0.04695299	0.001232974	38.081085
GROEIKRN	1.11964235	0.004277905	261.726787
log(P.EENGEZ80 + 1)	0.72933048	0.001331873	547.597616
log(PWON80.T34 + 1)	0.05863817	0.000819973	71.512317

(Dispersion Parameter for Binomial family taken to be 1)

Null Deviance: 2228956 on 23390 degrees of freedom

Residual Deviance: 1484345 on 23380 degrees of freedom

Number of Fisher Scoring Iterations: 7

Set 9 (s7rswt12.glm)

*** Generalized Linear Model ***

```
Call: glm(formula = YWT12.H ~ log(PBOSGEB81 + 1) + log(PNATNTR81 + 1) +  
WEGEN83 + RAIL80 + log(PSTAT80 + 1) + log(POPAF83 + 1) + GROEIKRN + log(  
P.EENGEZ80 + 1) + log(PWON80.T34 + 1), family = binomial(link =  
logit), data = totaalwonen80, weights = WYWT12.H, subset = RANDSTAD == 1,  
na.action = na.omit, control = list(epsilon = 0.001, maxit = 50, trace = F))
```

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-29.35398	-4.897279	-2.407198	-0.9145256	105.0484

Coefficients:

	Value	Std. Error	t value
(Intercept)	-6.95657341	0.0214228095	-324.72741
log(PBOSGEB81 + 1)	-0.01738684	0.0010407921	-16.70539
log(PNATNTR81 + 1)	-0.01873070	0.0013592626	-13.78004
WEGEN83	-0.34136491	0.0036216613	-94.25644
RAIL80	-0.57130621	0.0063297426	-90.25742
log(PSTAT80 + 1)	-0.13380112	0.0013284514	-100.71962
log(POPAF83 + 1)	0.04737694	0.0012275935	38.59334
GROEIKRN	1.11859592	0.0042652326	262.25907
log(P.EENGEZ80 + 1)	0.72922855	0.0013316696	547.60471
log(PWON80.T34 + 1)	0.05859544	0.0008198482	71.47108

(Dispersion Parameter for Binomial family taken to be 1)

Null Deviance: 2228956 on 23390 degrees of freedom

Residual Deviance: 1484356 on 23381 degrees of freedom

Number of Fisher Scoring Iterations: 7

Set 10 (s8rswt12.glm)

*** Generalized Linear Model ***

```
Call: glm(formula = YWT12.H ~ log(PBOSGEB81 + 1) + log(POVOWTR81 + 1) +  
WEGEN83 + RAIL80 + log(PSTAT80 + 1) + log(POPAF83 + 1) + GROEIKRN + log(
```

Ruimtelijke ontwikkelingen woningbouw 1980 - 1995

```
P.EENGEZ80 + 1) + log(PWON80.T34 + 1), family = binomial(link = logit),
data = totaalwonen80, weights = WYWT12.H, subset = RANDSTAD == 1, na.action
= na.omit, control = list(epsilon = 0.001, maxit = 50, trace = F))
```

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-29.26107	-4.911326	-2.40947	-0.9164918	105.2022

Coefficients:

	Value	Std. Error	t value
(Intercept)	-6.78938355	0.0355445671	-191.01044
log(PBOSGEB81 + 1)	-0.02010832	0.0011555940	-17.40085
log(POVOWTR81 + 1)	-0.02980984	0.0028738785	-10.37268
WEGEN83	-0.34134218	0.0036229958	-94.21545
RAIL80	-0.57329629	0.0063384115	-90.44794
log(PSTAT80 + 1)	-0.13539715	0.0013263800	-102.08022
log(POPAF83 + 1)	0.04611923	0.0012281572	37.55157
GROEIKRN	1.12227027	0.0042686119	262.91223
log(P.EENGEZ80 + 1)	0.72930504	0.0013315138	547.72622
log(PWON80.T34 + 1)	0.05914035	0.0008184104	72.26246

(Dispersion Parameter for Binomial family taken to be 1)

Null Deviance: 2228956 on 23390 degrees of freedom

Residual Deviance: 1484438 on 23381 degrees of freedom

Number of Fisher Scoring Iterations: 7

Set 11 (s9rswt12.glm)

*** Generalized Linear Model ***

```
Call: glm(formula = YWT12.H ~ log(PBOSREC81 + 1) + log(PNATNTR81 + 1) +
WEGEN83 + RAIL80 + log(PSTAT80 + 1) + log(POPAF83 + 1) + GROEIKRN +
log(P.EENGEZ80 + 1) + log(PWON80.T34 + 1), family = binomial(link =
logit), data = totaalwonen80, weights = WYWT12.H, subset = RANDSTAD == 1,
na.action = na.omit, control = list(epsilon = 0.001, maxit = 50, trace = F))
```

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-29.35416	-4.904739	-2.422981	-0.9157418	105.3075

Coefficients:

	Value	Std. Error	t value
(Intercept)	-6.76458324	0.0216591803	-312.31945
log(PBOSREC81 + 1)	-0.03032903	0.0009132269	-33.21083
log(PNATNTR81 + 1)	-0.03024642	0.0014211850	-21.28253
WEGEN83	-0.34545856	0.0036288708	-95.19726
RAIL80	-0.58246815	0.0063547985	-91.65800
log(PSTAT80 + 1)	-0.13643681	0.0013277548	-102.75754
log(POPAF83 + 1)	0.03964133	0.0012398352	31.97307
GROEIKRN	1.12578956	0.0042682803	263.75717
log(P.EENGEZ80 + 1)	0.73196086	0.0013356629	548.01318
log(PWON80.T34 + 1)	0.06165853	0.0008201107	75.18318

(Dispersion Parameter for Binomial family taken to be 1)

Null Deviance: 2228956 on 23390 degrees of freedom

Residual Deviance: 1483564 on 23381 degrees of freedom

Number of Fisher Scoring Iterations: 7

Set 12 (s10rswt12.glm)

*** Generalized Linear Model ***


```
Call: glm(formula = YWT12.H ~ log(PBOSREC81 + 1) + WEGEN83 + RAIL80 +
log(PSTAT80 + 1) + log(POPAF83 + 1) + GROEIKRN + log(P.EENGEZ80 + 1) +
log(PWON80.T34 + 1), family = binomial(link = logit), data = totaalwonen80,
weights = WYWT12.H, subset = RANDSTAD == 1, na.action = na.omit, control =
list(epsilon = 0.001, maxit = 50, trace = F))
```

