

Toetsing waarde agrarisch vastgoed

Onderzoek naar een regressiemodel



ir. ing. J.M. Geerlings RT

Masterscriptie
Master of Science in Real Estate
1^e beoordelaar drs. A.R. Marquard
27 maart 2019

Titel	Toetsing waarde agrarisch vastgoed <i>Onderzoek naar een regressiemodel</i>
Plaats en datum	Venlo, 27 maart 2019
Auteur	ir. ing. J.M. Geerlings RT
1 ^e beoordelaar	drs. A.R. Marquard
2 ^e beoordelaar	drs. M. Paelinck
Scriptie	Masterscriptie
Opleiding	Master of Science in Real Estate (MSRE)
Specialisatie	Real Estate Valuation (REV)
Instituut	Amsterdam School of Real Estate (ASRE)

Voorwoord

Geheel volgens planning ligt voor u de masterscriptie 'Toetsing waarde agrarisch vastgoed'. Deze scriptie vormt de afronding van mijn studie Master of Science in Real Estate aan de Amsterdam School of Real Estate.

Binnen de vastgoed- en taxatiebranche is agrarisch vastgoed een 'ondergeschoven' kindje. Voor de waardering van woningen zijn steeds betere geautomatiseerde waarderingsmodellen beschikbaar en voor bedrijfsmatig vastgoed zijn de waarderingsmethodieken in de loop der jaren verder doorontwikkeld. Alleen voor agrarisch vastgoed lijkt sprake van stilstand. Daarom waag ik in deze scriptie een poging om de regressiemethode toe te passen als toetsingsinstrument voor de waarde van gebouwd agrarisch vastgoed.

Het opbouwen van de dataset heeft de nodige tijd en inspanning gekost. Daarom dank aan Bernd Veldman die de ruwe transactiedata voor dit onderzoek vanuit de organisatie beschikbaar heeft gesteld. Dit heeft het onderzoek wel versneld. Ook dank aan Hans Cramer, Anja van Son en Judith Gielen van de organisatie voor het faciliteren van mijn studie. De kennisverbreding en -verdieping komt zeker van pas tijdens mijn werk.

Speciale dank aan Douglas Konadu voor zijn hulp en ondersteuning bij mijn ontdekkingstocht door STATA. Door hem en Arthur ben ik wel enthousiast geworden over statistiek. Bovenal gaat mijn dank uit naar mijn scriptiebegeleiders, in het bijzonder Arthur Marquard, voor de positieve feedback tijdens onze contactmomenten en op het concept. Hierdoor heb ik mijn scriptie tijdig en met een goed gevoel kunnen afronden. De ASRE en haar docenten dank ik voor het uitgebreide studieprogramma. Ik heb er een hoop van opgestoken.

En tot slot gaat mijn grootste dank uit naar mijn vrouw AnneMarie voor alle ruimte die zij mij de afgelopen twee jaar heeft gegeven om deze studie te doen. Nu deze is afgerond kan ik uit mijn 'studiebubble' komen en de tijd weer anders gebruiken.

Ik wens u veel plezier met het lezen van deze scriptie.

Han Geerlings

Venlo, maart 2019

Samenvatting

Door het Nederlands Register Vastgoed Taxateurs wordt aanbevolen om bij een waardering gebruik te maken van ten minste twee waarderingsmethoden. Voor de waardering van agrarisch vastgoed wordt de directe vergelijkingsmethode sectorbreed gehanteerd. Als tweede methode zou men de gecorrigeerde vervangingswaardemethode kunnen gebruiken. Volgens jurisprudentie van de Hoge Raad doet zich tussen deze twee methodieken een cirkelredenering voor. De waarde op basis van de vergelijkingsmethode kan niet meer zijn dan op basis van de gecorrigeerde vervangingswaarde methode. En de waarde op basis van de gecorrigeerde vervangingswaardemethode kan niet meer zijn dan op basis van de vergelijkingsmethode. Deze scriptie heeft daarom tot doel te onderzoeken in hoeverre het mogelijk is om met een regressiemodel de waarde van gebouwd agrarisch vastgoed te toetsen.

Literatuuronderzoek laat zien dat met name onderzoek is gedaan naar factoren die de prijs van landbouwgrond verklaren. Deze onderzoeken zijn gebaseerd op de klassieke locatietheorieën van Ricardo en Von Thünen, waarbij de afstand van de landbouwgrond tot de centrale marktplaats een belangrijke factor is. Uit andere internationale en nationale onderzoeken bleek dat ook bevolkingsdichtheid, bodemproductiviteit, bedrijfstype en bedrijfsoppervlakte van invloed zijn op de prijs van landbouwgrond. Voor gebouwd agrarisch vastgoed zijn volgens de gangbare waarderingspraktijk met name de oppervlakte, kwaliteit en doelmatigheid van de gebouwen van invloed op de prijs.

Op basis van het literatuuronderzoek is de keus gemaakt welke variabelen aan de dataset met transacties worden toegevoegd voor dit onderzoek. Voor het onderzoek is een dataset samengesteld met 132 variabelen en 283 waarnemingen uit 2016 en 2017. Uiteindelijk bleken slechts 28 van deze variabelen van toegevoegde waarde voor het onderzoek. Na eliminatie van waarnemingen c.q. uitschieters in variabelen en storingsterm zijn regressieanalyses uitgevoerd. De resultaten waren het best wanneer de variabelen logaritmisch waren getransformeerd, omdat de verklarende kracht van het regressiemodel op basis van de getransformeerde dataset groter is dan op basis van de originele dataset.

De getransformeerde variabelen bebouwde oppervlakte en bouwjaarklassen van de bedrijfsgebouwen, oppervlakte van de bouwkaavel, bouwjaarklassen van de woning (i.c. aanwezigheid van een woning) en de afstand tot Amsterdam waren significant en vormden de basis van het te formuleren regressiemodel. Uiteindelijk heeft dit geresulteerd in een regressiemodel met een verklarende kracht (R^2) van bijna 75%. Daarnaast zijn nog vier regressiemodellen gevonden met een verklarende kracht van meer dan 72%, waarbij één of meer van de variabelen bevolkingsdichtheid, geografische ligging in landsdelen, afstand tot dichtstbij gelegen top 10 stad en aanwezigheid intensief landbouwbedrijf significant is of zijn.

De 95% betrouwbaarheidsinterval van de variabelen is breed, dat leidt tot de conclusie dat bij de toetsing van een waarde van agrarisch vastgoed met dit model deze veelal significant zal zijn. De validiteit en betrouwbaarheid van het model worden daarom als matig tot gemiddeld beoordeeld.

Met deze scriptie is een eerste stap gezet in de richting van een regressiemodel als toetsing van de waarde van agrarisch vastgoed. Maar verder onderzoek zal nodig zijn om tot een betrouwbaarder model te komen.

Inhoudsopgave

Voorwoord.....	iii
Samenvatting	iv
1 Inleiding.....	1
1.1 De aanleiding voor het onderzoek	1
1.2 Het onderwerp van de scriptie	2
1.3 Centrale vraag en eindproduct.....	2
1.4 Afbakening onderzoek	2
1.5 De onderzoeksmethode en het onderzoeksmodel	3
1.6 Leeswijzer.....	4
2 Theorie.....	5
2.1 Klassieke locatietheorieën	5
2.1.1 Ricardo	5
2.1.2 Von Thünen	5
2.1.3 De klassieke locatietheorie in Nederland.....	7
2.2 Overzicht bestaande literatuur.....	7
2.3 De agrarische sector.....	9
2.3.1 Ontwikkeling van de agrarische sector	9
2.3.2 Agrarische bedrijfstypen.....	10
2.3.3 Toekomstige ontwikkelingen	12
2.4 Waardebepalende factoren bij agrarisch vastgoed	12
2.5 Hypothese waardebepalende factoren voor het onderzoek.....	13
3 Methode en data	15
3.1 Taxatieleer agrarische vastgoedobjecten.....	15
3.1.1 Comparatieve benadering.....	15
3.1.2 Inkomstenbenadering	16
3.1.3 Kostenbenadering.....	16
3.2 Omschrijving regressiemethode	17
3.2.1 Enkelvoudige regressie.....	17
3.2.2 Meervoudige regressie	17
3.3 Omschrijving data.....	18
3.4 Bewerking data.....	18
3.4 Variabelen dataset	20
3.5 Verwachte verbanden	22
4 Analyse	23
4.1 Beschrijvende Statistiek.....	23
4.2 Voorwaarden meervoudige regressie	25
4.2.1 Outliers.....	26
4.2.2 Multicollineariteit	28
4.2.3 Normale verdeling van de storingsterm	29
4.2.4 Lineariteit.....	30

4.3 Regressieanalyses.....	31
4.3.1 Originele dataset.....	32
4.3.2. Getransformeerde dataset.....	33
4.4 Regressiemodel	35
5 Conclusie.....	38
5.1 Conclusie	38
5.2 Reflectie	39
5.3 Aanbevelingen	40
Literatuurlijst.....	41
Databronnen.....	44
Bijlage 1 - Correlatie testen	45
Bijlage 2 - Afdruk regressieanalyses o.b.v. originele dataset	47
Bijlage 3 - Afdruk regressieanalyses o.b.v. getransformeerde dataset	48

1 Inleiding

1.1 De aanleiding voor het onderzoek

De vastgoed- en financiële crisis van 2008 heeft de wetgever en de financieel toezichthouders laten beseffen dat de Nederlandse economie afhankelijk is van de conditie van de vastgoed- en bouwmarkt. Die conditie kan betrouwbaar worden gemeten aan de hand van kwalitatief hoogwaardige vastgoedtaxaties die door vastgoedtaxateurs zijn opgesteld (Van Arnhem, Berkhout & Ten Have, 2015). De crisis heeft ertoe geleid dat in 2012 het Platform Taxateurs en Accountants is opgericht. Het Platform Taxateurs en Accountants (2013) had ten doel het benoemen van verbeterpunten in het taxatieproces en het opstellen van transparantere taxatierapporten (p. 5), naar wat blijkt met name voor commercieel vastgoed¹. Dit heeft geresulteerd in 28 concrete aanbevelingen, waaronder het gebruik van een twee waarderingmethode voor toetsing van de geschatte waarde van de eerste (hoofd)methode.

Voor bedrijfsmatig vastgoed is het redelijk uitgekristalliseerd. Daar hanteert men de DCF-methode als hoofdmethode en bijvoorbeeld de BAR-methode of de vergelijkingsmethode als tweede toetsende methode (p. 28).

Door het Nederlands Register Vastgoed Taxateurs (2018, p. 8) is de aanbeveling om bij een waardering gebruik te maken van ten minste twee waarderingmethoden zowel in de praktijkhandreiking voor de Kamer Bedrijfsmatig Vastgoed als in de praktijkhandreiking voor de Kamer Landelijk en Agrarisch Vastgoed opgenomen.

Bij de waardering van agrarisch vastgoed is een toetsing van de geschatte waarde niet ontwikkeld. Dat blijkt uit een zaak dat diende voor het Hof Arnhem-Leeuwarden (2015), waarbij het Hof heeft aangegeven dat voor de bepaling van de WOZ-waarde van een agrarisch vastgoed de landelijke taxatiewijzer niet moet worden gebruikt, als bruikbare transactiegegevens beschikbaar zijn.

Uit deze casus kan worden geconcludeerd dat de geschatte WOZ-waarde niet is getoetst aan een tweede methode, in dat geval de directe vergelijkingsmethode.

Bij de waardering van agrarisch vastgoed maakt de taxateur gebruik van kengetallen afkomstig uit referentietransacties, die leiden tot een geschatte waarde van het object. Of dit een reële waarde is, daar lijkt nu geen toetsingsmogelijkheid voor. Mogelijk dat de gecorrigeerde vervangingswaarde-methode als tweede methode uitkomst kan bieden. Men kan zich afvragen of dat wel het geval is. Immers in het zogenaamde Pieperschuurarrest, in het kader van de Wet WOZ, heeft de Hoge Raad (1995) bepaald dat de waarde in het economisch verkeer (lees: marktwaarde) voor courante onroerende zaken in beginsel niet afwijkt van de gecorrigeerde vervangingswaarde. Hieruit kan worden geconcludeerd dat de taxateur bij de gecorrigeerde vervangingswaarde methode naast technische verouderingscorrecties ook functionele verouderingscorrecties moet opnemen om tot een geschatte waarde te komen, die in overeenstemming is met de marktwaarde bepaald op basis van de vergelijkingsmethode.

In de gedachte van de Hoge Raad schuilt voor courante onroerende zaken een cirkelredenering. De waarde op basis van de vergelijkingsmethode kan niet meer zijn dan op basis van de gecorrigeerde vervangingswaarde methode. En de waarde op basis van de gecorrigeerde vervangingswaarde methode kan niet meer zijn dan op basis van de vergelijkingsmethode (Van der Burg, 2016). Deze cirkelredenering is thans nog niet doorbroken, omdat een andere methode ontbreekt om de waarde van in dit geval agrarisch vastgoed te toetsen met een andere waarderingmethode.

¹ Zoals woning, bedrijfspand, kantoor, grond en winkel voor belegging, eigen gebruik of ontwikkeling (p. 22).

1.2 Het onderwerp van de scriptie

Uit voorgaande paragraaf blijkt bij de waardering van agrarisch vastgoed geen bruikbare tweede waarderingsmethodiek te zijn om de geschatte waarde te toetsen aan referentietransacties.

Het doel van deze scriptie is om te komen tot een eenvoudig regressiemodel om deze toetsing wel te kunnen uitvoeren. Hiervoor wordt onderzocht welke factoren invloed hebben op de waarde van agrarisch vastgoed. Op basis van referentietransacties zal vervolgens op statistische wijze worden bepaald welke van de gevonden factoren het meeste invloed hebben op de waarde, om daarna een eenvoudig regressiemodel te formuleren.

Het onderzoek is relevant voor de taxatiebranche omdat het kan bijdragen aan het verhogen van de kwaliteit van de taxaties van agrarisch vastgoed. Daarnaast kan het mogelijk nieuwe inzichten geven in waardebepalende factoren van dit type vastgoed.

1.3 Centrale vraag en eindproduct

Om bovengenoemde cirkelredenering tussen de vergelijkingsmethode en de gecorrigeerde vervangingswaarde methode te doorbreken staat bij deze scriptie de volgende vraag centraal:

In hoeverre is het mogelijk om met een regressiemodel de waarde van gebouwd agrarisch vastgoed op een valide en betrouwbare manier te toetsen?

Om de centrale vraag te kunnen beantwoorden en om aan de doelstelling van de scriptie te kunnen voldoen, dienen de volgende deelvragen nader te worden onderzocht:

1. Welke aspecten zouden volgens de literatuur van invloed kunnen zijn op de waardering van agrarisch vastgoed?
 - a. Welke ruimtelijke verdeling van agrarisch vastgoed kan worden onderkend?
 - b. Welke methodieken zijn bruikbaar voor de waardering van agrarisch vastgoed?
 - c. Welke factoren zijn van invloed op de waarde van agrarisch vastgoed?
2. Hoe zou voor gebouwd agrarisch vastgoed de regressiemethode kunnen worden vorm gegeven?
 - a. Hoe werkt de regressiemethode als waarderingsmethodiek?
 - b. Hoe en welke data over agrarisch vastgoed worden voor het onderzoek verzameld?
3. Hoe kan een regressiemodel voor transacties van gebouwd agrarisch vastgoed eruit zien?
 - a. Welke factoren hebben een sterke correlatie met de waarde?
 - b. Welk ontwerp van een regressievergelijking leidt tot het beste resultaat?

Na beantwoording van deze deelvragen is het eindproduct van deze scriptie een overzicht van waardebepalende variabelen en hun invloed op de waarde, én een eenvoudig bruikbaar regressiemodel, waarmee men een geschatte waarde van gebouwd agrarisch vastgoed die is bepaald met behulp van een andere methodiek kan toetsen.

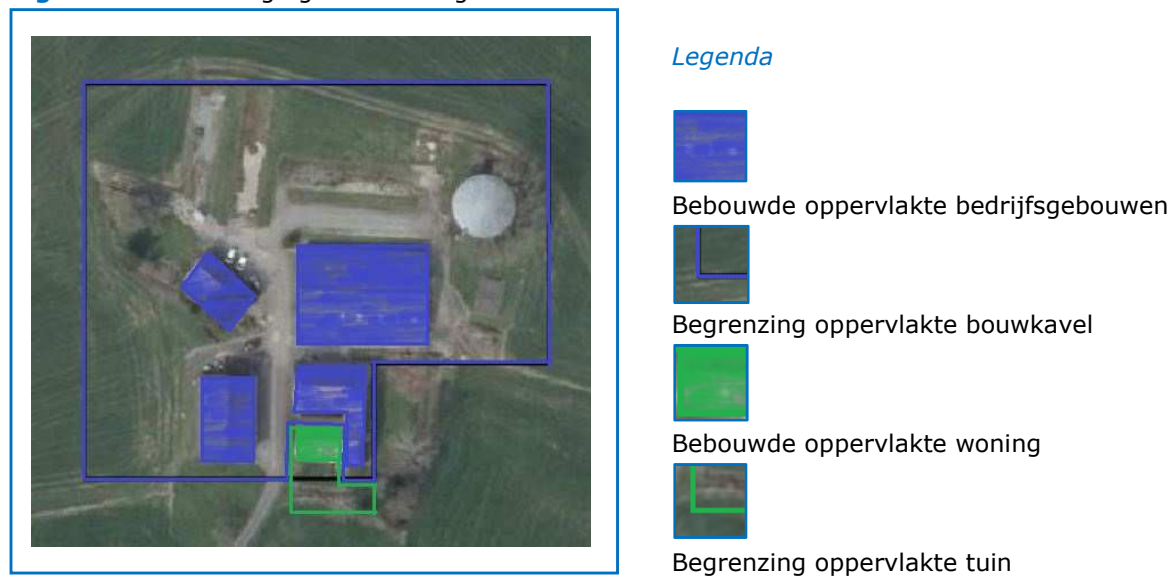
1.4 Afbakening onderzoek

In het onderzoek wordt beoordeeld of uit de beschikbare data een regressiemodel kan worden afgeleid die bruikbaar is als tweede toetsende waarderingsmethode. Buiten de scope van het onderzoek valt of en hoe vaak een tweede waarderingsmethodiek wordt gebruikt bij het toetsen van een waardering van gebouwd agrarisch vastgoed.

Met 'gebouwd agrarisch vastgoed' (hierna: agrarisch vastgoed) wordt in dit onderzoek bedoeld de agrarische bouwkaavel met daarop aanwezige bebouwing en werken (zie figuur 1). De agrarische bouwkaavel omvat de ondergrond van de aanwezige gebouwen, het omliggende erf en eventueel aanwezige tuin bij de bedrijfswoning. Onder bebouwing wordt verstaan de agrarische

bedrijfsgebouwen en eventueel aanwezige bedrijfswoning. Onder werken worden zaken verstaan zoals erfverharding, mest- en/of sleufsilo, e.d.. De landbouw- c.q. cultuurgrond ('onbebouwd agrarisch vastgoed') is in dit onderzoek buiten beschouwing gelaten.

Figuur 1: Afbakening agrarisch vastgoed



Bron: eigen bewerking

Het onderzoek is gebaseerd op transacties van agrarische bedrijven en gebouwen uit 2016 en 2017 die zijn gekocht door een agrarisch ondernemer. Deze gegevens, afkomstig uit de Basisregistratie Kadaster (BRK), zullen worden aangevuld met gegevens uit databronnen zoals de Basisregistratie Adressen en Gebouwen (BAG) en het internet, om zo voor elk agrarisch vastgoed de relevante objectkenmerken beschikbaar te krijgen. Uitgangspunt voor de uitvoering van de regressieanalyses is, naast de objectkenmerken, de totale koopsom voor het agrarisch vastgoed als vermeld in de leveringsakte minus de koopsom voor de landbouwgrond.

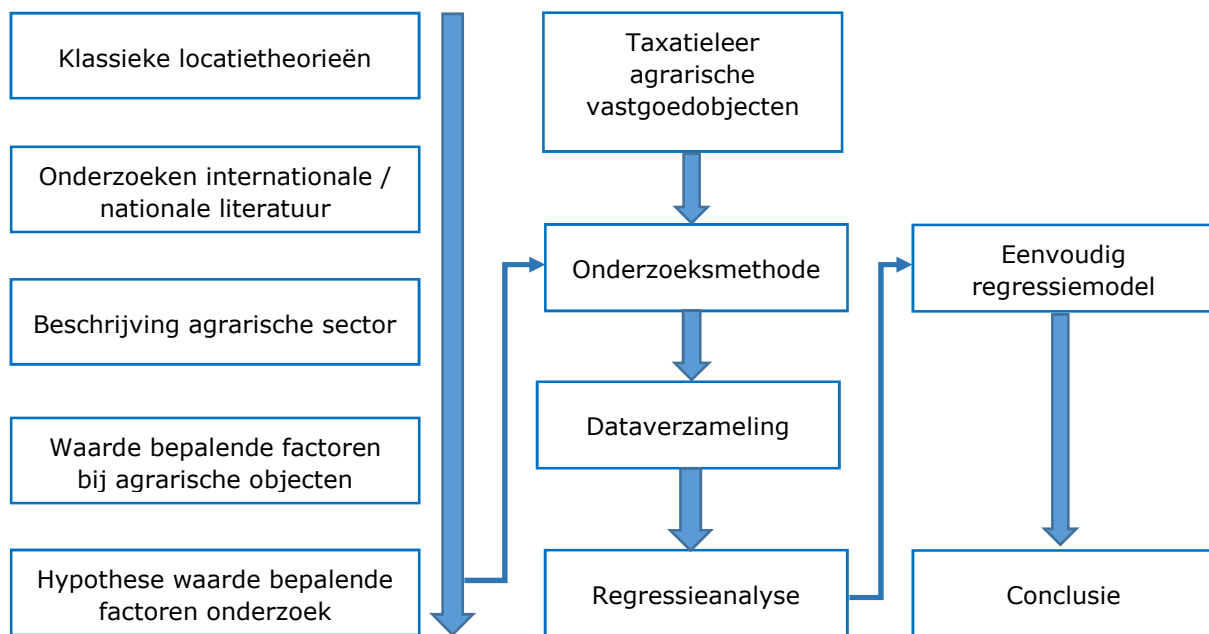
1.5 De onderzoeksmethode en het onderzoeksmodel

Het onderzoek voor de scriptie is gebaseerd enerzijds op een literatuurstudie en anderzijds op een kwalitatief onderzoek. De literatuurstudie zal beschrijvend van aard zijn. Aan de hand van onder andere de klassieke locatietheorieën, onderzoeken en studies beschreven in internationale en nationale literatuur, een beschrijving van de agrarische sector en waardebepalende factoren bij agrarisch vastgoed worden de waardebepalende factoren geformuleerd die opgenomen kunnen worden in het conceptueel model.

In het conceptueel model is een verwachting opgenomen over de mogelijke variabelen voor een regressiemodel die de waarde van agrarisch vastgoed zo goed mogelijk kunnen schatten. Afhankelijk van de sterkte van de samenhang van de variabelen kan de centrale vraag c.q. hypothese van dit onderzoek al dan niet worden gefalsificeerd.

Om de hypothese te kunnen toetsen wordt aan de hand van transactiegegevens van agrarisch vastgoed kwantitatief onderzoek gedaan. Met behulp van statistiek wordt onderzocht welk regressiemodel leidt tot het beste resultaat op basis van de gehanteerde dataset. Aan de hand van statistische toetsen kan de hypothese al dan niet worden gefalsificeerd.

Figuur 2: Onderzoeksmodel



Bron: eigen bewerking

1.6 Leeswijzer

In hoofdstuk 2 wordt het theoretisch kader voor het onderzoek uitgewerkt. Daarbij wordt aandacht besteed aan klassieke locatietheorieën, zoals van Von Thünen, en uitkomsten van internationale en nationale onderzoeken die in zekere mate zijn gebaseerd op deze locatietheorie. Ook wordt stilgestaan bij de ontwikkeling van de agrarische sector in Nederland, evenals bij bekende waardebepalende factoren bij agrarisch vastgoed. Aan de hand van het geschetste theoretisch kader zullen de waardebepalende factoren voor dit onderzoek worden gedefinieerd.

Een beschrijving van de taxatieleer en waarderingssystemen leidt hoofdstuk 3 in die verder in het teken staat van de regressie als statistische methode. Aan de hand van een beschrijving van de databronnen en de bewerkingen van de gegevens wordt inzichtelijk hoe voor dit onderzoek is gekomen tot bruikbare variabelen. Op basis van de waardebepalende factoren gevonden in hoofdstuk 2 wordt tot slot een verwachting uitgesproken over mogelijke verbanden tussen de variabelen.

In hoofdstuk 4 wordt de beschrijvende statistiek gepresenteerd, evenals de voorwaarden voor meervoudige regressie. Door uitvoering van regressieanalyses wordt duidelijk welke variabelen een sterke correlatie hebben en significant zijn. Uiteindelijk doel is het formuleren van een significant regressiemodel met een zo groot mogelijk verklarende kracht. Daarvoor zijn de regressieanalyses en de resultaten verder uitgeschreven in dit hoofdstuk.

In hoofdstuk 5 worden de centrale vraag en de deelvragen beantwoord. Daarnaast wordt een reflectie op het uitgevoerde onderzoek weergegeven, als mede aanbevelingen voor eventueel vervolgonderzoek.

2 Theorie

2.1 Klassieke locatietheorieën

De eerste locatietheorieën zijn gebaseerd op de klassieke vestigingsplaatsleer en zijn ontwikkeld door niet-geografen. In deze theorieën speelt de landbouwproductie en de afstand tot de centrale marktplaats een belangrijke rol. Ricardo en Von Thünen waren een van de eersten die hier onderzoek naar deden en over publiceerden.

2.1.1 Ricardo

In Engeland ontstond in 1815 een controverse over de graanwetten. Als gevolg van een daling van de graanprijzen besloot het parlement om het tarief voor geïmporteerd graan te verhogen. Dat leidde tot publieke verontwaardiging in het land en vormde de aanleiding voor David Ricardo (1772-1823) om zijn "essay over de invloed van een lage graanprijs op de winst uit voorraad" (1815) te publiceren. Hierin beargumenteerde hij dat de verhoging van het tarief op graanimport tot gevolg had dat de grondbezitters de huurprijzen voor landbouwgrond konden verhogen, terwijl de winsten van de fabrikanten werden verlaagd (Sprengler, 2018). Dit wordt weergegeven in de bekende frase van Ricardo:

'Corn is not high because rents are high, but rents are high because corn is high'

Volgens Wyatt (2013, p.7) houdt deze theorie van Ricardo in dat de grondhuur volledig door de vraag wordt bepaald, omdat de landbouwgrond eenmalig als geheel in gebruik wordt gegeven voor de teelt van slechts één gewas (i.c. graan).

In 1817 analyseerde Ricardo in zijn werk "Principes van politieke economie en belastingen" de wetmatigheden die de verdeling bepalen van alles wat geproduceerd zou kunnen worden door 'de drie klassen', te weten de grondbezitters, de arbeiders en de kapitaalbezitters. Een van de conclusies van zijn verdelingstheorie is dat de winst omgekeerd evenredig is met de lonen, omdat deze stijgen of dalen in overeenstemming met de benodigde kosten. Een andere conclusie is dat bij bevolkingsgroei de grondhuur (kan) stijgen, als gevolg van de hogere kosten van het verbouwen van meer voedsel voor de grotere bevolking. Hij bleef waakzaam voor snelle bevolkingsgroei die de lonen tot het bestaansminimum kon drukken en hij veronderstelde dat er geen tot weinig werkloosheid was. Een lage inkomensbesteding zou immers zowel de winst als de kapitaalvorming beperken door de marge van de teelt te vergroten (Sprengler, 2018).

Naast de vruchtbaarheid van de grond erkende Ricardo ook dat grond dichtbij de centrale handels- of arbeidsmarkt voordelen met zich meebracht. Daarbij ging hij ervan uit dat deze grond dezelfde gewasopbrengst zal opleveren als grond dat verder weg is gelegen, maar lagere transport- en arbeidskosten heeft waardoor meer inkomsten worden geneert. De gebruikswaarde van grond dichtbij de centrale markt is dus groter en deze waarde wordt via een concurrerende bieding van de gebruiker overgedragen naar de grondeigenaar in de vorm van grondhuur (Wyatt, 2013, p. 15).

In de volgende paragraaf zal worden ingegaan op de toepassing van de (grond)huurtheorie van Ricardo in een ruimtelijke context door de Duitse grootgrondbezitter Von Thünen, waarbij hij de relatie aantoonde tussen het vermogen om huur te betalen voor een stuk landbouwgrond en de afstand tot de centrale markt waarop de landbouwproducten worden verhandeld.

2.1.2 Von Thünen

Het boek *Der isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie* van Johann Heinrich von Thünen (1783-1850) is een standaardwerk dat gaat over de geografie en de economie gebaseerd op de agrarische sector (Paul, 2009). Vanaf 1811 tot en met 1820 hield Von Thünen een boekhouding bij van wat de productie van boter, rogge en hout op zijn landgoed had gekost en wat het had opgebracht. Door zijn eigen geïsoleerde staat te creëren kon hij de interne werking van de economie bestuderen. Zo kon hij de factoren identificeren die van invloed waren op de locatie van

verschillende soorten landbouw rond een centrale stad. Uit zijn boekhouding over de productie en de kosten leidde hij zijn model af, waarbij hij de volgende beperkende veronderstellingen hanteerde:

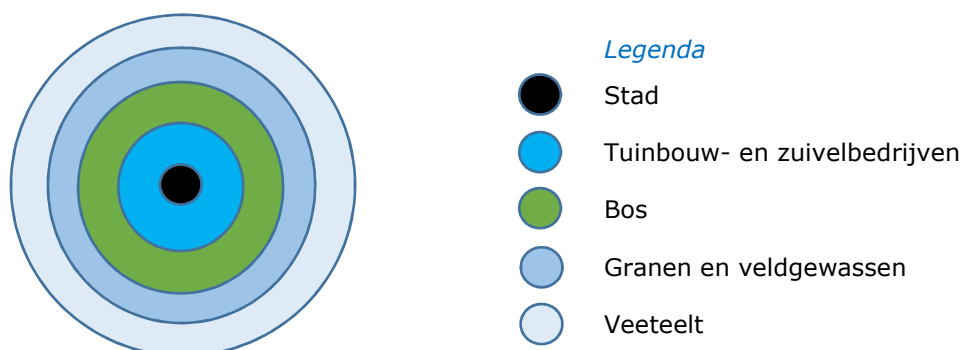
- de stad ligt centraal in een geïsoleerde staat die zelfvoorzienend is en geen externe invloeden heeft;
- de geïsoleerde staat wordt omringd door een onbezette wildernis;
- het land van de geïsoleerde staat is volledig vlak en heeft geen rivieren of bergen om het terrein te onderbreken;
- de bodemkwaliteit en het klimaat zijn overal in de geïsoleerde staat consistent;
- boeren in de geïsoleerde staat transporteren hun eigen goederen via ossenkar, over het land, rechtstreeks naar de centrale stad. Daarom zijn er geen wegen;
- boeren handelen om de winst te maximaliseren.

Paul (2009) beschrijft dat Von Thünen uit de waarnemingen ontdekte dat het optimale grondgebruik leidde tot een patroon van zes concentrische ringen rond de stad. Naarmate de afstand tot de centrale plaats toenam, veranderde de aard van de producten. Dit had onder andere te maken met bederf van de producten (Strootman Landschapsarchitecten, 2018, p.20) en de dure houtproductie door het transport (Bieleman, 1990, p.539). In de theorie van Von Thünen bepalen derhalve de factoren houdbaarheid, vervoerbaarheid en de opbrengst per hectare de afstand van een landbouwproduct tot de markt. Hierdoor ontstaat een onderlinge en geografische ordening van teelten ten opzichte van de centrale stad (Veen, De Jong, Jansma, Van der Schans & Lassu, 2013, p.4).

Het model van Von Thünen beschreven in zijn boek *Der isolierte Staat* ziet er als volgt uit. In de eerste ring rond de centrale stad konden het best gewassen met de hoogste opbrengst worden geteeld, zoals groenten. Daarnaast was de eerste ring ook geschikt voor intensieve veehouderij, zoals melkproductie. Houtproductie in bossen was in de tweede ring het meest rendabel. Niet vreemd gezien het feit dat in die tijd hout belangrijk was als energiebron en als bouw materiaal. Akkerbouw waarvoor geen braakperiode nodig was kon het beste in de derde ring plaatsvinden. De vierde ring was geschikt voor afwisselend akkerbouw en weideveeteelt. Het drieslagstelsel, de extensieve vorm van graanteelt in combinatie met braak, kon in de vijfde ring worden gebruikt. De zesde ring was alleen geschikt voor extensieve veehouderij (Paul, 2009. P.198).

Vaak ziet men in de literatuur het model van Von Thünen weergegeven met vier ringen (zie figuur 3). In een beschrijving van het model door Bieleman (1990, p. 539) zijn de eerste twee ringen, gericht op tuinbouw en zuivel respectievelijk op bosbouw, hetzelfde. De derde ring omvat daarentegen een drietal akkerbouwstelsels die van binnen naar buiten steeds extensiever worden, vergelijkbaar met de eerder genoemde derde, vierde en vijfde ring van Von Thünen. Ten slotte de vierde ring met extensieve veehouderij dat gericht is op het fokken van slachtvee.

Figuur 3: Model van Von Thünen (volgens de beschrijving door Bieleman)



Bron: eigen bewerking

2.1.3 De klassieke locatietheorie in Nederland

Kan de theorie van Von Thünen, ontwikkeld op zijn eigen landgoed, worden getransponeerd op de schaal van Nederland? Voor West-Nederland beschrijft Van der Ploeg de ontwikkeling van een bijzonder cultuurlandschap. De veengebieden werden ontgonnen door boeren. De gewassen die verbouwd werden op hun land werden verhandeld in nabij gelegen steden. Als gevolg van de graanbouw, die veel mineralen uit het verterende veen opnamen, vernatte het gebied waardoor boeren met akkerbouw moesten stoppen. De grond was alleen nog geschikt voor weidebouw, omdat deze minder kwetsbaar is voor een hoog grondwaterpeil. De overschakeling van akkerbouw naar weidebouw was mogelijk door de relatief gunstige prijzen voor zuivel in vergelijking met graan. Rond de steden in het veenweidegebied ontstonden consumptiemelkgebieden en verderop vond de kaasproductie en het vetmesten van varkens plaats. Door de boeren werd niet alleen gebruik gemaakt van de stad als afzetmarkt, maar men onttrok uit de stad mest en organisch afval, dat als goedkoop veevoer werd gebruikt. Als gevolg daarvan ontstond een intensieve rundveehouderij met name rond steden met drankstokerijen en levensmiddelenindustrieën (2002, p. 201). Door een groeiende bevolking, een internationale markt en de introductie van nieuwe transprottechnieken kende West-Nederland in de 16^e en 17^e eeuw een grote specialisatie in de landbouw. Boeren produceerden niet alleen voor de eigen stedelijke markt, maar ook voor de internationale markt (Bieleman, 1990, p. 537).

Op de zandgronden daarentegen was toen sprake van een gesloten economie. Maar door de toenemende vraag naar granen, vlees en wol door een snel groeiende bevolking in West-Nederland raakte de economie op de zandgronden steeds meer vervlochten met de groeiende en stedelijke economie in West-Nederland (Bieleman, 1990, p. 538).

In het dunbevolkte Drenthe waren de extensieve bedrijfstvormen. De rundveehouderij weidde hun vee op de heidesteppe. Na drie tot vijf jaar werden de dieren naar het westen van Nederland gedreven, waar ze werden vetgeweid en vervolgens geslacht. Door de vraag van de expanderende lakennijverheid in het westen werd de extensieve schapenhouderij uitgebreid voor de wolproductie. Daarnaast werd de akkerbouw intensiever met de overgang naar een zomer- en wintergewas op basis van een tweeslagsysteem zonder braak (Bieleman, 1990, pp. 541-543).

Op de centrale zandgronden (Veluwe, Salland) maakte de extensieve veehouderij plaats voor de meer intensieve akkerbouw als gevolg van de toenemende vraag naar broodgraan. Door stijgende graanprijzen is met name de teelt van boekweit, dat goed groeit op arme grond, uitgebreid. Ook het aantal paarden en schapen nam op onder andere de Veluwe af. Het paard had alleen nog maar de functie van trekdier (Bieleman, 1990, pp. 544-543).

Op de zuidelijke zandgronden met een hogere bevolkingsdichtheid was sprake van zeer intensief bedreven landbouw. Dat blijkt uit zware bemesting en intensieve bewerking van de grond dat leidde tot bol-liggende percelen. In Brabant was de stoppelweide veel eerder verdwenen dan op de centrale zandgronden en in Drenthe, en was men overgegaan op vruchtwisseling met 'groenbraak'. Daarbij werd het land na de graanoogst ingezaaid met een nagewas (spurrie) dat later werd ondergeploegd. Naast de reguliere graanteelten nam de teelt van oliehoudende gewassen toe, zoals vlas. Door de teelt van spurrie waren de zuidelijke zandgronden de eerste die grote hoeveelheden boter voor de markt produceerden. Hieruit blijkt al vroeg dat in Brabant de bedrijfsvoering veel intensiever was dan op de overige zandgronden (Bieleman, 1990, pp. 548-550).

