



Centraal Bureau voor de Statistiek

Programma Impact ICT

Onderzoeksrapport nr. 1

Medegefinancierd door het Ministerie van Economische Zaken en Prima

Marktplaats en statistiek

Andries Kuipers, Remco Mocking, Rik Schürmann

Kennisgeving:

De in dit rapport weergegeven opvattingen zijn die van de auteurs en komen niet noodzakelijk overeen met het beleid van het Centraal Bureau voor de Statistiek.

Projectnummer: 210305/03

Datum: 15 april 2013

MARKTPLAATS EN STATISTIEK

Samenvatting

Het CBS heeft van Marktplaats.nl een dataset ontvangen met het aantal op Marktplaats geplaatste advertenties en bijbehorende kenmerken voor de jaren 2006 tot en met 2011. Geanticipeerd gebruik van deze dataset was (1) het schatten van de omzet in tweedehands goederen, (2) het verrijken van de gegevens met data van het CBS zodat nieuwe statistische informatie zou ontstaan en (3) het onderzoeken van een mogelijke samenhang met conjunctuurindicatoren. Daarnaast was het ook een proces van trial-and-error om te ontdekken wat het CBS met de data zou kunnen bij het maken van 'officiële' statistieken. Dit project was onderdeel van het programma Impact ICT en dan vooral het deel dat ingaat op de mogelijkheden om te komen tot een betere beschrijving van de interneteconomie.

Een eerste reflex was de enorme dataset te structureren en 'gaaf te maken', dat wil zeggen gegevens die niet plausibel leken, te verwijderen. Uiteindelijk is dit maar op zeer bescheiden schaal gedaan en is besloten de data vooral voor zichzelf te laten spreken. In hoofdstuk 2 zijn de van Marktplaats ontvangen data gepresenteerd. Dit roept een beeld op van een hele grote, hele heterogene en heel scheef verdeelde dataset waar maar moeilijk patronen in te ontdekken zijn. Statistische kengetallen zoals een gemiddelde waarde van een variabele zijn slechts zeer ten dele in staat een dergelijke dataset tot leven te brengen. In hoofdstuk 3 zijn de Marktplaatsdata op postcodeniveau gekoppeld aan data van het CBS. Dit leidt tot nieuwe statistische informatie over de regionale spreiding van het gebruik van Marktplaats en, op basis van de karakteristieken van de postcodegebieden, tot inzicht in de karakteristieken van de 'Marktplaatsgebruiker'. Een schatting van de omzet in tweedehands goederen is niet gemaakt. Er ontbraken te veel cruciale gegevens zodat dit met te veel arbitraire aannames gepaard zou moeten gaan. De Marktplaatsdata bleken wel een samenhang met de conjunctuur te vertonen. Zowel met de algemene conjunctuurcyclus als met het consumentenvertrouwen en de detailhandelsomzet. Hierbij hing het aanbod op Marktplaats (advertenties) positief samen met de conjunctuur en de vraag (biedingen) negatief.

Trefwoorden: Marktplaats, C2C-handel, internet, tweedehands goederen, online handel.

Inhoud

1.	Inleiding.....	- 1 -
2.	De dataset van Marktplaats	- 3 -
2.1	Datamanagement	- 3 -
2.2	Advertenties en hun kenmerken	- 4 -
2.3	Advertentiecategorieën.....	- 8 -
2.4	Ordering op basis van profielen.....	- 15 -
2.5	Zonnebrillen en Opels	- 21 -
2.6	Conclusies	- 23 -
3.	Data van Marktplaats: koppeling met regionale CBS-data	- 24 -
3.1	Koppeling op postcode.....	- 24 -
3.2	Advertenties en adverteerders naar regio	- 25 -
3.3	Karakteristieken Marktplaatsgebruiker	- 27 -
3.4	Adverteerder versus bidder.....	- 30 -
3.5	Regressiemodel	- 38 -
3.6	Conclusies	- 41 -
4.	Data van Marktplaats; omzet C2C-handel.....	- 42 -
4.1	Kan een schatting van de C2C-handel worden gemaakt?	- 42 -
5.	Data van Marktplaats; conjunctuur	- 46 -
5.1	Introduction and general analysis.....	- 48 -
5.2	Single category analysis: Consumer Electronics.....	- 57 -
6.	Wetenswaardigheden.....	- 64 -
7.	Slotbeschouwing	- 68 -