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-29.27684	-4.906857	-2.423824	-0.9161102	105.6747

Coefficients:

	Value	Std. Error	t value
(Intercept)	-7.02037990	0.0180196877	-389.59498
log(PBOSREC81 + 1)	-0.02344555	0.0008537952	-27.46039
WEGEN83	-0.34348735	0.0036271368	-94.69931
RAIL80	-0.58014368	0.0063534640	-91.31140
log(PSTAT80 + 1)	-0.13722021	0.0013262099	-103.46795
log(POPAF83 + 1)	0.04098646	0.0012405023	33.04021
GROEIKRN	1.12648636	0.0042664184	264.03560
log(P.EENGEZ80 + 1)	0.73087098	0.0013336982	548.00326
log(PWON80.T34 + 1)	0.06201094	0.0008194532	75.67355

(Dispersion Parameter for Binomial family taken to be 1)

Null Deviance: 2228956 on 23390 degrees of freedom

Residual Deviance: 1484012 on 23382 degrees of freedom

Number of Fisher Scoring Iterations: 7

RANDSTAD – WONINGTYPEN 3/4/5

Startset

*** Generalized Linear Model ***

```
Call: glm(formula = YWT345.H ~ log(NEW.PLANDB + 1) + log(NEW.PINDUS + 1) +
log(NEW.PTRANP + 1) + log(NEW.PKENNI + 1) + log(PBOSGEB81 + 1) +
log(PBOSREC81 + 1) + log(PDRGNTR81 + 1) + log(PNATNTR81 + 1) + log(POVOWTR81
+ 1) + WEGEN83 + RAIL80 + log(PSTAT80 + 1) + log(POPAF83 + 1) + GROEIKRN +
log(P.EENGEZ80 + 1) + log(PWON80.T34 + 1), family = binomial(link = logit),
data = totaalwonen80, weights = WYWT345.H, subset = RANDSTAD == 1, na.action
= na.omit, control = list(epsilon = 0.001, maxit = 50, trace = F))
```

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-28.00093	-2.672253	-1.481426	-0.7918773	92.8921

Coefficients:

	Value	Std. Error	t value
(Intercept)	-17.982144444	0.220996132	-81.368594
log(NEW.PLANDB + 1)	-0.201591402	0.012734024	-15.830927
log(NEW.PINDUS + 1)	-0.334731785	0.024184939	-13.840505
log(NEW.PTRANP + 1)	-0.476189477	0.018960102	-25.115344
log(NEW.PKENNI + 1)	1.560783790	0.013782149	113.246761
log(PBOSGEB81 + 1)	0.177616363	0.003283719	54.090004
log(PBOSREC81 + 1)	0.008017103	0.003283340	2.441753
log(PDRGNTR81 + 1)	-0.118244220	0.002378176	-49.720541
log(PNATNTR81 + 1)	-0.024516874	0.004920435	-4.982664
log(POVOWTR81 + 1)	0.132808378	0.007323752	18.133926
WEGEN83	-0.350794670	0.005218982	-67.215151
RAIL80	0.074881512	0.006415410	11.672131
log(PSTAT80 + 1)	-0.110975640	0.001424512	-77.904295
log(POPAF83 + 1)	0.122085764	0.002032324	60.072003
GROEIKRN	0.837673437	0.007544155	111.036088
log(P.EENGEZ80 + 1)	0.564990915	0.002251317	250.960174

Ruimtelijke ontwikkelingen woningbouw 1980 - 1995

log(PWON80.T34 + 1) 0.389297719 0.001181343 329.538138

(Dispersion Parameter for Binomial family taken to be 1)

Null Deviance: 1542588 on 23390 degrees of freedom

Set 2 (s1rswt345.glm)

*** Generalized Linear Model ***

Call: glm(formula = YWT345.H ~ log(NEW.PINDUS + 1) + log(NEW.PTRANP + 1) + log(NEW.PKENNI + 1) + log(PBOSGEB81 + 1) + log(PBOSREC81 + 1) + log(PDRGNTR81 + 1) + WEGEN83 + RAIL80 + log(PSTAT80 + 1) + log(POPAF83 + 1) + GROEIKRN + log(P.EENGEZ80 + 1) + log(PWON80.T34 + 1), family = binomial(link = logit), data = totaalwonen80, weights = WYWT345.H, subset = RANDSTAD == 1, na.action = na.omit, control = list(epsilon = 0.001, maxit = 50, trace = F))

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-28.79852	-2.673194	-1.489228	-0.7903958	92.48905

Coefficients:

	Value	Std. Error	t value
(Intercept)	-20.409879486	0.183387667	-111.2936318
log(NEW.PINDUS + 1)	-0.039885808	0.022908083	-1.7411238
log(NEW.PTRANP + 1)	-0.703371909	0.017467892	-40.2665599
log(NEW.PKENNI + 1)	1.624840723	0.013224087	122.8697808
log(PBOSGEB81 + 1)	0.197387883	0.003118285	63.3001441
log(PBOSREC81 + 1)	0.002826738	0.003282718	0.8610968
log(PDRGNTR81 + 1)	-0.149414374	0.002060406	-72.5169441
WEGEN83	-0.349344267	0.005210477	-67.0465018
RAIL80	0.068232037	0.006414134	10.6377622
log(PSTAT80 + 1)	-0.110740997	0.001419678	-78.0043290
log(POPAF83 + 1)	0.127094993	0.002045315	62.1395744
GROEIKRN	0.848670890	0.007544427	112.4897815
log(P.EENGEZ80 + 1)	0.564285177	0.002252554	250.5090706
log(PWON80.T34 + 1)	0.388998211	0.001165132	333.8662190

(Dispersion Parameter for Binomial family taken to be 1)