Het voorgaande maakt duidelijk dat de landbouw in de verschillende regio's van Nederland het gevolg was van een krachtenveld waarin de afstand tot de belangrijkste steden in het Nederland en ook het Vlaanderen van de 16^e eeuw een grote rol speelde (Bieleman, 1990, p. 552).

2.2 Overzicht bestaande literatuur

In de bestaande academische literatuur lijkt niets te zijn geschreven over waardebepalende factoren en regressiemodellen voor de waardering van gebouwd agrarisch vastgoed. Over onbebouwde landbouwgronden daarentegen is wel het een en ander internationaal en nationaal gepubliceerd. In deze paragraaf worden een aantal van deze onderzoeken beknopt weergegeven.

In de Verenigde Staten (VS) zijn verschillende onderzoeken gedaan naar de prijs van landbouwgrond. De aanleidingen hiervan zijn zeer divers. Zo is een forse investering van tijd en geld gemoeid bij grondtransacties, dat vaak resulteert in een aanzienlijke persoonlijke interactie tussen verkoper en koper. Door Tsoodle, Golden en Featherstone (2006) is hiernaar onderzoek gedaan om de invloed van persoonlijke relaties op de verkoopprijzen van landbouwgrond in Kansas (VS) te kwantificeren. Daarnaast is onderzocht of de invloed van sociaal kapitaal² is veranderd in de loop van de tijd. Voor het hedonische model zijn onder andere de volgende verklarende variabelen gebruikt: perceelgrootte, grondverbeteringen (wel/geen irrigatie), bodemproductiviteit, type toegangswegen, bevolkingsdichtheid en verkoopomstandigheden. Uit het onderzoek is naar voren gekomen dat transacties tussen gelieerde partijen plaatsvonden met een korting van 43% op de verkoopprijs per acre³. De effecten van sociaal kapitaal zijn toegenomen, waarbij de korting op verkopen tussen gelieerde partijen met 2% per jaar toenam.

Door Huang, Miller, Sherrick en Go'mez (2006) is onderzoek uitgevoerd naar factoren die de waarde van landbouwgrond in Illinois (VS) beïnvloeden. Om de waarde van landbouwgronden te schatten is gebruik gemaakt van een hedonisch model. Uit het onderzoek komt naar voren dat de waarde van landbouwgrond toeneemt met de bodemproductiviteit, de bevolkingsdichtheid en het persoonlijk inkomen. De waarde blijkt af te nemen met de perceelgrootte, het landelijke karakter van de staat Illinois en de afstand tot Chicago en andere grote steden met meer dan 50.000 inwoners. Bovendien heeft de waarde van de landbouwgrond een negatieve relatie met het aantal varkensbedrijven per vierkante mijl, terwijl de bedrijfsomvang positief (maar niet statistisch significant) gerelateerd lijkt te zijn aan de waarde van landbouwgrond.

Sinds de zeventiger jaren is de prijs van landbouwgrond in het Midwesten van de Verenigde Staten ongekend gestegen. Dit was het gevolg van een combinatie van lage rentetarieven, een toegenomen vraag naar granen en kleinere voorraden. Bij Stokes en Cox (2014) leidde dit tot de vraag of er al dan niet speculatie is met landbouwgrond en in hoeverre de speculatie heeft bijgedragen aan de stijging van de landbouwgrondprijs. In hun paper beschrijven ze dat de graanprijzen en -opbrengsten van de geteelde gewassen de belangrijkste beïnvloedende elementen zijn van huurprijzen en waarden van de landbouwgrond. De hogere graanprijzen leiden vaak tot meer winstgevendheid van de agrarische bedrijfsvoering. Dit geeft aan de boer het economische signaal dat uitbreiding van hun (productie)activiteiten waarschijnlijk winstgevend zou zijn, waardoor ze hogere bedragen kunnen bieden voor te verwerven landbouwgrond.

Dat hoge graanprijzen tot speculatie zou leiden is in het begin van onderzoek middels een aanname uitgesloten. De landbouwgrond neemt dan weliswaar in waarde toe als de graanprijzen stijgen, maar benadert de intrinsieke of agrarische waarde van de grond die wordt gevonden met de kapitalisatieformule.

Vervolgens zijn speculaties inzichtelijk gemaakt door de aannames te wijzigen. Een investeerder die anticipeert op hogere graanprijzen zou de landbouwgrond mogelijk boven de agrarische grondprijs kunnen kopen, als hiermee tussen de aankoop en de verkoop op een later tijdstip voldoende winst kan worden gerealiseerd. Bij een graanprijs van \$ 3 per bushel⁴ is een verwaarloosbaar bedrag te wijten aan speculatie. Is de graanprijs \$ 4 per bushel is slechts 2% van de waarde van de landbouwgrond speculatief. De speculatieve component neemt verder toe naarmate de graanprijs toeneemt; 12% bij \$ 6, 20% bij \$ 7 tot 33% bij \$ 8 per bushel.

Het anticiperen op hogere graanprijzen verhoogt dus niet alleen de agrarische waarde van landbouwgrond, maar verhoogt ook het gedeelte van die waarde dat kan worden toegerekend aan speculatie.

² Sociaal kapitaal wordt door Robison, Myers en Siles (1999) gedefinieerd als 'de sympathie of het gevoel van verplichting van een persoon of groep voor een andere persoon of groep'. Deze definitie veronderstelt dat relaties tussen individuen de voorwaarden en waarschijnlijkheid van transacties kunnen wijzigen.

³ Amerikaanse oppervlaktemaat: 1 acre is gelijk aan 40,4687 are

⁴ Amerikaanse inhoudsmaat: 1 bushel is gelijk aan ca. 0,03524 m³

Op nationaal niveau zijn ook verschillende onderzoeken uitgevoerd naar factoren die invloed hebben de prijs van landbouwgrond. Cottelaar, Gardebroek en Luijt (2008) hebben onderzocht of marktmacht de prijs van landbouwgrond beïnvloedt. Hiervoor zijn hedonische prijsmodellen aangepast, waarbij onderscheid is gemaakt in landelijke en stedelijke landbouwgrondmarkten om marktmacht van zowel de verkopers als de kopers inzichtelijk te maken. Met behulp van het model is de transactieprijs per hectare verklaard met behulp van variabelen zoals onder andere: bedrijfstype en -omvang, ruimtelijke ordening, verpachting, familierelaties, aantal verkopers en kopers op de plaatselijk markt, persoonlijke kenmerken van verkopers en kopers (waaronder de ondernemingsvorm), perceelgrootte, bodemtype en afstanden tot het gekochte perceel en ander grondgebruik (glastuinbouw, snelweg, water, natuur-, recreatie-, industrie- of woongebied).

Uit de resultaten van het onderzoek blijkt dat op landelijke grondmarkten sprake is van marktmacht tussen het aanwezig aantal verkopers en kopers binnen een straal van 6,7 km van het verkochte perceel. De prijzen voor landbouwgrond in stedelijke gebieden worden echter gedomineerd door speculatie-effecten, omdat boeren niet langer de enige kopers van grond zijn.

Ten aanzien van de persoonlijke kenmerken van verkopers en kopers hebben de variabelen verkoper met kinderen, leeftijd van de koper en koper met een persoonlijke onderneming een significant effect op de grondprijs. Deze variabelen beïnvloeden de onderhandelingspositie en hebben dus een rechtstreeks effect op wat kopers bereid zijn te betalen of wat verkopers bereid zijn te accepteren.

Tot slot blijken variabelen zoals bodemtype en bedrijfsomvang niet van grote invloed te zijn bij de bepaling van de prijs van landbouwgrond. Daarnaast worden wel negatieve externaliteiten voor de landbouw onderkend wanneer de landbouwgrond in meer verstedelijkt gebied is gelegen of in de nabijheid van een snelweg. Wanneer de grond echter is gelegen in de nabijheid van recreatieve gebieden of glastuinbouw heeft dit in zowel in landelijke als in verstedelijkte gebieden een positief effect op de grondprijs.

De ontwikkeling van de landbouwsector onder druk van de wereldmarkt, stedelijke gebieden en nationale concurrentie heeft Kroes (2013) doen besluiten een onderzoek te doen binnen de specifieke regionale grondmarkten naar de invloed van de ontwikkeling van de landbouwsector op de grondprijs. In de hedonische prijsmethode die is gebruikt voor het onderzoek was de ruimtelijke component een belangrijk element. Om de ontwikkelingen in beeld te kunnen brengen is gewerkt met een dataset met gegevens uit 2000 en met gegevens uit 2010. Om de grondprijs te kunnen verklaren zijn de volgende variabelen gehanteerd: grondoppervlakte transactie, bevolking, bevolkingsgroei, afstanden tot natuur, tuinbouwgebied en snelweg, grondgebruik, concurrentievermogen van de landbouwsector en landbouwgebiedsgroep.

De resultaten van het onderzoek laten een significant effect zien van de verschillende landbouwgebiedsgroepen op de prijs voor landbouwgrond. Deze grondprijs wordt met name beïnvloed door het concurrentievermogen van de tuinbouwsector en de stedelijke druk als gevolg van bevolkingstoename.

Ondanks dat de hierboven aangehaalde literatuur een andere focus heeft dan dit onderzoek, namelijk waardebepalende factoren ten aanzien van de prijs van landbouwgrond, geeft het wel enig zicht op factoren die mogelijk bij dit onderzoek kunnen worden toegepast. Alvorens deze mogelijke factoren te benoemen (zie paragraaf 2.5) wordt in de volgende paragraaf een beschrijving op hoofdlijnen gegeven van de agrarische sector, zodat de lezer zich een beeld kan vormen van de context.

2.3 De agrarische sector

2.3.1 Ontwikkeling van de agrarische sector

Nederland heeft zich in de loop der eeuwen ontwikkeld tot een vooruitstrevende agrarische natie. In de vooroorlogse periode (voor 1940) was het platteland grotendeels agrarisch in gebruik. Akkerbouwbedrijven vond men op de kleigebieden en het traditionele gemengde bedrijf met akkers, hooilanden en zeer extensief gebruikte woeste gronden op de zandgronden. Grote delen woeste grond konden tussen 1850 en 1900 worden omgezet in landbouwgrond door de intrede van kunstmest (Gies, Nieuwenhuizen, Naeff & Van Och, 2016).

Bieleman (2008) beschrijft dat na de oorlog, onder leiding van Mansholt, alles was gericht op het op orde brengen van de voedselvoorziening. Door de overheid werd geïnvesteerd in onderzoek, onderwijs en schaalvergroting om boeren zo een economisch bestaan te geven. De agrarische bedrijven ontwikkelden zich destijds tot gezinsbedrijven.

Na de wederopbouw was de periode van schaalvergroting en rationalisatie van de landbouw (1965-1993). Om het inkomensniveau te verbeteren werd ingezet op kostprijsreductie, omdat de prijzen voor producten niet stegen (Bieleman, 2008). Arbeid en in mindere mate grond werden vervangen door kapitaal en aangekochte productiemiddelen (Vermeulen, 2007). Agrarische bedrijfsgebouwen werden op de goedkoopste en efficiëntste wijze gebouwd. Daarom werden in deze periode met name bijna alle stallen bedekt met asbesthoudende golfplaten⁵ (Gies, Nieuwenhuizen & Smidt, 2014). De overheid bleef de agrarische sector stimuleren door middel van onderzoek, voorlichting en onderwijs. Daarnaast voerde de overheid ruilverkavelingen met bedrijfsverplaatsingen uit om zo de landbouwstructuur te verbeteren. Het traditionele gemengd bedrijf op de zandgronden ontwikkelde zich tot een intensieve veehouderij. Ook de groenteteelt onder glas ontwikkelde zich in deze periode tot een belangrijke sector (Gies et al., 2016).

In de negentiger jaren kwamen de eerste gevolgen van de groei van de agrarische sector aan het licht. Binnen Europa ontstond overproductie (boterberg, wijnplas e.d.) en milieuproblemen (stikstof, fosfaat) als gevolg van de intensivering van de voedselproductie. Om dit het hoofd te bieden werden milieu- en productieregulerende maatregelen genomen. Dit ging gepaard met forse investeringen die gecompenseerd werden door bedrijfsvergroting en/of meer specialisatie. De overheid maakte het mogelijk dat bedrijven zich op de bestaande locatie ontwikkelden. Maar ook dierziekten, zoals de varkenspest in 1997, deden hun intrede. Dit leidde tot een herstructurering van de intensieve veehouderij waarbij bedrijven zijn gestopt en stallen zijn gesloopt (Gies et al., 2016). Tot 2010 is circa 3 miljoen m² stallen gesloopt in de provincies Utrecht, Overijssel, Gelderland, Noord-Brabant en Limburg in het kader van de Regeling Beëindiging Veehouderijtakken en de Ruimte voor Ruimte-Regeling (Gies et al., 2014).

2.3.2 Agrarische bedrijfstypen

De agrarische sector kan worden onderverdeeld in verschillende bedrijfstypen. Het CBS hanteert hierbij onderstaande verdeling in 8 hoofdbedrijfstypen (zie tabel 1). Deze kunnen nader worden gegroepeerd tot grondgebonden bedrijven (extensieve landbouw), zoals akkerbouwbedrijven, tuinbouwbedrijven, blijvende teeltbedrijven, graasdierbedrijven en gecombineerde bedrijven, respectievelijk niet-grondgebonden bedrijven (intensieve landbouw), zoals glastuinbouwbedrijven, hokdierbedrijven en veeteeltcombinaties.

Tabel 1: Aantal landbouwbedrijven

	2002		2007		2012		2017	
	aantal	%	Aantal	%	Aantal	%	aantal	%
Akkerbouwbedrijven	14.130	16%	12.144	16%	12.016	17%	10.685	20%
Tuinbouwbedrijven (*)	15.221	17%	12.043	16%	9.394	14%	7.078	13%
Blijvende teeltbedrijven	2.151	3%	1.859	2%	1.728	3%	1.612	3%
Graasdierbedrijven	42.338	47%	38.694	50%	36.682	53%	27.557	50%
Hokdierbedrijven	8.943	10%	7.389	10%	5.661	8%	4.650	8%
Gewascombinaties	1.140	1%	941	1%	833	1%	1.119	2%
Veeteeltcombinaties	2.379	3%	1.567	2%	819	1%	581	1%
Gewas/vee combinaties	2.907	3%	2.104	3%	1.677	3%	1.558	3%
Totaal alle bedrijven	89.479		76.741		68.810		54.840	

(*) waaronder glastuinbouwbedrijven

Bron: CBS, 2018a

⁵ I.v.m. het asbestverbod moeten eigenaren voor 2025 asbestdaken op gebouwen (laten) verwijderen (bron: www.infomil.nl).

Akkerbouwbedrijven

In 2017 telde Nederland ruim 10.000 akkerbouwbedrijven tegenover ongeveer 14.000 in 2002. Het areaal akkerbouwgrond bedroeg in 2002 466.000 hectare en 444.000 hectare in 2017 (CBS, 2018a). De akkerbouw is het sterkst grondgebonden en heeft daardoor de minste mogelijkheden om de productie te vergroten door vervanging van grond. De productiviteit van de grond is in de loop der jaren wel opgevoerd. Kwam rond 1950 kwam ongeveer 4 ton tarwe van een hectare, tegenwoordig wordt bijna 8,5 ton gerealiseerd (Vermeulen, 2007). Een gemiddeld akkerbouwbedrijf heeft ongeveer 60 hectare landbouwgrond. Ruim 1.000 bedrijven hebben meer dan 100 hectare landbouwgrond (Agrimatie, 2018a). De meest geteelde akkerbouwgewassen zijn graan, aardappelen, suikerbieten en uien (Van Gellicum, 2015).

Tuinbouwbedrijven en blijvende teeltbedrijven

Deze hoofdbedrijfstypen omvatten vollegrondsgroenteteelt, bloembollenteelt, boomkwekerij, fruitteelt, champignonenteelt en glastuinbouw. De bloembollenteelt is met name geconcentreerd in de Duin- en bollenstreek, terwijl de boomkwekerij in verschillende regio's in Nederland plaatsvindt. De fruitteelt treft men aan in de Betuwe, Zeeland en de kop van Noord-Holland. De belangrijkste productielocaties in de glastuinbouw liggen in het Westland, Oostland, gebieden rond Aalsmeer, Rijnsburg, Bleiswijk en Venlo (Van Gellicum, 2015).

In Nederland werd in 2017 in circa 7.000 hectare tuinbouwkassen groenten (ca. 5.000 ha), snijbloemen (ca. 1.100 ha), sier- en vaste planten (ca. 800 ha) en fruit geteeld (ca. 180 ha) (CBS, 2018b). Een gemiddeld glastuinbouwbedrijf met groenteteelt heeft een oppervlakte van ruim 3 hectare en bijna 20% van deze bedrijven heeft een oppervlakte groter dan 5 hectare. In de snijbloemen, de sier- en vaste planten is een gemiddeld bedrijf 1,6 hectare groot en 16% van deze bedrijven heeft een oppervlakte groter dan 3 hectare (Van Gellicum, 2015).

Graasdierbedrijven

Onder dit hoofdbedrijfstype vallen met name melkveehouderij (ca. 50% van het totale aantal graasdierbedrijven), maar ook rundveehouderij (met uitzondering van kalverhouderij), schapen- en geitenhouderij. In 2017 telde Nederland ruim 27.500 graasdierbedrijven tegenover ruim 42.000 in 2002. Het areaal weidegrond bedroeg in 2002 1.086.000 hectare en 1.029.000 hectare in 2017 (CBS, 2018a). Een melkveehouderij heeft gemiddeld 55 hectare landbouwgrond bij het bedrijf (Agrimatie, 2018b). De verwachting is dat de trend van bedrijfsbeëindigingen aan de ene kant en schaalvergroting aan de andere kant doorzetten, waardoor het aantal melkveehouderijen naar verwachting zal afnemen naar 10.000 in 2020 (Van Gellicum, 2015). Per bedrijf is het gemiddelde aantal melkkoeien in 2018 toegenomen tot 103 stuks ten opzichte van 57 stuks in 2000. In 2017 stopte als gevolg van de stoppersregeling, om de fosfaatproductie terug te brengen, met name melkveehouderijen die qua omvang de helft zijn van het gemiddelde (Agrimatie, 2018b).

Hokdierbedrijven

Dit hoofdbedrijfstype ziet met name op de varkenshouderij, met fokvarkens-, vleesvarkens- en gesloten varkensbedrijven, en de pluimveehouderij, met vleeskuikens- en leghennenbedrijven. Daarnaast hoort ook de kalverhouderij tot dit bedrijfstype.

In 1997 bereikte de varkensstapel (biggen, fokvarkens en vleesvarkens) zijn maximale omvang van 15,2 miljoen dieren (Vermeulen, 2007). Inmiddels is dat geslonken tot circa 12,4 miljoen dieren in 2017 (CBS, 2018c) onder invloed van het mestbeleid (Vermeulen, 2007). In 2017 waren er 2.960 varkensbedrijven en dit aantal krimpt nog steeds. Op fokvarkensbedrijven zijn gemiddeld 790 fokvarkens aanwezig. Vleesvarkens- en gesloten varkensbedrijven hebben gemiddeld 2.300 vleesvarkens (Agrimatie, 2018c).

De pluimveestapel (legghennen en vleeskuikens) omvatte bijna 105 miljoen dieren in 1999. De uitbraak van vogelpest in 2003 heeft een sterke inkrimping van de pluimveestapel tot gevolg gehad (Vermeulen, 2007). Inmiddels zijn er in 2017 iets minder dan 95 miljoen dieren (CBS, 2018d). Nederland telde in 2017 ongeveer 460 vleeskuikenbedrijven, met gemiddeld 80.000 vleeskuikens,

en 630 leghennenbedrijven, met gemiddeld 47.000 leghennen. Ook in de pluimveehouderij neemt het aantal bedrijven nog steeds af (Agrimatie, 2018d)

De kalverhouderij omvatte in 2017 ongeveer 1.150 bedrijven met gemiddeld 950 dieren. De afgelopen jaren is het aantal bedrijven afgenomen, maar de gemiddelde veebezetting is daarentegen toegenomen (Agrimatie, 2018e).

2.3.3 Toekomstige ontwikkelingen

Als gevolg van verhoging van de arbeidsefficiëntie door de voortschrijdende technische ontwikkelingen blijft volgens Gies et al. (2014) schaalvergroting een onomkeerbare trend voor de meeste agrarische bedrijven. In de periode 2000–2012 zijn ruim 22.000 agrarische bedrijven gestopt. Als deze trend doorzet naar 2030, dan zullen nog tussen de 16.000 en 24.000 agrarische bedrijven stoppen. Met name in regio's in Zuidwest-Friesland, Noordoostpolder, Overijssel, Gelderland, Oost-Brabant en Noord-Limburg komt dan veel agrarische bebouwing leeg te staan.

De belangrijkste ontwikkelingen in de akkerbouw zijn schaalvergroting en innovatie. Zo wordt met precisielandbouw per vierkante meter de behoefte van gewassen bepaald. Door de hoge grondprijzen voor akkerbouwgrond en de vergrijzing zullen eigendom en exploitatie van landbouwgrond steeds meer van elkaar gescheiden zijn (Rabobank, 2018a).

Binnen de glastuinbouwsector is sprake van clustervorming, waarbij bedrijven zijn gebundeld in concentratiegebieden. Dit geldt met name voor de concentratiegebieden in Zuid-Holland waar bedrijven over het algemeen goed presteren met een goede ligging. Glastuinbouwbedrijven die stoppen zullen over het algemeen worden gekocht door ondernemers die uit willen breiden (Brans, Silvis & Voskuilen, 2018).

Als gevolg van de begrenzing van het aantal dieren door middel van fosfaatrechten worden marginale bedrijven verkocht buiten de melkveehouderij, bijvoorbeeld aan boomkwekers. De gebouwen brengen dan relatief weinig op, maar dit wordt gecompenseerd door de opbrengst van de landbouwgrond en de fosfaatrechten (Brans et al., 2018). Verder blijft er een groeiende vraag naar zuivelproducten als gevolg van groei van de wereldbevolking en het groeiend besteedbaar inkomen in ontwikkel(en)de landen. Bij de melkproductie blijven duurzaamheidsitems een belangrijke rol spelen, zoals behoud van weidegang, dierenwelzijn en – gezondheid en mineralenkringlopen (Rabobank, 2018b).

De intensieve veehouderij zit in een spagaat tussen enerzijds het omschakelen naar een duurzame, diervriendelijke en transparante productiemethode en anderzijds het internationale verdienmodel. De vraag naar varkensvlees en eieren blijft naar verwachting de komende decennia stijgen. Maar als gevolg van verscherpte wet- en regelgeving zal het aantal intensieve veehouderijbedrijven de komende jaren afnemen (Rabobank, 2018c).

2.4 Waardebepalende factoren bij agrarisch vastgoed

Bij grondgebonden woningen is veel ervaring opgedaan met geautomatiseerde waardebepaling. Hieruit komt naar voren dat bijvoorbeeld locatie, grootte, woningtype, aanwezige bijgebouwen en bouwperiode belangrijke waardebepalende factoren c.q. woningkenmerken zijn (Van Gool et al., 2013, p.319) (De Groot, Marlet, Teulings & Vermeulen, 2010, p.22).

Zoals uit het literatuuronderzoek in paragraaf 2.2 blijkt is alleen onderzoek verricht naar waarde-bepalende factoren bij landbouwgronden. Factoren die hierbij naar voren komen zijn bijvoorbeeld bodemtype, bodemproductiviteit, grondverbetering, perceelgrootte, ligging, afstanden tot stad, woon-, industrie- of glastuinbouwgebied, snelweg of natuur, toegangswegen, bevolkingsdichtheid en ruimtelijke ordening. Dit komt in grote lijnen overeen met de waardebepalende factoren die gangbaar zijn in de agrarische taxatiepraktijk. Vermeulen (2007, pp. 392-394) beschrijft dat bij de

waardebepaling van landbouwgrond van belang zijn het bodemtype, de ontwatering, de ontsluiting, de verkaveling, aantal en ligging van de percelen en geschiktheid voor teelten.

Bij de agrarische bedrijfsgebouwen zijn daarentegen van belang de ligging, bouwaard, kwaliteit, constructie, doelmatigheid, functionaliteit, algehele staat van onderhoud, bedrijfsoppervlakte en economische omstandigheden (Vermeulen, 2007, pp. 395-397). Bij de bepaling van de hoogst toelaatbare pacht prijs voor agrarische bedrijfsgebouwen speelt naast bedrijfstype de doelmatigheid een rol. De doelmatigheid is hierbij gekoppeld aan een leeftijdscategorie van de bedrijfsgebouwen (Silvis, Van der Meer & Voskuilen, 2018, p. 19). In het kader van de Wet WOZ wordt de waardering van agrarisch vastgoed gebaseerd op (arche)type, bouwjaarklasse, grootte, kwaliteit, onderhoudstoestand en doelmatigheid van het bedrijfsgebouw (VNG, 2018, p. 19).

2.5 Hypothese waardebepalende factoren voor het onderzoek

In voorgaande paragrafen zijn verschillende factoren beschreven die van invloed zijn op de waarde van landbouwgrond. Dit zijn onder andere de afstand tot het industrie- en stedelijke gebied (vroeger de centrale marktplaats genoemd), de bevolkingsdichtheid, de bodemproductiviteit, het bedrijfstype en de bedrijfsoppervlakte. Bij de waardering van agrarische bedrijfsgebouwen wordt met name gelet op de kwaliteit, doelmatigheid, functionaliteit en de oppervlakte van de gebouwen. Voor het onderzoek naar een regressiemodel is de verwachting dat deze factoren op een of andere manier van invloed zijn op de koopsom c.q. waarde van het agrarisch vastgoed.

Landbouwgrond gelegen in de periferie van de stad heeft dikwijls een hogere waarde dan landbouwgrond elders in het buitengebied. Mogelijk dat de agrarische bouwkael en de ondergrond van de woning en tuin op kortere afstand van de stad daarom ook een hogere waarde vertegenwoordigen. Een vergelijkbare redenering geldt voor gebieden met een grote bevolkingsdichtheid. Door meer stedelijk druk zal de koopsom voor het agrarisch vastgoed hoger zijn.

De bodemproductiviteit heeft mede invloed op de (gewas-)opbrengsten en dus op de prijs die een agrarisch ondernemer kan betalen voor het agrarisch vastgoed. Door aan te sluiten op de indeling van Nederland in landbouwgebieden⁶, wordt bij het onderzoek een impliciete vertaling naar de bodemproductiviteit verondersteld. Een landbouwgebied is immers een groep burgerlijke gemeenten met zoveel mogelijk gelijksoortige type landbouwbedrijven. Refererend aan Gies et al. (2016) zullen dit bijvoorbeeld met name akkerbouwbedrijven zijn in een landbouwgebied met overwegend kleigebieden, terwijl in een landbouwgebied met zandgronden dit overwegend intensieve veehouderijbedrijven zullen zijn. In het eerste geval is het bodemtype en de bodemproductiviteit van de omliggende landbouwgrond belangrijker voor de agrarische bedrijfsvoering dan in het tweede geval.

Om het bedrijfstype inzichtelijk te krijgen wordt aangesloten bij de acht categorieën die het CBS (2018a) heeft gedefinieerd, waarbij mogelijk bepaalde bedrijfstypen moeten worden samengevoegd als het aantal waarnemingen hiervan beperkt is.

Voor dit onderzoek is het ondoenlijk is om alle gegevens te verzamelen over de kwaliteit, doelmatigheid, functionaliteit en onderhoudstoestand van de agrarische bedrijfsgebouwen. Daarom wordt voor dit onderzoek verondersteld dat dit tot uitdrukking komt in het (gemiddelde) bouwjaar of bouwjaarklassen van de bedrijfsgebouwen en de woningen.

Hiervoor wordt gerefereerd aan de koppeling van de doelmatigheid aan de leeftijdscategorie van de bedrijfsgebouwen bij de bepaling van de pacht prijs voor de bedrijfsgebouwen (Silvis et al., 2018). Verder wordt door Gies, Nieuwenhuizen, Naeff, Vleemingh en Paulissen (2016) een verscheidenheid

⁶ Een landbouwgebied is een groep gemeenten, binnen eenzelfde provincie, met (zoveel mogelijk) gelijksoortige type landbouwbedrijven, zoals in 1991 door de toenmalige Adviescommissie Landbouwstatistiek is vastgesteld. Nederland telt 66 landbouwgebieden en 14 groepen van landbouwgebieden (CBS, 2018d).

aan bouwtypen onderkend. Zo waren in de vooroorlogse periode het woon- en bedrijfsgedeelte in één gebouw geïntegreerd of met elkaar verbonden. Voorbeelden hiervan de stolpboerderij in Noord-Holland, de kop-hals-rompboerderij in Friesland en de langgevelboerderij in Brabant en Noord-Limburg. In de wederopbouwperiode van 1940 tot 1965 werden verwoeste boerderijen zodanig herbouwd dat die een voor die tijd moderne bedrijfsvoering toelieten. In de ingepolderde Noordoostpolder en Oostelijk Flevoland werden destijds door de overheid ook voor die tijd moderne boerderijen gerealiseerd. Streekgebonden bouwen was in de periode van schaalvergroting en rationalisatie niet meer aan de orde. De agrarisch bedrijfsgebouwen werden op de goedkoopste en efficiëntste wijze gebouwd. In deze periode werden bijna alle stallen bedekt met asbesthoudende golfplaten (Gies et al., 2014). In de jaren negentig vonden in de agrarische sector veel bedrijfsvergrotingen en/of specialisaties plaats, waarvoor grotere en modernere bedrijfsgebouwen werden gebouwd.

Voor de woning bij het agrarisch bedrijf gelden vergelijkbare bouwperiodes als voor de bedrijfsgebouwen. Het bouwjaar van een woning kenmerkt volgens Liebrechts (2011) de kwaliteit van de woning door de voorschriften die op dat moment golden en de bouwtechniek die werd gehanteerd.

Als meest basale gegevens worden bovendien de bebouwde oppervlakte van zowel de agrarische bedrijfsgebouwen als van de woningen en de oppervlakte van de bouwkael en de tuin in dit onderzoek betrokken.

Samenvattend is de verwachting dat afstand tot een grote stad, bevolkingsdichtheid, geografische ligging in Nederland, bedrijfstypen, bouwjaar/-klassen, bebouwde oppervlakte van de bedrijfsgebouwen en woningen en de oppervlakte van de bouwkael en tuin verklarende variabelen kunnen zijn voor de afhankelijke variabele, te weten koopsom c.q. waarde van het agrarisch vastgoed. In het volgende hoofdstuk zullen deze variabelen nader worden beschreven en toegelicht.

3 Methode en data

Alvorens wordt ingegaan op de methode en de onderliggende dataset wordt eerst stil gestaan bij de taxatieleer ten aanzien van agrarische vastgoedobjecten, zodat de lezer zich een beeld kan vormen van de context met betrekking tot dit onderzoek.

3.1 Taxatieleer agrarische vastgoedobjecten

Over waarderingsmethodieken wordt in de literatuur veel geschreven. Volgens Van Gool, Jager, Theebe en Weisz (2013) is hierbij van belang op welke wijze de taxateur de waarde van het object bepaalt en of dat aansluit bij het gehanteerde waardebegrip. Met waardebegrip wordt bedoeld het doel van de te bepalen waarde. Bijvoorbeeld de marktwaarde als mogelijk opbrengstwaarde bij verkoop van het object.

In de Europese Taxatiestandaarden (EVS) wordt gesteld dat de taxatiemethodiek is gebaseerd op de werking van de markt. Dat heeft tot gevolg dat elke methodiek moet worden getoetst aan de economische basisprincipes van de echte wereld (TEGoVA, 2016, p. 312).

In de basis worden drie taxatiebenaderingen onderscheiden, namelijk de comparatieve benadering, de inkomstenbenadering en de kostenbenadering. In tabel 2 is aan de hand van de EVS en de beschrijving door Van Gool et al. (2013) een nadere onderverdeling van deze benaderingen voor de Nederlandse taxatiebranche weergegeven.

Tabel 2: Taxatiebenadering met waarderingsmethodieken

Comparatieve benadering	Inkomstenbenadering	Kostenbenadering
Directe vergelijkingsmethode	Huurfactormethode	Residuele waarde methode
Regressiemethode	Bruto-aanvangsrendement methode (BAR-methode)	Gecorrigeerde vervangingswaarde methode (GVW-methode)
	Netto-aanvangsrendement methode (NAR-methode)	
	Contante waarde methode (DCF-methode)	

Bron: eigen bewerking

3.1.1 Comparatieve benadering

Bij de comparatieve benadering vindt de waardebepaling plaats door het vastgoedobject te vergelijken met transacties van identieke of vergelijkbare vastgoedobjecten die zijn verhandeld op dezelfde markt, en is daarom goed bruikbaar voor de bepaling van de marktwaarde (van Gool et al., 2013, p. 318). Deze benadering kent twee methodieken, namelijk de directe vergelijkingsmethode en de regressiemethode.

Directe vergelijkingsmethode

Bij de directe vergelijkingsmethode wordt het vastgoedobject vergeleken met soortgelijke objecten in dezelfde straat, buurt of omgeving waarvan de recente transactieprijs bekend is. Als de vergelijkingsobjecten verschillen in grootte of oppervlakte kan geen directe vergelijking plaatsvinden. Dan wordt de transactieprijs uitgedrukt in een bedrag per kubieke meter en/of vierkante meter en vervolgens vermenigvuldigd met de grootte en/of oppervlakte van het te waarden vastgoedobject. Als de kenmerken van de vastgoedobjecten te veel van elkaar verschillen, dienen correcties toegepast te worden (van Gool et al., 2013, p. 318).

In de landbouw wordt de directe vergelijkingsmethode sectorbreed toegepast bij de waardering van losse percelen landbouwgrond en van landbouwbedrijven. De objectkenmerken van landbouwbedrijven verschillen in de praktijk echter erg veel waardoor het lastig is deze vergelijkbaar te maken. De transactieprijs van een landbouwbedrijf wordt daarom gesplitst in één of meer bestanddelen van het vastgoedobject, zoals de waarde van de bedrijfsgebouwen, de waarde de bouwkaavel en de waarde van de landbouwgrond. Deze kengetallen vormen dan de input voor de waardering van het vastgoedobject (Van Gellicum, 2015, pp. 518-519).

Regressiemethode

Bij deze methode wordt op basis van vele objectkenmerken en de transactiepreizen via een regressieanalyse een samenhang geschat tussen een aantal variabelen en de waarde van de vastgoedobjecten. Wanneer in het regressiemodel de kenmerken van het te waarderen object worden ingevoerd, wordt hiervan met een zekere mate van nauwkeurigheid de marktwaarde automatisch bepaald (Van Gool et al., 2013, p. 319). Bij vastgoedobjecten in de landbouw wordt de regressiemethode niet toegepast (Van Gellicum, 2015, p. 518).

In paragraaf 3.2. zal de regressiemethode nader worden beschreven.

3.1.2 Inkomstenbenadering

De inkomstenbenadering wordt gebruikt om toekomstige huur- c.q. kasstromen om te zetten naar een enkele kapitaalswaarde. Deze benadering kent vier methodieken die hierna kort worden beschreven.

Bij de huurfactormethode wordt de feitelijke bruto(huur) inkomsten vermenigvuldigd met een factor, die overeenkomt met recentelijk verkochte vergelijkbare vastgoedobjecten in dezelfde markt en onder dezelfde omstandigheden.

Bij de bruto aanvangsrendement (BAR-) methode wordt de bruto markthuur, gecorrigeerd voor eventueel op de markt gangbare verhuurincentives, gedeeld door het bruto aanvangsrendement⁷. Om de marktwaarde te bepalen dienen nog op basis van de contante waarde correcties plaats te vinden voor de over- of onderhuur, (aanvangs)leegstand en achterstallig onderhoud en tot slot de correctie voor kosten koper.

De netto aanvangsrendement (NAR-) methode is vergelijkbaar met de BAR-methode waarbij op de bruto markthuur de exploitatiekosten in mindering worden gebracht. Bij zowel de BAR- als de NAR-methode dient het betreffende aanvangsrendement overeen te komen met recentelijk verkochte vergelijkbare panden in dezelfde markt en onder dezelfde omstandigheden.

Bij de discounted-cash-flow (DCF-) methode worden de toekomstige kasstromen die uit de exploitatie en uit de verkoop van het vastgoedobject komen contant gemaakt tegen een disconteringsvoet (v.o.n.) die is afgeleid uit transacties van vergelijkbare vastgoedobjecten in dezelfde markt. Voor de uiteindelijke marktwaarde dient nog voor de kosten koper te worden gecorrigeerd (Van Gool et al., 2013).

Van bovengenoemde methodieken wordt alleen de huurfactormethode en de DCF-methode toegepast bij de waardering van verpachte en in erfpacht uitgegeven vastgoedobjecten in de landbouw, zoals landbouwgrond en hoeven (Van Gellicum, 2015, p.520).

3.1.3 Kostenbenadering

De kostenbenadering⁸, gebaseerd op het principe dat een koper niet meer voor een vastgoedobject zal betalen dan wat het zal kosten om een gelijkwaardig object te verkrijgen door aankoop of bouw, kent de volgende twee methodieken.

Bij de residuele waarde methode is de marktwaarde van het te realiseren vastgoedobject het uitgangspunt waarop alle te maken kosten in mindering wordt gebracht, zoals de bouwkosten, de eventuele sloopkosten en de kosten bouwrijp maken van het terrein en de (verwachte) winstmarge van de projectontwikkelaar. Wat resteert is de marktwaarde van de onbebouwde bouwrijpe grond, het residu (Van Gool et al., 2013). Deze methode wordt bij de waardering van vastgoedobjecten in de landbouw gebruikt in geval van functiewijzing, zoals bij ruimte-voor-ruimte regelingen en bij transformaties op het agrarische bedrijfserf (Van Gellicum, 2015, p. 521).

⁷ Het bruto aanvangsrendement is gedefinieerd als inclusief transactiekosten en overdrachtsbelasting (vrij op naam).

⁸ De kostenbenadering wordt met name gebruikt voor het waarderen van specifieke vastgoedobjecten waar geen markt voor is.

Bij de gecorrigeerde vervangingswaarde methode wordt uitgegaan van de vervangingswaarde van de grond en de opstallen. Vervolgens dient een correctie plaats te vinden voor technische en economische veroudering van het opstal. De som van de grondwaarde en de gecorrigeerde opstalwaarde(n) geeft de marktwaarde van het vastgoedobject weer (Van Gool, 2013). Bij de waardering van vastgoedobjecten in de landbouw wordt de gecorrigeerde vervangingswaarde methode bij de waardering van landbouwbedrijfsgebouwen algemeen toegepast. Door de lineaire afschrijving wordt inzichtelijk in welke mate de bedrijfsgebouwen nog voldoen aan de laatste eisen van de techniek. Daarnaast kunnen aanvullende economische correcties noodzakelijk zijn wanneer de landbouwsector langdurig wordt geraakt door bijvoorbeeld exportbeperkingen of aanhoudende hoge energieprijzen (Van Gellicum, 2015, p. 522).

In dit onderzoek zal verder geen aandacht meer worden besteed aan de kosten- en inkomstenbenadering.

3.2 Omschrijving regressiemethode

Door Van Arnhem, Berkhout en Ten Have (2013) wordt beschreven dat de regressiemethode naast de directe vergelijkingsmethode een comparatieve benadering is. Het is een techniek voor het analyseren van gegevens waarin mogelijk sprake is van een samenhang tussen variabelen, bijvoorbeeld een bepaald kenmerk van het vastgoedobject en de waarde. Bij de regressiemethode wordt onderscheid gemaakt in de enkelvoudige en de meervoudige regressie.

3.2.1 Enkelvoudige regressie

Bij een enkelvoudige regressie kan de waarde van het vastgoedobject (afhankelijke variabele; W) worden gespecificeerd vanuit één objectkenmerk (onafhankelijke variabele; O). Op basis van een lineaire samenhang (correlatie) tussen twee variabelen kan een voorspelling van W vanuit O worden gedaan. De correlatiecoëfficiënt, gelegen tussen -1 en 1, geeft de sterkte van het verband tussen de twee variabelen weer. Bij 1 (of -1) is sprake van een perfect positief (of negatief) lineair verband en bij 0 is geen sprake van een verband.

Een enkelvoudige regressie ziet er als volgt uit: $W_i = bO_i + a$

waarin:

W_i = voorspelde waarde van waarneming i op variabele W

O_i = waarde van waarneming i op variabele O

a = regressieconstante

b = regressiecoëfficiënt

De regressiecoëfficiënt en de regressieconstante worden bepaald met behulp van het kleinste-kwadratencriterium. De regressiecoëfficiënt (b) van de variabele O geeft het marginale effect van een eenheidsverandering in O op W weer. De regressieconstante (a) is de waarde van W als de waarde van variabele O gelijk is aan nul, en waarmee de regressievergelijking kloppend wordt gemaakt.

3.2.2 Meervoudige regressie

Bij dit onderzoek zal gebruik worden gemaakt van een meervoudige regressieanalyse, omdat wordt aangenomen dat de waarde van het agrarisch vastgoed afhankelijk is van meerdere verklarende factoren. Bij een meervoudige regressie gaat het om welke onafhankelijke varianties (objectkenmerken) samenhangen met de waarde en het vastgoedobject (Van Arnhem et al., 2013). Door gebruik te maken van de partiële correlatie kunnen de verschillende objectkenmerken ieder afzonderlijk worden gewogen.

Een meervoudige regressie ziet er als volgt uit: $W_i = b_1O_1 + b_2O_2 + \dots + b_nO_n + a$
waarin:

W_i = voorspelde waarde van waarneming i op variabele W
 $O_{1..n}$ = waarde van variabelen O_1 tot en met O_n
 a = regressieconstante
 $b_{1..n}$ = partiële regressiecoëfficiënten

De regressiecoëfficiënt geeft aan hoe groot het aandeel is dat de variabele bijdraagt aan de verklaarde variantie. De verklarende variantie, uitgedrukt in R-kwadraat (R^2), geeft de voorspellende kracht van de meervoudige regressie weer.

Wanneer twee variabelen min of meer hetzelfde meten doet zich multicollineariteit voor en dit kan de interpretatie van de relaties en de stabiliteit van de regressiecoëfficiënten en –constante verstoren (zie paragraaf 4.2.2).

3.3 Omschrijving data

Voor het onderzoek is gebruik gemaakt van de volgende (geo)databronnen die betrekking hebben op de jaren 2016 en 2017:

- Transacties van kadastrale percelen, uit (kopie van) de Basisregistratie Kadaster.
- Kadastrale percelen, uit (kopie van) de Basisregistratie Kadaster.
- Kadastrale leveringsakte, uit de openbare registers van het Kadaster.
- Bouwjaar gebouwen, uit de Basisregistratie Adressen en Gebouwen.
- Bebouwde oppervlakte gebouwen, uit de Basisregistratie Adressen en Gebouwen, uit (kopie van) de Basisregistratie Kadaster en/of van luchtfoto Google maps.
- Oppervlakte bouwkwavel, uit PDOK viewer.
- Landsdelen, uit Regioatlas.
- Landbouwgebiedsgroepen, uit Statline.
- Landbouwgebieden, uit Statline.
- Burgerlijke gemeenten, uit Statline.
- Afstanden tot plaatsen, uit ANWB-routeplanner.
- Inwoners per burgerlijke gemeente, uit Statline.
- Economische activiteit, uit intern systeem van de organisatie.

Door een combinatie te maken van deze (geo)data is het mogelijk om van de bebouwing en van de bouwkwavel de volgende eigenschappen te achterhalen:

- Type bedrijf,
- Type gebouw (woning of bedrijfsgebouwen),
- Omvang gebouwen in m^2 ,
- Bouwjaar gebouwen,
- Omvang bouwkwavel in m^2 ,
- Ligging (landsdeel, landbouwgebiedsgroep en landbouwgebied),
- Afstand tot stadscentrum (Amsterdam of dichtstbij gelegen top 10 stad).

3.4 Bewerking data

Aan de hand van een query op de Basisregistratie Kadaster (BRK) is van de jaren 2016 en 2017 een dataset gemaakt. In de dataset zijn de transacties opgenomen die als cultuuromschrijving 'bebouwd' de omschrijving woning, boerderij of overig agrarisch heeft en als cultuuromschrijving 'onbebouwd' de omschrijving *akkerbouwland*, *grasland*, *tuinbouw onder glas*, *tuinbouw in de open grond*, *boomgaarden*, *bloembollengrond* of *boomkwekerijen*. Dit resulteerde in een dataset van 5.617 transacties.

Van de transacties in de dataset is weergegeven welk object het betreft (objectomschrijving, adres, kadastrale gegevens en kadastrale oppervlakte), de totale verkochte oppervlakte, de totale koopsom, de verkoper(s) en de koper(s). Vanuit een intern systeem van de organisatie zijn aan de verkoper(s) en koper(s) in de dataset de economische activiteit(en) gekoppeld, waaruit blijkt of de (ver)koper een agrarische onderneming heeft of niet.

Aan de hand van de aanduiding van de kadastrale gemeente zijn met behulp van koppeltabellen de burgerlijke gemeenten, provincies, landsdelen, landbouwgebieden en landbouwgebiedsgroepen aan de dataset toegevoegd. Vervolgens zijn aan de hand van de burgerlijke gemeente het totaal aantal inwoners en het aantal inwoners per vierkante kilometer van die gemeente aan de dataset gekoppeld.

Daar het onderzoek ziet op de transacties van agrarisch vastgoed dat in agrarische gebruik is, is gekeken of de koper een economische activiteit verricht die ziet op een agrarische onderneming. Economische activiteiten die hiertoe worden gerekend of aanverwant zijn, zijn de bedrijfstypen: akkerbouwbedrijf, paardenfokkerij, rundvee-, kleinvee- en varkenshouderij, pluimvee-/gevogeltehouderij, groentekwekerij, fruitkwekerij, sierplantenkwekerij, bloembollenbedrijf, snijbloemenkwekerij, tuin-/bloemzadenkwekerij, hoveniersbedrijf of loonwerkbedrijf. Omdat het merendeel van de kopers geen agrarische onderneming drijft is de dataset gereduceerd tot 569 transacties. Van de overgebleven transacties blijkt dat het grootste deel (88%) van de kopers één agrarische onderneming drijft, 10% drijft twee, 1,5% drie en 0,5% vier agrarische ondernemingen. Dat betreffen dan gemengde bedrijven (gewas-/veecombinatie of veecombinaties) en/of bedrijven met een hoveniers- of loonwerkbedrijf.

Van deze transacties waren 107 transacties niet bruikbaar omdat hiervan de koopsom niet is geregistreerd bij het kadaster (koopsom=0). Dit doet zich onder andere voor bij bedrijfsoverdrachten binnen de familiesfeer en bij de vestiging van een opstalrecht ten behoeve van een nutsvoorziening. Van de overige transacties is de leveringsakte bij het kadaster geraadpleegd. Naar aanleiding van de kadastrale recherche bleken nog eens 179 transacties niet bruikbaar. Reden hiervoor zijn bijvoorbeeld dat het object is gekocht op de veiling, het object is gekocht als woonhuis en onttrokken wordt aan het agrarisch gebruik, sprake is van een bedrijfsoverdracht binnen de familie, interne reorganisatie van het bedrijf, sprake is van vestiging van recht van erfpacht en/of van het recht van opstal of het betrof bouwgrond. Uiteindelijk bleken 283 transacties bruikbaar voor dit onderzoek. De dataset heeft hiermee voldoende waarnemingen om regressieanalyses uit te kunnen voeren. De kritische betrouwbaarheidsgrens om betrouwbare regressieanalyses uit te voeren ligt rond de circa 120 waarnemingen.

Van deze 283 transacties c.q. waarnemingen zijn de volgende objectgegevens verzameld: de bebouwde oppervlakte van de bedrijfsgebouwen en de woning, de oppervlakte van de bouwkaavel en de tuin, het gewogen gemiddelde bouwjaar van de bedrijfsgebouwen⁹ en het bouwjaar van de woningen, de afstand tot Amsterdam en tot de dichtstbij gelegen top 10 stad.

Aanvankelijk was de bedoeling om aan de objectgegevens uit de BKR gegevens uit de Basisregistratie Adressen en Gebouwen (BAG) te koppelen. De interne gegevensleverancier gaf echter aan dat de gegevens uit de BAG niet betrouwbaar genoeg zijn om te gebruiken voor de bepaling van de bebouwde oppervlakten en dat daarom geen koppeling heeft plaatsgevonden. Het bleek niet mogelijk om op een andere geautomatiseerde wijze de gewenste gegevens van de aanwezige bebouwing op de agrarische bouwkaavel te verkrijgen. Daarom zijn met behulp van geografische databronnen en luchtfoto's op internet van het agrarisch vastgoed objectgegevens verzameld door deze te raadplegen en/of in te meten. Het gaat hier om de bebouwde oppervlakte van de bedrijfsgebouwen en de woning, de oppervlakte van de bouwkaavel en de tuin en het gewogen gemiddelde bouwjaar van de bedrijfsgebouwen en het bouwjaar van de woning.

⁹ Voorbeeld: $\{(\text{bouwjaar gebouw 1} \times \text{oppervlakte gebouw 1}) + (\text{bouwjaar gebouw 2} \times \text{oppervlakte gebouw 2})\}$
/ totale oppervlakte gebouwen 1+2

Wanneer bij een transactie van agrarisch vastgoed geen woning aanwezig was, zijn in de dataset de variabelen bouwjaarklassen en bebouwde oppervlakte van de woning en de oppervlakte van de tuin op 0 gesteld. Hierdoor kennen deze variabelen meer waarnemingen, hetgeen een positieve uitwerking heeft op de regressieanalyses.

De bouwjaren zijn gecategoriseerd in twee klassen, één met drie lange perioden en één met zeven korte perioden. De lange perioden sluiten op hoofdlijnen aan bij de eerder beschreven perioden van de bedrijfsvergroting/specialisatie in de negentiger jaren, de schaalvergroting/rationalisatie daarvoor en de overige periode van wederopbouw en vooroorlogse periode. De andere bouwklasse is opgedeeld in overwegend bouwperiodes van 10 jaren als mede de vooroorlogse periode. Bij het definiëren van de bouwperiodes is het jaar 1993, waarin het gebruik van asbest als bouw materiaal is verboden, het ijkpunt geweest.

Voor het onderzoek is de afstand beperkt tot de dichtstbij gelegen stad uit de top 10 van grote steden als weergegeven in onderstaande tabel. De afstand van het verkochte agrarisch vastgoed tot de dichtstbij gelegen top 10 stad is bepaald met behulp van de ANWB-routeplanner.

Tabel 3: Top 10 steden in Nederland

	Stad	Aantal inwoners
1.	Amsterdam	845.000
2.	Rotterdam	635.000
3.	Den Haag	525.000
4.	Utrecht	343.000
5.	Eindhoven	227.000
6.	Tilburg	214.000
7.	Almere	201.000
8.	Groningen	203.000
9.	Breda	182.000
10.	Nijmegen	174.000

Bron: CBS (2018e)

Tot slot is de totale koopsom als vermeld in de leveringsakte gecorrigeerd voor eventueel mee verkochte cultuurgrond, roerende zaken en productierechten, zodat de koopsom voor het agrarisch vastgoed resteert om te kunnen dienen als afhankelijke variabele in dit onderzoek.

3.4 Variabelen dataset

In de dataset is een onderscheid gemaakt tussen afhankelijke en onafhankelijke variabelen. Verschillende variabelen zijn in zowel originele vorm als in getransformeerde vorm (LN_) in de dataset opgenomen. In onderstaande tabellen volgt een omschrijving en codering van en een toelichting op de variabelen.

Tabel 4: Afhankelijke variabele

Omschrijving	Codering	Toelichting
Koopsom bedrijfs- en woongedeelte	KS_BD_HS LN_ksb_h	Betreft de koopsom van een transactie van gebouwd agrarisch vastgoed, bestaande uit de bedrijfsgebouwen en werken alsmede de agrarische bouw kavel en indien aanwezig de bedrijfswoning met tuin. Of te wel de totale koopsom als vermeld in de leveringsakte minus de koopsom voor de landbouwgrond.

Tabel 5: Onafhankelijke variabelen

Omschrijving	Codering	Toelichting
Oppervlakte bedrijfsgebouwen	GEB_OPP LN_opgeb	De totale bebouwde oppervlakte van de agrarische bedrijfsgebouwen ingemeten op basis van de panden in de Basisregistratie Adressen en Gebouwen (BAG) en/of op basis van een luchtfoto.
Bouwjaar bedrijfsgebouwen	GEB_BWJR	Het gewogen gemiddelde bouwjaar van de agrarische bedrijfsgebouwen ¹⁰ op basis van de bouwjaar en de oppervlakten van de gebouwen (panden) volgens de BAG en of luchtfoto.
Oppervlakte bouwkaavel	BKV_OPP LN_opbkv	De oppervlakte van de agrarische bouwkaavel (excl. oppervlakte grond c.q. tuin bij de bedrijfswoning) ingemeten zoals weergegeven op de website ruimtelijkeplannen.nl, of indien onbekend ingemeten op basis van een luchtfoto.
Oppervlakte woning	HS_OPP LN_ophs	De bebouwde oppervlakte van de agrarische bedrijfswoning ingemeten op basis van de panden in de BAG en/of op basis van een luchtfoto.
Bouwjaar woning	HS_BWJR	Bouwjaar van de agrarische bedrijfswoning volgens de BAG.
Oppervlakte tuin	OPP_TN LN_optn	De oppervlakte van de grond c.q. tuin bij de agrarische bedrijfswoning zoals vermeld in de leveringsakte. Indien deze oppervlakte niet in de akte staat vermeld, is deze ingemeten op basis van een luchtfoto.
Afstand tot Amsterdam	Afs_Adam LN_Adam	De kortste afstand (in km) over de weg van het agrarisch vastgoed naar Amsterdam volgens de ANWB-routeplanner.
Afstand tot top 10 stad	Afs_tp10 LN_top10	De kortste afstand (in km) over de weg van het agrarisch vastgoed naar de dichtstbij gelegen top 10 stad volgens de ANWB-routeplanner.
Inwoners gemeente	Inw_Gem LN_inw	Totaal aantal inwoners in de burgerlijke gemeente waarin het agrarisch vastgoed is gelegen.
Inwoners per km ²	Inw_Gkm2 LN_inkm2	Het gemiddeld aantal inwoners per vierkante kilometer van de burgerlijke gemeente waarin het agrarisch vastgoed is gelegen.
Aanwezigheid bedrijfswoning	WHS_j_n	Dummy aanduiding of bij de transactie wel (1) of geen (0) een woning betrokken is.
Aanwezigheid intensief bedrijf	INT_j_n	Dummy aanduiding of het agrarisch vastgoed een niet-gebonden (intensief) bedrijf (1) of een grondgebonden (extensief) bedrijf (0) betreft.
Aanwezigheid glastuinbouw	INT_glas	Dummy aanduiding of bij het agrarisch vastgoed wel (1) of geen (0) sprake is van een glastuinbouwbedrijf.
Aanwezigheid varkensbedrijf	INT_vark	Dummy aanduiding of bij het agrarisch vastgoed wel (1) of geen (0) sprake is van een varkenshouderijbedrijf.
Aanwezigheid overig intensief bedrijf	INT_over	Dummy aanduiding of bij het agrarisch vastgoed wel (1) of geen (0) sprake is van overige intensief bedrijf, zoals een pluimvee-, geiten- of kalverhouderijbedrijf e.d..
Aanwezigheid extensief veebedrijf	EXT_vee	Dummy aanduiding of bij het agrarisch vastgoed wel (1) of geen (0) sprake is van een paarden- of melkveehouderijbedrijf.
Aanwezigheid akkerbouw	EXT_akke	Dummy aanduiding of bij het agrarisch vastgoed wel (1) of geen (0) sprake is van een akkerbouwbedrijf.
Aanwezigheid overig extensief bedrijf	EXT_over	Dummy aanduiding of bij het agrarisch vastgoed wel (1) of geen (0) sprake is van een overig extensief bedrijf, zoals een vollegrondstuinbouwbedrijf, fruit- en boomkwekerij e.d..
Bouwjaarclassen (lange tijdvakken)	G_BJ_KL3 H_BJ_KL3	Categorische aanduiding in drie bouwjaarclassen voor zowel de bedrijfsgebouwen (G_) als de bedrijfswoning (H_) te weten: • 1979 en ouder • 1980 tot en met 1993 • 1994 en jonger

¹⁰ Voorbeeld: $\{(\text{bouwjaar gebouw 1} \times \text{oppervlakte gebouw 1}) + (\text{bouwjaar gebouw 2} \times \text{oppervlakte gebouw 2})\}$
/ totale oppervlakte gebouwen 1+2

Bouwjaarklassen (korte tijdvakken)	G_BJ_KL7 H_BJ_KL7	Categorische aanduiding in zeven bouwjaarklassen voor zowel de bedrijfsgebouwen (G_) als de bedrijfswoning (H_) te weten: • 1944 en ouder • 1945 tot en met 1963 • 1964 tot en met 1973 • 1974 tot en met 1983 • 1984 tot en met 1993 • 1993 tot en met 2003 • 2004 en jonger
Geografische ligging	DEEL_NED	Categorische aanduiding in vier landsdelen voor de geografische ligging van het agrarisch vastgoed. Het betreffen de landsdelen: - Noord-Nederland (provincies Groningen, Friesland en Drenthe) - Oost-Nederland (provincies Overijssel, Gelderland en Flevoland) - West-Nederland (provincies Noord-Holland, Zuid-Holland en Utrecht) - Zuid-Nederland (provincies Zeeland, Noord-Brabant en Limburg)

3.5 Verwachte verbanden

Naar aanleiding van het uitgevoerde literatuuronderzoek (hoofdstuk 2) en de verzamelde dataset wordt in deze paragraaf de mogelijk te verwachten verbanden beschreven alvorens in hoofdstuk 4 verder te gaan met de data- en regressieanalyse.

Volgens de theorieën van Ricardo en Von Thünen heeft de afstand tot de stad invloed op de waarde van landbouwgrond. Voor dit onderzoek zou dit kunnen betekenen dat voor agrarisch vastgoed op kortere afstand van een stad meer is betaald dan wanneer het verder van de stad is gelegen.

In het verlengde hiervan zou hetzelfde kunnen gelden voor de bevolkingsdichtheid. Naar mate meer mensen per vierkante kilometer in een burgerlijke gemeente wonen zal de druk op het landelijk gebied om uit te kunnen breiden en/of te kunnen wonen toenemen dat leidt tot een prijsopdrijvend c.q. speculatief effect. Uit niet bruikbare transacties is gebleken dat kleine marginale agrarische bedrijfjes door particulieren worden gekocht om te gebruiken als woning.

Voor de geografische ligging is eenzelfde vergelijk te maken. In de landsdelen waar veel steden zijn en dus mensen wonen, zoals West-Nederland, zou daarom meer voor het agrarisch vastgoed betaald kunnen zijn. Wellicht dat ook de ligging in bepaalde landbouwgebieden van invloed is op de koopsom van het agrarisch vastgoed.

Naar analogie van de invloed van bodemproductiviteit op de prijs van landbouwgrond zou het bedrijfstype van invloed kunnen zijn op de koopsom van het agrarisch vastgoed. De (mogelijke) bedrijfsresultaten van een bedrijf(stype) bepaalt immers mede wat voor het agrarisch vastgoed kan worden betaald.

De bedrijfsresultaten hangen af van een efficiënte bedrijfsvoering. Refererend aan de bedrijfsvergroting/specialisatie is dit het best te realiseren in grote moderne bedrijfsgebouwen. Daarom is de verwachting dat door de koper meer is betaald voor agrarisch vastgoed met een grote bebouwde oppervlakte aan bedrijfsgebouwen, en van een recente bouwjaarklasse, op een grotere bouwkaavel, dan voor een kleinere bebouwde oppervlakte en/of van een (veel) oudere bouwjaarklasse op een kleinere bouwkaavel.

Op basis van ervaringen bij geautomatiseerde waardbepaling van woningen is de verwachting dat ook bij de (bedrijfs)woning met tuin de koopsom mede wordt bepaald door de bebouwde oppervlakte van de woning, de bouwjaarklasse en de oppervlakte van de tuin. Daarbij zouden ook nog de afstand tot een stad en/of de geografische ligging van invloed kunnen zijn.

4 Analyse

Voordat begonnen kan worden met het onderzoek naar een regressiemodel voor de waardering van agrarisch vastgoed zal eerst een beschrijving worden gegeven van de belangrijkste kenmerken van de dataset. Daarna zal worden gecontroleerd of de dataset voldoet aan de voorwaarden voor een meervoudige regressie. Voor de analyses en de regressie is gebruik gemaakt van het programma STATA.

4.1 Beschrijvende Statistiek

Het onderzoek is uiteindelijk gebaseerd op 211 waarnemingen zoals weergegeven in onderstaande tabel. Bij alle transacties is sprake van bedrijfsgebouwen met bijbehorende bouwkaavel, waarvan in 100 gevallen in de leveringsakte een uitsplitsing van de koopsom in bedrijfsgebouwen en bouwkaavel is opgenomen. Bij 164 transacties is een woning met ondergrond/tuin betrokken, waarvan in 32 gevallen een uitsplitsing van de koopsom in woning en ondergrond/tuin in de leveringsakte is opgenomen. En bij 150 transacties is ook cultuurgrond betrokken.

Met de uitsplitsingen is in het onderzoek niets gedaan, omdat die het aantal bruikbare waarnemingen voor de regressieanalyses vermindert waardoor de verklarende kracht van het model afneemt.

Tabel 6: Samenstelling transacties

Samenstelling	Aantal	Percentage
Bedrijfsgedeelte + woongedeelte + cultuurgrond	127	60%
Bedrijfsgedeelte + woongedeelte	37	18%
Bedrijfsgedeelte	24	11%
Bedrijfsgedeelte + cultuurgrond	23	11%
TOTAAL	211	100%

Bron: eigen bewerking

Naast de samenstelling van de transacties is het interessant om te weten welke bedrijfstypen het betreffen en waar deze in Nederland zijn gelegen. Uit tabel 7 blijkt dat de waarnemingen redelijk over de intensieve en extensieve landbouwbedrijven zijn verdeeld. Wat opvalt is dat na alle correcties voor uitschieters er weinig glastuinbouwbedrijven onderdeel uitmaken van de populatie en dat veehouderijbedrijven juist een groot aandeel (75%) vormt. Verder blijkt dat ruim een derde van de waarnemingen ligt in de provincies Noord-Brabant en Limburg en dat in de provincies Friesland en Drenthe, maar ook in de provincies Overijssel en Gelderland een vijfde van de waarnemingen ligt. Zoals hierboven reeds is aangegeven is bij 164 waarnemingen een woning aanwezig.

Tabel 7: Onderverdeling waarnemingen

Bedijfstype		Geografische ligging			
INT_glas	9	Groningen	8	Utrecht	7
INT_vark	34	Friesland	26	Noord-Holland	11
INT_over	55	Drenthe	17	Zuid-Holland	13
<i>INTENSIEF</i>	<u>98</u>	<i>Nrd_NED</i>	<u>51</u>	<i>West_NED</i>	<u>31</u>
EXT_vee	70	Overijssel	13	Zeeland	5
EXT_akke	19	Flevoland	6	Noord-Brabant	55
EXT_over	24	Gelderland	28	Limburg	22
<i>EXTENSIEF</i>	<u>113</u>	<i>Oost_NED</i>	<u>47</u>	<i>Zuid_NED</i>	<u>82</u>
TOTAAL	211	WHS_j_n (ja)	164	TOTAAL_NED	211

Bron: eigen bewerking

Bij de veredeling van de gegevens zijn kenmerken van het agrarisch vastgoed aan de dataset toegevoegd, zoals bouwjaar, omvang van de bebouwing, bouwkaavel en tuin. In onderstaande tabel is een statistische samenvatting weergegeven van de variabelen die betrekking hebben op het agrarisch vastgoed.

Tabel 8: Beschrijvende statistiek objectgegevens

Variabele	Observaties	Gemiddelde	Std. Deviatie	Minimum	Maximum
OPP_TOT	211	63715.38	119684.4	555	685765
CULT_OPP	150	79070.67	133594.4	455	671115
GEB_BWJR	211	1980.697	18.50078	1879	2013
GEB_OPP	211	1768.73	1479.43	127	8480
BKV_OPP	211	6531.758	3999.214	280	17800
HS_BWJR	164	1951.915	36.90701	1819	2009
HS_OPP	164	127.9024	48.13098	50	288
TN_OPP	164	1250.884	682.6037	200	3360
Afs_Adam	211	122.3839	43.93165	10	219
Afs_tp10	211	46.28436	24.46036	9	115
Inw_Gem	221	42641.12	62771.25	6657	844947
Inw_Gkm2	211	357.6872	457.1074	69	5111

Bron: STATA

Uit verschillende variabelen met betrekking tot het agrarisch vastgoed blijkt uit tabel 8 dat in de agrarische sector sprake is van heterogeniteit. Dit is onder andere het gevolg van de hoeveelheid cultuurgrond dat aanwezig is bij het object. Zo kent de totale oppervlakte (TOT_OPP) een erg brede spreiding en varieert van 555 m² bij een transactie van enkel een stal met ondergrond tot circa 69 hectare bij een transactie van een geheel melkveebedrijf. Ook bij de oppervlakte van de bouwkaavel (BKV_OPP) en de oppervlakte ondergrond van de woning en de tuin (TN_OPP) is een brede spreiding te zien. De grote uitschieters bij de bouwkaavel betroffen transacties van grote glastuinbouwbedrijven en bij de ondergrond van de woning en de tuin betrof het onder andere een transactie van een glastuinbouwcomplex met drie woningen die waren gesommeerd in de dataset. Deze uitschieters zijn al verwijderd uit de dataset waarvan de beschrijvende statistiek is weergegeven.

Naast de spreiding in de grondoppervlaktevariabelen kennen de agrarische bedrijfsgebouwen en de woningen eveneens een spreiding ten aanzien van de bebouwde oppervlakte en het (gemiddelde) bouwjaar. De bebouwde oppervlakte van de bedrijfsgebouwen varieerde van een stalletje van 127 m² bij een transactie van een voormalige agrarische bedrijf met woning tot kassen van 63.500 m² bij een transactie van een glastuinbouwbedrijf. Bij de woningen varieerde de bebouwde oppervlakte van 50 m² voor een woning tot de gesommeerde oppervlakte van 373 m² voor drie woningen bij de transactie van een glastuinbouwcomplex. Deze hoge uitschieters zijn al verwijderd uit de dataset waarvan de beschrijvende statistiek is weergegeven.

De bouwjaren laten bij de bedrijfsgebouwen en de woningen een vergelijkbare spreiding zien. De woning heeft vaak het oudste bouwjaar, omdat in het verleden een boerderij naast de agrarische functie ook een woonfunctie voor het boerengezin had. Tegenwoordig zijn er ook agrarische bedrijven zonder woning of koopt een agrarisch ondernemer alleen de bedrijfsgebouwen en de cultuurgrond en blijft de woning achter bij een derde. Het gemiddelde bouwjaar van de agrarische bedrijfsgebouwen (GEB_BWJR) is ongeveer 30 jaar hoger dan het gemiddelde bouwjaar van de woningen (HS_BWJR), omdat in de loop der jaren op de bouwkaavel agrarische bedrijfsgebouwen zijn bijgebouwd of oude gebouwen zijn vervangen door nieuwe modernere gebouwen.

Ten aanzien van de afstandsvariabelen kan worden geconstateerd dat het agrarisch vastgoed opgenomen in de dataset gemiddeld dicht bij een top 10 stad (Afs_tp10) is gelegen dan bij de stad Amsterdam (Afs_Adam). Dat is logisch gezien de geografische spreiding van deze steden.

Uit het aantal inwoners van de burgerlijke gemeente (Inw_Gem) en het gemiddelde aantal inwoners per vierkante kilometer (Inw_Gkm2) is af te leiden dat het agrarisch vastgoed gemiddeld is gelegen in een burgerlijke gemeente met circa 42.500 inwoners (bijvoorbeeld gemeente De Ronde Venen) of in een burgerlijke gemeente met circa 360 inwoners per vierkante kilometer (bijvoorbeeld gemeente Elburg).

Tabel 9 geeft een indruk van het gemiddelde en de bandbreedte van de koopsommen op basis van de transacties in de dataset. Weergegeven zijn de totale koopsom (KS_TOT) volgens de leveringsakte en de eventuele onderverdeling in de koopsom voor de cultuurgrond (KS_CULT), de bedrijfsgebouwen met de bouwkafe (KS_BEDR) en/of de woning en de tuin (KS_HS_TN), als mede de optelling van de koopsommen van het bedrijfs- en woongedeelte (KS_BD_HS).

Tabel 9: Beschrijvende statistiek transactiegegevens

Variabele	Observaties	Gemiddelde	Std. Deviatie	Minimum	Maximum
KS_TOT	211	719667.8	771355.3	28732	4800000
KS_CULTG	150	480046.5	781731.3	1820	4045000
KS_BD_HS	211	379340.3	203168.4	28732	1000000
KS_BEDR	211	199808.3	147258.4	15000	750000
KS_HS_TN	164	229776.6	92644.11	52500	500000

Bron: STATA

Bij dit onderzoek wordt alleen gebruik gemaakt van de totale koopsom voor het agrarisch vastgoed als vermeld in de leveringsakte minus de koopsom voor de landbouw- c.q. cultuurgrond. Dit geeft immers de werkelijke prijs van het agrarisch vastgoed weer die is overeengekomen tussen de verkoper en de koper. Met een eventuele verdere uitsplitsing van de koopsom in de leveringsakte in bedrijfsgebouwen, bouwkafe, woning en/of tuin is geen rekening gehouden. Dit omdat de uitsplitsing van de koopsom in de leveringsakte veelal een fiscale achtergrond heeft (Van Gellicum, 2015). Zo kan bij een verkooptransactie in combinatie met de staking van de onderneming aan de woning met tuin, als dat privé vermogen is, een hogere waarde zijn toegekend. De waarde van de bedrijfsgebouwen en de bouwkafe is dan lager, waardoor hierover door de verkoper minder belasting in het kader van de staking wordt betaald. Bij een uitsplitsing in de leveringsakte kan de toegekende prijs aan de verschillende onderdelen onjuist zijn, terwijl de totale koopprijs tussen de onafhankelijke verkoper en koper vast staat en als juist wordt beschouwd.

4.2 Voorwaarden meervoudige regressie

Alvorens kan worden begonnen met het zoeken en opstellen van een meervoudig regressiemodel dient eerst te worden gecontroleerd of de dataset met 283 waarnemingen voldoet aan de voorwaarden voor meervoudige regressie. Dan gaat het om het lokaliseren van outliers, het uitsluiten van multicollineariteit, een normale verdeling van de storing en lineariteit tussen de afhankelijke en onafhankelijke variabelen. Om een beeld te krijgen van de hierna beschreven bewerkingen op de dataset om deze aan de voorwaarden van meervoudige regressie te laten voldoen, wordt in tabel 10 een onderverdeling van de waarnemingen weergegeven naar bedrijfstype, geografische ligging en aanwezigheid van een woning.

Tabel 10: Onderverdeling waarnemingen bij aanvang

Bedijfstype		Geografische ligging			
INT_glas	35	Groningen	10	Utrecht	9
INT_vark	52	Friesland	32	Noord-Holland	16
INT_over	66	Drenthe	21	Zuid-Holland	29
INTENSIEF	153	Nrd_NED	63	West_NED	54
EXT_veen	83	Overijssel	16	Zeeland	5
EXT_akke	22	Flevoland	8	Noord-Brabant	70
EXT_over	25	Gelderland	39	Limburg	28
EXTENSIEF	130	Oost_NED	63	Zuid_NED	103
TOTAAL	283	WHS_j_n (ja)	214	TOTAAL_NED	283

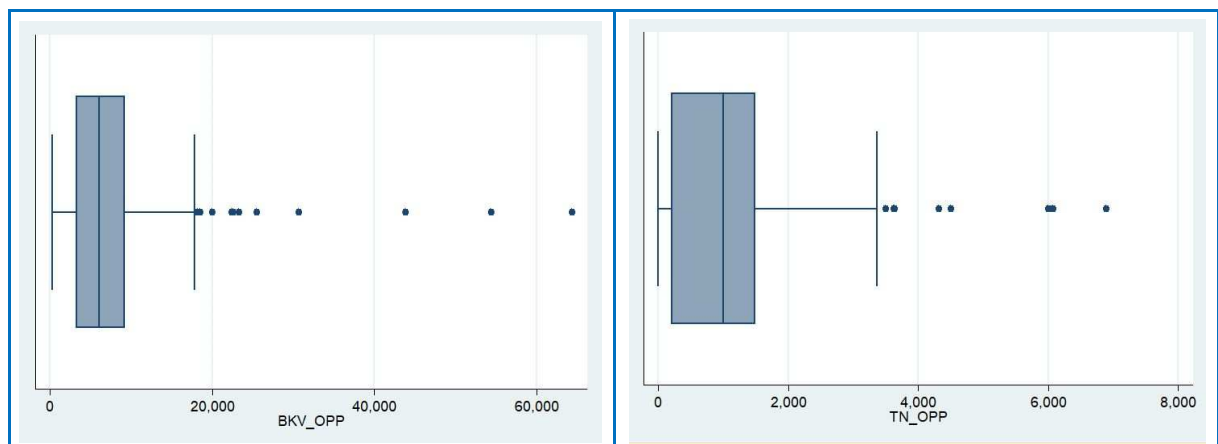
Bron: eigen bewerking

4.2.1 Outliers

Outliers zijn waarnemingen die erg ver van het regressiemodel liggen en dus grote residuen hebben. Omdat outliers van bovenmatige invloed zijn op de uiteindelijke regressie is het van belang om de extreme uitschieters te bekijken en in de meeste gevallen te verwijderen uit de dataset voordat de regressievergelijking wordt opgesteld.