Null Deviance: 1542588 on 23390 degrees of freedom

Residual Deviance: 748106.8 on 23377 degrees of freedom

Number of Fisher Scoring Iterations: 9

Analysis of Deviance Table

Binomial model

Response: YWT345.H

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid. Dev
NULL			23390	1542588
log(NEW.PINDUS + 1)	1	171248.7	23389	1371340
log(NEW.PTRANP + 1)	1	12581.6	23388	1358758
log(NEW.PKENNI + 1)	1	57519.8	23387	1301238
log(PBOSGEB81 + 1)	1	6904.2	23386	1294334
log(PBOSREC81 + 1)	1	3.0	23385	1294331
log(PDRGNTR81 + 1)	1	86.2	23384	1294245
WEGEN83	1	819.8	23383	1293425
RAIL80	1	21658.5	23382	1271767
log(PSTAT80 + 1)	1	62125.7	23381	1209641
log(POPAF83 + 1)	1	542.5	23380	1209098
GROEIKRN	1	71.1	23379	1209027
log(P.EENGEZ80 + 1)	1	334715.4	23378	874312
log(PWON80.T34 + 1)	1	126205.0	23377	748107

Set 3 (s2rswt345.glm)

*** Generalized Linear Model ***

Call: glm(formula = YWT345.H ~ log(NEW.PKENNI + 1) + log(PBOSGEB81 + 1) + WEGEN83 + RAIL80 + log(PSTAT80 + 1) + log(POPAF83 + 1) + GROEIKRN + log(P.EENGEZ80 + 1) + log(PWON80.T34 + 1), family = binomial(link = logit), data = totaalwonen80, weights = WYWT345.H, subset = RANDSTAD == 1, na.action = na.omit, control = list(epsilon = 0.001, maxit = 50, trace = F))

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-29.58275	-2.672666	-1.519108	-0.84853	91.03512

Coefficients:

	Value	Std. Error	t value
(Intercept)	-20.35237162	0.079613088	-255.641028
log(NEW.PKENNI + 1)	0.99620545	0.006225301	160.025255
log(PBOSGEB81 + 1)	0.06367019	0.001770735	35.956921
WEGEN83	-0.36310434	0.005216239	-69.610373
RAIL80	0.04501650	0.006396852	7.037289
log(PSTAT80 + 1)	-0.12140934	0.001389668	-87.365740
log(POPAF83 + 1)	0.12231780	0.001996785	61.257360
GROEIKRN	0.73839239	0.007311624	100.988833
log(P.EENGEZ80 + 1)	0.53694752	0.002217596	242.130402
log(PWON80.T34 + 1)	0.38642018	0.001164921	331.713675

(Dispersion Parameter for Binomial family taken to be 1)
 Null Deviance: 1542588 on 23390 degrees of freedom
 Residual Deviance: 755992.6 on 23381 degrees of freedom
 Number of Fisher Scoring Iterations: 9

Set 4 (s3rswt345.glm)

*** Generalized Linear Model ***

Call: glm(formula = YWT345.H ~ log(NEW.PINDUS + 1) + log(PBOSGEB81 + 1) + WEGEN83 + log(PSTAT80 + 1) + log(POPAF83 + 1) + GROEIKRN + log(P.EENGEZ80 + 1) + log(PWON80.T34 + 1), family = binomial(link = logit), data = totaalwonen80, weights = WYWT345.H, subset = RANDSTAD == 1, na.action = na.omit, control = list(epsilon = 0.001, maxit = 50, trace = F))

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-29.80902	-2.705865	-1.657008	-0.9874927	92.19016

Coefficients:

	Value	Std. Error	t value
(Intercept)	-17.34388563	0.103040503	-168.321049
log(NEW.PINDUS + 1)	0.78417704	0.008731548	89.809622
log(PBOSGEB81 + 1)	0.01386458	0.001756965	7.891211
WEGEN83	-0.40818851	0.005236141	-77.955987
log(PSTAT80 + 1)	-0.13131755	0.001190249	-110.327819
log(POPAF83 + 1)	0.07831506	0.001827545	42.852611
GROEIKRN	0.71662768	0.007310789	98.023296
log(P.EENGEZ80 + 1)	0.50566354	0.002184814	231.444648
log(PWON80.T34 + 1)	0.42741369	0.001159945	368.477647

(Dispersion Parameter for Binomial family taken to be 1)
 Null Deviance: 1542588 on 23390 degrees of freedom
 Residual Deviance: 775733.5 on 23382 degrees of freedom
 Number of Fisher Scoring Iterations: 9

Set 5 (s4rswt345.glm)

*** Generalized Linear Model ***

Call: glm(formula = YWT345.H ~ log(NEW.PKENNI + 1) + log(PBOSGEB81 + 1) + WEGEN83 + log(PSTAT80 + 1) + log(POPAF83 + 1) + GROEIKRN + log(P.EENGEZ80 + 1) + log(PWON80.T34 + 1), family = binomial(link = logit), data = totaalwonen80, weights = WYWT345.H, subset = RANDSTAD == 1, na.action = na.omit, control = list(epsilon = 0.001, maxit = 50, trace = F))

Deviance Residuals:

Min	1Q	Median	3Q	Max
-29.57799	-2.674263	-1.519547	-0.8476288	90.88825

Coefficients:

	Value	Std. Error	t value
(Intercept)	-20.31552319	0.079465095	-255.65342
log(NEW.PKENNI + 1)	0.99631726	0.006226768	160.00551
log(PBOSGEB81 + 1)	0.06432879	0.001768411	36.37661
WEGEN83	-0.36414939	0.005213673	-69.84508
log(PSTAT80 + 1)	-0.12624009	0.001206852	-104.60280
log(POPAF83 + 1)	0.12235766	0.001997121	61.26701
GROEIKRN	0.73664999	0.007306630	100.81939
log(P.EENGEZ80 + 1)	0.53711828	0.002217199	242.25080
log(PWON80.T34 + 1)	0.38656817	0.001164770	331.88380

(Dispersion Parameter for Binomial family taken to be 1)

Null Deviance: 1542588 on 23390 degrees of freedom

Residual Deviance: 756041.4 on 23382 degrees of freedom

Number of Fisher Scoring Iterations: 9

INTERMEDIAR/PERIFERIE - WONINGTYPEN 1/2

Startset (flinwt12.glm)