Het onderzoek naar uitschieters is beperkt tot het vervaardigen van boxplots. Gekozen is om dit alleen te doen voor de variabelen in de dataset die zien op de kenmerken van het vastgoedobject, te weten GEB_OPP, BKV_OPP, HS_OPP en TN_OPP (zie figuur 4 en 5). Daarnaast is ook gekeken naar uitschieters bij de koopsom voor het bedrijfs- en woongedeelte (KS_BD_HS) en voor alleen de woning met tuin (KS_HS_TN) (zie figuur 6).

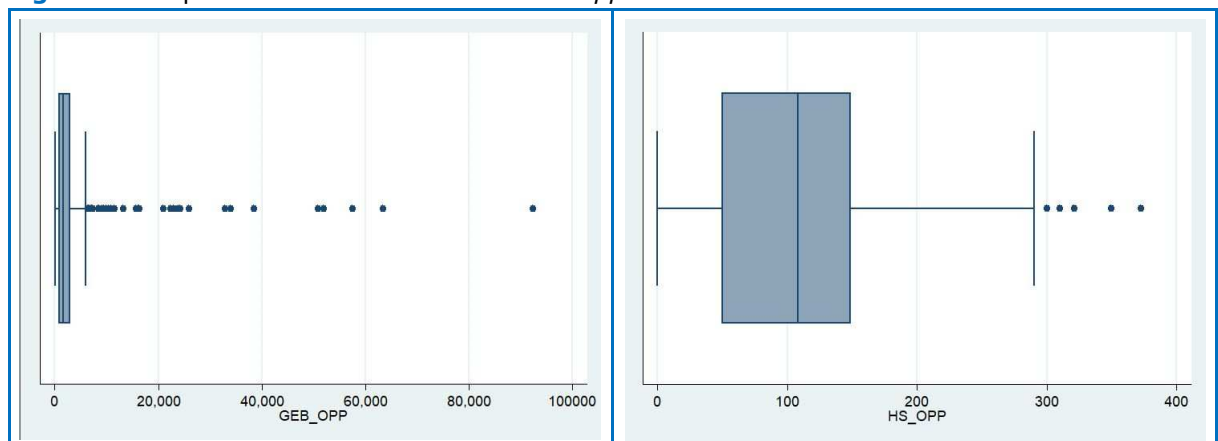
Figuur 4: Boxplots voor de variabelen *Grondoppervlakte*



Bron: STATA

Bij de bouwkeizers (BVK_OPP) vallen de uitschieters met een oppervlakte boven 18.000 m² op. Deze uitschieters blijken bouwkeizers van grote intensieve glastuinbouwbedrijven en varkenshouderijen te zijn. De uitschieters bij ondergrond woning en tuin (TN_OPP) met een oppervlakte boven 3.500 m² blijken met name onderdeel uit te maken van glastuinbouwbedrijven en veehouderijen.

Figuur 5: Boxplots voor de variabelen *Gebouwoppervlakte*

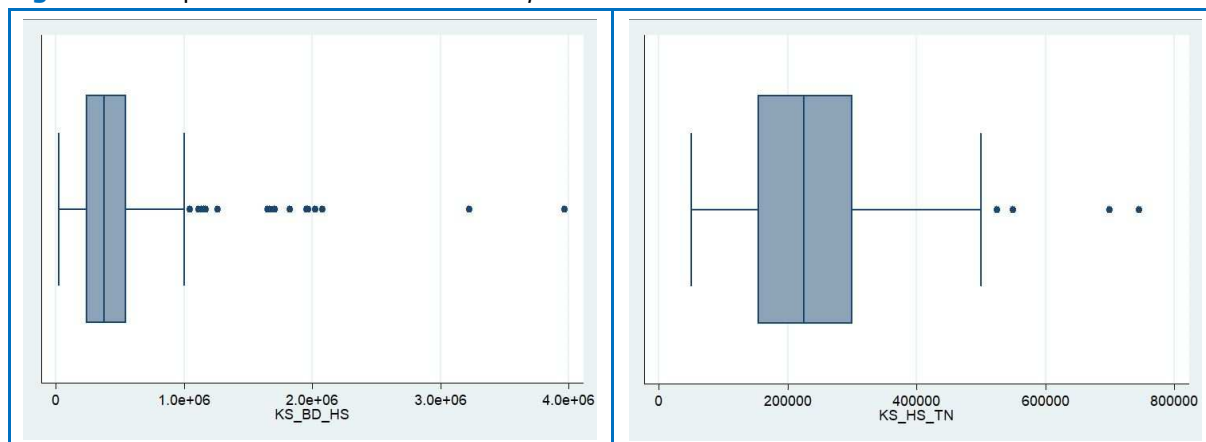


Bron: STATA

De uitschieters bij de bedrijfsgebouwen (GEB_OPP) zijn bijna allemaal toe te rekenen aan de kassen met een bebouwde oppervlakte van 8.500 m² en meer van de glastuinbouwbedrijven. Bij de woningen (HS_OPP) vallen de uitschieters van 300 m² en meer op. Dit blijken de eerder aangehaalde transactie van het glastuinbouwcomplex met de gesommeerde bebouwde oppervlakte van de drie woningen en de grote woningen bij veehouderijbedrijven te betreffen.

Voor uitschieters in bouwjaar bij de bedrijfsgebouwen en de woning zijn geen boxplots vervaardigd, omdat het bouwjaar als een categorie in de regressie zal worden meegenomen. Daardoor is het probleem van uitschieters uitgesloten.

Figuur 6: Boxplots voor de variabelen *Koopsom*



Bron: STATA

Naast de uitschieters op basis van objectkenmerken is ook gekeken naar uitschieters in koopsommen om tot een evenwichtige database te komen. Bij de koopsom voor het bedrijfs- en woongedeelte (KS_BD_HS) vallen de uitschieters boven € 1.000.000 op. Deze uitschieters blijken wederom grote intensieve glastuinbouwbedrijven en veehouderijen te betreffen. Hetzelfde geldt voor de uitschieters bij de koopsom voor alleen de woning met tuin (KS_HS_TN).

Uit de bovenstaande boxplots wordt geconstateerd dat meeste uitschieters zich voordoen bij de variabelen die zien op de gebouwoppervlakte. Daarom zijn op basis van de bebouwde oppervlakte van de bedrijfsgebouwen (GEB_OPP) 28 uitschieters en op basis van de bebouwde oppervlakte van de woningen (HS_OPP) 7 uitschieters uit de dataset verwijderd. Daarnaast zijn bij de oppervlakte van de bouwkael (BKV_OPP) nog eens 11 uitschieters en bij de oppervlakte ondergrond woning met tuin (TN_OPP) 7 uitschieters verwijderd. Tot slot zijn 16 uitschieters ten aanzien van de koopsom voor het bedrijfs- en woongedeelte (KS_BD_HS) en 5 uitschieters ten aanzien van de koopsom voor de woning met tuin (KS_HS_TN) uit de dataset verwijderd.

Op basis van de boxplots voor grondoppervlakte, gebouwoppervlakte en koopsom is sprake van 76 uitschieters. Bij verscheidene waarnemingen kwamen echter meerdere uitschieters voor, waardoor effectief 52 waarnemingen uit de database zijn verwijderd. In tabel 11 wordt een onderverdeling van de verwijderde uitschieters gemaakt naar bedrijfstype, geografische ligging en aanwezigheid van een woning.

Tabel 11: Onderverdeling uitschieters boxplots

Bedijfstype		Geografische ligging			
INT_glas	23	Groningen	0	Utrecht	1
INT_vark	14	Friesland	2	Noord-Holland	3
INT_over	6	Drenthe	4	Zuid-Holland	14
INTENSIEF	43	Nrd_NED	6	West_NED	18
EXT_vee	8	Overijssel	2	Zeeland	0
EXT_akke	1	Flevoland	1	Noord-Brabant	15
EXT_over	0	Gelderland	4	Limburg	6
EXTENSIEF	9	Oost_NED	7	Zuid_NED	21
TOTAAL	52	WHS_j_n (ja)	41	TOTAAL_NED	52

Bron: eigen bewerking

Na deze correctie omvat de dataset 231 waarnemingen en is nog voldoende van omvang om betrouwbare uitspraken te doen ten aanzien van de significantie van de variabelen en het regressiemodel.

4.2.2 Multicollineariteit

Multicollineariteit is een statistisch fenomeen waarin twee of meer variabelen in een regressiemodel sterk zijn gecorreleerd. Deze variabelen meten min of meer hetzelfde, waardoor de betrouwbaarheid van de regressie afneemt (Van Arnhem et al., 2013). Met behulp van de Pearson Correlatie Coëfficiënt en de Spearman Rangorde Correlatie is voor de dataset met 231 waarnemingen de onderlinge correlatie van alle variabelen inzichtelijk gemaakt. Dit is gedaan om multicollineariteit tussen zowel de originele variabelen als de getransformeerde variabelen zo veel mogelijk uit te sluiten (zie bijlage 1). In de eerste kolom van de tabel kan worden afgelezen welke onafhankelijke variabelen een hoge correlatie hebben met de (eerst genoemde) afhankelijke variabele (zie ook tabel 14 in paragraaf 4.3). Vervolgens is in de opvolgende kolommen gekeken of voor die onafhankelijke variabele sprake is van multicollineariteit ten opzichte van een andere onafhankelijke variabele. Van multicollineariteit tussen twee onafhankelijke variabelen is sprake bij een correlatie van 0,7 of hoger, dan wel -0,7 of lager (University of Amsterdam, 2012). In die gevallen dient slechts één van de twee onafhankelijke variabelen in de regressie te worden opgenomen. In onderstaande tabel is weergegeven tussen welke onafhankelijke variabelen sprake is van multicollineariteit op basis van de dataset.

Tabel 12: Multicollineariteit

ORIGINELE VARIABLEN		GETRANSFORMEERDE VARIABLEN	
Bij afhankelijke variabele	KS_BD_HS	Bij afhankelijke variabele	LN_ksb_h
<i>heeft onafhankelijke variabele</i>	multicollineariteit i.c.m.	<i>heeft onafhankelijke variabele</i>	multicollineariteit i.c.m.
GEB_BWJR	G_BJ_KL3 en G_BJ_KL7	GEB_BWJR	G_BJ_KL3 en G_BJ_KL7
G_BJ_KL3	GEB_BWJR en G_BJ_KL7	G_BJ_KL3	GEB_BWJR en G_BJ_KL7
G_BJ_KL7	GEB_BWJR en G_BJ_KL3	G_BJ_KL7	GEB_BWJR en G_BJ_KL3
HS_BWJR	H_BJ_KL3 en H_BJ_KL7	HS_BWJR	H_BJ_KL3 en H_BJ_KL7
H_BJ_KL3	HS_BWJR en H_BJ_KL7	H_BJ_KL3	HS_BWJR en H_BJ_KL7
H_BJ_KL7	HS_BWJR en H_BJ_KL3	H_BJ_KL7	HS_BWJR en H_BJ_KL3
DEEL_NED	Nrd_NED en Zuid_NED	DEEL_NED	Nrd_NED en Zuid_NED
Nrd_NED	DEEL_NED	Nrd_NED	DEEL_NED
Zuid_NED	DEEL_NED	Zuid_NED	DEEL_NED
Inw_Gem	Inw_Gkm2		
Inw_Gkm2	Inw_Gem		

Bron: eigen bewerking

De meeste gevallen van multicollineariteit kunnen worden verklaard uit het feit dat de waarnemingen van de onafhankelijke variabelen, zoals bouwjaar en landsdelen, zijn verwerkt tot categorische groepen. Het is dan logisch om bij het opstellen van een regressie uit een groep gelijksoortige variabelen slechts één variabele in het model op te nemen. Ook voor de variabelen die het bevolkingsaantal of de bevolkingsdichtheid van een burgerlijke gemeente weergeven is sprake van multicollineariteit.

Bij de originele variabelen lijkt sprake van enige multicollineariteit tussen de bebouwde oppervlakte van de woning (HS_OPP) en de oppervlakte van de tuin. Een aannemelijke verklaring daarvoor is dat bij een grotere woning vaak een grotere tuin aanwezig is.

Bij de getransformeerde variabelen lijkt daarentegen juist enige multicollineariteit tussen de oppervlaktevariabelen van de bedrijfsgebouwen (LN_opgeb) en de bouwkael (LN_opbkv) aan de

orde. Een aannemelijke verklaring daarvoor is dat voor bedrijfsgebouwen met een groter bebouwd oppervlak ook meer grond nodig is.

4.2.3 Normale verdeling van de storingsterm

Voor een meervoudige regressie is het van belang dat de storingsterm¹¹ normaal is verdeeld. Dit is gecontroleerd door voor de regressie met de grootste verklarende kracht voor alle waarnemingen een residu c.q. storing berekenen en deze af te zetten tegen de normale verdeling. Uit de *Normaal Q-Q plot* en de *Kernal density estimate grafiek* is gebleken dat de residuen niet volledig de normale verdeling volgen. Met name aan de uiteinden van de verdeling zijn enkele afwijkingen geconstateerd. Zowel de positieve als de negatieve uitschieters aan de uiteinden van de verdeling zijn alsnog verwijderd uit de dataset. Het betreffen in totaal nog eens 20 uitschieters. In tabel 13 is weer een onderverdeling gemaakt naar bedrijfstype, geografische ligging en aanwezigheid van een woning.

Tabel 13: Onderverdeling uitschieters storingsterm

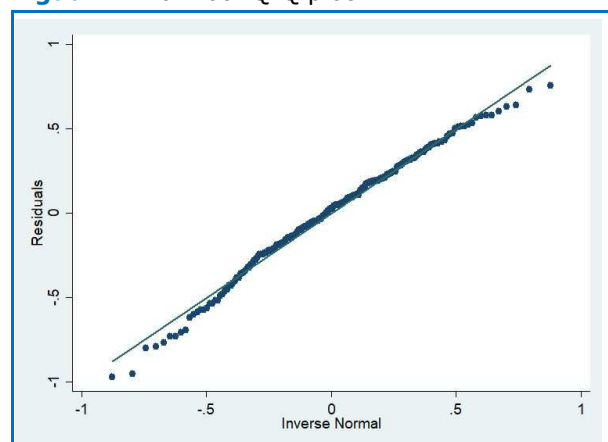
Bedijfstype		Geografische ligging			
INT_glas	3	Groningen	2	Utrecht	1
INT_vark	4	Friesland	4	Noord-Holland	2
INT_over	5	Drenthe	0	Zuid-Holland	2
INTENSIEF	12	Nrd_NED	6	West_NED	5
EXT_vee	5	Overijssel	1	Zeeland	0
EXT_akke	2	Flevoland	1	Noord-Brabant	0
EXT_over	1	Gelderland	7	Limburg	0
EXTENSIEF	8	Oost_NED	9	Zuid_NED	0
TOTAAL	20	WHS_j_n (ja)	9	TOTAAL_NED	20

Bron: eigen bewerking

Na deze correctie omvat de dataset 211 waarnemingen en dat is nog steeds voldoende om betrouwbare uitspraken te doen ten aanzien van de significantie van de variabelen en het regressiemodel.

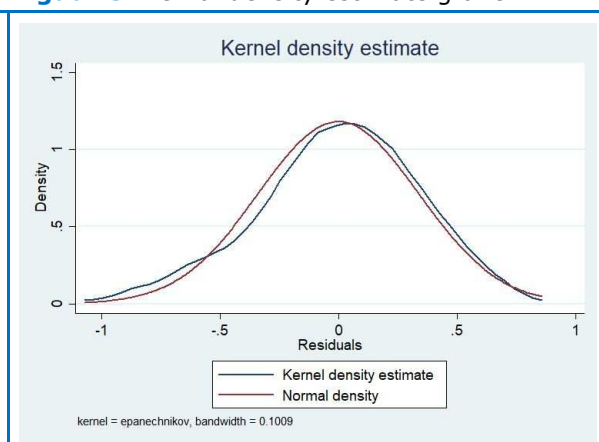
Na deze correctie is nogmaals de storingsterm voor de regressie met de grootste verklarende kracht berekend en afgezet tegen de normale verdeling. Uit de figuren 7 en 8 wordt geconstateerd dat de storingsterm nog steeds niet volledig de normale verdeling volgt. Maar omdat in figuur 7 de residuen voor het overgrote deel in redelijke mate de normaal verdeling volgen, wordt voor dit onderzoek uitgegaan van een heteroscedastisch verdeelde storingsterm. Standaardfouten in het regressiemodel zijn hiervoor gecorrigeerd door in STATA gebruik te maken van de robuuste regressie.

Figuur 7: Normaal Q-Q plot



Bron: STATA

Figuur 8: Kernal density estimate grafiek



Bron: STATA

¹¹ Met storingsterm wordt de verdeling van de residuen bedoeld.

4.2.4 Lineariteit

Voor een goede (meervoudige) regressieanalyse is het van belang dat het model lineair is. In de praktijk blijkt dat bij vastgoed verschillende redenen zijn om aan te nemen dat de relatie tussen de afhankelijke en onafhankelijke variabelen wel eens niet lineair zou kunnen zijn (Marquard, z.j.) (Francke, 2018).

Op basis van de wet van de afnemende meeropbrengst is sprake van proportionaliteit ten aanzien van de omvang van de bebouwde oppervlakte van de bedrijfsgebouwen en de woning, maar ook van omvang van de oppervlakte van de bouwkaal en de tuin. Op een gegeven moment leidt een extra vierkante meter gebouw en/of grond tot minder toegevoegde waarde dan van de eerste vierkante meters. Daarnaast kan de geografische spreiding van de waarnemingen over Nederland leiden tot multiplicatieve effecten, zoals de invloed van de locatie. Een agrarisch vastgoed in landsdeel A is x euro duurder dan in landsdeel B.

Om aan het gebrek van lineariteit tussen de afhankelijke en onafhankelijke variabele op basis van bovengenoemde redenen tegemoet te komen en dit aldus op te heffen, is bij dit onderzoek gebruik gemaakt van de logaritmische transformatie (LN-transformatie). Het gebruik van een logaritmische transformatie leidt bovendien tot een betere normaal verdeling van de waarnemingen en de storingsterm en tot minimalisatie van absolute en relatieve verschillen van de residuen. Bij het minimaliseren van de som van de kleinste kwadraten telt een residu van € 10.000 op een koopsom van € 100.000 even hard mee als een residu van € 10.000 op een koopsom van €100.000. In het eerste geval is de afwijking 10% en in het tweede geval slechts 1%.

Bij dit onderzoek is ook de logaritmische transformatie gehanteerd, omdat de coëfficiënten dan elasticiteiten zijn die veranderingen in percentages weergeven, de verdeling van de variabelen een betere normaal verdeling krijgt en de variabelen met een exponentieel karakter door transformatie lineair worden.

4.3 Regressieanalyses

Op basis van de uitgevoerde controles, zoals beschreven in de voorgaande sub paragrafen, wordt ervan uitgegaan dat voor dit onderzoek aan de voorwaarden voor een meervoudige regressie is voldaan. Aan de hand van de dataset kan nu een meervoudige regressievergelijking worden gemodelleerd.

Eerst is aan de hand van de Pearson Correlatie Coëfficiënt en de Spearman Rangorde Correlatie bepaald welke originele en getransformeerde variabelen de grootste correlatie hebben met de originele dan wel de getransformeerde afhankelijke variabele, KS_BD_HS respectievelijk LN_ksb_h. Een grote correlatie draagt immers bij aan de verklarende kracht van het regressiemodel. De correlatie van deze variabelen met de afhankelijke variabele wordt in onderstaande tabel weergegeven.

Tabel 14: Correlatie

Variabele	Correlatie met KS_BD_HS	Correlatie met LN_ksb_h
GEB_BWJR	0.2427	0.1663
G_BJ_KL3	0.3692	0.3692
G_BJ_KL7	0.3943	0.3943
HS_BWJR	0.2446	0.2488
H_BJ_KL3	0.1668	0.1668
H_BJ_KL7	0.2612	0.2612
WHS_j_n	.	.
INT_j_n	0.1089	0.1089
DEEL_NED	0.1957	0.1957
OPP_TOT / LN_optot	0.2557	0.3418
OPP_GEB / LN_opgeb	0.4063	0.4384
OPP_BKV / LN_opbkv	0.4799	0.4669
OPP_HS / LN_ophs	0.5458	0.2926
OPP_TN / LN_optn	0.3964	0.1162
Afs_Adam / LN_Adam	-0.2271	-0.1861
Afs_tp10 / LN_top10	-0.1281	-0.0374
Inw_Gem / LN_inw	0.0073	0.0118
Inw_Gkm2 / LN_Gkm2	-0.0085	-0.0107

Bron: STATA

Uit de tabel kan worden geconstateerd dat de omvang van de bedrijfsgebouwen (OPP_GEB/LN_opgeb) en de bouwkael (OPP_BKV/LN_opbkv) de grootste invloed heeft op de waarde van het agrarisch vastgoed. Daarna hebben de bouwjaarklassen van de bedrijfsgebouwen (G_BJ_KL3 en G_BJ_KL7) ongeveer evenveel invloed.

De afstand tot de grote stad (Afs_Adam/LN_Adam en Afs_tp10/LN_top10) heeft een negatieve correlatie. Dit betekent dat de waarde van het agrarisch vastgoed lager wordt naar mate het verder van de stad is gelegen. In zekere zin is dat vergelijkbaar met de speculatie-effecten die leiden tot een hogere prijs voor landbouwgrond rond stedelijke gebieden (Cotteleer et al., 2008). De geografische ligging van het agrarisch vastgoed in de verschillende landsdelen (DEEL_NED) heeft eveneens een negatieve correlatie. Wanneer wordt ingezoomd op de specifieke landsdelen dan blijkt dit met name te worden veroorzaakt door de landsdelen Noord-Nederland en Zuid-Nederland.

Met betrekking tot de variabelen die zien op de bevolking van de burgerlijke gemeente waarin het agrarisch vastgoed is gelegen blijkt voor het inwoneraantal (Inw_Gem resp. LN_inw) een positieve correlatie en voor de bevolkingsdichtheid (Inw_Gkm2 resp. LN_Gkm2) een negatieve.

Voor de aanwezigheid van een woning (met tuin) kon geen correlatie coëfficiënt worden bepaald, omdat tussen de aanwezigheid van een woning en de afhankelijke variabele, de (getransformeerde) koopsom voor het bedrijfs- en woongedeelte, geen verband aanwezig is. De variabelen oppervlakte huis (OPP_HS) en oppervlakte tuin (OPP_TN) hebben wel een sterke correlatie met de koopsom agrarisch vastgoed (KS_BD_HS). Wat betreft de bouwjaaren van de woning is de correlatie het sterkst bij de bouwjaarklassen voor woningen met de kortere perioden (H_GB_KL7).

De overige variabelen, zoals landbouwgebied, hebben geen significante invloed op de waarde. Dit in tegenstelling tot de door Kroes (2013) gevonden significante invloed van het landbouwgebied op de prijs van landbouwgrond. Een verklaring hiervoor kan zijn dat de dataset voor dit onderzoek waarschijnlijk beduidend kleiner is dan de database waarmee onderzoek is gedaan naar de prijs van landbouwgrond waardoor per landbouwgebied meer waarnemingen beschikbaar zijn.

Op basis van de 231 waarnemingen zijn voor dit onderzoek verschillende modellen opgesteld en uitgetoetst om te komen tot een regressie met een zo groot mogelijke verklarende kracht (R^2). In de volgende sub paragrafen is een beschrijving van de zoektocht naar een regressiemodel weergegeven, eerst op basis van de originele dataset en vervolgens op basis van de getransformeerde dataset.

4.3.1 Originele dataset

Het agrarisch vastgoed, zoals gedefinieerd voor dit onderzoek, bestaat uit de bebouwde oppervlakte van de bedrijfsgebouwen, de oppervlakte van de bouwkael en de eventuele aanwezigheid van een woning. Daarom zijn de eerste lineaire regressies gemodelleerd aan de hand van de variabelen GEB_OPP, BKV_OPP en WHS_j_n. Dit leidde tot een model met een verklarende kracht (R^2) van 0.4246.

Omdat bij de aanwezigheid van een woning sprake is van multicollineariteit met de variabelen bouwjaarklasse (H_BJ_KL3 en H_BJ_KL7) en bebouwde oppervlakte van de woning (HS_OPP) is de variabele WHS_j_n hierdoor vervangen. Met bouwjaarklassen neemt de verklarende kracht van het model toe tot 0.4409 respectievelijk 0.4640. De verklarende kracht neemt iets toe (0.4262) met variabele bebouwde oppervlakte woning, maar dan is de variabele GEB_OPP niet meer significant. Wanneer de oppervlakte van de tuin (TN_OPP) of H_BJ_KL3 als extra variabele aan het model wordt toegevoegd neemt de verklarende kracht wel toe, maar zijn één of meer variabele(n) niet meer significant. Uit het voorgaande wordt geconcludeerd dat de invloed van de aanwezigheid van een woning op de waarde van het agrarisch vastgoed in een model het beste tot uitdrukking komt door de variabele H_BJ_KL7 te gebruiken. De verklarende kracht van het model is dan 0.4640.

Vervolgens is getracht de verklarende kracht te verbeteren door telkens een andere variabele aan het model toe te voegen naast de variabelen GEB_OPP, BKV_OPP en H_BJ_KL7. Het toevoegen van de variabele DEEL_NED (R^2 0.5032), Afs_Adam (R^2 0.5320), Inw_Gem (R^2 0.4640), Inw_Gkm2 (R^2 0.4670), OPP_TOT (R^2 0.4684) of G_BJ_KL3 (R^2 0.5058) leidde weliswaar tot een hogere verklarende kracht van het model, maar dan bleek de toegevoegde variabele vaak niet significant te zijn. Alleen in combinatie met de variabele afstand tot de dichtstbij gelegen top 10 stad (Afs_tp10) is de verklarende kracht toegenomen tot 0.4911 en waren bovendien alle variabelen significant.

Aan dit model met de variabele GEB_OPP, BKV_OPP, H_BJ_KL3 en Afs_tp10 is vervolgens telkens weer een andere variabele toegevoegd. Na toevoeging van één of meer variabelen zoals G_BJ_KL3, Afs_Adam, DEEL_NED, Inw_Gem, Inw_Gkm2, INT_j_n, INT_glas, INT_vark, INT_over, EXT_veen, EXT_akke of EXT_over waren uiteindelijk telkens één of meer variabelen niet meer significant. Dit betrof dan vaak de laatst toegevoegde variabele en/of de variabele GEB_OPP.

Op basis van 231 waarnemingen, die reeds met behulp van boxplots waren gecorrigeerd voor uitschieters, heeft een regressiemodel met de variabelen GEB_OPP, BKV_OPP, H_BJ_KL7 en Afs_tp10 momenteel de grootste verklarende kracht (R^2 0.4911).

Met behulp van STATA zijn voor dit model de residuen van de waarnemingen berekend en de storingsterm bepaald. Uit de *Normaal Q-Q plot* en de *Kernal density estimate grafiek* is gebleken dat de residuen niet volledig de normale verdeling volgen. Om de verklarende kracht van het model verder te verbeteren zijn zowel de positieve als de negatieve uitschieters aan de uiteinden van de verdeling uit de dataset verwijderd. Na deze correctie omvat de dataset voor de lineaire regressie 215 waarnemingen.

Na de correctie van 16 afwijkende residuen is de verklarende kracht van het model op basis van de variabelen GEB_OPP, BKV_OPP, H_BJ_KL7 en Afs_tp10 toegenomen van 0.4911 tot 0.6414. Deze toename is reden geweest om voorgaande modellen waarvan niet alle variabelen significant waren opnieuw door te rekenen. De doorrekening laat zien dat bij drie modellen alle variabelen alsnog significant blijken te zijn en de verklarende kracht enorm is toegenomen (tabel 15).

Tabel 15: Verklarende kracht model met afhankelijke variabele KS_BD_HS

Variabelen	R ²	Opmerking
GEB_OPP, BKV_OPP, H_BJ_KL7, Afs_Adam	0.6683	Eerder niet significant
GEB_OPP, BKV_OPP, H_BJ_KL7, G_BJ_KL3	0.6431	Eerder niet significant
GEB_OPP, BKV_OPP, H_BJ_KL7, Afs_tp10, G_BJ_KL3	0.6599	Eerder niet significant

Bron: STATA

Voorgaande analyses leidt tot de conclusie dat de afhankelijke variabele koopsom agrarisch vastgoed (KS_BD_HS) door lineaire regressie het best wordt verklaard (R^2 0.6683) aan de hand van de variabelen GEB_OPP, BKV_OPP, H_BJ_KL7 en Afs_Adam.

Voor de resultaten van uitgevoerde lineaire regressieanalyses wordt verwezen naar bijlage 2.

4.3.2. Getransformeerde dataset

De eerste lineaire regressies zijn wederom gemodelleerd aan de hand van de getransformeerde variabelen LN_opgeb, LN_opbkv en de variabele WHS_j_n. De verklarende kracht (R^2) van dit model is 0.5350. Uit de toename van de verklarende kracht ten opzichte van het eerste lineaire model op basis van de originele dataset (R^2 0.4246) (zie paragraaf 4.3.1) blijkt dat sprake is van proportionaliteit in bebouwde oppervlakte van de bedrijfsgebouwen en de in de oppervlakte van de bouwkaavel.

Vervolgens is ook hier de variabele WHS_j_n vervangen door de variabelen bouwklasse (H_BJ_KL3 en H_BJ_KL7) en de getransformeerde variabele LN_ophs. Met bouwjaarklassen neemt de verklarende kracht van het model toe tot 0.5442 respectievelijk 0.5563. Maar met de getransformeerde variabele bebouwde oppervlakte woning (LN_ophs) neemt de verklarende kracht van het model fors af tot 0.1836! Dit is komt doordat het aantal getransformeerde variabelen (172) minder is dan het aantal originele variabelen (231). Bij de originele variabelen is bij afwezigheid van een woning de bebouwde oppervlakte woning en ook de oppervlakte tuin weergegeven als 0. Maar de natuurlijke logaritme van 0 wordt in de getransformeerde dataset weergegeven als een leeg veld, waardoor voor de variabele LN_ophs en ook LN_optn dus geen waarneming beschikbaar is. Uit het voorgaande wordt geconcludeerd dat de invloed van de aanwezigheid van een woning op de waarde van het agrarisch vastgoed ook met de getransformeerde dataset het beste tot uitdrukking komt door de variabele H_BJ_KL7 te gebruiken. De verklarende kracht van het model is dan 0.5563.

Vervolgens is getracht de verklarende kracht te verbeteren door telkens een andere (getransformeerde) variabele aan het lineaire regressiemodel toe te voegen naast de variabelen LN_opgeb, LN_opbkv en H_BJ_KL7. Het toevoegen van de variabele DEEL_NED (R^2 0.5897), LN_inw (R^2 0.5566), of LN_optot (R^2 0.5571) leidde weliswaar tot een hogere verklarende kracht van het model, maar dan bleek de toegevoegde variabele vaak niet significant te zijn. Alleen in combinatie met variabele LN_Adam (R^2 0.6062), LN_top10 (R^2 0.5756), LN_inkm2 (R^2 0.5688), INT_j_n (R^2

0.5650) of G_BJ_KL3 (R^2 0.5830) is de verklarende kracht van het model toegenomen en waren bovendien alle variabelen significant.

Aan de modellen met de variabelen LN_opgeb, LN_opbkv, H_BJ_KL7 en één van de significante extra variabele LN_Adam, LN_top10, LN_inkm2, INT_j_n of G_BJ_KL3 is vervolgens telkens weer een andere (getransformeerde) variabele toegevoegd. Na toevoeging van één of meer variabelen zoals DEEL_NED, LN_Adam, LN_top10, LN_inw, LN_inkm2, LN_optot, INT_j_n, INT_glas, INT_vark, INT_over, EXT_veen, EXT_akke of EXT_over waren uiteindelijk telkens één of meer variabelen niet meer significant. Dit betrof dan vaak de laatst toegevoegde variabele. Alleen wanneer de basis van het model, LN_opgeb, LN_opbkv en H_BJ_KL7, werden aangevuld met de volgende combinatie van variabelen nam de verklarende kracht van het model toe en bleven alle variabelen significant: G_BJ_KL3 en LN_inkm2 (R^2 0.6000), DEEL_NED en INT_j_n (R^2 0.6002) en LN_top10, DEEL_NED en INT_j_n (R^2 0.6105).

Op basis van 231 waarnemingen, die reeds met behulp van boxplots waren gecorrigeerd voor uitschieters, heeft een regressiemodel met de variabelen LN_opgeb, LN_opbkv, H_BJ_KL7, LN_top10, DEEL_NED en INT_j_n momenteel de grootste verklarende kracht (R^2 0.6105).

Met behulp van STATA zijn voor ook dit model de residuen van de waarnemingen berekend en de storingsterm bepaald. Uit de *Normaal Q-Q plot* en de *Kernal density estimate grafiek* is gebleken dat de residuen niet volledig de normale verdeling volgen. Om de verklarende kracht van het model verder te verbeteren zijn zowel de positieve als de negatieve uitschieters aan de uiteinden van de verdeling uit de dataset verwijderd. Na deze correctie omvat de dataset voor de lineaire regressie 211 waarnemingen.

Na de correctie van 20 afwijkende residuen is de verklarende kracht van het model op basis van de variabelen LN_opgeb, LN_opbkv, H_BJ_KL7, LN_top10, DEEL_NED en INT_j_n toegenomen van 0.6105 tot 0.7397. Ook hier is de verklarende kracht van het model toegenomen en reden om de voorgaande modellen waarvan variabelen al dan niet significant waren opnieuw door te rekenen. De doorrekening laat zien dat bij vier modellen, waarbij aanvankelijk niet alle variabelen significant waren, deze nu wel significant blijken te zijn en de verklarende kracht enorm is toegenomen. In onderstaande tabel is een overzicht van de betreffende variabelen en de verklarende kracht (R^2) weergegeven.

Tabel 16: Verklarende kracht model met afhankelijke variabele LN_ksb_h

Variabelen	R^2	Opmerking
LN_opgeb, LN_opbkv, H_BJ_KL7, DEEL_NED	0.7192	Eerder niet significant
LN_opgeb, LN_opbkv, H_BJ_KL7, LN_Adam	0.7333	
LN_opgeb, LN_opbkv, H_BJ_KL7, LN_top10	0.7050	
LN_opgeb, LN_opbkv, H_BJ_KL7, LN_inkm2	0.7057	
LN_opgeb, LN_opbkv, H_BJ_KL7, INT_j_n	0.6969	
LN_opgeb, LN_opbkv, H_BJ_KL7, G_BJ_KL3	0.7100	
LN_opgeb, LN_opbkv, H_BJ_KL7, G_BJ_KL3, DEEL_NED	0.7346	Eerder niet significant
LN_opgeb, LN_opbkv, H_BJ_KL7, G_BJ_KL3, LN_Adam	0.7489	Eerder niet significant
LN_opgeb, LN_opbkv, H_BJ_KL7, G_BJ_KL3, LN_top10	0.7228	Eerder niet significant
LN_opgeb, LN_opbkv, H_BJ_KL7, G_BJ_KL3, LN_inkm2	0.7293	
LN_opgeb, LN_opbkv, H_BJ_KL7, DEEL_NED, INT_j_n	0.7282	

Bron: STATA

Voorgaande analyses leidt tot de conclusie dat de getransformeerde variabele koopsom agrarisch vastgoed (LN_ksb_h) door lineaire regressie het best wordt verklaard (R^2 0.7489) aan de hand van de variabelen LN_opgeb, LN_opbkv, H_BJ_KL7, G_BJ_KL3 en LN_Adam.

Voor de resultaten van uitgevoerde lineaire regressieanalyses wordt verwezen naar bijlage 3.

4.4 Regressiemodel

De doelstelling van deze scriptie is om te komen tot een eenvoudig regressiemodel als mede een overzicht van waardebepalende variabelen. In voorgaande sub paragrafen is onderzocht welk model de grootste verklarende kracht (R^2) heeft. Het beste regressiemodel op basis van de originele dataset heeft een R^2 van 0.6683 (figuur 9), maar het beste regressiemodel op basis van de getransformeerde dataset met een R^2 van 0.7489 (figuur 10) scoort beter. In beide modellen zijn de gehanteerde variabelen allemaal significant.