*** Generalized Linear Model ***

Call: glm(formula = ywt12.h ~ log(NEW.PLANDB83 + 1) + log(NEW.PINDUS83 + 1) + log(NEW.PTRANP83 + 1) + log(NEW.PKENNI83 + 1) + log(PBOSGEB81 + 1) + log(PBOSREC81 + 1) + log(PDRGNTR81 + 1) + log(PNATNTR81 + 1) + log(POVOWTR81 + 1) + WEGEN83 + RAIL80 + log(PSTAT80 + 1) + log(POPAF83 + 1) + GROEIKRN + log(P.EENGEZ80 + 1) + log(PWON80.T345 + 1), family = binomial(link = logit), data = totaalwonen80, weights = wywt12.h, subset = RANDSTAD == 0, na.action = na.omit, control = list(epsilon = 0.001, maxit = 50, trace = F))

Deviance Residuals:

Min	1Q	Median	3Q	Max
-34.79787	-2.473775	-1.349781	-0.5700266	88.17247

Coefficients:

	Value	Std. Error	t value
(Intercept)	-11.224381471	0.0294071001	-381.689505
log(NEW.PLANDB83 + 1)	-0.064289052	0.0034003592	-18.906547
log(NEW.PINDUS83 + 1)	-0.105267497	0.0045202983	-23.287733
log(NEW.PTRANP83 + 1)	0.193884456	0.0053424346	36.291405
log(NEW.PKENNI83 + 1)	0.233984220	0.0048480742	48.263333
log(PBOSGEB81 + 1)	0.085754545	0.0012813483	66.925243
log(PBOSREC81 + 1)	0.002707919	0.0007358600	3.679938
log(PDRGNTR81 + 1)	-0.031041091	0.0008385776	-37.016359

Ruimtelijke ontwikkelingen woningbouw 1980 - 1995

log(PNATNTR81 + 1)	0.081702549	0.0012077930	67.646153
log(POVOWTR81 + 1)	-0.011084335	0.0014762694	-7.508341
WEGEN83	-0.105469218	0.0025214228	-41.829247
RAIL80	-0.364068201	0.0043516366	-83.662364
log(PSTAT80 + 1)	-0.114950488	0.0010146342	-113.292542
log(POPAF83 + 1)	0.053941173	0.0008963661	60.177613
GROEIKRN	1.486541176	0.0047464645	313.189148
log(P.EENGEZ80 + 1)	0.825570417	0.0008837181	934.201104
log(PWON80.T345 + 1)	0.092596509	0.0007821357	118.389319

(Dispersion Parameter for Binomial family taken to be 1)
Null Deviance: 5093441 on 116228 degrees of freedom
Residual Deviance: 2975159 on 116212 degrees of freedom
Number of Fisher Scoring Iterations: 8

Correlation of Coefficients:

	(Intercept)	log(NEW.PLANDB83 + 1)	log(NEW.PINDUS83 + 1)	
1)				
log(NEW.PLANDB83 + 1)	-0.2858563			
log(NEW.PINDUS83 + 1)	-0.4588068	-0.2569985		
log(NEW.PTRANP83 + 1)	0.3472815	-0.0400538	-0.4336382	
log(NEW.PKENNI83 + 1)	-0.1949014	-0.0134205	-0.1532162	
log(PBOSGEB81 + 1)	0.2359232	-0.4177624	-0.2841614	
log(PBOSREC81 + 1)	-0.0546305	0.2714037	-0.0999320	
log(PDRGNTR81 + 1)	-0.0175164	0.1061224	0.1832753	
log(PNATNTR81 + 1)	-0.1952872	0.0372009	-0.0494240	
log(POVOWTR81 + 1)	-0.1549574	-0.2876751	0.2790776	
WEGEN83	-0.1244733	-0.0035910	0.0446434	
RAIL80	-0.1239757	-0.0276052	-0.0022909	
log(PSTAT80 + 1)	-0.2910487	-0.0293108	0.0045425	
log(POPAF83 + 1)	-0.3679399	-0.0424272	0.0865489	
GROEIKRN	-0.0106266	-0.0230183	0.1842438	
log(P.EENGEZ80 + 1)	-0.1531933	0.0429435	-0.0803190	
log(PWON80.T345 + 1)	-0.0187466	0.0741681	0.0036718	
		log(NEW.PTRANP83 + 1)	log(NEW.PKENNI83 + 1)	
log(PBOSGEB81 + 1)				
log(NEW.PLANDB83 + 1)				
log(NEW.PINDUS83 + 1)				
log(NEW.PTRANP83 + 1)				
log(NEW.PKENNI83 + 1)	-0.7131497			
log(PBOSGEB81 + 1)	0.0802919	0.1075752		
log(PBOSREC81 + 1)	-0.0868798	0.0232626	-0.1760270	
log(PDRGNTR81 + 1)	0.1098277	-0.2852797	-0.5711197	
log(PNATNTR81 + 1)	-0.0023760	0.1233846	-0.0492931	
log(POVOWTR81 + 1)	-0.0992109	-0.1456493	0.2359423	
WEGEN83	-0.0170500	0.0151522	-0.0251887	
RAIL80	0.0067660	0.0063574	0.0111332	
log(PSTAT80 + 1)	0.0048047	0.0209783	-0.0118400	
log(POPAF83 + 1)	0.0267807	-0.0105967	-0.0204057	
GROEIKRN	0.0166203	-0.2041514	0.0499259	
log(P.EENGEZ80 + 1)	-0.0119854	0.0034600	0.0540548	
log(PWON80.T345 + 1)	0.0264707	-0.0637939	-0.0349629	
		log(PBOSREC81 + 1)	log(PDRGNTR81 + 1)	log(PNATNTR81 + 1)
log(POVOWTR81 + 1)				
log(NEW.PLANDB83 + 1)				
log(NEW.PINDUS83 + 1)				
log(NEW.PTRANP83 + 1)				
log(NEW.PKENNI83 + 1)				
log(PBOSGEB81 + 1)				
log(PBOSREC81 + 1)				
log(PDRGNTR81 + 1)	-0.1685582			
log(PNATNTR81 + 1)	0.0198995	-0.1813514		
log(POVOWTR81 + 1)	0.0359465	-0.0365079	-0.4456127	

Ruimtelijke ontwikkelingen woningbouw 1980 - 1995

```

WEGEN83          0.0164344          0.0105290          0.0234920
0.0152907          RAIL80          -0.0217773          0.0031514
0.0395449          -0.0025835
log(PSTAT80 + 1)  -0.0657743          0.0086455          -0.0077865
0.0018140
log(POPAF83 + 1)          0.0150717          0.0747493          -0.0092760
0.0662895
GROEIKRN          -0.1617665          0.0125391          -0.1948991
0.0595171
log(P.EENGEZ80 + 1) -0.0476090          -0.0098106          0.0309892
0.0007885
log(PWON80.T345 + 1) -0.0195840          0.0196655          0.0009320
-0.0049015
          WEGEN83          RAIL80 log(PSTAT80 + 1) log(POPAF83 +
1) GROEIKRN
log(NEW.PLANDB83 + 1)
log(NEW.PINDUS83 + 1)
log(NEW.PTRANP83 + 1)
log(NEW.PKENNI83 + 1)
  log(PBOSGEB81 + 1)
  log(PBOSREC81 + 1)
  log(PDRGNTR81 + 1)
  log(PNATNTR81 + 1)
  log(POVOWTR81 + 1)
          WEGEN83
          RAIL80 -0.0035413
          log(PSTAT80 + 1) 0.0139648 0.3982191
          log(POPAF83 + 1) 0.1914755 0.0200716 -0.1085639
          GROEIKRN 0.0162625 0.0273025 0.1455545 0.0188872
log(P.EENGEZ80 + 1) -0.1272041 -0.0378819 0.1434185 0.0851731 -
0.0053460
  log(PWON80.T345 + 1) 0.0139940 -0.0032981 0.1253697 0.0811320
0.0354736
          log(P.EENGEZ80 + 1)
log(NEW.PLANDB83 + 1)
log(NEW.PINDUS83 + 1)
log(NEW.PTRANP83 + 1)
log(NEW.PKENNI83 + 1)
  log(PBOSGEB81 + 1)
  log(PBOSREC81 + 1)
  log(PDRGNTR81 + 1)
  log(PNATNTR81 + 1)
  log(POVOWTR81 + 1)
          WEGEN83
          RAIL80
          log(PSTAT80 + 1)
          log(POPAF83 + 1)
          GROEIKRN
          log(P.EENGEZ80 + 1)
          log(PWON80.T345 + 1) -0.3700717

```

Analysis of Deviance Table
 Binomial model
 Response: ywt12.h

Terms added sequentially (first to last)

	Df	Deviance	Resid.	Df	Resid.	Dev
NULL			116228		5093441	
log(NEW.PLANDB83 + 1)	1	57118	116227		5036323	
log(NEW.PINDUS83 + 1)	1	211948	116226		4824375	
log(NEW.PTRANP83 + 1)	1	39100	116225		4785275	
log(NEW.PKENNI83 + 1)	1	10599	116224		4774676	
log(PBOSGEB81 + 1)	1	2334	116223		4772341	

Ruimtelijke ontwikkelingen woningbouw 1980 - 1995

log(PBOSREC81 + 1)	1	5459	116222	4766883
log(PDRGNTR81 + 1)	1	110	116221	4766773
log(PNATNTR81 + 1)	1	183	116220	4766590
log(POVOWTR81 + 1)	1	134	116219	4766456
WEGEN83	1	51568	116218	4714888
RAIL80	1	23252	116217	4691636
log(PSTAT80 + 1)	1	193972	116216	4497664
log(POPAF83 + 1)	1	7606	116215	4490057
GROEIKRN	1	60689	116214	4429368
log(P.EENGEZ80 + 1)	1	1440809	116213	2988559
log(PWON80.T345 + 1)	1	13400	116212	2975159

Set 2 (f3in12.glm)

*** Generalized Linear Model ***

Call: glm(formula = YWT12.H ~ log(NEW.PINDUS + 1) + log(PBOSGEB81 + 1) + WEGEN83 + RAIL80 + log(PSTAT80 + 1) + log(POPAF83 + 1) + GROEIKRN + log(P.EENGEZ80 + 1) + log(PWON80.T34 + 1), family = binomial(link = logit), data = totwon80, weights = WYWT12.H, subset = RANDSTAD == 0, na.action = na.omit, control = list(epsilon = 0.001, maxit = 50, trace = F))

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-38.23401	-2.505055	-1.344516	-0.5689217	89.27769

Coefficients:

	Value	Std. Error	t value
(Intercept)	-11.12498799	0.0235546487	-472.30541
log(NEW.PINDUS + 1)	0.25875003	0.0023176440	111.64356
log(PBOSGEB81 + 1)	0.04323294	0.0008704047	49.66993
WEGEN83	-0.11412125	0.0025222135	-45.24647
RAIL80	-0.38508849	0.0043537737	-88.44936
log(PSTAT80 + 1)	-0.11832380	0.0010065821	-117.55008
log(POPAF83 + 1)	0.04699594	0.0008776056	53.55018
GROEIKRN	1.69323981	0.0043823928	386.37336
log(P.EENGEZ80 + 1)	0.82611510	0.0008869564	931.40442
log(PWON80.T34 + 1)	0.09760752	0.0007788742	125.31873

(Dispersion Parameter for Binomial family taken to be 1)

Null Deviance: 5093441 on 116228 degrees of freedom
Residual Deviance: 2993355 on 116219 degrees of freedom
Number of Fisher Scoring Iterations: 8

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid. Dev
NULL			116228	5093441
log(NEW.PINDUS + 1)	1	245302	116227	4848140
log(PBOSGEB81 + 1)	1	18052	116226	4830087
WEGEN83	1	57090	116225	4772998
RAIL80	1	24051	116224	4748946
log(PSTAT80 + 1)	1	201641	116223	4547306
log(POPAF83 + 1)	1	10773	116222	4536533
GROEIKRN	1	89472	116221	4447061
log(P.EENGEZ80 + 1)	1	1438750	116220	3008310
log(PWON80.T34 + 1)	1	14955	116219	2993355

Set 3 (s4inwt12.glm)

*** Generalized Linear Model ***

Call: glm(formula = YWT12.H ~ log(NEW.PINDUS + 1) + log(PBOSGEB81 + 1) + log(PNATNTR81 + 1) + WEGEN83 + RAIL80 + log(POPAF83 + 1) + log(PSTAT80 + 1) + GROEIKRN + log(P.EENGEZ80 + 1) + log(PWON80.T34 + 1), family =