Figuur 9: Regressiemodel o.b.v. originele dataset

Linear regression		Number of obs		=	215	
		F(10, 204)		=	39.27	
		Prob > F		=	0.0000	
		R-squared		=	0.6683	
		Root MSE		=	1.2e+05	
KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	31.98685	9.593281	3.33	0.001	13.07215	50.90154
BKV_OPP	16.40543	3.238902	5.07	0.000	10.01942	22.79145
H_BJ_KL7						
1	162197.8	20519.2	7.90	0.000	121740.9	202654.7
2	215419.4	29100.95	7.40	0.000	158042.2	272796.6
3	142795.6	28113.47	5.08	0.000	87365.32	198225.8
4	289351.2	31443.68	9.20	0.000	227354.9	351347.5
5	243326.9	39671.03	6.13	0.000	165109.1	321544.7
6	319433.4	37820.38	8.45	0.000	244864.5	394002.4
7	301098.6	51759.11	5.82	0.000	199047.2	403150
Afs_Adam	-1010.348	187.1435	-5.40	0.000	-1379.332	-641.3648
_cons	169841.8	27223.01	6.24	0.000	116167.2	223516.3

Bron: STATA

Het lineaire regressiemodel op basis van de originele dataset is met de bijbehorende coëfficiënten als volgt te beschrijven:

$$KS_BD_HS = 31.98685 * GEB_OPP + 16.40543 * BKV_OPP + 162197.8 * H_BJ_KL7_1 + 215419.4 * H_BJ_KL7_2 + 142795.6 * H_BJ_KL7_3 + 289351.2 * H_BJ_KL7_4 + 243326.9 * H_BJ_KL7_5 + 319433.4 * H_BJ_KL7_6 + 301098.6 * H_BJ_KL7_7 - 1010.348 * Afs_Adam + 169841.8$$

Figuur 10: Regressiemodel o.b.v. getransformeerde dataset

Linear regression		Number of obs	=	211
		F(12, 198)	=	38.01
		Prob > F	=	0.0000
		R-squared	=	0.7489
		Root MSE	=	.34821

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
LN_opgeb	.1125436	.047172	2.39	0.018	.0195197 .2055676
LN_opbkv	.2486562	.0506562	4.91	0.000	.1487614 .348551
H_BJ_KL7					
1	.8770617	.0812956	10.79	0.000	.7167454 1.037378
2	1.109312	.096092	11.54	0.000	.9198168 1.298807
3	.9281171	.0968333	9.58	0.000	.7371601 1.119074
4	1.151474	.0953269	12.08	0.000	.9634872 1.33946
5	1.103054	.0992202	11.12	0.000	.9073901 1.298718
6	1.125542	.1396941	8.06	0.000	.8500628 1.401021
7	1.12291	.1358885	8.26	0.000	.8549355 1.390884
G_BJ_KL3					
2	.1310813	.0616972	2.12	0.035	.0094133 .2527493
3	.2456793	.0732367	3.35	0.001	.1012553 .3901034
LN_Adam	-.2921293	.0637611	-4.58	0.000	-.4178672 -.1663913
_cons	10.2189	.394219	25.92	0.000	9.441491 10.9963

Bron: STATA

Het lineaire regressiemodel op basis van de getransformeerde dataset is met de bijbehorende coëfficiënten als volgt te beschrijven:

$$LN_{kb_h} = 0.1125436 * LN_{opgeb} + 0.2486562 * LN_{opbkv} + 0.8770617 * H_BJ_KL7_1 + 1.109312 * H_BJ_KL7_2 + 0.9281171 * H_BJ_KL7_3 + 1.1551474 * H_BJ_KL7_4 + 1.103054 * H_BJ_KL7_5 + 1.125542 * H_BJ_KL7_6 + 1.12291 * H_BJ_KL7_7 + 0.1310813 * G_BJ_KL_2 + 0.2456793 * G_BJ_KL3 - 0.2921293 * LN_Adam + 10.2189$$

Naast dit regressiemodel zijn met verschillende andere getransformeerde variabelen nog vier regressies te modelleren die een verklarende kracht groter dan 0.7290 hebben. In dat geval worden de variabelen LN_opgeb, LN_opbkv, H_BJ_KL7 aangevuld met:

- DEEL_NED, LN_top10 en INT_j_n of;
- LN_Adam of;
- G_BJ_KL3 en DEEL_NED of;
- G_BJ_KL3 en LN_inkm2.

Om vervolgens met behulp van bovenstaand regressiemodel een geschatte waarde te kunnen toetsen dient de getransformeerde oppervlakte van de bedrijfsgebouwen en de bouwkael en de getransformeerde afstand tot de stad Amsterdam te worden ingevuld. Voor de dummy variabelen H_BJ_KL7 en G_BJ_KL3 dient indien van toepassing een 1 ingevuld te worden, anders een 0. Na alle vermenigvuldigingen en optellingen resulteert de getransformeerde waarde van het gebouwde agrarisch vastgoed. Deze waarde geeft de punt op de regressielijn weer. Door deze te verminderen respectievelijk te vermeerderen met 1,96 maal Root MSE (0.34821) wordt met 95% betrouwbaarheid de bandbreedte van de waarde weergegeven. Vervolgens dienen deze getransformeerde waarden middels de exponentiele functie¹² naar normale (originele) waarden te worden teruggerekend.

¹² $\exp(x) = e^x$, waarbij $e \approx 2,718282$

Wanneer de geschatte waarde op basis van een andere waarderingsmethode binnen deze bandbreedte is gelegen, kan worden gesteld dat de geschatte waarde betrouwbaar is.

Uit de twee gevonden regressiemodellen kan worden afgeleid dat de hieronder genoemde variabelen c.q. factoren van invloed zijn (geweest) op de koopsom respectievelijk de waarde van agrarisch vastgoed:

- Bebouwde oppervlakte van de bedrijfsgebouwen
- Oppervlakte van de bouwkaavel
- Bouwjaarklasse van de woning (met de korte perioden) mede als indicatie van de aanwezigheid van een woning
- Afstand tot de stad Amsterdam

Uit de overige regressies met een grote verklarende kracht kan tevens worden afgeleid dat één of meer van onderstaande variabelen c.q. factoren van invloed zijn (geweest) op de koopsom respectievelijk de waarde van agrarisch vastgoed:

- Bouwjaarklasse van de bedrijfsgebouwen
- Afstand tot de dichtstbij gelegen top 10 stad
- (Geografische) ligging in landsdeel
- Bevolkingsdichtheid van de burgerlijke gemeente
- Aanwezigheid intensief landbouwbedrijf

5 Conclusie

5.1 Conclusie

De kwaliteit van een waardering op basis van de vergelijkingsmethode kan worden verbeterd door gebruik te maken van een tweede waarderingsmethode voor de toetsing van de geschatte waarde. De gecorrigeerde vervangingswaarde methode zou hier geschikt voor kunnen zijn, ware het niet dat de waarde op basis van de gecorrigeerde vervangingswaarde methode niet meer kan zijn dan op basis van de vergelijkingsmethode. Om de cirkelredenering tussen de vergelijkingsmethode en de gecorrigeerde vervangingswaarde methode te doorbreken stond bij dit onderzoek de volgende vraag centraal:

In hoeverre het mogelijk is om met een regressiemodel de waarde van gebouwd agrarisch vastgoed op een valide en betrouwbare manier te toetsen?

De conclusie van dit onderzoek is, met enig voorbehoud, dat het mogelijk is om met behulp van een meervoudig regressiemodel de waarde van agrarisch vastgoed te toetsen. Uit de verklarende kracht (R^2) van het regressiemodel op basis van de getransformeerde dataset kan worden geconcludeerd dat de gehanteerde variabelen bijna 75% van de waarde verklaren. Uit de lage t-waarde van deze variabelen wordt echter geconcludeerd dat de 95% betrouwbaarheidsinterval breed is. Daardoor zal bij de toetsing van een waarde van agrarisch vastgoed op basis van zijn objectkenmerken, deze waarde veelal significant zijn. De validiteit en betrouwbaarheid van het model wordt daarom als matig tot gemiddeld beoordeeld. Dat is de reden om enig voorbehoud te stellen.

Op basis van de uitgevoerde regressieanalyses op de getransformeerde dataset zijn, naast het hierboven aangehaalde model, nog vier regressiemodellen te formuleren die ruim 72% van de variantie van de totale waarde van het agrarisch vastgoed verklaren. De variabelen die de grootste relatieve invloed hebben op de waarde van het agrarisch vastgoed zijn:

- Bouwjaarklasse van de woning (8 dummy variabelen) (i.c. aanwezigheid van een woning)
- Oppervlakte van de bouwkaavel
- Oppervlakte van de bedrijfsgebouwen
- Bouwjaarklasse van de bedrijfsgebouwen (3 dummy variabelen)
- Afstand tot Amsterdam
- Ligging in landsdeel (4 dummy variabelen)
- Afstand tot dichtstbij gelegen top 10 stad
- Bevolkingsdichtheid burgerlijke gemeente
- Intensieve bedrijfsvoering (2 dummy variabelen)

Op basis van de dataset is het mogelijk gebleken om een regressiemodel te formuleren en kan de hypothese derhalve worden aangenomen. Hierdoor wordt de cirkelredenering die schuilt in de gedachte van de Hoge Raad doorbroken. De waarde van agrarisch vastgoed kan nu ook worden getoetst met behulp van de regressiemethode in plaats van de gecorrigeerde vervangingswaardemethode.

Om het antwoord op de centrale vraagstelling te kunnen geven zijn voor dit onderzoek een drietal deelvragen geformuleerd, die in de hierna volgende alinea's kort zullen worden beantwoord.

Deelvraag 1: Welke aspecten zouden volgens de literatuur van invloed kunnen zijn op de waardering van agrarisch vastgoed?

Uit geraadpleegde literatuur over onderzoeken die gericht waren op factoren die de prijs van landbouwgrond verklaarden, blijkt het te gaan om variabelen zoals de afstand tot het industrie- en stedelijk gebied, bevolkingsdichtheid, bodemproductiviteit, bedrijfstype en bedrijfsoppervlakte. Uit de taxatieleer blijkt dat het voor gebouwd agrarisch vastgoed gaat om kenmerken zoals oppervlakte, bouwjaar, bouwaard, kwaliteit en doelmatigheid van de bedrijfsgebouwen en de woning als mede de omvang van de bouwkaavel en de tuin.

Deelvraag 2: Hoe zou voor agrarisch vastgoed de regressiemethode kunnen worden vorm geven?

Omdat meerdere variabelen invloed hebben op de waarde van het agrarisch vastgoed is bij dit onderzoek gebruik gemaakt van een meervoudige regressie. Na uitvoering van de Pearson Correlatie Coëfficiënt en de Spearman Rangorde Correlatie testen en de regressieanalyses kan worden geconcludeerd dat de variabelen omvang en bouwjaarklassen bedrijfsgebouwen, omvang bouwkwavel, bouwjaarklasse van de woning in combinatie met de aanwezigheid van een woning en de afstand tot Amsterdam significant zijn. Op basis van deze getransformeerde variabelen is een regressiemodel geformuleerd met een verklarende kracht bijna 75%.

Deelvraag 3: Hoe kan een regressiemodel voor transacties van gebouwd agrarisch vastgoed eruit zien?

Het regressiemodel ziet op basis van de genoemde getransformeerde variabelen met de bijbehorende coëfficiënten er als volgt uit:

$$LN_kb_h = 0.1125436 * LN_opgeb + 0.2486562 * LN_opbkv + 0.8770617 * H_BJ_KL7_1 + 1.109312 * H_BJ_KL7_2 + 0.9281171 * H_BJ_KL7_3 + 1.1551474 * H_BJ_KL7_4 + 1.103054 * H_BJ_KL7_5 + 1.125542 * H_BJ_KL7_6 + 1.12291 * H_BJ_KL7_7 + 0.1310813 * G_BJ_KL_2 + 0.2456793 * G_BJ_KL3 - 0.2921293 * LN_Adam + 10.2189$$

5.2 Reflectie

Bij deze scriptie is geworsteld met een wetenschappelijke onderbouwing van het theoretisch kader op basis waarvan een selectie van verklarende variabelen voor de waarde van agrarisch vastgoed kon worden samengesteld. Tijdens de zoektocht naar literatuur over waardebepalende factoren van agrarisch vastgoed kwamen alleen artikelen en studies naar voren over waardebepalende factoren van landbouwgrond. In eerste instantie was dit niet hoopgevend, omdat deze waardebepalende factoren betrekking hebben op het locatieaspect. Maar refererend aan de klassieke locatietheorieën is dat bij uitstek een geschikte factor om in de dataset op te nemen. Door ook uit de taxatiemethodiek van agrarisch vastgoed waardebepalende factoren af te leiden is uiteindelijk een plausibele dataset ontstaan.

Zoals bij elk kwantitatief onderzoek is een uitgebreide dataset met vele waarnemingen en vele variabelen onontbeerlijk. Het is tegengevallen om deze gegevens uit openbare bronnen te verkrijgen en met elkaar te koppelen. De kwaliteit van de gegevensbronnen bleken niet altijd even betrouwbaar, waardoor het meer tijd en energie heeft gevergd om de gewenste vastgoeddata te verzamelen. Daarom zijn aanvankelijk alleen waarnemingen uit 2017 in het onderzoek betrokken, terwijl aanvankelijk waarnemingen uit de periode 2014 tot en met 2017 waren beoogd. Voor het concept van deze scriptie is een dataset met 160 waarnemingen opgebouwd. Hieruit resulteerde een regressiemodel met een verklarende kracht van 56% en met een breed betrouwbaarheidsinterval. Omdat de verwachting is dat een toename van het aantal waarnemingen zal leiden tot betrouwbaardere waarden, is voor de definitieve versie van de scriptie de dataset uitgebreid met transacties uit 2016 tot 283 waarnemingen. Als gevolg hiervan is de verklarende kracht toegenomen tot bijna 75% en is de betrouwbaarheidsinterval iets smaller geworden (iets hogere t-waarde van de variabelen).

De koppeling met locatiedata, zoals landsdelen, provincies, landbouwgebieden en landbouwgebiedsgroepen, was relatief eenvoudig. De verwachting was dat elk landbouwgebied zijn eigen significante invloed heeft op de waarde van het agrarisch vastgoed. Dit bleek niet het geval. Doordat sprake is van 66 (extra) variabelen en bij elk van variabele te weinig waarnemingen aanwezig waren, werden geen significante variabelen onderscheiden. Verder zijn duidelijk verschillen waarneembaar op landsdeelniveau. Zo zijn in Zuid-Nederland beduidend meer waarnemingen gelegen dan in West-Nederland.

Daarnaast bleken de dummy variabelen die zien op het bedrijfstype geen van allen significant te zijn. Waarschijnlijk is dit het gevolg van de onevenwichtige verdeling van de bedrijfstypen binnen de categorieën intensieve en extensieve landbouwbedrijven. Glastuinbouwbedrijven zijn in de dataset ondervertegenwoordigd, terwijl paarden- en melkveehouderijbedrijven oververtegenwoordigd zijn.

De verwachting dat de afstand tot de stad een verklarende variabele is van de koopsom blijkt op basis van deze dataset juist te zijn. Verder blijken de bevolkingsdichtheid, de geografische ligging in landsdelen en/of de aanwezigheid intensief landbouwbedrijf een bijdrage te leveren aan de verklaring van de koopsom van agrarisch vastgoed. Deze variabelen vergroten de verklarende kracht van het regressiemodel die in de basis bestaat uit de variabelen bouwjaarklassen en bebouwde oppervlakte van de bedrijfsgebouwen, oppervlakte van de bouwkael en bouwjaarklassen van de woning.

Op basis van de uitgevoerde regressieanalyses in combinatie met deze dataset lijkt een plausibel model gevonden om de waarde van agrarisch vastgoed te toetsen. De betrouwbaarheid van het model zal waarschijnlijk beter worden wanneer hier meer waarnemingen en beter gemeten variabelen aan ten grondslag liggen. De gevonden regressiemodellen zijn niet geschikt om de waarde van glastuinbouwbedrijven met meer dan 8.500 m² kassen en/of bedrijven met een bouwkael groter dan 18.000 m² te toetsen, gelet op het feit dat de verwijderde uitschieters juist deze bedrijven c.q objectkenmerken betroffen. Gezien de uiteindelijke samenstelling van de dataset zijn de uitkomsten het best toepasbaar op veehouderijbedrijven in de provincies Friesland, Drenthe, Overijssel, Gelderland, Noord-Brabant en Limburg.

5.3 Aanbevelingen

Op basis van deze scriptie kunnen een aantal aanbevelingen voor vervolgonderzoek worden gedaan. De belangrijkste aanbeveling sluit hierbij aan op de laatste alinea van de vorige paragraaf.

- Meer waarnemingen in de dataset opnemen door de transactieperiode uit te breiden over nog meer jaren in plaats van waarnemingen uit slechts twee jaar.
- In de dataset de eenheid van de woning meten in kubieke meters en/of vierkante meters gebruiksoppervlak in plaats van bebouwde oppervlakte.
- In de dataset de afstand tot de dichtstbij gelegen stad verder verfijnen door ook andere grote steden op te nemen, zoals steden als Leeuwarden, Zwolle, Enschede, Maastricht e.d..
- Glastuinbouwbedrijven niet in de dataset opnemen, omdat de bebouwde oppervlakte van de bedrijfsgebouwen fors afwijkt van andere agrarische bedrijven.
- Het uitvoeren van een regressieanalyse op een deelsegment van de agrarische sector, bijvoorbeeld de paarden- en melkveehouderijbedrijven, om een specifiek regressiemodel te vinden. Dit is alleen mogelijk wanneer hiervoor voldoende (120 en meer) waarnemingen beschikbaar zijn.

Literatuurlijst

- Agrimatie. (2018a). *Economische resultaten*. Opgehaald van [www.agrimatie.nl](https://www.agrimatie.nl/ThemaResultaat.aspx?subpubID=2232&themaID=2272&indicatorID=2100§orID=2233):
<https://www.agrimatie.nl/ThemaResultaat.aspx?subpubID=2232&themaID=2272&indicatorID=2100§orID=2233>
- Agrimatie. (2018b). *Economisch resultaat*. Opgehaald van [Argimatie.nl](https://www.agrimatie.nl/themaresultaat.aspx?subpubid=2291&themaID=2272&indicatorid=2100§orid=2245):
<https://www.agrimatie.nl/themaresultaat.aspx?subpubid=2291&themaID=2272&indicatorid=2100§orid=2245>
- Agrimatie. (2018c). *Economisch resultaat*. Opgehaald van [Agrimatie.nl](https://www.agrimatie.nl/PublicatiePage.aspx?subpubID=2291§orID=2255&themaID=2272&indicatorID=2100):
<https://www.agrimatie.nl/PublicatiePage.aspx?subpubID=2291§orID=2255&themaID=2272&indicatorID=2100>
- Agrimatie. (2018d). *Economisch resultaat*. Opgehaald van [agrimatie.nl](https://www.agrimatie.nl/ThemaResultaat.aspx?subpubID=2232&themaID=2272&indicatorID=2100§orID=2258):
<https://www.agrimatie.nl/ThemaResultaat.aspx?subpubID=2232&themaID=2272&indicatorID=2100§orID=2258>
- Agrimatie. (2018e). *Economische resultaten*. Opgehaald van [www.agrimatie.nl](https://www.agrimatie.nl/ThemaResultaat.aspx?subpubID=2232&themaID=2272&indicatorID=2100§orID=2257):
<https://www.agrimatie.nl/ThemaResultaat.aspx?subpubID=2232&themaID=2272&indicatorID=2100§orID=2257>
- Arnhem, P.C. van, Berkhout, T.M. & Have, G.G.M. ten. (2013). *Taxatieleer Vastgoed 1*. Groningen/Houten: Noordhoff Uitgevers bv.
- Arnhem, P.C. van, Berkhout, T.M. & Have, G.G.M. ten. (2015). *Taxatieleer Vastgoed 1*. Groningen/Houten: Noordhoff Uitgevers bv.
- Bieleman, J. (1990). De verscheidenheid van de landbouw op de Nederlandse zandgronden tijdens 'de lange zestiende eeuw'. In *BMGN*, 105 afl. 4 (pp. 537-552).
- Bieleman, J. (2008). *Boeren in Nederland. Geschiedenis van de landbouw 1500-2000*. Amsterdam: Uitgeverij Boom.
- Brans, S., Silvis, H. & Voskuilen, M. (2018). *NVM Agrarisch & Landelijk Vastgoedbericht*. Nieuwegein: Nederlandse Vereniging van MAKelaars en Taxateurs in onroerende goederen NVM.
- Burg, van der. (2016). Heffingsambtenaar maakt de waarde van een in de commerciële sfeer geëxploiteerde onroerende zaak niet aannemelijk. *NDFR*.
- CBS. (2018a). *StatLine*. Opgehaald van Landbouw: gewassen, dieren en grondgebruik naar hoofdbedrijfstype, regio:
<https://opendata.cbs.nl/statline/#/CBS/nl/dataset/80783ned/table?dl=15236>
- CBS. (2018b). *StatLine*. Opgehaald van Landbouw: gewassen, dieren, grondgebruik en arbied op nationaal niveau:
<https://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=81302NED&D1=336-352&D2=12-17&HDR=T&STB=G1&VW=T>
- CBS. (2018c). *Landbouw; gewassen, dieren en grondgebruik naar regio*. Opgehaald van www.cbs.nl:
<https://statline.cbs.nl/Statweb/selection/?VW=T&DM=SLNL&PA=80780ned&D1=598,636&D2=0,13&D3=0,5,10,17-18&HDR=G1,G2&STB=T>
- CBS. (2018d). Opgehaald van https://www.cbs.nl/-/media/_excel/2017/48/land%20en%20tuinbouwbedrijven%20%20aantal%20bedrijven%20standaardopbrengsten%20en%20kengetallen%20per%20bedrijfstype.xls
- CBS. (2018e). *Regionale kerncijfers Nederland*. Opgehaald van www.cbs.nl:
<https://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=70072NED&D1=0,56&D2=71,82,154,229,276,281,452,486,573,643,654,691,715&D3=21-23&HDR=T&STB=G1,G2&VW=T>
- Cotteleer, G., Gardbroek, C. & Luijt, J. (2008). Market power in a GIS-based hedonic price model of local farmland markets. *Land Economics*, 573-592.
- Francke, M.K. (2018). Sheets Modelmatig waarderen. *Module 2: Geavanceerde waarderingsmethoden*. Amsterdam: Amsterdam School of Real Estate.
- Gellicum, A. van. (2015). Landelijk Vastgoed. In P.C. van Arnhem, T.M. Berkhout & G.G.M. ten Have, *Taxatieleer Vastgoed 2* (pp. 494-528). Groningen/Houten: Noordhoff uitgevers bv.
- Gies, T.J.A., Nieuwenhuizen, W. & Smidt, R.A. (2014). *Vrijkomende agrarische bebouwing in het landelijk gebied*. Utrecht: InnovatieNetwerk.

- Gies, T.J.A., Nieuwenhuizen, W., Naeff, H.S.D. & Och, R.A.F. van. (2016). *Leegstand agrarisch vastgoed Noord-Brabant*. Wageningen: Alterra Wageningen UR.
- Gies, T.J.A., Nieuwenhuizen, W., Naeff, H.S.D., Vleemigh, I. & Paulissen, M. (2016). *Landelijk gebied en Leegstand*. Wageningen: Wageningen Environmental Research.
- Gool, P. van, Jager, P., Theebe, M. & Weisz, R. (2013). *Onroerend goed als belegging*. Groningen/houten: Noordhoff Uitgevers bv.
- Groot, H. de, Marlet, G., Teulings C. & Vermeulen, W. (2010). *Stad en Land*. Den Haag: Centraal Planbureau.
- Huang, H., Miller, G.Y., Sherrick, B.J. & Go'mes, M.I. (2006). Factors influencing Illinois farmland values. *American Agricultural Economics Association*, 458-470.
- Hof Arnhem-Leeuwarden, ECLI:NL:GHARL:2015:637 (2015).
- Hoge Raad, ECLI:NL:HR:1995:AA1634 (1995).
- Kroes, R.G. (2013). Regional differences of the Dutch rural land market. *Influence of agricultural zones on land prices*. Wageningen: Wageningen UR / Landbouw Economisch Instituut.
- Liebrechts, M. (2011). *Woningbouwtypen als ordeningsprincipe (1)*. Opgehaald van kennisbank Bestaande Woningbouw:
<http://www.bestaandewoningbouw.nl/woningkwaliteit-in-de-tijd-periode-tot-en-met-1960-woningtypen-als-ordeningsprincipe-1/print/>
- Marquard, A.R. (z.j.). *Syllabus Methoden en technieken voor Mass Appraisal*. Amsterdam: Amsterdam School of Real Estate.
- Nederlands Register Vastgoed Taxateurs. (2018). *Praktijkhandreiking Kamer Landelijk en Agrarisch Vastgoed*. Opgehaald van de website van het NRVt:
<https://www.nrvt.nl/regelgeving/kamer-en-reglementen/>
- Paul, L. (2009). Op zoek naar Von Thünen. In B. J. M. van Egmond, *Bijzonder onderzoek: Een ontdekkingsreis door de Bijzondere Collecties van de Universiteitsbibliotheek Utrecht* (pp. 196-201). Utrecht: Universiteitsbibliotheek Utrecht.
- Platform Taxateurs en Accountants. (2013). *Goed gewaardeerd vastgoed. 28 Aanbevelingen voor taxeren en taxatierapporten*. Opgehaald van de website de IVBN:
<https://www.ivbn.nl/viewer/file.aspx?fileinfoID=596>
- Ploeg, B. van der. (2002). Toekomst voor een historisch landschap? Het veenweidegebied tussen wereldmarkt en Randstad. *Landschap*, 201-213.
- Rabobank. (2018a). *Granen en groente*. Opgehaald van www.rabobank.nl:
<https://www.rabobank.nl/bedrijven/cijfers-en-trends/akkerbouw/granen-groente/>
- Rabobank. (2018b). *Trends en ontwikkelingen*. Opgehaald van www.rabobank.nl:
<https://www.rabobank.nl/bedrijven/cijfers-en-trends/veehouderij/>
- Rabobank. (2018c). *Varkensstudie 2018: Nederlandse varkenshouderij van de toekomst*. Opgehaald van www.rabobank.nl: <https://www.rabobank.nl/bedrijven/cijfers-en-trends/veehouderij/varkensstudie-toekomst-varkenshouderij/>
- Silvis, H.J., Meer, R.W. van der & Voskuilen, M.J. (2018). *Pachtnormen 2018*. Wageningen: Wageningen Economic Research.
- Sprengler, J. (2018). *David Ricardo*. Opgehaald van Encyclopaedia Britannica:
<https://www.britannica.com/print/article/502193>
- Stokes, J.R. & Cox, A.T. (2014). The speculative value of farm real estate. *Journal of Real Estate Research*, 169-186.
- Strootman Landschapsarchitecten. (2018). *Ruimtelijke perspectieven Laag Holland. Ontwerpend onderzoek naar de relatie tussen veenbehoud en landbouw*. z.p.
- TEGoVA. (2016). *Europese Taxatiestandaarden*. België: Gullis nv/sa.
- Tsoodle, L.J., Golden, B.B. & Featherstone, A.M. (2006). Factors influencing Kansas agricultural farm land values. *Land Economics*, 124-139.
- University of Amsterdam. (2012). Formulas and tables. *Mathematics & Statistics 2012-2013*. Amsterdam: Department of Quantitative Economics, Faculty of Economics and Business.
- Veen, E., Jong, D. de, Jansma, J.E., Schans, J.W. van der & Lassu, A. (2013). *Notitie Stedennetwerk Stadslandbouw. Mogelijkheden voor het (tijdelijk) inzetten van grond voor stadslandbouw*. Wageningen: Stichting Dienst landbouwkundig Onderzoek (DLO) onderzoeksinstituut Praktijkonderzoek Plant & Omgeving.

Vermeulen, I. (2007). *Agrarisch vastgoed*. Doetinchem: Reed Business b.v.
VNG. (2018). *Taxatiewijzer en kengetallen Deel 20B, Agrarische bedrijfsgebouwen*. Den Haag:
Vereniging Nederlandse Gemeenten.
Wyatt, P. (2013). *Property Valuation*. Chichester: John Wiley & Sons, Ltd.

Databronnen

PDOK viewer

Ten behoeve van:

- Bouwjaar gebouwen > laag Basisregistraties Adressen en Gebouwen; Panden [wms].
- Bebouwde oppervlakte gebouwen > laag Basisregistraties Adressen en Gebouwen; Panden [wms].
- Bebouwde oppervlakte gebouwen > laag Kadastrale Kaart; Kadastrale kaart [wms].
- Kadastrale percelen > laag Kadastrale Kaart; Kadastrale kaart [wms].
- Bouwvlak > laag Ruimtelijke plannen; Bouwvlak [wms] en Enkelbestemming [wms].

Website: <https://pdokviewer.pdok.nl/>

ANWB

Ten behoeve van:

- Afstanden tot plaatsen

Website:

<https://www.anwb.nl/verkeer/routeplanner/kilometers-berekenen>

Google Maps

Ten behoeve van:

- Bebouwde oppervlakte gebouwen > luchtfoto.

Website: <https://www.google.nl/maps/@52.2379264,6.197248,14z>

StatLine

Ten behoeve van:

- Landbouwgebiedsgroepen
- Landbouwgebieden
- Burgerlijke gemeenten

Website: <http://statline.cbs.nl/Statweb/carto/?LA=NL>

Ten behoeve van:

- Inwoners per burgerlijke gemeente; aantal.
- Inwoners per burgerlijke gemeente; aantal inwoners per km².

Website: <https://statline.cbs.nl/Statweb/publication/?DM=SLNL>

Regioatlas

Ten behoeve van:

- Landsdelen

Website: https://www.regioatlas.nl/kaarten#_nuts1regioslandsdelen

Interne systemen

Ten behoeve van:

- Transacties van kadastrale percelen > (kopie) Basisregistratie Kadaster.
- Leveringsakte > (kopie) Openbare registers van het Kadaster.
- Economische activiteit > intern systeem van de organisatie (BvR).

Bijlage 1 - Correlatie testen

Resultaten van de Pearson Correlatie Coëfficiënt en de Spearman Rangorde Correlatie testen verwerkt in één overzicht.

De correlatie tussen twee continue variabelen zijn lichtblauw gemarkeerd (Pearson Correlatie Coëfficiënt) en de correlatie tussen een continue variabele en een categorische of dummy variabele of tussen twee dummy variabelen of tussen twee categorische variabelen zijn niet (wit) gemarkeerd (Spearman Rangorde Correlatie).

Correlatie testen op basis van originele dataset

	KS_BD_HS	OPP_TOT	GEB_BWJR	G_BJ_KL3	G_BJ_KL7	GEB_OPP	BKV_OPP	HS_BWJR	H_BJ_KL3	H_BJ_KL7	HS_OPP	TN_OPP	Afs_Adam	Afs_tp10
KS_BD_HS	1.0000													
OPP_TOT	0.2557	1.0000												
GEB_BWJR	0.2427	0.0580	1.0000											
G_BJ_KL3	0.3692	0.0960	0.9286	1.0000										
G_BJ_KL7	0.3943	0.1541	0.9817	0.9155	1.0000									
GEB_OPP	0.4063	0.0527	0.2531	0.2468	0.2634	1.0000								
BKV_OPP	0.4799	0.3943	0.1054	0.0945	0.1270	0.5660	1.0000							
HS_BWJR	0.2446	-0.0064	0.3405	0.3303	0.3611	0.0544	-0.0349	1.0000						
H_BJ_KL3	0.1668	0.0894	0.3945	0.3789	0.3853	-0.0098	-0.0866	0.7690	1.0000					
H_BJ_KL7	0.2612	0.0616	0.3689	0.3402	0.3688	0.0634	-0.0421	0.9714	0.7689	1.0000				
HS_OPP	0.5458	0.0722	0.0028	0.1684	0.2312	0.2836	0.2689	0.1678	0.1356	0.2311	1.0000			
TN_OPP	0.3964	-0.0616	-0.0142	0.1428	0.1667	0.1842	0.1882	0.1061	0.1446	0.1792	0.6879	1.0000		
Afs_Adam	-0.2271	0.0192	-0.1330	-0.0738	-0.0970	-0.0352	0.1418	-0.0860	-0.0067	-0.0770	-0.0247	-0.0115	1.0000	
Afs_tp10	-0.1281	0.1216	-0.2119	-0.1696	-0.1992	-0.0314	0.1524	-0.2698	-0.1714	-0.2643	0.0501	-0.0040	0.3180	1.0000
Inw_Gem	0.0073	0.1074	-0.0017	-0.0647	-0.0404	0.0447	-0.0477	-0.0411	-0.0492	0.0276	0.1080	0.0008	-0.2590	-0.0705
Inw_Gkm2	-0.0085	-0.0175	0.0474	-0.0281	0.0298	0.0578	-0.1253	0.0116	0.0261	0.1001	-0.0143	-0.1026	-0.3815	-0.2906
WHS_i_n														
INT_i_n	0.1089	-0.2419	-0.0310	-0.0440	-0.0347	0.2979	0.0531	0.0946	0.0258	0.0736	0.2227	0.1880	-0.0617	-0.0952
INT_glas	-0.0133	-0.1526	-0.0356	-0.0645	-0.0371	0.1345	-0.0856	0.0898	0.0786	0.0760	0.0996	-0.0306	-0.2120	-0.1091
INT_vark	-0.0356	-0.0323	0.0253	0.0434	0.0456	0.2435	0.1144	0.1621	0.1283	0.1660	0.1540	0.0939	0.0631	-0.0346
INT_over	0.1569	-0.1768	-0.0396	-0.0561	-0.0596	0.0753	0.0041	-0.0662	-0.1107	-0.0868	0.0795	0.1475	-0.0267	-0.0300
EXT_veen	-0.1202	0.3483	-0.0524	-0.0791	-0.0720	-0.0269	0.1619	-0.1512	-0.0993	-0.1532	-0.2169	-0.2177	0.0867	0.1456
EXT_akke	0.0617	-0.0766	0.0647	0.1073	0.0875	-0.0966	-0.0603	-0.0415	-0.0113	-0.0262	-0.0169	0.0978	0.0722	0.0577
EXT_over	-0.0540	-0.0865	0.0795	0.1047	0.0962	-0.3942	-0.3163	0.1360	0.1387	0.1616	-0.0090	-0.0662	-0.1157	-0.1422
DEEL_NED	0.1957	-0.1263	0.1322	0.1173	0.1490	-0.1665	-0.1118	0.2448	0.1040	0.2232	0.1553	0.0371	-0.1828	-0.4449
Nrd_NED	-0.2431	0.0619	-0.1023	-0.0704	-0.1252	0.0667	0.1405	-0.2784	-0.0751	-0.2506	-0.2834	-0.1023	0.4414	0.2915
Oost_NED	0.0868	0.0545	-0.0437	-0.0889	-0.0291	0.1654	-0.0326	0.1258	-0.0308	0.1207	0.2361	0.1462	-0.2702	0.2485
West_NED	0.0706	0.0778	0.0395	0.0717	0.0368	-0.0596	-0.0875	-0.0569	0.0007	-0.0797	0.0258	-0.0672	-0.4874	-0.1600
Zuid_NED	0.1032	-0.1622	0.1098	0.0986	0.1202	-0.1732	-0.0435	0.1892	0.0990	0.1831	0.0363	0.0102	0.1629	-0.3935

	Inw_Gem	Inw_Gkm2	WHS_i_n	INT_i_n	INT_glas	INT_vark	INT_over	EXT_veen	EXT_akke	EXT_over	DEEL_NED	Nrd_NED	Oost_NED	West_NED	Zuid_NED
KS_BD_HS															
OPP_TOT															
GEB_BWJR															
G_BJ_KL3															
G_BJ_KL7															
GEB_OPP															
BKV_OPP															
HS_BWJR															
H_BJ_KL3															
H_BJ_KL7															
HS_OPP															
TN_OPP															
Afs_Adam															
Afs_tp10															
Inw_Gem	1.0000														
Inw_Gkm2	0.7915	1.0000													
WHS_i_n															
INT_i_n	-0.0002	0.1580		1.0000											
INT_glas	0.1105	0.1111		0.2183	1.0000										
INT_vark	0.0340	0.1085		0.4574	-0.0889	1.0000									
INT_over	-0.0770	0.0392		0.6499	-0.1263	-0.2646	1.0000								
EXT_veen	-0.0002	-0.1431		-0.6994	-0.1527	-0.3199	-0.4546	1.0000							
EXT_akke	-0.0161	-0.1655		-0.3021	-0.0660	-0.1382	-0.1964	-0.2374	1.0000						
EXT_over	0.0180	0.1377		-0.2808	-0.0613	-0.1284	-0.1825	-0.2207	-0.0953	1.0000					
DEEL_NED	-0.0500	0.3812		0.1582	0.0813	0.2261	-0.0434	-0.2996	0.0135	0.2211	1.0000				
Nrd_NED	-0.0857	-0.4545		-0.2084	-0.1300	-0.2015	-0.0113	0.3398	-0.0248	-0.1879	-0.8155	1.0000			
Oost_NED	0.1816	0.0832		0.1278	0.0170	0.0461	0.0979	-0.1473	0.0462	-0.0246	-0.1880	-0.3644	1.0000		
West_NED	0.0982	0.2645		-0.0316	0.1928	-0.1121	-0.0294	0.0576	-0.0583	0.0189	0.1293	-0.2354	-0.2153	1.0000	
Zuid_NED	-0.1518	0.1738		0.1034	-0.0249	0.2269	-0.0583	-0.2285	0.0216	0.1880	0.8577	-0.4561	-0.4172	-0.2695	1.0000

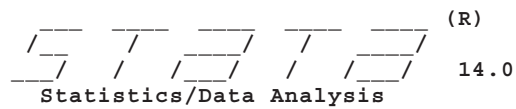
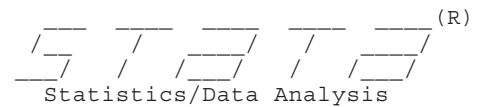
Correlatie testen op basis van getransformeerde dataset

	LN_ksb_h	LN_optot	GEB_BWJR	G_BJ_KL3	G_BJ_KL7	LN_opgeb	LN_opbkv	HS_BWJR	H_BJ_KL3	H_BJ_KL7	LN_ophs	LN_optn	LN_Adam	LN_top10
LN_ksb_h	1.0000													
LN_optot	0.3418	1.0000												
GEB_BWJR	0.1663	0.1021	1.0000											
G_BJ_KL3	0.3692	0.0960	0.9286	1.0000										
G_BJ_KL7	0.3943	0.1541	0.9817	0.9155	1.0000									
LN_opgeb	0.4384	0.2289	0.2524	0.2468	0.2634	1.0000								
LN_opbkv	0.4669	0.4039	0.0461	0.0945	0.1270	0.6664	1.0000							
HS_BWJR	0.2488	0.0289	0.3405	0.3303	0.3611	0.0218	-0.0833	1.0000						
H_BJ_KL3	0.1668	0.0894	0.3945	0.3789	0.3853	-0.0098	-0.0866	0.7690	1.0000					
H_BJ_KL7	0.2612	0.0616	0.3689	0.3402	0.3688	0.0634	-0.0421	0.9714	0.7689	1.0000				
LN_ophs	0.2926	-0.0402	0.2676	0.1684	0.2312	0.2887	0.0510	0.2045	0.1356	0.2311	1.0000			
LN_optn	0.1162	-0.2229	0.1848	0.1428	0.1667	0.1179	-0.0435	0.1508	0.1446	0.1792	0.3659	1.0000		
LN_Adam	-0.1861	-0.0077	-0.1079	-0.0738	-0.0970	-0.0019	0.0864	-0.0738	-0.0067	-0.0770	-0.1214	-0.0047	1.0000	
LN_top10	-0.0374	-0.1442	-0.1921	-0.1696	-0.1992	0.0997	0.2565	-0.2625	-0.1714	-0.2643	-0.0860	-0.0991	0.3410	1.0000
LN_inw	0.0118	-0.0475	0.0020	-0.0647	-0.0404	0.0583	-0.0220	-0.0131	-0.0492	0.0276	0.1375	0.0161	-0.2893	0.0453
LN_inkm2	-0.0107	-0.2044	0.0767	-0.0281	0.0298	-0.0014	-0.1398	0.0723	0.0261	0.1001	0.1630	0.0003	-0.4846	-0.4125
WHS_i_n														
INT_i_n	0.1089	-0.2419	-0.0310	-0.0440	-0.0347	0.2979	0.0531	0.0946	0.0258	0.0736	0.2227	0.1880	-0.0617	-0.0952
INT_glas	-0.0133	-0.1526	-0.0356	-0.0645	-0.0371	0.1345	-0.0856	0.0898	0.0786	0.0760	0.0996	-0.0306	-0.2120	-0.1091
INT_vark	-0.0356	-0.0323	0.0253	0.0434	0.0456	0.2435	0.1144	0.1621	0.1283	0.1660	0.1540	0.0939	0.0631	-0.0346
INT_over	0.1569	-0.1768	-0.0396	-0.0561	-0.0596	0.0753	0.0041	-0.0662	-0.1107	-0.0868	0.0795	0.1475	-0.0267	-0.0300
EXT_veen	-0.1202	0.3483	-0.0524	-0.0791	-0.0720	-0.0269	0.1619	-0.1512	-0.0993	-0.1532	-0.2169	-0.2177	0.0867	0.1456
EXT_akke	0.0617	-0.0766	0.0647	0.1073	0.0875	-0.0966	-0.0603	-0.0415	-0.0113	-0.0262	-0.0169	0.0978	0.0722	0.0577
EXT_over	-0.0540	-0.0865	0.0795	0.1047	0.0962	-0.3942	-0.3163	0.1360	0.1387	0.1616	-0.0090	-0.0662	-0.1157	-0.1422
DEEL_NED	0.1957	-0.1263	0.1322	0.1173	0.1490	-0.1665	-0.1118	0.2448	0.1040	0.2232	0.1553	0.0371	-0.1828	-0.4449
Nrd_NED	-0.2431	0.0619	-0.1023	-0.0704	-0.1252	0.0667	0.1405	-0.2784	-0.0751	-0.2506	-0.2834	-0.1023	0.4414	0.2915
Oost_NED	0.0868	0.0545	-0.0437	-0.0889	-0.0291	0.1654	-0.0326	0.1258	-0.0308	0.1207	0.2361	0.1462	-0.2702	0.2485
West_NED	0.0706	0.0778	0.0395	0.0717	0.0368	-0.0596	-0.0875	-0.0569	0.0007	-0.0797	0.0258	-0.0672	-0.4874	-0.1600
Zuid_NED	0.1032	-0.1622	0.1098	0.0986	0.1202	-0.1732	-0.0435	0.1892	0.0990	0.1831	0.0363	0.0102	0.1629	-0.3935

	LN_inw	LN_inkm2	WHS_i_n	INT_i_n	INT_glas	INT_vark	INT_over	EXT_veen	EXT_akke	EXT_over	DEEL_NED	Nrd_NED	Oost_NED	West_NED	Zuid_NED
LN_ksb_h															
LN_optot															
GEB_BWJR															
G_BJ_KL3															
G_BJ_KL7															
LN_opaeb															
LN_opbkv															
HS_BWJR															
H_BJ_KL3															
H_BJ_KL7															
LN_ophs															
LN_optn															
LN_Adam															
LN_top10															
LN_inw	1.0000														
LN_inkm2	0.5379	1.0000													
WHS_i_n															
INT_i_n	-0.0002	0.1580		1.0000											
INT_glas	0.1105	0.1111		0.2183	1.0000										
INT_vark	0.0340	0.1085		0.4574	-0.0889	1.0000									
INT_over	-0.0770	0.0392		0.6499	-0.1263	-0.2646	1.0000								
EXT_veen	-0.0002	-0.1431		-0.6994	-0.1527	-0.3199	-0.4546	1.0000							
EXT_akke	-0.0161	-0.1655		-0.3021	-0.0660	-0.1382	-0.1964	-0.2374	1.0000						
EXT_over	0.0180	0.1377		-0.2808	-0.0613	-0.1284	-0.1825	-0.2207	-0.0953	1.0000					
DEEL_NED	-0.0500	0.3812		0.1582	0.0613	0.2261	-0.0434	-0.2996	0.0135	0.2211	1.0000				
Nrd_NED	-0.0857	-0.4545		-0.2084	-0.1300	-0.2015	-0.0113	0.3398	-0.0248	-0.1879	-0.8155	1.0000			
Oost_NED	0.1816	0.0832		0.1278	0.0170	0.0461	0.0979	-0.1473	0.0462	-0.0246	-0.1880	-0.3644	1.0000		
West_NED	0.0982	0.2645		-0.0316	0.1928	-0.1121	-0.0294	0.0576	-0.0583	0.0189	0.1293	-0.2354	-0.2153	1.0000	
Zuid_NED	-0.1518	0.1738		0.1034	-0.0249	0.2269	-0.0583	-0.2285	0.0216	0.1880	0.8577	-0.4561	-0.4172	-0.2695	1.0000

Bijlage 2 - Afdruk regressieanalyses o.b.v. originele dataset

Afdruk van de resultaten van de met STATA uitgevoerde regressieanalyses. Tussen de regressies zijn de genomen stappen weergegeven. Onder elke regressie staat een korte conclusie met betrekking tot de verklarende kracht van het model en de significantie van de variabelen.



Copyright 1985-2015 StataCorp LP
 StataCorp
 4905 Lakeway Drive
 College Station, Texas 77845 USA
 800-STATA-PC <http://www.stata.com>
 979-696-4600 stata@stata.com
 979-696-4601 (fax)

Single-user Stata perpetual license:
 Serial number: 301406227217
 Licensed to: sr
 sr

Notes:

1. Unicode is supported; see [help unicode advice](#).

```
1 . doedit "C:\Users\Gebruiker\Documents\MSRE\definitief\Databestand 20162017_TOT_e
> x_uitschieters_Lin .do"

2 . do "C:\Users\GEBRUI~1\AppData\Local\Temp\STD00000000.tmp"

3 . clear all

4 . import excel "C:\Users\Gebruiker\Documents\MSRE\definitief\Databestand 20162017
> _TOT_ex_uitschieters.xls", sheet("Sheet1")firstrow

5 . set more off, permanently
(set more preference recorded)

6 .
7 .
8 . ** REGRESSIES lineair-lineair o.b.v. 231 waarnemingen **
9 . * Basis gebouwd agrarisch vastgoed => GEB_OPP + BKV_OPP + WONING (WHS_j_n, H_BJ
> _KL3, H_BJ_KL7 of HS_OPP en TN_OPP)
10 .
11 . regress KS_BD_HS GEB_OPP BKV_OPP WHS_j_n, vce(robust)
```

Linear regression	Number of obs	=	231
	F(3, 227)	=	63.28
	Prob > F	=	0.0000
	R-squared	=	0.4246
	Root MSE	=	1.6e+05

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	27.45635	10.85145	2.53	0.012	6.07391	48.8388
BKV_OPP	13.24982	3.804552	3.48	0.001	5.753067	20.74657
WHS_j_n	202468.3	22289.7	9.08	0.000	158547.1	246389.5
_cons	80611.8	19705.95	4.09	0.000	41781.81	119441.8

```
12 . * R2 0.4246 alle significant
13 .
14 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL3, vce(robust)
```

Linear regression	Number of obs	=	231
	F(5, 225)	=	39.22
	Prob > F	=	0.0000
	R-squared	=	0.4409
	Root MSE	=	1.6e+05

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	25.55297	10.54848	2.42	0.016	4.766534	46.33942
BKV_OPP	13.87263	3.714938	3.73	0.000	6.552107	21.19315
H_BJ_KL3						
1	183686	23752.11	7.73	0.000	136881	230491.1
2	249275.8	37536.04	6.64	0.000	175308.7	323242.9
3	262776.4	44885.08	5.85	0.000	174327.5	351225.3
_cons	80307.82	19552.07	4.11	0.000	41779.23	118836.4

15 . * R2 0.4409 alle significant

16 .

17 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7, vce(robust)

Linear regression	Number of obs	=	231
	F(9, 221)	=	23.19
	Prob > F	=	0.0000
	R-squared	=	0.4640
	Root MSE	=	1.6e+05

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	26.43169	10.73818	2.46	0.015	5.269351	47.59403
BKV_OPP	13.29	3.676375	3.61	0.000	6.044764	20.53524
H_BJ_KL7						
1	152170.7	26379.73	5.77	0.000	100182.7	204158.8
2	213957.4	38316.82	5.58	0.000	138444.3	289470.5
3	162933.6	40229.38	4.05	0.000	83651.29	242215.9
4	264756.5	36971.4	7.16	0.000	191894.9	337618.1
5	254178.3	45864.95	5.54	0.000	163789.7	344567
6	254736	55863.57	4.56	0.000	144642.5	364829.5
7	291571.6	37691.19	7.74	0.000	217291.5	365851.8
_cons	81859.51	19550.48	4.19	0.000	43330.28	120388.7

18 . * R2 0.4640 alle significant

19 .

20 . regress KS_BD_HS GEB_OPP BKV_OPP HS_OPP, vce(robust)

Linear regression	Number of obs	=	231
	F(3, 227)	=	49.91
	Prob > F	=	0.0000
	R-squared	=	0.4262
	Root MSE	=	1.6e+05

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	16.1888	10.2451	1.58	0.115	-3.99886	36.37645
BKV_OPP	15.41057	3.494881	4.41	0.000	8.524011	22.29712
HS_OPP	1291.54	149.5986	8.63	0.000	996.7609	1586.32
_cons	113161.7	20413.49	5.54	0.000	72937.57	153385.9

```
21 . * R2 0.4262 1 niet significant (GEB_OPP)
22 .
23 . regress KS_BD_HS GEB_OPP BKV_OPP HS_OPP TN_OPP, vce(robust)
```

```
Linear regression                                Number of obs   =      231
                                                F(4, 226)       =      38.51
                                                Prob > F        =      0.0000
                                                R-squared       =      0.4270
                                                Root MSE       =      1.6e+05
```

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	16.2974	10.24881	1.59	0.113	-3.898038	36.49284
BKV_OPP	15.38244	3.497544	4.40	0.000	8.490469	22.2744
HS_OPP	1208.724	205.8222	5.87	0.000	803.1483	1614.3
TN_OPP	10.19619	19.72955	0.52	0.606	-28.68121	49.07359
_cons	111437.3	21186.37	5.26	0.000	69689.18	153185.4

```
24 . * R2 0.4270 2 niet significant (GEB_OPP, TN_OPP)
25 .
26 . regress KS_BD_HS GEB_OPP BKV_OPP HS_OPP TN_OPP i.H_BJ_KL3, vce(robust)
```

```
Linear regression                                Number of obs   =      231
                                                F(7, 223)       =      29.28
                                                Prob > F        =      0.0000
                                                R-squared       =      0.4581
                                                Root MSE       =      1.6e+05
```

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	19.65178	10.51679	1.87	0.063	-1.073214	40.37678
BKV_OPP	14.40581	3.667038	3.93	0.000	7.179332	21.63229
HS_OPP	717.1258	232.244	3.09	0.002	259.4521	1174.8
TN_OPP	-6.267197	21.45596	-0.29	0.770	-48.54958	36.01518
H_BJ_KL3						
1	102076.6	39738.97	2.57	0.011	23764.65	180388.6
2	161398.6	52210.15	3.09	0.002	58510.22	264287
3	180334.2	59182.24	3.05	0.003	63706.14	296962.2
_cons	86050.44	19661.7	4.38	0.000	47303.92	124796.9

```
27 . * R2 0.4581 2 niet significant (GEB_OPP, TN_OPP)
28 .
29 . regress KS_BD_HS GEB_OPP BKV_OPP HS_OPP i.H_BJ_KL3, vce(robust)
```

```
Linear regression                                Number of obs   =      231
                                                F(6, 224)       =      32.68
                                                Prob > F        =      0.0000
                                                R-squared       =      0.4578
                                                Root MSE       =      1.6e+05
```

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	19.54641	10.47625	1.87	0.063	-1.098207	40.19102
BKV_OPP	14.44271	3.642007	3.97	0.000	7.265727	21.61968
HS_OPP	688.3869	230.9058	2.98	0.003	233.3613	1143.412
H_BJ_KL3						
1	97925.29	35080.82	2.79	0.006	28794.63	167055.9
2	156711.2	48389.56	3.24	0.001	61354.16	252068.2
3	175381.6	54733.92	3.20	0.002	67522.35	283240.9
_cons	86022.07	19646.54	4.38	0.000	47306.39	124737.8


```

30 . * R2 0.4578 1 niet significant (GEB_OPP)
31 .
32 .
33 . ** Voor aanwezigheid woning verder met H_BJ_KL7 **
34 . * Extra variabele toevoegen aan de beste regressie:
35 .
36 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 i.DEEL_NED, vce(robust)

```

```

Linear regression                                Number of obs   =      231
                                                F(12, 218)      =      20.40
                                                Prob > F        =      0.0000
                                                R-squared       =      0.5032
                                                Root MSE       =      1.5e+05

```

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	23.21989	10.50977	2.21	0.028	2.506126	43.93365
BKV_OPP	15.6432	3.788502	4.13	0.000	8.176415	23.10998
H_BJ_KL7						
1	183804.9	27050.62	6.79	0.000	130490.7	237119.1
2	226478.6	37895.79	5.98	0.000	151789.6	301167.6
3	177537.4	39862.31	4.45	0.000	98972.59	256102.3
4	269812.2	37225.48	7.25	0.000	196444.3	343180.1
5	259845.9	42820.33	6.07	0.000	175451	344240.7
6	289774.6	52491.77	5.52	0.000	186318.3	393230.9
7	303395.3	38209.5	7.94	0.000	228088	378702.7
DEEL_NED						
2	74712.58	31892.56	2.34	0.020	11855.37	137569.8
3	137740.8	31354.95	4.39	0.000	75943.21	199538.5
4	87503.35	25992.4	3.37	0.001	36274.78	138731.9
_cons	-13437.26	30865.76	-0.44	0.664	-74270.77	47396.25

```

37 . * R2 0.5032 1 niet significant (constante)
38 .
39 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 Afs_Adam, vce(robust)

```

```

Linear regression                                Number of obs   =      231
                                                F(10, 220)     =      26.41
                                                Prob > F        =      0.0000
                                                R-squared       =      0.5320
                                                Root MSE       =      1.5e+05

```

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	19.41236	11.3152	1.72	0.088	-2.887709	41.71242
BKV_OPP	16.93228	3.72576	4.54	0.000	9.589532	24.27503
H_BJ_KL7						
1	156475.3	24556.49	6.37	0.000	108079.2	204871.3
2	204977.1	35579.47	5.76	0.000	134856.9	275097.3
3	162143.5	34553.86	4.69	0.000	94044.6	230242.5
4	254111.1	36469.32	6.97	0.000	182237.1	325985
5	238176.3	41864.67	5.69	0.000	155669.2	320683.4
6	279187.4	57666.68	4.84	0.000	165537.6	392837.2
7	281713.5	61323.09	4.59	0.000	160857.6	402569.4
Afs_Adam	-1292.208	217.5452	-5.94	0.000	-1720.947	-863.4687
_cons	230507.4	33282.75	6.93	0.000	164913.6	296101.3

```

40 . * R2 0.5320 1 niet significant (GEB_OPP)
41 .
42 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 Afs_tp10, vce(robust)

```

```

Linear regression                               Number of obs   =      231
                                                F(10, 220)       =      21.89
                                                Prob > F         =      0.0000
                                                R-squared        =      0.4911
                                                Root MSE        =      1.5e+05

```

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	22.44389	10.53261	2.13	0.034	1.686159	43.20162
BKV_OPP	15.45194	3.738967	4.13	0.000	8.083159	22.82071
H_BJ_KL7						
1	166375.8	27258.62	6.10	0.000	112654.3	220097.2
2	215888.2	37236.02	5.80	0.000	142503.3	289273.2
3	167630.1	38588.82	4.34	0.000	91579.05	243681.2
4	255807.8	36929.61	6.93	0.000	183026.7	328588.9
5	248877	42326.19	5.88	0.000	165460.3	332293.7
6	250689.2	56391.41	4.45	0.000	139552.7	361825.7
7	292999.1	46629.91	6.28	0.000	201100.7	384897.6
Afs_tp10	-1450.902	436.4069	-3.32	0.001	-2310.975	-590.8289
_cons	140127.1	26350.04	5.32	0.000	88196.31	192057.9

```

43 . * R2 0.4911 alle significant
44 .
45 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 Inw_Gem, vce(robust)

```

```

Linear regression                               Number of obs   =      231
                                                F(10, 220)       =      20.94
                                                Prob > F         =      0.0000
                                                R-squared        =      0.4640
                                                Root MSE        =      1.6e+05

```

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	26.42235	10.70353	2.47	0.014	5.32778	47.51692
BKV_OPP	13.29404	3.647469	3.64	0.000	6.105587	20.48249
H_BJ_KL7						
1	152139.7	26687.08	5.70	0.000	99544.65	204734.8
2	213948.7	38366.74	5.58	0.000	138335.3	289562.1
3	162923.9	40313.36	4.04	0.000	83474.11	242373.7
4	264730.9	37087.98	7.14	0.000	191637.7	337824.1
5	254176.5	45973.31	5.53	0.000	163572.1	344781
6	254725	55913.25	4.56	0.000	144530.8	364919.1
7	291601.7	37819.13	7.71	0.000	217067.6	366135.9
Inw_Gem	.0020138	.091697	0.02	0.982	-.1787031	.1827307
_cons	81778.42	19798.19	4.13	0.000	42760.03	120796.8

```

46 . * R2 0.4640 1 niet significant (Inw_Gem)

```

```
47 .
48 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 Inw_Gkm2, vce(robust)
```

```
Linear regression                                Number of obs   =      231
                                                F(10, 220)      =      21.56
                                                Prob > F         =      0.0000
                                                R-squared        =      0.4670
                                                Root MSE        =      1.6e+05
```

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	24.90816	10.94756	2.28	0.024	3.332644	46.48367
BKV_OPP	13.92619	3.720818	3.74	0.000	6.59318	21.2592
H_BJ_KL7						
1	153559.6	26418.98	5.81	0.000	101492.9	205626.3
2	217314.1	38345.45	5.67	0.000	141742.6	292885.5
3	166082.9	40378.32	4.11	0.000	86505.11	245660.8
4	265191.3	37001.75	7.17	0.000	192268.1	338114.6
5	255019.2	45433.99	5.61	0.000	165477.6	344560.7
6	257234.1	57053.56	4.51	0.000	144792.6	369675.6
7	297599.5	38765.99	7.68	0.000	221199.3	373999.8
Inw_Gkm2	26.64165	19.6855	1.35	0.177	-12.15464	65.43794
_cons	69606.19	21350.92	3.26	0.001	27527.68	111684.7

```
49 . * R2 0.4670 1 niet significant (Inw_Gkm2)
50 .
51 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 INT_j_n, vce(robust)
```

```
Linear regression                                Number of obs   =      231
                                                F(10, 220)      =      21.03
                                                Prob > F         =      0.0000
                                                R-squared        =      0.4641
                                                Root MSE        =      1.6e+05
```

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	27.23236	11.24313	2.42	0.016	5.07434	49.39038
BKV_OPP	13.15038	3.794383	3.47	0.001	5.672389	20.62837
H_BJ_KL7						
1	151493.5	26455.57	5.73	0.000	99354.74	203632.3
2	214452.3	38332.04	5.59	0.000	138907.3	289997.3
3	162469.7	40658.08	4.00	0.000	82340.57	242598.9
4	264485.7	37004.4	7.15	0.000	191557.2	337414.1
5	254968.2	45941.61	5.55	0.000	164426.2	345510.2
6	253332.8	56246.05	4.50	0.000	142482.8	364182.8
7	288978	39082.55	7.39	0.000	211953.9	366002.1
INT_j_n	-5421.267	21728.69	-0.25	0.803	-48244.28	37401.75
_cons	84158.73	22162.4	3.80	0.000	40480.95	127836.5

```
52 . * R2 0.4641 1 niet significant (INT_j_n)
53 .
```

```
54 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 OPP_TOT, vce(robust)
```

```
Linear regression                                Number of obs   =      231
                                                F(10, 220)      =      23.41
                                                Prob > F         =      0.0000
                                                R-squared        =      0.4684
                                                Root MSE        =      1.6e+05
```

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	29.2253	11.34502	2.58	0.011	6.866483	51.58412
BKV_OPP	11.20455	4.472189	2.51	0.013	2.390731	20.01836
H_BJ_KL7						
1	149991.9	26557.25	5.65	0.000	97652.72	202331.1
2	214295.9	38187.77	5.61	0.000	139035.2	289556.6
3	165549.2	41381.44	4.00	0.000	83994.47	247104
4	260074.4	36890.78	7.05	0.000	187369.8	332778.9
5	250685.4	44774.14	5.60	0.000	162444.2	338926.5
6	246192.2	56022	4.39	0.000	135783.7	356600.7
7	284925.6	35445.35	8.04	0.000	215069.7	354781.5
OPP_TOT	.1369532	.0944083	1.45	0.148	-.0491071	.3230136
_cons	83925.95	19908.66	4.22	0.000	44689.84	123162.1

```
55 . * R2 0.4684 1 niet significant (OPP_TOT)
```

```
56 .
```

```
57 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 ib3.G_BJ_KL3, vce(robust)
```

```
Linear regression                                Number of obs   =      231
                                                F(11, 219)      =      23.65
                                                Prob > F         =      0.0000
                                                R-squared        =      0.5058
                                                Root MSE        =      1.5e+05
```

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	19.38595	10.66114	1.82	0.070	-1.625617	40.39752
BKV_OPP	13.06956	3.599235	3.63	0.000	5.975984	20.16313
H_BJ_KL7						
1	178958.7	26849.22	6.67	0.000	126042.8	231874.7
2	251254.8	36254.8	6.93	0.000	179801.9	322707.8
3	196609.2	40725.41	4.83	0.000	116345.3	276873.1
4	274166.5	34879.99	7.86	0.000	205423.1	342909.9
5	258597.3	45036.97	5.74	0.000	169835.9	347358.6
6	213777.9	57240.38	3.73	0.000	100965.3	326590.4
7	221145.5	45355.4	4.88	0.000	131756.6	310534.4
G_BJ_KL3						
1	-125070.2	31415.53	-3.98	0.000	-186985.6	-63154.72
2	-74280.03	30501.73	-2.44	0.016	-134394.5	-14165.53
_cons	159465.8	30905.24	5.16	0.000	98556.05	220375.6

```

58 . * R2 0.5058 1 niet significant (GEB_OPP)
59 .
60 .
61 . * Nog extra variabele(n) toevoegen:
62 .
63 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 Afs_tp10 ib3.G_BJ_KL3 , vce(robust)

```

```

Linear regression                                Number of obs   =      231
                                                F(12, 218)      =      22.07
                                                Prob > F        =      0.0000
                                                R-squared       =      0.5278
                                                Root MSE       =      1.5e+05

```

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	16.273	10.42692	1.56	0.120	-4.277478	36.82348
BKV_OPP	15.04383	3.630092	4.14	0.000	7.889266	22.1984
H_BJ_KL7						
1	189859.4	27029.13	7.02	0.000	136587.6	243131.3
2	250183.9	35786.07	6.99	0.000	179652.9	320714.9
3	198257.5	38518.92	5.15	0.000	122340.3	274174.6
4	265385.8	35106.63	7.56	0.000	196193.9	334577.7
5	254182.5	41814.44	6.08	0.000	171770.2	336594.8
6	211798.3	58723.36	3.61	0.000	96060.13	327536.5
7	225826.4	52382.95	4.31	0.000	122584.5	329068.2
Afs_tp10	-1310.532	415.7758	-3.15	0.002	-2129.987	-491.0772
G_BJ_KL3						
1	-117473.8	30560.41	-3.84	0.000	-177705.5	-57242.15
2	-71876.04	29815.96	-2.41	0.017	-130640.5	-13111.58
_cons	208186.7	35393.28	5.88	0.000	138429.9	277943.5

```

64 . * R2 0.5278 1 niet significant (GEB_OPP)
65 .
66 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 Afs_tp10 Afs_Adam, vce(robust)

```

```

Linear regression                                Number of obs   =      231
                                                F(11, 219)      =      23.54
                                                Prob > F        =      0.0000
                                                R-squared       =      0.5410
                                                Root MSE       =      1.4e+05

```

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	17.77109	11.15526	1.59	0.113	-4.214314	39.7565
BKV_OPP	17.83957	3.788552	4.71	0.000	10.37288	25.30625
H_BJ_KL7						
1	164541.6	25491.63	6.45	0.000	114301.2	214781.9
2	207099.3	35297.58	5.87	0.000	137532.9	276665.8
3	165047.8	34376.14	4.80	0.000	97297.42	232798.2
4	249880.4	36528.29	6.84	0.000	177888.5	321872.4
5	236709.8	40240.23	5.88	0.000	157402.1	316017.5
6	274135.6	57512.71	4.77	0.000	160786.4	387484.9
7	283627.7	65106.17	4.36	0.000	155312.9	411942.6
Afs_tp10	-871.0441	442.0837	-1.97	0.050	-1742.327	.2388133
Afs_Adam	-1153.623	225.0635	-5.13	0.000	-1597.191	-710.0556
_cons	249546.2	35230.16	7.08	0.000	180112.7	318979.8

```
67 . * R2 0.5410 1 niet significant (GEB_OPP)
68 .
69 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 Afs_tp10 i.DEEL_NED, vce(robust)
```

```
Linear regression                                Number of obs    =      231
                                                F(13, 217)       =      19.07
                                                Prob > F         =      0.0000
                                                R-squared       =      0.5205
                                                Root MSE      =      1.5e+05
```

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	19.22281	10.52115	1.83	0.069	-1.513915	39.95954
BKV_OPP	17.40728	3.82632	4.55	0.000	9.865767	24.94879
H_BJ_KL7						
1	191248.6	27489.6	6.96	0.000	137067.8	245429.4
2	220338	36810.64	5.99	0.000	147785.9	292890.2
3	178222.5	37779.45	4.72	0.000	103760.9	252684.2
4	260969.5	36746.83	7.10	0.000	188543.1	333395.9
5	252372.7	40207.45	6.28	0.000	173125.5	331619.8
6	282254.8	54238.4	5.20	0.000	175353.2	389156.3
7	293308.7	49948.6	5.87	0.000	194862.2	391755.2
Afs_tp10	-1253.773	444.6816	-2.82	0.005	-2130.221	-377.3247
DEEL_NED						
2	82601.45	31162.18	2.65	0.009	21182.16	144020.7
3	118651.6	31678.8	3.75	0.000	56214.09	181089.2
4	68389.39	26623.04	2.57	0.011	15916.53	120862.2
_cons	50182.66	37563.69	1.34	0.183	-23853.73	124219

```
70 . * R2 0.5205 2 niet significant (GEB_OPP, constante)
71 .
72 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 i.DEEL_NED Afs_Adam, vce(robust)
```

```
Linear regression                                Number of obs    =      231
                                                F(13, 217)       =      20.83
                                                Prob > F         =      0.0000
                                                R-squared       =      0.5525
                                                Root MSE      =      1.4e+05
```

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	19.29569	10.62394	1.82	0.071	-1.643623	40.235
BKV_OPP	17.91755	3.767028	4.76	0.000	10.4929	25.3422
H_BJ_KL7						
1	174560.5	26225.17	6.66	0.000	122871.9	226249.2
2	208937.7	34420.34	6.07	0.000	141096.7	276778.7
3	164146.7	33764.58	4.86	0.000	97598.14	230695.2
4	247775.2	35735.79	6.93	0.000	177341.5	318208.9
5	243490.7	41060.21	5.93	0.000	162562.8	324418.6
6	283508.5	56984.88	4.98	0.000	171193.7	395823.2
7	288473.4	53054.21	5.44	0.000	183905.8	393040.9
DEEL_NED						
2	1772.547	34305.16	0.05	0.959	-65841.43	69386.52
3	-8046.358	44688.93	-0.18	0.857	-96126.29	80033.57
4	66437.97	25573.53	2.60	0.010	16033.65	116842.3
Afs_Adam	-1574.107	336.311	-4.68	0.000	-2236.961	-911.2526
_cons	228791.9	60809	3.76	0.000	108940	348643.8

73 . * R2 0.5525 2 niet significant (GEB_OPP, DEEL_NED)

74 .

75 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 Afs_tp10 Inw_Gem, vce(robust)

Linear regression	Number of obs	=	231
	F(11, 219)	=	20.07
	Prob > F	=	0.0000
	R-squared	=	0.4912
	Root MSE	=	1.5e+05

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	22.59662	10.44234	2.16	0.032	2.016284	43.17695
BKV_OPP	15.38785	3.699949	4.16	0.000	8.095789	22.67992
H_BJ_KL7						
1	166999.9	27652.9	6.04	0.000	112500	221499.8
2	216054.8	37285.33	5.79	0.000	142570.8	289538.8
3	167826	38703.92	4.34	0.000	91546.13	244105.8
4	256234.4	37048.04	6.92	0.000	183218.1	329250.8
5	248887.3	42415.07	5.87	0.000	165293.3	332481.2
6	250872.5	56331.26	4.45	0.000	139851.7	361893.3
7	292459.3	46719.73	6.26	0.000	200381.5	384537.1
Afs_tp10	-1457.072	441.7522	-3.30	0.001	-2327.702	-586.4426
Inw_Gem	-.0365699	.0959343	-0.38	0.703	-.2256427	.1525028
_cons	141847.6	27444.77	5.17	0.000	87757.93	195937.3

76 . * R2 0.4912 1 niet significant (Inw_Gem)

77 .

78 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 Afs_tp10 Inw_Gkm2, vce(robust)

Linear regression	Number of obs	=	231
	F(11, 219)	=	20.05
	Prob > F	=	0.0000
	R-squared	=	0.4913
	Root MSE	=	1.5e+05

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	22.1785	10.55189	2.10	0.037	1.382247	42.97475
BKV_OPP	15.55295	3.733901	4.17	0.000	8.193973	22.91193
H_BJ_KL7						
1	166409.1	27312.6	6.09	0.000	112579.9	220238.3
2	216608.2	37452.73	5.78	0.000	142794.3	290422.1
3	168248.7	38765.14	4.34	0.000	91848.22	244649.2
4	256082.6	36983.61	6.92	0.000	183193.3	328971.9
5	249171.5	42307.59	5.89	0.000	165789.3	332553.6
6	251333.3	56989.41	4.41	0.000	139015.4	363651.2
7	294332.4	46968.06	6.27	0.000	201765.1	386899.6
Afs_tp10	-1422.27	456.7791	-3.11	0.002	-2322.515	-522.0239
Inw_Gkm2	6.016969	17.55645	0.34	0.732	-28.58425	40.61818
_cons	136209.8	30035.93	4.53	0.000	77013.37	195406.3

79 . * R2 0.4913 niet significant (Inw_Gkm2)

80 .

81 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 Afs_tp10 INT_j_n, vce(robust)

Linear regression	Number of obs	=	231
	F(11, 219)	=	20.01
	Prob > F	=	0.0000
	R-squared	=	0.4914
	Root MSE	=	1.5e+05

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	23.57985	11.02937	2.14	0.034	1.842547	45.31715
BKV_OPP	15.25854	3.840296	3.97	0.000	7.68987	22.8272
H_BJ_KL7						
1	165449.6	27381.76	6.04	0.000	111484.1	219415.1
2	216604.4	37253.5	5.81	0.000	143183.2	290025.7
3	166979.7	38999.15	4.28	0.000	90117.97	243841.4
4	255390.6	36971.59	6.91	0.000	182524.9	328256.2
5	249993.2	42353.97	5.90	0.000	166519.6	333466.7
6	248663.1	57100.39	4.35	0.000	136126.5	361199.7
7	289282.5	47729.68	6.06	0.000	195214.2	383350.8
Afs_tp10	-1455.548	436.8558	-3.33	0.001	-2316.528	-594.5682
INT_j_n	-7777.962	21625.39	-0.36	0.719	-50398.47	34842.55
_cons	143612.4	28919.77	4.97	0.000	86615.72	200609.1

82 . * R2 0.4914 1 niet significant (INT_j_n)

83 .