Ruimtelijke ontwikkelingen woningbouw 1980 - 1995

```
binomial(link = logit), data = totaalwonen80, weights = WYWT12.H, subset =  
RANDSTAD == 0, na.action = na.omit, control = list(epsilon = 0.001, maxit =  
50, trace = F))
```

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-38.25077	-2.493096	-1.352222	-0.5657455	88.94977

Coefficients:

	Value	Std. Error	t value
(Intercept)	-11.81059046	0.0255506946	-462.24146
log(NEW.PINDUS + 1)	0.28298343	0.0023319912	121.34841
log(PBOSGEB81 + 1)	0.03520182	0.0008827435	39.87774
log(PNATNTR81 + 1)	0.07236123	0.0010474162	69.08546
WEGEN83	-0.10723159	0.0025238180	-42.48784
RAIL80	-0.37140986	0.0043642585	-85.10263
log(POPAF83 + 1)	0.04993489	0.0008824536	56.58642
log(PSTAT80 + 1)	-0.11796565	0.0010068736	-117.16034
GROEIKRN	1.63244323	0.0044598546	366.03059
log(P.EENGEZ80 + 1)	0.82802218	0.0008864783	934.05802
log(PWON80.T34 + 1)	0.09820695	0.0007792782	126.02296

(Dispersion Parameter for Binomial family taken to be 1)

Null Deviance: 5093441 on 116228 degrees of freedom

Residual Deviance: 2988618 on 116218 degrees of freedom

Number of Fisher Scoring Iterations: 8

Set 4 (s5inwt12.glm)

*** Generalized Linear Model ***

```
Call: glm(formula = YWT12.H ~ log(NEW.PINDUS + 1) + log(PNATNTR81 + 1) +  
WEGEN83 + RAIL80 + log(POPAF83 + 1) + log(PSTAT80 + 1) + GROEIKRN +  
log(P.EENGEZ80 + 1) + log(PWON80.T34 + 1), family = binomial(link = logit),  
data = totaalwonen80, weights = WYWT12.H, subset = RANDSTAD == 0, na.action  
= na.omit, control = list(epsilon = 0.001, maxit = 50, trace = F))
```

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-37.88452	-2.487008	-1.35217	-0.5567117	88.63508

Coefficients:

	Value	Std. Error	t value
(Intercept)	-12.05746193	0.0247777405	-486.62476
log(NEW.PINDUS + 1)	0.33564432	0.0019353476	173.42844
log(PNATNTR81 + 1)	0.07815618	0.0010383538	75.26932
WEGEN83	-0.10427669	0.0025226513	-41.33615
RAIL80	-0.37165883	0.0043607412	-85.22836
log(POPAF83 + 1)	0.05039110	0.0008818045	57.14544
log(PSTAT80 + 1)	-0.11620267	0.0010076726	-115.31789
GROEIKRN	1.62075860	0.0044446078	364.65728
log(P.EENGEZ80 + 1)	0.82517116	0.0008839263	933.52933
log(PWON80.T34 + 1)	0.09844486	0.0007790240	126.36948

(Dispersion Parameter for Binomial family taken to be 1)

Null Deviance: 5093441 on 116228 degrees of freedom

Residual Deviance: 2990209 on 116219 degrees of freedom

Number of Fisher Scoring Iterations: 8

INTERMEDIAR/PERIFERIE - WONINGTYPEN 3/4/5

Startset (f1inwt345.glm)

Ruimtelijke ontwikkelingen woningbouw 1980 - 1995

*** Generalized Linear Model ***

Call: glm(formula = ywt345.h ~ log(NEW.PLANDB83 + 1) + log(NEW.PINDUS83 + 1) + log(NEW.PTRANP83 + 1) + log(NEW.PKENNI83 + 1) + log(PBOSGEB81 + 1) + log(PBOSREC81 + 1) + log(PDRGNTR81 + 1) + log(PNATNTR81 + 1) + log(POVOWTR81 + 1) + WEGEN83 + RAIL80 + log(PSTAT80 + 1) + log(POPAF83 + 1) + GROEIKRN + log(P.EENGEZ80 + 1) + log(PWON80.T345 + 1), family = binomial(link = logit), data = totaalwonen80, weights = wywt345.h, subset = RANDSTAD == 0, na.action = na.omit, control = list(epsilon = 0.001, maxit = 50, trace = F))

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-30.20356	-0.8113082	-0.4324789	-0.2517992	88.56906

Coefficients:

	Value	Std. Error	t value
(Intercept)	-13.805223993	0.065074915	-212.143558
log(NEW.PLANDB83 + 1)	-0.478609502	0.008087232	-59.180878
log(NEW.PINDUS83 + 1)	0.109237355	0.010440439	10.462908
log(NEW.PTRANP83 + 1)	0.140087235	0.012724207	11.009506
log(NEW.PKENNI83 + 1)	0.307953868	0.010584647	29.094392
log(PBOSGEB81 + 1)	0.089501053	0.003404741	26.287182
log(PBOSREC81 + 1)	-0.015157408	0.001838367	-8.245037
log(PDRGNTR81 + 1)	-0.042667698	0.002156828	-19.782612
log(PNATNTR81 + 1)	0.129282365	0.003138586	41.191274
log(POVOWTR81 + 1)	0.047675998	0.003562729	13.381875
WEGEN83	-0.147747011	0.005826142	-25.359320
RAIL80	-0.278123030	0.008440601	-32.950620
log(PSTAT80 + 1)	-0.174134111	0.001645495	-105.824761
log(POPAF83 + 1)	-0.002720204	0.001534081	-1.773181
GROEIKRN	1.145319768	0.009767226	117.261524
log(P.EENGEZ80 + 1)	0.874782970	0.002939500	297.595837
log(PWON80.T345 + 1)	0.484351331	0.001525257	317.553976

(Dispersion Parameter for Binomial family taken to be 1)

Null Deviance: 1496231 on 116228 degrees of freedom

Residual Deviance: 615910.2 on 116212 degrees of freedom

Number of Fisher Scoring Iterations: 12

Correlation of Coefficients:

	(Intercept)	log(NEW.PLANDB83 + 1)	log(NEW.PINDUS83 + 1)
log(NEW.PLANDB83 + 1)	-0.2494828		
log(NEW.PINDUS83 + 1)	-0.4806611	-0.2545892	
log(NEW.PTRANP83 + 1)	0.2705889	-0.0593321	-0.4452090
log(NEW.PKENNI83 + 1)	-0.1491905	0.0214906	-0.1411527
log(PBOSGEB81 + 1)	0.2629062	-0.4443730	-0.2470885
log(PBOSREC81 + 1)	-0.0741548	0.2324142	-0.0645560
log(PDRGNTR81 + 1)	-0.0172943	0.1016128	0.1880703
log(PNATNTR81 + 1)	-0.1752803	-0.0027296	-0.0537922
log(POVOWTR81 + 1)	-0.1745165	-0.3909299	0.3593110
WEGEN83	-0.1521330	-0.0140759	0.0215903
RAIL80	-0.0828093	-0.0510891	-0.0136091
log(PSTAT80 + 1)	-0.1459915	-0.0546453	-0.0570181
log(POPAF83 + 1)	-0.3465162	0.0034136	0.0465681
GROEIKRN	0.0263277	-0.0949096	0.2181972
log(P.EENGEZ80 + 1)	-0.2090273	0.0024540	-0.0813653
log(PWON80.T345 + 1)	0.0816708	0.1095945	0.0108940
log(NEW.PTRANP83 + 1)		log(NEW.PKENNI83 + 1)	
log(PBOSGEB81 + 1)			
log(NEW.PLANDB83 + 1)			
log(NEW.PINDUS83 + 1)			
log(NEW.PTRANP83 + 1)			
log(NEW.PKENNI83 + 1)			-0.6944639

Ruimtelijke ontwikkelingen woningbouw 1980 - 1995

log (PBOSGEB81 + 1)	0.0669962		0.0771216	
log (PBOSREC81 + 1)	-0.1852674		0.1152560	-0.1592772
log (PDRGNTR81 + 1)	0.0468136		-0.2166991	-0.6134556
log (PNATNTR81 + 1)	0.1725327		-0.0788628	-0.1086838
log (POVOWTR81 + 1)	-0.1799642		-0.0812339	0.2633257
WEGEN83	0.0031254		0.0261431	-0.0279199
RAIL80	0.0058284		0.0097426	0.0115915
log (PSTAT80 + 1)	0.0467664		-0.0151248	0.0152861
log (POPAF83 + 1)	0.0088157		0.0398166	-0.0647948
GROEIKRN	-0.0746253		-0.1099079	0.1700982
log (P.EENGEZ80 + 1)	0.0153903		-0.0047837	0.0317644
log (PWON80.T345 + 1)	0.0168823		-0.1045750	-0.0284518
	log (PBOSREC81 + 1)	log (PDRGNTR81 + 1)	log (PNATNTR81 + 1)	
1) log (POVOWTR81 + 1)				
log (NEW.PLANDB83 + 1)				
log (NEW.PINDUS83 + 1)				
log (NEW.PTRANP83 + 1)				
log (NEW.PKENNI83 + 1)				
log (PBOSGEB81 + 1)				
log (PBOSREC81 + 1)				
log (PDRGNTR81 + 1)	-0.2257682			
log (PNATNTR81 + 1)	0.0336513		-0.0692374	
log (POVOWTR81 + 1)	0.0781189		-0.0882891	-0.4259677
WEGEN83	0.0068184		0.0115361	0.0157522
0.0482589				
RAIL80	-0.0332808		0.0083743	0.0591570
-0.0171670	log (PSTAT80 + 1)		-0.0330997	0.0155291
0.0416027	-0.0059184			
log (POPAF83 + 1)	-0.0511119		0.0661733	0.0561976
0.0194443				
GROEIKRN	-0.2082020		-0.1102082	-0.1964769
0.0849068				
log (P.EENGEZ80 + 1)	-0.0347046		-0.0068686	0.0352306
0.0170019				
log (PWON80.T345 + 1)	-0.0290235		0.0362070	0.0366941
-0.0654980				
	WEGEN83	RAIL80	log (PSTAT80 + 1)	log (POPAF83 + 1)
GROEIKRN				
log (NEW.PLANDB83 + 1)				
log (NEW.PINDUS83 + 1)				
log (NEW.PTRANP83 + 1)				
log (NEW.PKENNI83 + 1)				
log (PBOSGEB81 + 1)				
log (PBOSREC81 + 1)				
log (PDRGNTR81 + 1)				
log (PNATNTR81 + 1)				
log (POVOWTR81 + 1)				
WEGEN83				
RAIL80	-0.0080732			
log (PSTAT80 + 1)	0.0042667	0.5411193		
log (POPAF83 + 1)	0.3105854	0.0014755	-0.0554417	
GROEIKRN	-0.0033340	0.0396129	0.0647102	-0.0641784
log (P.EENGEZ80 + 1)	-0.0452550	-0.0214610	0.0940551	0.0707929
0.0548278				
log (PWON80.T345 + 1)	0.0296866	0.0379551	0.1343878	0.1001999
0.0359129				
	log (P.EENGEZ80 + 1)			
log (NEW.PLANDB83 + 1)				
log (NEW.PINDUS83 + 1)				
log (NEW.PTRANP83 + 1)				
log (NEW.PKENNI83 + 1)				
log (PBOSGEB81 + 1)				
log (PBOSREC81 + 1)				

Ruimtelijke ontwikkelingen woningbouw 1980 - 1995

```

log(PDRGNTR81 + 1)
log(PNATNTR81 + 1)
log(POVOWTR81 + 1)
      WEGEN83
      RAIL80
log(PSTAT80 + 1)
log(POPAF83 + 1)
      GROEIKRN
log(P.EENGEZ80 + 1)
log(PWON80.T345 + 1) -0.5757230