84 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL3 INT_glas INT_vark INT_over EXT_vee
> EXT_akke, vce(robust)

Linear regression	Number of obs	=	231
	F(10, 220)	=	20.87
	Prob > F	=	0.0000
	R-squared	=	0.4677
	Root MSE	=	1.6e+05

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	29.93891	10.93515	2.74	0.007	8.387857	51.48996
BKV_OPP	14.78413	3.68243	4.01	0.000	7.526778	22.04149
H_BJ_KL3						
1	175806.9	23321.89	7.54	0.000	129844	221769.9
2	247084.5	36683.77	6.74	0.000	174788	319381.1
3	258347.6	43299.88	5.97	0.000	173012	343683.3
INT_glas	-86842.63	61139.75	-1.42	0.157	-207337.2	33651.93
INT_vark	-81755.29	36632.17	-2.23	0.027	-153950.2	-9560.4
INT_over	-2069.999	31962.11	-0.06	0.948	-65061.11	60921.12
EXT_vee	-57816.85	28263.94	-2.05	0.042	-113519.6	-2114.121
EXT_akke	7465.804	45833.71	0.16	0.871	-82863.52	97795.13
_cons	108533.9	25820.51	4.20	0.000	57646.68	159421.1


```

85 . * R2 0.4677 3 niet significant (INT_glas, INT_over, EXT_akke)
86 .
87 .
88 . ** Verder met de beste regressie **
89 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 Afs_tpl0, vce(robust)

```

```

Linear regression                                Number of obs    =      231
                                                F(10, 220)        =      21.89
                                                Prob > F           =      0.0000
                                                R-squared          =      0.4911
                                                Root MSE          =      1.5e+05

```

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	22.44389	10.53261	2.13	0.034	1.686159	43.20162
BKV_OPP	15.45194	3.738967	4.13	0.000	8.083159	22.82071
H_BJ_KL7						
1	166375.8	27258.62	6.10	0.000	112654.3	220097.2
2	215888.2	37236.02	5.80	0.000	142503.3	289273.2
3	167630.1	38588.82	4.34	0.000	91579.05	243681.2
4	255807.8	36929.61	6.93	0.000	183026.7	328588.9
5	248877	42326.19	5.88	0.000	165460.3	332293.7
6	250689.2	56391.41	4.45	0.000	139552.7	361825.7
7	292999.1	46629.91	6.28	0.000	201100.7	384897.6
Afs_tpl0	-1450.902	436.4069	-3.32	0.001	-2310.975	-590.8289
_cons	140127.1	26350.04	5.32	0.000	88196.31	192057.9

```

90 . * R2 0.4911 alle significant
91 .
92 .
93 . * STORING ANALYSE:
94 . predict storing, resid
95 . swilk storing

```

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
storing	231	0.98763	2.094	1.713	0.04339

```

96 . qnorm storing
97 . kdensity storing, normal
98 .
99 . ** storing ligt redelijk langs de normaal verdeling
100 .
101 . ** Dataset opgeslagen:
102 . export excel using "C:\Users\Gebruiker\Documents\MSRE\definitief\Databestand 20
    > 162017_TOT_storing.xls", firstrow(variables)
    file C:\Users\Gebruiker\Documents\MSRE\definitief\Databestand 20162017_TOT_storin
    > g.xls saved
103 . ** uitschieters (pos/neg) storing aan de uiteinden verwijderd. 16 waarnemingen
    > van de 231 = 7%

```

```

104 . ** opgeslagen in (...) \Databestand 20162017_TOT_ex_uitschieters_storing
105 .
106 . clear all

107 . import excel "C:\Users\Gebruiker\Documents\MSRE\definitief\Databestand 20162017
> _TOT_ex_uitschieters_storing.xls", sheet("Sheet1")firstrow

108 . set more off, permanently
(set more preference recorded)

109 .
110 .
111 . ** Vervolg REGRESSIES lineair-lineair o.b.v. 215 waarnemingen **
112 . ** Verder met de eerdere beste regressie:
113 .
114 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 Afs_tp10, vce(robust)

```

```

Linear regression                                Number of obs   =      215
                                                F(10, 204)      =     39.34
                                                Prob > F        =     0.0000
                                                R-squared       =     0.6414
                                                Root MSE       =     1.2e+05

```

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	33.31163	9.374476	3.55	0.000	14.82834	51.79491
BKV_OPP	15.7989	3.188015	4.96	0.000	9.513212	22.08458
H_BJ_KL7						
1	169954.8	22859.83	7.43	0.000	124883	215026.7
2	222993.8	29989.23	7.44	0.000	163865.2	282122.4
3	147552.2	32219.93	4.58	0.000	84025.45	211079
4	294345.8	30883.94	9.53	0.000	233453.1	355238.4
5	252475.4	40994.55	6.16	0.000	171648.1	333302.8
6	298924.8	34671.53	8.62	0.000	230564.3	367285.3
7	311397.4	42207.98	7.38	0.000	228177.6	394617.2
Afs_tp10	-1135.507	393.5806	-2.89	0.004	-1911.515	-359.4996
_cons	96964.7	22169.79	4.37	0.000	53253.39	140676

```

115 . * R2 0.6414 alle significant (R2 was 0.4911)
116 .
117 .
118 . * Testen met enkele andere (niet) significant regressies:
119 .
120 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 Afs_Adam, vce(robust)

```

```

Linear regression                                Number of obs   =      215
                                                F(10, 204)      =     39.27
                                                Prob > F        =     0.0000
                                                R-squared       =     0.6683
                                                Root MSE       =     1.2e+05

```

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	31.98685	9.593281	3.33	0.001	13.07215	50.90154
BKV_OPP	16.40543	3.238902	5.07	0.000	10.01942	22.79145
H_BJ_KL7						
1	162197.8	20519.2	7.90	0.000	121740.9	202654.7
2	215419.4	29100.95	7.40	0.000	158042.2	272796.6
3	142795.6	28113.47	5.08	0.000	87365.32	198225.8
4	289351.2	31443.68	9.20	0.000	227354.9	351347.5
5	243326.9	39671.03	6.13	0.000	165109.1	321544.7
6	319433.4	37820.38	8.45	0.000	244864.5	394002.4
7	301098.6	51759.11	5.82	0.000	199047.2	403150
Afs_Adam	-1010.348	187.1435	-5.40	0.000	-1379.332	-641.3648

_cons	169841.8	27223.01	6.24	0.000	116167.2	223516.3
-------	----------	----------	------	-------	----------	----------

121 . * R2 0.6683 alle significant (R2 was 0.5320; 1 niet significant)

122 .

123 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 ib3.G_BJ_KL3, vce(robust)

Linear regression	Number of obs	=	215
	F(11, 203)	=	39.03
	Prob > F	=	0.0000
	R-squared	=	0.6431
	Root MSE	=	1.2e+05

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	31.54699	9.479601	3.33	0.001	12.85588	50.2381
BKV_OPP	14.16121	3.095387	4.57	0.000	8.057982	20.26445
H_BJ_KL7						
1	174539	23067.07	7.57	0.000	129057.2	220020.8
2	241256.2	30071.48	8.02	0.000	181963.7	300548.7
3	165785.9	34101.76	4.86	0.000	98546.82	233025
4	305564.9	29664.78	10.30	0.000	247074.3	364055.5
5	259782.8	43572.1	5.96	0.000	173870.8	345694.7
6	262477.1	37218.22	7.05	0.000	189093.3	335861
7	258401.7	40241.75	6.42	0.000	179056.3	337747.1
G_BJ_KL3						
1	-85374.61	24145.13	-3.54	0.001	-132982	-37767.21
2	-56638.69	23908.39	-2.37	0.019	-103779.3	-9498.064
_cons	107564.5	24646.11	4.36	0.000	58969.34	156159.7

124 . * R2 0.6431 alle significant (R2 was 0.5058; 1 niet significant)

125 .

126 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 Afs_tp10 ib3.G_BJ_KL3 , vce(robust)

Linear regression	Number of obs	=	215
	F(12, 202)	=	37.61
	Prob > F	=	0.0000
	R-squared	=	0.6599
	Root MSE	=	1.2e+05

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	28.54985	9.250684	3.09	0.002	10.30956	46.79014
BKV_OPP	15.91125	3.075455	5.17	0.000	9.847135	21.97536
H_BJ_KL7						
1	185196.2	23530.45	7.87	0.000	138799.4	231593
2	240408.7	29683.58	8.10	0.000	181879.3	298938.1
3	168705.3	32729.72	5.15	0.000	104169.6	233241.1
4	298248.5	29527.64	10.10	0.000	240026.6	356470.4
5	256253.8	40746.76	6.29	0.000	175910.3	336597.4
6	261379.1	38037.27	6.87	0.000	186378.1	336380.1
7	261184.2	46076.47	5.67	0.000	170331.7	352036.7
Afs_tp10	-1075.687	387.6152	-2.78	0.006	-1839.978	-311.3963
G_BJ_KL3						
1	-80999.13	24064.03	-3.37	0.001	-128448	-33550.22
2	-55892.59	23383.97	-2.39	0.018	-102000.6	-9784.603
_cons	148610.8	28418.48	5.23	0.000	92575.85	204645.7

```

127 . * R2 0.6599 alle significant (R2 was 0.5278; 1 niet significant)
128 .
129 .
130 . *****
131 . **  BESTE REGRESSIES  **
132 . *****
133 .
134 . * 1 *
135 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 Afs_Adam, vce(robust)

```

```

Linear regression                                Number of obs   =      215
                                                F(10, 204)      =      39.27
                                                Prob > F         =      0.0000
                                                R-squared        =      0.6683
                                                Root MSE        =      1.2e+05

```

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	31.98685	9.593281	3.33	0.001	13.07215	50.90154
BKV_OPP	16.40543	3.238902	5.07	0.000	10.01942	22.79145
H_BJ_KL7						
1	162197.8	20519.2	7.90	0.000	121740.9	202654.7
2	215419.4	29100.95	7.40	0.000	158042.2	272796.6
3	142795.6	28113.47	5.08	0.000	87365.32	198225.8
4	289351.2	31443.68	9.20	0.000	227354.9	351347.5
5	243326.9	39671.03	6.13	0.000	165109.1	321544.7
6	319433.4	37820.38	8.45	0.000	244864.5	394002.4
7	301098.6	51759.11	5.82	0.000	199047.2	403150
Afs_Adam	-1010.348	187.1435	-5.40	0.000	-1379.332	-641.3648
_cons	169841.8	27223.01	6.24	0.000	116167.2	223516.3

```

136 . * R2 0.6683 alle significant (R2 was 0.5320; 1 niet significant)
137 .
138 . * 2 *
139 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 Afs_tp10 ib3.G_BJ_KL3 , vce(robust)

```

```

Linear regression                                Number of obs   =      215
                                                F(12, 202)      =      37.61
                                                Prob > F         =      0.0000
                                                R-squared        =      0.6599
                                                Root MSE        =      1.2e+05

```

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	28.54985	9.250684	3.09	0.002	10.30956	46.79014
BKV_OPP	15.91125	3.075455	5.17	0.000	9.847135	21.97536
H_BJ_KL7						
1	185196.2	23530.45	7.87	0.000	138799.4	231593
2	240408.7	29683.58	8.10	0.000	181879.3	298938.1
3	168705.3	32729.72	5.15	0.000	104169.6	233241.1
4	298248.5	29527.64	10.10	0.000	240026.6	356470.4
5	256253.8	40746.76	6.29	0.000	175910.3	336597.4
6	261379.1	38037.27	6.87	0.000	186378.1	336380.1
7	261184.2	46076.47	5.67	0.000	170331.7	352036.7
Afs_tp10	-1075.687	387.6152	-2.78	0.006	-1839.978	-311.3963
G_BJ_KL3						
1	-80999.13	24064.03	-3.37	0.001	-128448	-33550.22
2	-55892.59	23383.97	-2.39	0.018	-102000.6	-9784.603
_cons	148610.8	28418.48	5.23	0.000	92575.85	204645.7

```

140 . * R2 0.6599 alle significant (R2 was 0.5278; 1 niet significant)
141 .
142 . * 3 *
143 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 ib3.G_BJ_KL3, vce(robust)

```

```

Linear regression                               Number of obs   =       215
                                                F(11, 203)      =       39.03
                                                Prob > F        =       0.0000
                                                R-squared       =       0.6431
                                                Root MSE       =       1.2e+05

```

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	31.54699	9.479601	3.33	0.001	12.85588	50.2381
BKV_OPP	14.16121	3.095387	4.57	0.000	8.057982	20.26445
H_BJ_KL7						
1	174539	23067.07	7.57	0.000	129057.2	220020.8
2	241256.2	30071.48	8.02	0.000	181963.7	300548.7
3	165785.9	34101.76	4.86	0.000	98546.82	233025
4	305564.9	29664.78	10.30	0.000	247074.3	364055.5
5	259782.8	43572.1	5.96	0.000	173870.8	345694.7
6	262477.1	37218.22	7.05	0.000	189093.3	335861
7	258401.7	40241.75	6.42	0.000	179056.3	337747.1
G_BJ_KL3						
1	-85374.61	24145.13	-3.54	0.001	-132982	-37767.21
2	-56638.69	23908.39	-2.37	0.019	-103779.3	-9498.064
_cons	107564.5	24646.11	4.36	0.000	58969.34	156159.7

```

144 . * R2 0.6431 alle significant (R2 was 0.5058; 1 niet significant)
145 .
146 . * 4 *
147 . regress KS_BD_HS GEB_OPP BKV_OPP i.H_BJ_KL7 Afs_tp10, vce(robust)

```

```

Linear regression                               Number of obs   =       215
                                                F(10, 204)     =       39.34
                                                Prob > F        =       0.0000
                                                R-squared       =       0.6414
                                                Root MSE       =       1.2e+05

```

KS_BD_HS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
GEB_OPP	33.31163	9.374476	3.55	0.000	14.82834	51.79491
BKV_OPP	15.7989	3.188015	4.96	0.000	9.513212	22.08458
H_BJ_KL7						
1	169954.8	22859.83	7.43	0.000	124883	215026.7
2	222993.8	29989.23	7.44	0.000	163865.2	282122.4
3	147552.2	32219.93	4.58	0.000	84025.45	211079
4	294345.8	30883.94	9.53	0.000	233453.1	355238.4
5	252475.4	40994.55	6.16	0.000	171648.1	333302.8
6	298924.8	34671.53	8.62	0.000	230564.3	367285.3
7	311397.4	42207.98	7.38	0.000	228177.6	394617.2
Afs_tp10	-1135.507	393.5806	-2.89	0.004	-1911.515	-359.4996
_cons	96964.7	22169.79	4.37	0.000	53253.39	140676

```

148 . * R2 0.6414 alle significant (R2 was 0.4911)
149 .
150 .
151 . * STORING ANALYSE:
152 . predict storing2, resid
153 . swilk storing2

```

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
storing2	215	0.98790	1.923	1.510	0.06548

```

154 . qnorm storing2
155 . kdensity storing2, normal
156 .
157 . ** storing ligt redelijk langs de normaal verdeling
158 .
    end of do-file
159 .

```

Bijlage 3 - Afdruk regressieanalyses o.b.v. getransformeerde dataset

Afdruk van de resultaten van de met STATA uitgevoerde regressieanalyses. Tussen de regressies zijn de genomen stappen weergegeven. Onder elke regressie staat een korte conclusie met betrekking tot de verklarende kracht van het model en de significantie van de variabelen.

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1890498	.0563116	3.36	0.001	.0780842	.3000154
LN_opbkv	.1884215	.0652332	2.89	0.004	.0598754	.3169676
H_BJ_KL3						
1	.893044	.1008688	8.85	0.000	.6942756	1.091812
2	1.068007	.131962	8.09	0.000	.8079676	1.328046
3	1.111764	.1377444	8.07	0.000	.8403298	1.383198
_cons	8.929939	.4281454	20.86	0.000	8.086252	9.773627

15 . * R2 0.5442 alle significant

16 .

17 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7, vce(robust)

Linear regression	Number of obs	=	231
	F(9, 221)	=	24.37
	Prob > F	=	0.0000
	R-squared	=	0.5563
	Root MSE	=	.50715

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1863514	.0577271	3.23	0.001	.0725854	.3001174
LN_opbkv	.1873676	.0655582	2.86	0.005	.0581685	.3165668
H_BJ_KL7						
1	.8090146	.1051981	7.69	0.000	.6016949	1.016334
2	.9661009	.131626	7.34	0.000	.7066982	1.225504
3	.8574538	.1207606	7.10	0.000	.6194641	1.095443
4	1.113026	.1227679	9.07	0.000	.8710805	1.354972
5	1.056499	.1590623	6.64	0.000	.7430259	1.369972
6	1.068048	.1582812	6.75	0.000	.7561143	1.379981
7	1.276506	.1282335	9.95	0.000	1.023789	1.529223
_cons	8.957297	.4247635	21.09	0.000	8.120192	9.794403

18 . * R2 0.5563 alle significant

19 .

20 . regress LN_ksb_h LN_opgeb LN_opbkv LN_ophs, vce(robust)

Linear regression	Number of obs	=	172
	F(3, 168)	=	12.10
	Prob > F	=	0.0000
	R-squared	=	0.1836
	Root MSE	=	.41942

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.0966658	.0593397	1.63	0.105	-.0204818	.2138134
LN_opbkv	.1078282	.0541535	1.99	0.048	.0009192	.2147372
LN_ophs	.2969243	.0814992	3.64	0.000	.1360298	.4578187
_cons	9.818045	.5284891	18.58	0.000	8.774709	10.86138

```

21 . * R2 0.1836 1 niet significant (LN_opgeb; bovendien nog maar 172 waarnemingen d
    > oor minder woningen)
22 .
23 . regress LN_ksb_h LN_opgeb LN_opbkv LN_ophs LN_optn, vce(robust)

```

```

Linear regression                               Number of obs   =       172
                                                F(4, 167)         =       9.38
                                                Prob > F           =       0.0000
                                                R-squared          =       0.1839
                                                Root MSE          =       .42058

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.0955458	.0592971	1.61	0.109	-.0215227	.2126144
LN_opbkv	.1093963	.0547866	2.00	0.047	.0012327	.2175599
LN_ophs	.2877305	.0836686	3.44	0.001	.122546	.452915
LN_optn	.0177451	.069376	0.26	0.798	-.119222	.1547122
_cons	9.732355	.6889027	14.13	0.000	8.372274	11.09244

```

24 . * R2 0.1839 2 niet significant (LN_opgeb, LN_optn; bovendien nog maar 172 waarn
    > emingen door minder woningen)
25 .
26 .
27 . ** Voor aanwezigheid woning verder met H_BJ_KL7 **
28 . * Extra variabele toevoegen aan de beste regressie:
29 .
30 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.DEEL_NED, vce(robust)

```

```

Linear regression                               Number of obs   =       231
                                                F(12, 218)       =      22.31
                                                Prob > F           =       0.0000
                                                R-squared          =       0.5897
                                                Root MSE          =       .49106

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1871801	.056959	3.29	0.001	.0749193	.2994409
LN_opbkv	.2102305	.0661585	3.18	0.002	.0798384	.3406227
H_BJ_KL7						
1	.8908106	.1017363	8.76	0.000	.690298	1.091323
2	1.031378	.1312179	7.86	0.000	.7727598	1.289996
3	.9065918	.1184365	7.65	0.000	.6731647	1.140019
4	1.14143	.1229132	9.29	0.000	.8991796	1.38368
5	1.072081	.1468278	7.30	0.000	.7826971	1.361465
6	1.178892	.1466219	8.04	0.000	.8899146	1.46787
7	1.346193	.1380951	9.75	0.000	1.07402	1.618365
DEEL_NED						
2	.1576375	.0970142	1.62	0.106	-.0335683	.3488433
3	.4651127	.1001349	4.64	0.000	.2677563	.6624692
4	.2251242	.0783675	2.87	0.004	.0706693	.3795792
_cons	8.518345	.4313345	19.75	0.000	7.668225	9.368464

```

31 . * R2 0.5897 1 niet significant (DEEL_NED)
32 .
33 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 LN_Adam, vce(robust)

```

```

Linear regression                               Number of obs   =       231
                                                F(10, 220)      =       26.41
                                                Prob > F         =       0.0000
                                                R-squared        =       0.6062
                                                Root MSE        =       .47889

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1625283	.0597828	2.72	0.007	.0447079	.2803487
LN_opbkv	.2224067	.0673542	3.30	0.001	.0896648	.3551487
H_BJ_KL7						
1	.8267414	.0997806	8.29	0.000	.6300934	1.02339
2	.9727945	.1244234	7.82	0.000	.7275801	1.218009
3	.8759412	.1070999	8.18	0.000	.6648681	1.087014
4	1.111937	.1189098	9.35	0.000	.877589	1.346285
5	1.01978	.1451085	7.03	0.000	.7337993	1.30576
6	1.138385	.1691862	6.73	0.000	.8049518	1.471818
7	1.25603	.1356134	9.26	0.000	.9887629	1.523298
LN_Adam	-.3707803	.0759726	-4.88	0.000	-.5205075	-.2210531
_cons	10.57313	.5213672	20.28	0.000	9.545615	11.60064

```

34 . * R2 0.6062 alle significant
35 .
36 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 LN_top10, vce(robust)

```

```

Linear regression                               Number of obs   =       231
                                                F(10, 220)      =       26.08
                                                Prob > F         =       0.0000
                                                R-squared        =       0.5756
                                                Root MSE        =       .49713

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1725117	.0558594	3.09	0.002	.0624237	.2825997
LN_opbkv	.2321675	.0644385	3.60	0.000	.1051717	.3591634
H_BJ_KL7						
1	.8454926	.1056707	8.00	0.000	.6372362	1.053749
2	.9664217	.1287316	7.51	0.000	.7127168	1.220127
3	.8662353	.1165482	7.43	0.000	.6365414	1.095929
4	1.093112	.1218062	8.97	0.000	.8530556	1.333168
5	1.02559	.1492389	6.87	0.000	.7314692	1.319711
6	1.052914	.1585518	6.64	0.000	.7404393	1.365389
7	1.303363	.1404237	9.28	0.000	1.026615	1.580111
LN_top10	-.1957877	.064337	-3.04	0.003	-.3225835	-.068992
_cons	9.392806	.472282	19.89	0.000	8.46203	10.32358

```

37 . * R2 0.5756 alle significant

```

```
38 .
39 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 LN_inw, vce(robust)
```

```
Linear regression                                Number of obs    =      231
                                                F(10, 220)      =      22.00
                                                Prob > F         =      0.0000
                                                R-squared       =      0.5566
                                                Root MSE      =      .50813
```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1841559	.058703	3.14	0.002	.0684636	.2998482
LN_opbkv	.1894685	.0663392	2.86	0.005	.0587269	.3202101
H_BJ_KL7						
1	.8094105	.1053504	7.68	0.000	.6017854	1.017036
2	.9672128	.1322885	7.31	0.000	.7064978	1.227928
3	.8568521	.1207121	7.10	0.000	.618952	1.094752
4	1.111067	.1228805	9.04	0.000	.8688929	1.35324
5	1.057793	.159683	6.62	0.000	.7430887	1.372497
6	1.071805	.1600059	6.70	0.000	.756464	1.387145
7	1.282931	.1301925	9.85	0.000	1.026347	1.539515
LN_inw	.0188697	.0456226	0.41	0.680	-.0710437	.108783
_cons	8.759323	.6548351	13.38	0.000	7.468771	10.04988

```
40 . * R2 0.5566 1 niet significant (LN_inw)
41 .
42 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 LN_inkm2, vce(robust)
```

```
Linear regression                                Number of obs    =      231
                                                F(10, 220)      =      23.12
                                                Prob > F         =      0.0000
                                                R-squared       =      0.5688
                                                Root MSE      =      .50112
```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1688288	.0575468	2.93	0.004	.0554152	.2822424
LN_opbkv	.2114928	.0657208	3.22	0.001	.0819698	.3410158
H_BJ_KL7						
1	.8463998	.106498	7.95	0.000	.636513	1.056287
2	1.011354	.1331761	7.59	0.000	.7488903	1.273819
3	.8872249	.1217981	7.28	0.000	.6471845	1.127265
4	1.122079	.1218266	9.21	0.000	.8819821	1.362175
5	1.067023	.1537474	6.94	0.000	.7640165	1.370029
6	1.113219	.1721897	6.47	0.000	.7738665	1.452571
7	1.327956	.1374559	9.66	0.000	1.057057	1.598855
LN_inkm2	.1186981	.0439652	2.70	0.007	.0320512	.205345
_cons	8.1961	.5225692	15.68	0.000	7.166218	9.225983

```
43 . * R2 0.5688 alle significant
44 .
```

Number of obs	=	231
F(10, 220)	=	23.48
Prob > F	=	0.0000
R-squared	=	0.5650
Root MSE	=	.50332

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.2265871	.0629052	3.60	0.000	.1026132	.3505609
LN_opbkv	.1637077	.0660218	2.48	0.014	.0335915	.2938238
H_BJ_KL7						
1	.790592	.1033383	7.65	0.000	.5869323	.9942516
2	.9745082	.1293892	7.53	0.000	.7195072	1.229509
3	.8468167	.1213307	6.98	0.000	.6076974	1.085936
4	1.101234	.122317	9.00	0.000	.8601713	1.342297
5	1.084147	.1532404	7.07	0.000	.7821396	1.386154
6	1.029489	.1592447	6.46	0.000	.7156487	1.343329
7	1.205203	.1297986	9.29	0.000	.9493957	1.461011
INT_j_n	-.1514212	.0731589	-2.07	0.040	-.2956032	-.0072392
_cons	8.950143	.4121506	21.72	0.000	8.137874	9.762412

47 .

Number of obs	=	231
F(10, 220)	=	22.08
Prob > F	=	0.0000
R-squared	=	0.5571
Root MSE	=	.50783

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1887182	.0583411	3.23	0.001	.0737392	.3036971
LN_opbkv	.1743449	.0694222	2.51	0.013	.0375271	.3111626
H_BJ_KL7						
1	.7970094	.1063579	7.49	0.000	.5873986	1.00662
2	.9541775	.1316625	7.25	0.000	.6946964	1.213659
3	.8539888	.1212992	7.04	0.000	.6149316	1.093046
4	1.096403	.1246651	8.79	0.000	.8507119	1.342093
5	1.040734	.1591	6.54	0.000	.7271788	1.354289
6	1.048941	.1591872	6.59	0.000	.7352144	1.362668
7	1.246944	.1265655	9.85	0.000	.9975079	1.49638
LN_optot	.0182647	.0243719	0.75	0.454	-.0297675	.066297
_cons	8.878995	.4344918	20.44	0.000	8.022696	9.735294

50 .

Number of obs	=	231
F(11, 219)	=	23.51
Prob > F	=	0.0000
R-squared	=	0.5830
Root MSE	=	.49388

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1470259	.0588064	2.50	0.013	.031127	.2629248
LN_opbkv	.1948109	.0662522	2.94	0.004	.0642375	.3253843
H_BJ_KL7						
1	.8858443	.1058016	8.37	0.000	.6773247	1.094364
2	1.078323	.1273104	8.47	0.000	.8274121	1.329233
3	.9550708	.1267026	7.54	0.000	.7053583	1.204783
4	1.139887	.1174145	9.71	0.000	.9084802	1.371294
5	1.051924	.155879	6.75	0.000	.7447089	1.359139
6	.9578067	.1629835	5.88	0.000	.6365897	1.279024
7	1.090384	.1470205	7.42	0.000	.8006281	1.380141
G_BJ_KL3						
2	.1695432	.0793722	2.14	0.034	.013112	.3259744
3	.3543503	.1022878	3.46	0.001	.1527559	.5559447
_cons	8.991684	.4238075	21.22	0.000	8.156421	9.826947

```

52 . * R2 0.5830 alle significant
53 .
54 .
55 . * Nog extra variabele(n) toevoegen:
56 .
57 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.G_BJ_KL3 i.DEEL_NED, vce(robust
> )

```

```

Linear regression                                Number of obs    =      231
                                                F(14, 216)        =      21.33
                                                Prob > F          =      0.0000
                                                R-squared         =      0.6092
                                                Root MSE         =      .48146

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1492296	.0570743	2.61	0.010	.0367357	.2617234
LN_opbkv	.2158114	.0657818	3.28	0.001	.086155	.3454678
H_BJ_KL7						
1	.9535079	.1022502	9.33	0.000	.751972	1.155044
2	1.114926	.1284901	8.68	0.000	.861671	1.368181
3	.9865281	.1259264	7.83	0.000	.7383263	1.23473
4	1.162159	.1185867	9.80	0.000	.9284238	1.395894
5	1.062774	.1436822	7.40	0.000	.7795756	1.345973
6	1.076342	.1582536	6.80	0.000	.7644228	1.388261
7	1.162531	.1505398	7.72	0.000	.8658163	1.459247
G_BJ_KL3						
2	.1526074	.0797498	1.91	0.057	-.0045799	.3097948
3	.3080946	.0992597	3.10	0.002	.1124529	.5037363
DEEL_NED						
2	.1778146	.0956129	1.86	0.064	-.0106391	.3662684
3	.4194068	.10164	4.13	0.000	.2190735	.61974
4	.2025149	.0777049	2.61	0.010	.0493581	.3556718
_cons	8.596394	.4315038	19.92	0.000	7.745897	9.446891

```

58 . * R2 0.6092 2 niet significant (G_BJ_KL3, DEEL_NED)
59 .
60 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.G_BJ_KL3 LN_Adam, vce(robust)

```

```

Linear regression
Number of obs      =      231
F(12, 218)         =      26.13
Prob > F            =      0.0000
R-squared           =      0.6275
Root MSE           =      .46789

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1292473	.0601137	2.15	0.033	.010769	.2477257
LN_opbkv	.2275057	.0675597	3.37	0.001	.0943519	.3606595
H_BJ_KL7						
1	.8930309	.0997394	8.95	0.000	.6964539	1.089608
2	1.069987	.1202599	8.90	0.000	.8329656	1.307008
3	.9594677	.111663	8.59	0.000	.7393904	1.179545
4	1.135277	.1151698	9.86	0.000	.9082884	1.362266
5	1.021377	.1418989	7.20	0.000	.7417078	1.301047
6	1.030917	.1711468	6.02	0.000	.693603	1.368231
7	1.084682	.1566845	6.92	0.000	.7758713	1.393492
G_BJ_KL3						
2	.1401886	.0754892	1.86	0.065	-.0085935	.2889708
3	.3185076	.0966808	3.29	0.001	.127959	.5090563
LN_Adam	-.3514061	.0728217	-4.83	0.000	-.4949308	-.2078814
_cons	10.51732	.5072408	20.73	0.000	9.517598	11.51705

```

61 . * R2 0.6275 1 niet significant G_BJ_KL3)
62 .
63 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.G_BJ_KL3 LN_top10, vce(robust)

```

```

Linear regression
Number of obs      =      231
F(12, 218)         =      24.27
Prob > F            =      0.0000
R-squared           =      0.5986
Root MSE           =      .4857

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1377741	.0561733	2.45	0.015	.0270619	.2484864
LN_opbkv	.2349005	.0639881	3.67	0.000	.1087859	.3610151
H_BJ_KL7						
1	.9122565	.1052012	8.67	0.000	.7049148	1.119598
2	1.068659	.126703	8.43	0.000	.8189395	1.318379
3	.9540689	.1204364	7.92	0.000	.7167001	1.191438
4	1.119496	.117384	9.54	0.000	.8881432	1.350849
5	1.027179	.1468561	6.99	0.000	.7377394	1.316619
6	.9479865	.1664547	5.70	0.000	.61992	1.276053
7	1.123173	.157995	7.11	0.000	.8117796	1.434566
G_BJ_KL3						
2	.1491528	.0792547	1.88	0.061	-.0070506	.3053563
3	.3307651	.1010665	3.27	0.001	.1315726	.5299577
LN_top10	-.1765866	.0632292	-2.79	0.006	-.3012054	-.0519678
_cons	9.380489	.4748736	19.75	0.000	8.444558	10.31642

```

64 . * R2 0.5986 1 niet significant (G_BJ_KL3)
65 .
66 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.G_BJ_KL3 LN_inw, vce(robust)

```

```

Linear regression
Number of obs      =      231
F(12, 218)         =      21.68
Prob > F            =      0.0000
R-squared           =      0.5844
Root MSE           =      .49421

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1413195	.0599028	2.36	0.019	.0232566	.2593823
LN_opbkv	.1995546	.0666456	2.99	0.003	.0682024	.3309068
H_BJ_KL7						
1	.8885953	.1057628	8.40	0.000	.6801469	1.097044
2	1.083426	.1283227	8.44	0.000	.8305143	1.336338
3	.956101	.1264891	7.56	0.000	.706803	1.205399
4	1.136316	.1175951	9.66	0.000	.9045467	1.368084
5	1.054955	.1563477	6.75	0.000	.746809	1.363102
6	.9624392	.1641946	5.86	0.000	.6388272	1.286051
7	1.098659	.148712	7.39	0.000	.8055619	1.391757
G_BJ_KL3						
2	.1729355	.0789467	2.19	0.030	.017339	.328532
3	.3639468	.101822	3.57	0.000	.1632653	.5646283
LN_inw	.0405857	.0437115	0.93	0.354	-.0455656	.126737
_cons	8.566579	.6238924	13.73	0.000	7.336946	9.796212

```

67 . * R2 0.5844 1 niet significant (LN_inw)
68 .
69 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.G_BJ_KL3 LN_inkm2, vce(robust)

```

```

Linear regression
Number of obs      =      231
F(12, 218)         =      22.87
Prob > F            =      0.0000
R-squared           =      0.6000
Root MSE           =      .48487

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1245577	.0580309	2.15	0.033	.0101843	.2389311
LN_opbkv	.224508	.0650162	3.45	0.001	.0963672	.3526488
H_BJ_KL7						
1	.9329305	.106064	8.80	0.000	.7238883	1.141973
2	1.134634	.1287406	8.81	0.000	.8808989	1.38837
3	.9919371	.1268876	7.82	0.000	.7418537	1.242021
4	1.151122	.1174064	9.80	0.000	.9197254	1.382519
5	1.072401	.1485687	7.22	0.000	.7795857	1.365215
6	.9904522	.174904	5.66	0.000	.6457328	1.335172
7	1.122508	.1548983	7.25	0.000	.8172178	1.427798
G_BJ_KL3						
2	.1578744	.0780521	2.02	0.044	.004041	.3117077
3	.3869999	.101014	3.83	0.000	.1879107	.586089
LN_inkm2	.1401035	.0424713	3.30	0.001	.0563965	.2238104
_cons	8.091278	.5041042	16.05	0.000	7.097736	9.08482

70 . * R2 0.6000 alle significant

71 .

72 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.G_BJ_KL3 INT_j_n, vce(robust)

Linear regression

Number of obs = 231
F(12, 218) = 22.37
Prob > F = 0.0000
R-squared = 0.5872
Root MSE = .49256

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1783734	.0651529	2.74	0.007	.0499631	.3067837
LN_opbkv	.1777726	.0668173	2.66	0.008	.0460821	.3094631
H_BJ_KL7						
1	.8667178	.1046589	8.28	0.000	.660445	1.072991
2	1.074915	.1260546	8.53	0.000	.8264733	1.323357
3	.9393884	.1258336	7.47	0.000	.6913822	1.187395
4	1.129368	.117629	9.60	0.000	.8975324	1.361204
5	1.073041	.1522272	7.05	0.000	.7730152	1.373066
6	.9369623	.1658637	5.65	0.000	.6100607	1.263864
7	1.051882	.1476344	7.12	0.000	.7609088	1.342856
G_BJ_KL3						
2	.1528381	.0797634	1.92	0.057	-.004368	.3100441
3	.3288572	.101701	3.23	0.001	.1284142	.5293003
INT_j_n	-.1061821	.0730953	-1.45	0.148	-.2502461	.0378818
_cons	8.98335	.4173307	21.53	0.000	8.16083	9.805869

73 . * R2 0.5872 2 niet significant (G_BJ_KL3, INT_j_n)

74 .

75 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.G_BJ_KL3 LN_optot, vce(robust)

Linear regression

Number of obs = 231
F(12, 218) = 21.59
Prob > F = 0.0000
R-squared = 0.5832
Root MSE = .49495

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1482911	.0591709	2.51	0.013	.0316708	.2649114
LN_opbkv	.1901334	.0697411	2.73	0.007	.0526803	.3275866
H_BJ_KL7						
1	.880673	.1084085	8.12	0.000	.6670101	1.094336
2	1.072657	.1278378	8.39	0.000	.8207006	1.324613
3	.9525522	.1270374	7.50	0.000	.7021735	1.202931
4	1.133571	.1198773	9.46	0.000	.8973045	1.369838
5	1.047053	.1569529	6.67	0.000	.7377138	1.356392
6	.9506935	.166368	5.71	0.000	.6227979	1.278589
7	1.079911	.1531819	7.05	0.000	.7780038	1.381818
G_BJ_KL3						
2	.1660088	.0804703	2.06	0.040	.0074094	.3246081
3	.3521044	.1016312	3.46	0.001	.1517988	.5524099
LN_optot	.0065618	.0241705	0.27	0.786	-.0410759	.0541995
_cons	8.962875	.4306689	20.81	0.000	8.114067	9.811683

```

76 . * R2 0.5832 1 niet significant (LN_optot)
77 .
78 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.DEEL_NED LN_top10, vce(robust)

```

```

Linear regression
Number of obs      =      231
F(13, 217)         =      23.22
Prob > F           =      0.0000
R-squared          =      0.5992
Root MSE          =      .48642

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1710278	.057172	2.99	0.003	.0583442	.2837113
LN_opbkv	.2417779	.0659888	3.66	0.000	.1117168	.371839
H_BJ_KL7						
1	.9048424	.1020998	8.86	0.000	.7036082	1.106077
2	1.012763	.1295466	7.82	0.000	.7574321	1.268093
3	.9051778	.1146241	7.90	0.000	.6792587	1.131097
4	1.124656	.1218029	9.23	0.000	.8845873	1.364724
5	1.041335	.1417132	7.35	0.000	.7620249	1.320646
6	1.15548	.1511424	7.64	0.000	.857585	1.453375
7	1.336862	.1425626	9.38	0.000	1.055877	1.617846
DEEL_NED						
2	.1703189	.0953843	1.79	0.076	-.0176794	.3583172
3	.4054214	.1034237	3.92	0.000	.2015777	.609265
4	.1691538	.0829357	2.04	0.043	.0056913	.3326164
LN_top10	-.1493175	.0711257	-2.10	0.037	-.2895032	-.0091319
_cons	8.948884	.515752	17.35	0.000	7.932359	9.965408

```

79 . * R2 0.5992 1 niet significant (DEEL_NED)
80 .
81 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.DEEL_NED LN_Adam, vce(robust)

```

```

Linear regression
Number of obs      =      231
F(13, 217)         =      23.78
Prob > F           =      0.0000
R-squared          =      0.6189
Root MSE          =      .47434

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1774405	.0571778	3.10	0.002	.0647457	.2901354
LN_opbkv	.2345654	.0648444	3.62	0.000	.1067599	.3623708
H_BJ_KL7						
1	.8716682	.0995657	8.75	0.000	.6754285	1.067908
2	.9996767	.1231907	8.11	0.000	.7568731	1.24248
3	.885887	.1054888	8.40	0.000	.6779732	1.093801
4	1.099243	.1171663	9.38	0.000	.8683138	1.330173
5	1.037358	.1406371	7.38	0.000	.7601684	1.314547
6	1.151976	.1588774	7.25	0.000	.8388352	1.465116
7	1.3047	.1184227	11.02	0.000	1.071294	1.538106
DEEL_NED						
2	-.0217038	.1055963	-0.21	0.837	-.2298296	.1864219
3	.0286981	.1582074	0.18	0.856	-.2831218	.340518
4	.1858742	.0787669	2.36	0.019	.0306281	.3411203
LN_Adam	-.4278638	.1253686	-3.41	0.001	-.6749597	-.1807678
_cons	10.54429	.776794	13.57	0.000	9.013266	12.07532

```

82 . * R2 0.6189 1 niet significant (DEEL_NED)
83 .
84 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.DEEL_NED LN_inw, vce(robust)

```

```

Linear regression                                Number of obs      =          231
                                                F(13, 217)         =          20.57
                                                Prob > F           =          0.0000
                                                R-squared          =          0.5897
                                                Root MSE          =          .49215

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1864171	.0575684	3.24	0.001	.0729523	.2998819
LN_opbkv	.2111316	.066663	3.17	0.002	.0797417	.3425215
H_BJ_KL7						
1	.8909025	.1019624	8.74	0.000	.6899391	1.091866
2	1.032176	.1318702	7.83	0.000	.7722657	1.292087
3	.9062199	.1185152	7.65	0.000	.6726317	1.139808
4	1.140292	.1232407	9.25	0.000	.8973896	1.383194
5	1.072901	.1476564	7.27	0.000	.7818761	1.363925
6	1.180154	.147823	7.98	0.000	.8888008	1.471506
7	1.349989	.1398894	9.65	0.000	1.074273	1.625705
DEEL_NED						
2	.1555254	.0976536	1.59	0.113	-.0369455	.3479964
3	.4634459	.1013912	4.57	0.000	.2636083	.6632835
4	.225778	.0781379	2.89	0.004	.0717717	.3797843
LN_inw	.0088973	.0436644	0.20	0.839	-.0771632	.0949579
_cons	8.424322	.6286773	13.40	0.000	7.185227	9.663418

```

85 . * R2 0.5897 2 niet significant (DEEL_NED, LN_inw)
86 .
87 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.DEEL_NED LN_inkm2, vce(robust)

```

```

Linear regression                                Number of obs      =          231
                                                F(13, 217)         =          20.91
                                                Prob > F           =          0.0000
                                                R-squared          =          0.5912
                                                Root MSE          =          .49129

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1807419	.0576183	3.14	0.002	.0671788	.2943051
LN_opbkv	.2166435	.0663496	3.27	0.001	.0858713	.3474157
H_BJ_KL7						
1	.8949977	.102691	8.72	0.000	.6925983	1.097397
2	1.044087	.1327132	7.87	0.000	.7825147	1.305659
3	.9133992	.1195574	7.64	0.000	.6777567	1.149042
4	1.142539	.1226523	9.32	0.000	.9007969	1.384281
5	1.07451	.146266	7.35	0.000	.7862266	1.362794
6	1.184646	.1538175	7.70	0.000	.8814789	1.487814
7	1.362345	.1414979	9.63	0.000	1.083459	1.641231
DEEL_NED						
2	.1328387	.0990438	1.34	0.181	-.0623723	.3280497
3	.4184841	.1133037	3.69	0.000	.1951675	.6418006
4	.1975986	.0882237	2.24	0.026	.0237136	.3714837
LN_inkm2	.0456055	.0482655	0.94	0.346	-.0495236	.1407347
_cons	8.276179	.4994341	16.57	0.000	7.291817	9.260542

```

88 . * R2 0.5912 2 niet significant (DEEL_NED, LN_inkm2)
89 .
90 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.DEEL_NED INT_j_n, vce(robust)

```

```

Linear regression                                Number of obs    =        231
                                                F(13, 217)        =        22.08
                                                Prob > F           =        0.0000
                                                R-squared          =        0.6002
                                                Root MSE          =        .48583

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.2340383	.0625211	3.74	0.000	.1108121	.3572646
LN_opbkv	.1871718	.0663209	2.82	0.005	.0564561	.3178874
H_BJ_KL7						
1	.8866751	.1001341	8.85	0.000	.6893152	1.084035
2	1.03517	.1284124	8.06	0.000	.782075	1.288266
3	.8940757	.119921	7.46	0.000	.6577167	1.130435
4	1.123666	.1223841	9.18	0.000	.8824526	1.36488
5	1.107519	.1402182	7.90	0.000	.8311555	1.383883
6	1.138153	.146388	7.77	0.000	.8496286	1.426677
7	1.254365	.142884	8.78	0.000	.9727467	1.535983
DEEL_NED						
2	.2029251	.0985918	2.06	0.041	.008605	.3972452
3	.4713568	.1009189	4.67	0.000	.2724501	.6702636
4	.2794857	.0834873	3.35	0.001	.1149359	.4440355
INT_j_n	-.1736909	.07248	-2.40	0.017	-.3165458	-.0308359
_cons	8.434388	.4203518	20.07	0.000	7.605893	9.262883

```

91 . * R2 0.6002 alle significant
92 .
93 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.DEEL_NED LN_top10 INT_j_n, vce(
> robust)

```

```

Linear regression                                Number of obs    =        231
                                                F(14, 216)        =        23.57
                                                Prob > F           =        0.0000
                                                R-squared          =        0.6105
                                                Root MSE          =        .48063

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.2189271	.0618171	3.54	0.000	.0970853	.340769
LN_opbkv	.2191326	.0660112	3.32	0.001	.089024	.3492411
H_BJ_KL7						
1	.9011095	.1003262	8.98	0.000	.7033657	1.098853
2	1.01596	.1266534	8.02	0.000	.7663256	1.265595
3	.8921588	.1153823	7.73	0.000	.6647393	1.119578
4	1.105599	.1218412	9.07	0.000	.8654491	1.345749
5	1.076835	.1343391	8.02	0.000	.8120511	1.341618
6	1.112366	.1535462	7.24	0.000	.8097251	1.415006
7	1.241385	.1438493	8.63	0.000	.9578573	1.524913
DEEL_NED						
2	.2177227	.0975799	2.23	0.027	.0253919	.4100534
3	.4095469	.1047582	3.91	0.000	.2030677	.616026
4	.2232625	.0869422	2.57	0.011	.0518988	.3946262
LN_top10	-.1551755	.0706285	-2.20	0.029	-.2943848	-.0159661
INT_j_n	-.1798989	.0718454	-2.50	0.013	-.3215069	-.038291
_cons	8.878817	.5006062	17.74	0.000	7.892118	9.865515

```

94 . * R2 0.6105 alle significant
95 .
96 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.DEEL_NED LN_Adam INT_j_n, vce(r
> obust)

```

```

Linear regression                                Number of obs    =      231
                                                F(14, 216)         =      24.15
                                                Prob > F           =      0.0000
                                                R-squared          =      0.6301
                                                Root MSE          =      .46842

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.2255991	.0631279	3.57	0.000	.1011736	.3500247
LN_opbkv	.2110822	.0649066	3.25	0.001	.0831507	.3390136
H_BJ_KL7						
1	.8671973	.0979515	8.85	0.000	.6741342	1.06026
2	1.003233	.1195714	8.39	0.000	.7675573	1.23891
3	.8727665	.1064117	8.20	0.000	.6630282	1.082505
4	1.080481	.1172455	9.22	0.000	.8493891	1.311573
5	1.073478	.1321528	8.12	0.000	.8130042	1.333952
6	1.109715	.1597199	6.95	0.000	.794906	1.424524
7	1.209655	.1175176	10.29	0.000	.9780274	1.441283
DEEL_NED						
2	.0229665	.1063227	0.22	0.829	-.1865963	.2325293
3	.0303168	.158909	0.19	0.849	-.2828941	.3435277
4	.241436	.0837093	2.88	0.004	.0764443	.4064277
LN_Adam	-.4325825	.1271718	-3.40	0.001	-.6832391	-.1819259
INT_j_n	-.1789093	.0713167	-2.51	0.013	-.3194751	-.0383434
_cons	10.48016	.7720884	13.57	0.000	8.958365	12.00195

```

97 . * R2 0.6301 1 niet significant (DEEL_NED)
98 .
99 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.DEEL_NED LN_top10 INT_j_n i.G_B
> J_KL3, vce(robust)

```

```

Linear regression                                Number of obs    =      231
                                                F(16, 214)         =      22.36
                                                Prob > F           =      0.0000
                                                R-squared          =      0.6248
                                                Root MSE          =      .47394

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1765346	.0625919	2.82	0.005	.053159	.2999101
LN_opbkv	.228226	.064926	3.52	0.001	.1002497	.3562024
H_BJ_KL7						
1	.9549264	.1011748	9.44	0.000	.7554997	1.154353
2	1.086495	.1262145	8.61	0.000	.8377124	1.335278
3	.9624429	.119563	8.05	0.000	.7267709	1.198115
4	1.127612	.1186763	9.50	0.000	.8936878	1.361536
5	1.064877	.1334846	7.98	0.000	.8017642	1.32799
6	1.028846	.1683933	6.11	0.000	.6969245	1.360768
7	1.098093	.1605309	6.84	0.000	.7816687	1.414517
DEEL_NED						
2	.2238306	.0970117	2.31	0.022	.0326098	.4150514
3	.3701106	.1080131	3.43	0.001	.1572048	.5830164
4	.1935285	.0865718	2.24	0.026	.0228858	.3641711
LN_top10	-.1498772	.0702343	-2.13	0.034	-.2883169	-.0114375
INT_j_n	-.1409235	.0730849	-1.93	0.055	-.2849819	.003135
G_BJ_KL3						

2	.1239919	.0793237	1.56	0.120	-.0323638	.2803476
3	.2695594	.0997376	2.70	0.007	.0729654	.4661534
_cons	8.948467	.5124661	17.46	0.000	7.93834	9.958595

```

100 . * R2 0.6248 2 niet significant (INT_j_n, G_BJ_KL)
101 .
102 . regress LN_ksb_h LN_opgeb LN_opbkv WHS_j_n i.DEEL_NED INT_glas INT_vark INT_ove
> r EXT_vee EXT_akke, vce(robust)

```

```

Linear regression                                Number of obs    =      231
                                                F(11, 219)        =      24.88
                                                Prob > F          =      0.0000
                                                R-squared         =      0.6051
                                                Root MSE         =      .48065

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.2855424	.0642351	4.45	0.000	.1589442	.4121405
LN_opbkv	.1829016	.0658777	2.78	0.006	.0530662	.312737
WHS_j_n	1.005658	.0920972	10.92	0.000	.8241479	1.187169
DEEL_NED						
2	.2247501	.0910569	2.47	0.014	.0452901	.40421
3	.4761673	.1038859	4.58	0.000	.2714231	.6809115
4	.304825	.0847744	3.60	0.000	.1377469	.4719031
INT_glas	-.5198443	.1986877	-2.62	0.010	-.911429	-.1282596
INT_vark	-.4821517	.151032	-3.19	0.002	-.7798138	-.1844896
INT_ove	-.3354966	.1255824	-2.67	0.008	-.5830013	-.0879919
EXT_vee	-.3179268	.1191175	-2.67	0.008	-.5526901	-.0831635
EXT_akke	-.0965893	.1428145	-0.68	0.500	-.3780561	.1848775
_cons	8.305059	.4173179	19.90	0.000	7.482586	9.127532

```

103 . * R2 0.6051 1 niet significant (EXT_akke)
104 .
105 .
106 . ** Verder met de beste regressie **
107 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.DEEL_NED LN_top10 INT_j_n, vce(
> robust)

```

```

Linear regression                                Number of obs    =      231
                                                F(14, 216)        =      23.57
                                                Prob > F          =      0.0000
                                                R-squared         =      0.6105
                                                Root MSE         =      .48063

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.2189271	.0618171	3.54	0.000	.0970853	.340769
LN_opbkv	.2191326	.0660112	3.32	0.001	.089024	.3492411
H_BJ_KL7						
1	.9011095	.1003262	8.98	0.000	.7033657	1.098853
2	1.01596	.1266534	8.02	0.000	.7663256	1.265595
3	.8921588	.1153823	7.73	0.000	.6647393	1.119578
4	1.105599	.1218412	9.07	0.000	.8654491	1.345749
5	1.076835	.1343391	8.02	0.000	.8120511	1.341618
6	1.112366	.1535462	7.24	0.000	.8097251	1.415006
7	1.241385	.1438493	8.63	0.000	.9578573	1.524913
DEEL_NED						
2	.2177227	.0975799	2.23	0.027	.0253919	.4100534
3	.4095469	.1047582	3.91	0.000	.2030677	.616026
4	.2232625	.0869422	2.57	0.011	.0518988	.3946262
LN_top10	-.1551755	.0706285	-2.20	0.029	-.2943848	-.0159661

INT_j_n	-.1798989	.0718454	-2.50	0.013	-.3215069	-.038291
_cons	8.878817	.5006062	17.74	0.000	7.892118	9.865515

108 . * R2 0.6105 alle significant

109 .

110 .

111 . * STORING ANALYSE:

112 . predict storing, resid

113 . swilk storing

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
storing	231	0.97749	3.809	3.100	0.00097

114 . qnorm storing

115 . kdensity storing, normal

116 .

117 . ** storing ligt redelijk langs de normaal verdeling

118 .

119 . ** Dataset opgeslagen:

120 . export excel using "C:\Users\Gebruiker\Documents\MSRE\definitief\Databestand 20
> 162017_LN_storing.xls", firstrow(variables)
file C:\Users\Gebruiker\Documents\MSRE\definitief\Databestand 20162017_LN_storing
> .xls saved

121 . ** uitschieters (pos/neg) storing aan de uiteinden verwijderd. 20 waarnemingen
> van de 231 = 9%

122 . ** opgeslagen in (...) \Databestand 20162017_TOT_LN_ex_uitschieters_storing

123 .

124 . clear all

125 . import excel "C:\Users\Gebruiker\Documents\MSRE\definitief\Databestand 20162017
> _TOT_LN_ex_uitschieters_storing.xls", sheet("Sheet1")firstrow

126 . set more off, permanently
(set more preference recorded)

127 .

128 .

129 . ** Vervolg REGRESSIES lineair-lineair o.b.v. 211 waarnemingen **

130 . ** Verder met eerdere beste regressie:

131 .

132 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.DEEL_NED LN_top10 INT_j_n, vce(
> robust)

Linear regression	Number of obs	=	211
	F(14, 196)	=	32.73
	Prob > F	=	0.0000
	R-squared	=	0.7397
	Root MSE	=	.35635

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
LN_opgeb	.1810553	.0476345	3.80	0.000	.0871134 .2749971
LN_opbkv	.2462639	.0494641	4.98	0.000	.1487138 .343814
H_BJ_KL7					
1	.8761696	.0801566	10.93	0.000	.7180895 1.03425
2	1.046164	.0939412	11.14	0.000	.8608984 1.231429
3	.8623363	.0950867	9.07	0.000	.6748119 1.049861
4	1.127156	.0965286	11.68	0.000	.9367883 1.317524
5	1.115015	.1039951	10.72	0.000	.9099218 1.320108
6	1.18519	.1179594	10.05	0.000	.9525575 1.417822
7	1.192511	.1348428	8.84	0.000	.9265816 1.45844

DEEL_NED						
2	.2418738	.077573	3.12	0.002	.0888889	.3948586
3	.3395053	.0882404	3.85	0.000	.1654828	.5135278
4	.1458011	.0718995	2.03	0.044	.0040052	.2875971
LN_top10	-.1506254	.0565891	-2.66	0.008	-.2622271	-.0390238
INT_j_n	-.1544496	.0575738	-2.68	0.008	-.2679933	-.040906
_cons	8.955259	.3636601	24.63	0.000	8.23807	9.672448

```

133 . * R2 0.7397 alle significant (R2 was 0.6105)
134 .
135 .
136 . * Testen met enkele andere (niet) significant regressies:
137 .
138 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.DEEL_NED, vce(robust)

```

```

Linear regression                                Number of obs   =      211
                                                F(12, 198)      =      33.38
                                                Prob > F         =      0.0000
                                                R-squared        =      0.7192
                                                Root MSE        =      .36822

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
LN_opgeb	.1572868	.0443777	3.54	0.000	.0697733 .2448004
LN_opbkv	.2383864	.0507565	4.70	0.000	.1382937 .3384791
H_BJ_KL7					
1	.8619173	.0847071	10.18	0.000	.6948734 1.028961
2	1.069019	.1024794	10.43	0.000	.8669275 1.271111
3	.8721934	.1009703	8.64	0.000	.6730781 1.071309
4	1.161166	.099868	11.63	0.000	.9642247 1.358107
5	1.125312	.1137578	9.89	0.000	.9009797 1.349644
6	1.240967	.1151164	10.78	0.000	1.013956 1.467979
7	1.283511	.1234851	10.39	0.000	1.039996 1.527025
DEEL_NED					
2	.1899759	.0764818	2.48	0.014	.0391524 .3407994
3	.395035	.0830107	4.76	0.000	.2313365 .5587335
4	.1506063	.0635916	2.37	0.019	.0252026 .27601
_cons	8.557653	.343226	24.93	0.000	7.880805 9.2345

```

139 . * R2 0.7192 alle significant (R2 was 0.5897; 1 niet significant)
140 .
141 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 LN_Adam, vce(robust)

```

```

Linear regression                                Number of obs   =      211
                                                F(10, 200)      =      42.67
                                                Prob > F         =      0.0000
                                                R-squared        =      0.7333
                                                Root MSE        =      .35707

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
LN_opgeb	.1406448	.0463172	3.04	0.003	.049312 .2319776
LN_opbkv	.2463681	.050851	4.84	0.000	.1460951 .346641
H_BJ_KL7					
1	.8195731	.0805086	10.18	0.000	.6608185 .9783278
2	1.036492	.0942972	10.99	0.000	.8505478 1.222436
3	.8569205	.0905862	9.46	0.000	.6782938 1.035547
4	1.13637	.0956931	11.88	0.000	.9476736 1.325067
5	1.106913	.1010323	10.96	0.000	.9076877 1.306138
6	1.206393	.1368586	8.81	0.000	.9365219 1.476264
7	1.24594	.1223645	10.18	0.000	1.00465 1.48723

LN_Adam	-.3087684	.0661476	-4.67	0.000	-.4392046	-.1783322
_cons	10.24951	.4114284	24.91	0.000	9.438217	11.06081

142 . * R2 0.7333 alle significant (R2 was 0.6062)

143 .

144 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 LN_top10, vce(robust)

Linear regression	Number of obs	=	211
	F(10, 200)	=	36.30
	Prob > F	=	0.0000
	R-squared	=	0.7050
	Root MSE	=	.37558

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1523005	.0433838	3.51	0.001	.0667522	.2378489
LN_opbkv	.2450124	.0485141	5.05	0.000	.1493477	.3406771
H_BJ_KL7						
1	.8379929	.0865798	9.68	0.000	.6672666	1.008719
2	1.028839	.0980511	10.49	0.000	.8354919	1.222185
3	.8510647	.1001059	8.50	0.000	.6536662	1.048463
4	1.124215	.0984812	11.42	0.000	.9300198	1.318409
5	1.105608	.1120219	9.87	0.000	.8847122	1.326503
6	1.146017	.1162843	9.86	0.000	.9167166	1.375318
7	1.283628	.1283427	10.00	0.000	1.030549	1.536706
LN_top10	-.1603475	.0496073	-3.23	0.001	-.2581679	-.062527
_cons	9.312139	.3365774	27.67	0.000	8.648444	9.975835

145 . * R2 0.7050 alle significant (R2 was 0.5756)

146 .

147 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 LN_inkm2, vce(robust)

Linear regression	Number of obs	=	211
	F(10, 200)	=	39.06
	Prob > F	=	0.0000
	R-squared	=	0.7057
	Root MSE	=	.3751

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1455665	.0444903	3.24	0.001	.0570225	.2341105
LN_opbkv	.2372083	.0503718	4.71	0.000	.1378803	.3365364
H_BJ_KL7						
1	.8520879	.0858852	9.92	0.000	.6827313	1.021445
2	1.092594	.0981361	11.13	0.000	.8990798	1.286108
3	.8806379	.1032829	8.53	0.000	.6769748	1.084301
4	1.159315	.095859	12.09	0.000	.9702914	1.348339
5	1.153986	.1121807	10.29	0.000	.9327772	1.375195
6	1.215854	.1333917	9.11	0.000	.9528197	1.478889
7	1.323208	.1263819	10.47	0.000	1.073996	1.572419
LN_inkm2	.1213242	.0377073	3.22	0.002	.0469693	.1956791
_cons	8.135545	.4399909	18.49	0.000	7.267929	9.003162

148 . * R2 0.7057 alle significant (R2 was 0.5688)

149 .

150 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 INT_j_n, vce(robust)

Linear regression

Number of obs = 211
F(10, 200) = 38.91
Prob > F = 0.0000
R-squared = 0.6969
Root MSE = .38066

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1986503	.0503756	3.94	0.000	.0993149	.2979857
LN_opbkv	.1953005	.0491379	3.97	0.000	.0984057	.2921953
H_BJ_KL7						
1	.7935322	.0854229	9.29	0.000	.625087	.9619774
2	1.038775	.0982073	10.58	0.000	.8451204	1.23243
3	.8336122	.1053196	7.92	0.000	.6259329	1.041291
4	1.130221	.0991807	11.40	0.000	.934647	1.325795
5	1.157849	.117186	9.88	0.000	.9267699	1.388927
6	1.13092	.114506	9.88	0.000	.9051257	1.356714
7	1.205834	.1184564	10.18	0.000	.9722501	1.439417
INT_j_n	-.1268728	.0605052	-2.10	0.037	-.2461827	-.0075629
_cons	8.882203	.3183483	27.90	0.000	8.254453	9.509953

151 . * R2 0.6969 alle significant (R2 was 0.5650)

152 .

153 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.G_BJ_KL3, vce(robust)

Linear regression

Number of obs = 211
F(11, 199) = 34.66
Prob > F = 0.0000
R-squared = 0.7100
Root MSE = .3733

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1318575	.0462902	2.85	0.005	.0405752	.2231398
LN_opbkv	.2172695	.0493804	4.40	0.000	.1198934	.3146455
H_BJ_KL7						
1	.8750791	.0886749	9.87	0.000	.700216	1.049942
2	1.121529	.1027057	10.92	0.000	.9189978	1.32406
3	.9268765	.1123486	8.25	0.000	.70533	1.148423
4	1.159336	.0985115	11.77	0.000	.9650758	1.353596
5	1.128658	.1171842	9.63	0.000	.8975763	1.359741
6	1.076231	.1229598	8.75	0.000	.8337591	1.318702
7	1.129581	.1290184	8.76	0.000	.8751619	1.383999
G_BJ_KL3						
2	.1606361	.0648742	2.48	0.014	.032707	.2885652
3	.2776268	.078023	3.56	0.000	.1237689	.4314846
_cons	8.950062	.3113562	28.75	0.000	8.336081	9.564043

```

154 . * R2 0.7100 alle significant (R2 was 0.5830)
155 .
156 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.G_BJ_KL3 i.DEEL_NED, vce(robust
> )

```

```

Linear regression
Number of obs      =      211
F(14, 196)         =      30.54
Prob > F            =      0.0000
R-squared           =      0.7346
Root MSE           =      .35983

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1252239	.0450022	2.78	0.006	.0364732	.2139745
LN_opbkv	.2399961	.0502734	4.77	0.000	.1408497	.3391424
H_BJ_KL7						
1	.9182414	.0854248	10.75	0.000	.7497718	1.086711
2	1.136734	.1024278	11.10	0.000	.9347316	1.338736
3	.9441993	.1083638	8.71	0.000	.7304906	1.157908
4	1.175812	.098215	11.97	0.000	.9821181	1.369506
5	1.112544	.1111583	10.01	0.000	.8933239	1.331764
6	1.167817	.1246389	9.37	0.000	.9220112	1.413622
7	1.155244	.1333745	8.66	0.000	.892211	1.418278
G_BJ_KL3						
2	.1504192	.0629249	2.39	0.018	.0263225	.2745159
3	.2401796	.0752712	3.19	0.002	.0917342	.3886251
DEEL_NED						
2	.1988317	.0748052	2.66	0.009	.0513053	.3463581
3	.353252	.0830171	4.26	0.000	.1895305	.5169735
4	.1317239	.0620883	2.12	0.035	.0092769	.2541708
_cons	8.647135	.3232716	26.75	0.000	8.009598	9.284672

```

157 . * R2 0.7346 alle significant (R2 was 0.6092; 2 niet significant)
158 .
159 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.G_BJ_KL3 LN_Adam, vce(robust)

```

```

Linear regression
Number of obs      =      211
F(12, 198)         =      38.01
Prob > F            =      0.0000
R-squared           =      0.7489
Root MSE           =      .34821

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1125436	.047172	2.39	0.018	.0195197	.2055676
LN_opbkv	.2486562	.0506562	4.91	0.000	.1487614	.348551
H_BJ_KL7						
1	.8770617	.0812956	10.79	0.000	.7167454	1.037378
2	1.109312	.096092	11.54	0.000	.9198168	1.298807
3	.9281171	.0968333	9.58	0.000	.7371601	1.119074
4	1.151474	.0953269	12.08	0.000	.9634872	1.33946
5	1.103054	.0992202	11.12	0.000	.9073901	1.298718
6	1.125542	.1396941	8.06	0.000	.8500628	1.401021
7	1.12291	.1358885	8.26	0.000	.8549355	1.390884
G_BJ_KL3						
2	.1310813	.0616972	2.12	0.035	.0094133	.2527493
3	.2456793	.0732367	3.35	0.001	.1012553	.3901034
LN_Adam	-.2921293	.0637611	-4.58	0.000	-.4178672	-.1663913
_cons	10.2189	.394219	25.92	0.000	9.441491	10.9963

```

160 . * R2 0.7489 alle significant (R2 was 0.6275; 1 niet significant)
161 .
162 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.G_BJ_KL3 LN_top10, vce(robust)

```

```

Linear regression
Number of obs      =      211
F(12, 198)         =      31.97
Prob > F            =      0.0000
R-squared           =      0.7228
Root MSE           =      .36584

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1217103	.0440881	2.76	0.006	.0347677	.2086529
LN_opbkv	.2463318	.0481611	5.11	0.000	.1513572	.3413064
H_BJ_KL7						
1	.8981693	.0864746	10.39	0.000	.7276399	1.068699
2	1.108587	.1012458	10.95	0.000	.9089286	1.308246
3	.9286474	.1054739	8.80	0.000	.7206509	1.136644
4	1.141975	.0971292	11.76	0.000	.9504346	1.333515
5	1.100829	.1101346	10.00	0.000	.8836414	1.318016
6	1.067074	.1246749	8.56	0.000	.8212129	1.312935
7	1.152851	.1374966	8.38	0.000	.8817057	1.423997
G_BJ_KL3						
2	.1457302	.0653261	2.23	0.027	.0169061	.2745543
3	.2611427	.0762204	3.43	0.001	.1108347	.4114506
LN_top10	-.1462861	.048613	-3.01	0.003	-.2421518	-.0504203
_cons	9.322869	.3206315	29.08	0.000	8.690578	9.95516

```

163 . * R2 0.7228 alle significant (R2 was 0.5986; 1 niet significant)
164 .
165 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.G_BJ_KL3 LN_inkm2, vce(robust)

```

```

Linear regression
Number of obs      =      211
F(12, 198)         =      34.95
Prob > F            =      0.0000
R-squared           =      0.7293
Root MSE           =      .36154

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1080026	.0454075	2.38	0.018	.0184582	.1975471
LN_opbkv	.2452235	.0491151	4.99	0.000	.1483676	.3420793
H_BJ_KL7						
1	.9263398	.0855982	10.82	0.000	.7575386	1.095141
2	1.184818	.1001583	11.83	0.000	.9873039	1.382332
3	.9696356	.1092955	8.87	0.000	.754103	1.185168
4	1.178373	.0950783	12.39	0.000	.990877	1.365869
5	1.152765	.1079939	10.67	0.000	.9397991	1.365731
6	1.117018	.1378781	8.10	0.000	.84512	1.388916
7	1.168675	.1358341	8.60	0.000	.900808	1.436543
G_BJ_KL3						
2	.1477867	.0632394	2.34	0.020	.0230775	.2724958
3	.3051214	.0779263	3.92	0.000	.1514494	.4587933
LN_inkm2	.1335527	.0366362	3.65	0.000	.0613055	.2057999
_cons	8.10602	.4076384	19.89	0.000	7.30215	8.90989

```

166 . * R2 0.7293 alle significant (R2 was 0.6000)
167 .
168 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.DEEL_NED INT_j_n, vce(robust)

```

```

Linear regression                                Number of obs    =        211
                                                F(13, 197)        =        34.02
                                                Prob > F           =        0.0000
                                                R-squared          =        0.7282
                                                Root MSE          =        .36324

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1967342	.0489914	4.02	0.000	.1001192	.2933492
LN_opbkv	.2202331	.0502912	4.38	0.000	.1210548	.3194114
H_BJ_KL7						
1	.8601388	.0820701	10.48	0.000	.6982901	1.021987
2	1.068081	.0979231	10.91	0.000	.8749689	1.261193
3	.8619954	.1014313	8.50	0.000	.661965	1.062026
4	1.144179	.0986975	11.59	0.000	.9495399	1.338818
5	1.148382	.111844	10.27	0.000	.9278167	1.368947
6	1.208517	.1110309	10.88	0.000	.989555	1.427478
7	1.209784	.1229891	9.84	0.000	.9672396	1.452328
DEEL_NED						
2	.2254754	.0785549	2.87	0.005	.070559	.3803919
3	.4014641	.0844834	4.75	0.000	.2348562	.5680721
4	.1985886	.0663373	2.99	0.003	.0677661	.3294111
INT_j_n	-.1440218	.0581488	-2.48	0.014	-.2586959	-.0293477
_cons	8.474299	.3278853	25.85	0.000	7.827684	9.120915

```

169 . * R2 0.7282 alle significant (R2 was 0.6002)
170 .
171 .
172 . *****
173 . **  BESTE REGRESSIES  **
174 . *****
175 .
176 . * 1 *
177 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.G_BJ_KL3 LN_Adam, vce(robust)

```

```

Linear regression                                Number of obs    =        211
                                                F(12, 198)        =        38.01
                                                Prob > F           =        0.0000
                                                R-squared          =        0.7489
                                                Root MSE          =        .34821

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1125436	.047172	2.39	0.018	.0195197	.2055676
LN_opbkv	.2486562	.0506562	4.91	0.000	.1487614	.348551
H_BJ_KL7						
1	.8770617	.0812956	10.79	0.000	.7167454	1.037378
2	1.109312	.096092	11.54	0.000	.9198168	1.298807
3	.9281171	.0968333	9.58	0.000	.7371601	1.119074
4	1.151474	.0953269	12.08	0.000	.9634872	1.33946
5	1.103054	.0992202	11.12	0.000	.9073901	1.298718
6	1.125542	.1396941	8.06	0.000	.8500628	1.401021
7	1.12291	.1358885	8.26	0.000	.8549355	1.390884
G_BJ_KL3						
2	.1310813	.0616972	2.12	0.035	.0094133	.2527493
3	.2456793	.0732367	3.35	0.001	.1012553	.3901034
LN_Adam	-.2921293	.0637611	-4.58	0.000	-.4178672	-.1663913
_cons	10.2189	.394219	25.92	0.000	9.441491	10.9963

```

178 . * R2 0.7489 alle significant (R2 was 0.6275; 1 niet significant)
179 .
180 . * 2 *
181 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.DEEL_NED LN_top10 INT_j_n, vce(
> robust)

```

```

Linear regression                                Number of obs    =      211
                                                F(14, 196)        =      32.73
                                                Prob > F          =      0.0000
                                                R-squared         =      0.7397
                                                Root MSE         =      .35635

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1810553	.0476345	3.80	0.000	.0871134	.2749971
LN_opbkv	.2462639	.0494641	4.98	0.000	.1487138	.343814
H_BJ_KL7						
1	.8761696	.0801566	10.93	0.000	.7180895	1.03425
2	1.046164	.0939412	11.14	0.000	.8608984	1.231429
3	.8623363	.0950867	9.07	0.000	.6748119	1.049861
4	1.127156	.0965286	11.68	0.000	.9367883	1.317524
5	1.115015	.1039951	10.72	0.000	.9099218	1.320108
6	1.18519	.1179594	10.05	0.000	.9525575	1.417822
7	1.192511	.1348428	8.84	0.000	.9265816	1.45844
DEEL_NED						
2	.2418738	.077573	3.12	0.002	.0888889	.3948586
3	.3395053	.0882404	3.85	0.000	.1654828	.5135278
4	.1458011	.0718995	2.03	0.044	.0040052	.2875971
LN_top10	-.1506254	.0565891	-2.66	0.008	-.2622271	-.0390238
INT_j_n	-.1544496	.0575738	-2.68	0.008	-.2679933	-.040906
_cons	8.955259	.3636601	24.63	0.000	8.23807	9.672448

```

182 . * R2 0.7397 alle significant (R2 was 0.6105)
183 .
184 . * 3 *
185 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.G_BJ_KL3 i.DEEL_NED, vce(robust
> )

```

```

Linear regression                                Number of obs    =      211
                                                F(14, 196)        =      30.54
                                                Prob > F          =      0.0000
                                                R-squared         =      0.7346
                                                Root MSE         =      .35983

```

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1252239	.0450022	2.78	0.006	.0364732	.2139745
LN_opbkv	.2399961	.0502734	4.77	0.000	.1408497	.3391424
H_BJ_KL7						
1	.9182414	.0854248	10.75	0.000	.7497718	1.086711
2	1.136734	.1024278	11.10	0.000	.9347316	1.338736
3	.9441993	.1083638	8.71	0.000	.7304906	1.157908
4	1.175812	.098215	11.97	0.000	.9821181	1.369506
5	1.112544	.1111583	10.01	0.000	.8933239	1.331764
6	1.167817	.1246389	9.37	0.000	.9220112	1.413622
7	1.155244	.1333745	8.66	0.000	.892211	1.418278
G_BJ_KL3						
2	.1504192	.0629249	2.39	0.018	.0263225	.2745159
3	.2401796	.0752712	3.19	0.002	.0917342	.3886251
DEEL_NED						

2	.1988317	.0748052	2.66	0.009	.0513053	.3463581
3	.353252	.0830171	4.26	0.000	.1895305	.5169735
4	.1317239	.0620883	2.12	0.035	.0092769	.2541708
_cons	8.647135	.3232716	26.75	0.000	8.009598	9.284672

186 . * R2 0.7346 alle significant (R2 was 0.6092; 2 niet significant)
 187 .
 188 . * 4 *
 189 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 LN_Adam, vce(robust)

Linear regression	Number of obs	=	211
	F(10, 200)	=	42.67
	Prob > F	=	0.0000
	R-squared	=	0.7333
	Root MSE	=	.35707

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1406448	.0463172	3.04	0.003	.049312	.2319776
LN_opbkv	.2463681	.050851	4.84	0.000	.1460951	.346641
H_BJ_KL7						
1	.8195731	.0805086	10.18	0.000	.6608185	.9783278
2	1.036492	.0942972	10.99	0.000	.8505478	1.222436
3	.8569205	.0905862	9.46	0.000	.6782938	1.035547
4	1.13637	.0956931	11.88	0.000	.9476736	1.325067
5	1.106913	.1010323	10.96	0.000	.9076877	1.306138
6	1.206393	.1368586	8.81	0.000	.9365219	1.476264
7	1.24594	.1223645	10.18	0.000	1.00465	1.48723
LN_Adam	-.3087684	.0661476	-4.67	0.000	-.4392046	-.1783322
_cons	10.24951	.4114284	24.91	0.000	9.438217	11.06081

190 . * R2 0.7333 alle significant (R2 was 0.6062)
 191 .
 192 . * 5 *
 193 . regress LN_ksb_h LN_opgeb LN_opbkv i.H_BJ_KL7 i.G_BJ_KL3 LN_inkm2, vce(robust)

Linear regression	Number of obs	=	211
	F(12, 198)	=	34.95
	Prob > F	=	0.0000
	R-squared	=	0.7293
	Root MSE	=	.36154

LN_ksb_h	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LN_opgeb	.1080026	.0454075	2.38	0.018	.0184582	.1975471
LN_opbkv	.2452235	.0491151	4.99	0.000	.1483676	.3420793
H_BJ_KL7						
1	.9263398	.0855982	10.82	0.000	.7575386	1.095141
2	1.184818	.1001583	11.83	0.000	.9873039	1.382332
3	.9696356	.1092955	8.87	0.000	.754103	1.185168
4	1.178373	.0950783	12.39	0.000	.990877	1.365869
5	1.152765	.1079939	10.67	0.000	.9397991	1.365731
6	1.117018	.1378781	8.10	0.000	.84512	1.388916
7	1.168675	.1358341	8.60	0.000	.900808	1.436543
G_BJ_KL3						
2	.1477867	.0632394	2.34	0.020	.0230775	.2724958
3	.3051214	.0779263	3.92	0.000	.1514494	.4587933
LN_inkm2	.1335527	.0366362	3.65	0.000	.0613055	.2057999
_cons	8.10602	.4076384	19.89	0.000	7.30215	8.90989

```

194 . * R2 0.7293 alle significant (R2 was 0.6000)
195 .
196 .
197 . * STORING ANALYSE:
198 . predict storing, resid
199 . swilk storing

```

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
storing	211	0.99314	1.073	0.162	0.43556

```

200 . qnorm storing
201 . kdensity storing, normal
202 .
203 . ** storing ligt redelijk langs de normaal verdeling
204 .
    end of do-file
205 .

```