```

Analysis of Deviance Table

Binomial model

Response: ywt345.h

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid. Dev
NULL			116228	1496231
log(NEW.PLANDB83 + 1)	1	3515.0	116227	1492716
log(NEW.PINDUS83 + 1)	1	114854.9	116226	1377861
log(NEW.PTRANP83 + 1)	1	20449.9	116225	1357411
log(NEW.PKENNI83 + 1)	1	11834.9	116224	1345577
log(PBOSGEB81 + 1)	1	1806.4	116223	1343770
log(PBOSREC81 + 1)	1	1879.8	116222	1341890
log(PDRGNTR81 + 1)	1	4.9	116221	1341885
log(PNATNTR81 + 1)	1	-0.1	116220	1341885
log(POVOWTR81 + 1)	1	1161.4	116219	1340724
WEGEN83	1	12820.7	116218	1327903
RAIL80	1	23476.5	116217	1304427
log(PSTAT80 + 1)	1	103598.2	116216	1200829
log(POPAF83 + 1)	1	12025.2	116215	1188803
GROEIKRN	1	6810.6	116214	1181993
log(P.EENGEZ80 + 1)	1	454335.1	116213	727658
log(PWON80.T345 + 1)	1	111747.4	116212	615910

Set 2 (f3in345.glm)

*** Generalized Linear Model ***

```

Call: glm(formula = YWT345.H ~ log(NEW.PINDUS + 1) + log(PBOSREC81 + 1) +
log(PNATNTR81 + 1) + WEGEN83 + RAIL80 + log(PSTAT80 + 1) + GROEIKRN +
log(P.EENGEZ80 + 1) + log(PWON80.T34 + 1), family = binomial(link = logit),
data = totwon80, weights = WYWT345.H, subset = RANDSTAD == 0, na.action =
na.omit, control = list(epsilon = 0.001, maxit = 50, trace = F))

```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-30.15654	-0.8281144	-0.4533115	-0.2548649	87.72661

Coefficients:

	Value	Std. Error	t value
(Intercept)	-14.892466411	0.052050572	-286.115326
log(NEW.PINDUS + 1)	0.282779213	0.004945509	57.178996
log(PBOSREC81 + 1)	0.004942953	0.001628638	3.035023
log(PNATNTR81 + 1)	0.117600368	0.002648965	44.394831
WEGEN83	-0.162368181	0.005525516	-29.385164
RAIL80	-0.306316232	0.008490847	-36.076052
log(PSTAT80 + 1)	-0.181499428	0.001630785	-111.295774
GROEIKRN	1.203974698	0.009049184	133.047868
log(P.EENGEZ80 + 1)	0.877505008	0.002940328	298.437770
log(PWON80.T34 + 1)	0.506888981	0.001506342	336.503165

(Dispersion Parameter for Binomial family taken to be 1)

Ruimtelijke ontwikkelingen woningbouw 1980 - 1995

Null Deviance: 1496231 on 116228 degrees of freedom
Residual Deviance: 622221.4 on 116219 degrees of freedom
Number of Fisher Scoring Iterations: 12

Set 3 (s4inwt345.glm)

*** Generalized Linear Model ***

Call: glm(formula = YWT345.H ~ log(NEW.PKENNI + 1) + log(PBOSGEB81 + 1) + log(PNATNTR81 + 1) + WEGEN83 + RAIL80 + log(PSTAT80 + 1) + log(POPAF83 + 1) + GROEIKRN + log(P.EENGEZ80 + 1) + log(PWON80.T34 + 1), family = binomial(link = logit), data = totaalwonen80, weights = WYWT345.H, subset = RANDSTAD == 0, na.action = na.omit, control = list(epsilon = 0.001, maxit = 50, trace = F))

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-28.71696	-0.8230629	-0.4446225	-0.2522946	89.12899

Coefficients:

	Value	Std. Error	t value
(Intercept)	-15.015548071	0.050085636	-299.797493
log(NEW.PKENNI + 1)	0.354580837	0.004820361	73.558973
log(PBOSGEB81 + 1)	-0.010145823	0.001837553	-5.521376
log(PNATNTR81 + 1)	0.107190711	0.002698579	39.721174
WEGEN83	-0.150150640	0.005814006	-25.825676
RAIL80	-0.302014088	0.008462046	-35.690435
log(PSTAT80 + 1)	-0.178875590	0.001638022	-109.202174
log(POPAF83 + 1)	-0.001818096	0.001532675	-1.186224
GROEIKRN	1.074465634	0.009064069	118.541199
log(P.EENGEZ80 + 1)	0.877920888	0.002928017	299.834602
log(PWON80.T34 + 1)	0.494730942	0.001516811	326.165153

(Dispersion Parameter for Binomial family taken to be 1)

Null Deviance: 1496231 on 116228 degrees of freedom
Residual Deviance: 619781.1 on 116218 degrees of freedom
Number of Fisher Scoring Iterations: 12

Set 4 (s5inwt345.glm)

*** Generalized Linear Model ***

Call: glm(formula = YWT345.H ~ log(NEW.PKENNI + 1) + log(PNATNTR81 + 1) + WEGEN83 + RAIL80 + log(PSTAT80 + 1) + GROEIKRN + log(P.EENGEZ80 + 1) + log(PWON80.T34 + 1), family = binomial(link = logit), data = totaalwonen80, weights = WYWT345.H, subset = RANDSTAD == 0, na.action = na.omit, control = list(epsilon = 0.001, maxit = 50, trace = F))

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-28.72112	-0.8232791	-0.4456366	-0.2523614	88.95631

Coefficients:

	Value	Std. Error	t value
(Intercept)	-15.0211899	0.045542634	-329.82699
log(NEW.PKENNI + 1)	0.3448183	0.004347493	79.31429
log(PNATNTR81 + 1)	0.1057314	0.002670713	39.58920
WEGEN83	-0.1492398	0.005527564	-26.99920
RAIL80	-0.3027039	0.008462922	-35.76825
log(PSTAT80 + 1)	-0.1792784	0.001633639	-109.74178
GROEIKRN	1.0825625	0.008903945	121.58235
log(P.EENGEZ80 + 1)	0.8780407	0.002919433	300.75727
log(PWON80.T34 + 1)	0.4953888	0.001507295	328.66087

```
(Dispersion Parameter for Binomial family taken to be 1 )  
Null Deviance: 1496231 on 116228 degrees of freedom  
Residual Deviance: 619813.3 on 116220 degrees of freedom  
Number of Fisher Scoring Iterations: 12
```