1. Inleiding

Marktplaats is een dochteronderneming van eBay en is bekend van de website Marktplaats.nl; een website voor het bij elkaar brengen van vraag en aanbod van (tweedehands) goederen. Het CBS heeft een dataset van Marktplaats ontvangen om zich een beeld te kunnen vormen van de aard van de gegevens die Marktplaats uit hoofde van zijn bedrijfsmodel zoal vastlegt. En, om zich een beeld te vormen in hoeverre deze gegevens gebruikt zouden kunnen worden bij het maken van ‘officiële’ statistieken. Dit project past binnen het programma Impact ICT en dan met name het onderdeel dat gaat over het beter en gedetailleerder beschrijven van de interneteconomie. Het programma wordt deels gefinancierd door het Ministerie van Economische Zaken en deels bemenst vanuit ontwikkelcapaciteit van het CBS zelf.

De van Marktplaats ontvangen dataset omvat alle advertenties en de daarbij behorende kenmerken over de periode 2006-2011. Een van de geanticipeerde gebruiksmogelijkheden van de data van Marktplaats was het maken van een schatting van de omzet in tweedehands goederen tussen particulieren (consumer to consumer ofwel C2C-handel). Op zich geen nieuw verschijnsel, maar door internet en intermediairs zoals Marktplaats en anderen heeft hier een enorme schaalvergroting plaatsgevonden. Daarnaast is gezocht naar mogelijkheden om gegevens van Marktplaats te koppelen aan gegevens van het CBS zodat nieuwe statistische informatie zou kunnen ontstaan. Het gaat hier uitdrukkelijk niet om het koppelen van gegevens op persoonsniveau. De van Marktplaats ontvangen dataset bevat namelijk geen tot op de individuele persoon herleidbare informatie. Een andere open vraag was in hoeverre de data van Marktplaats gerelateerd zouden kunnen worden aan conjunctuurindicatoren zoals die door het CBS worden samengesteld en eventueel vooruit zouden lopen op deze conjunctuurindicatoren. Ook is er een relatie met de duurzaamheid van de economie. Consumptiegoederen gaan namelijk langer mee door de toegenomen mogelijkheden om ze ‘door te geven’ van het ene aan het andere huishouden. Voor een deel was het echter ook een proces van trial-and-error om te ontdekken wat het CBS met de grote hoeveelheid gegevens van Marktplaats zou kunnen.

De volgende onderwerpen zijn in dit rapport uiteindelijk daadwerkelijk uitgewerkt. Hoofdstuk 2 dat vooraf gaat aan genoemde hoofdstukken, beoogt een beeld te geven van de data van Marktplaats. Wat zijn de patronen? Is het plausibel en te interpreteren? Zitten er curieuze zaken in? In hoofdstuk 3 zijn gegevens van Marktplaats op basis van postcode gekoppeld aan gegevens die het CBS op postcodeniveau beschikbaar heeft. Dit biedt mogelijkheden om gegevens van Marktplaats (a) zeer gedetailleerd naar regio weer te geven en (b) aard en omvang van het gebruik van Marktplaats en de Marktplaatsgebruiker te relateren aan de kenmerken van het postcodegebied. In hoofdstuk 4 wordt ingegaan op de haken-en-ogen om met de gegevens van Marktplaats een schatting te kunnen maken van de C2C-omzet via Marktplaats. Hoofdstuk 5 gaat in op de relatie tussen Marktplaatsdata en conjunctuur.

Werkwijze

Er gaat veel werk schuil achter de analyses die in dit rapport worden getoond. Wat het CBS in handen kreeg was een enorme hoeveelheid data op een detailniveau dat vreemd is aan veel van de statistische informatie die het CBS zelf verzamelt. Het ging hier in eerste aanleg om een bestand van 590 miljoen records met gegevens van op Marktplaats geplaatste advertenties per week, dag, uur en minuut,

onderverdeeld in honderden advertentiecategorieën deels voorzien van vraagprijzen en hoogste biedingen, bijna altijd voorzien van een postcode. Een eerste reflex was om de data ‘gaaf te maken’, dat wil zeggen zaken die raar zijn, niet plausibel lijken, uit de dataset te verwijderen. Uiteindelijk is dit slechts op zeer bescheiden schaal gedaan. Ten eerste was het in zekere zin onbegonnen werk en ten tweede werd het al gauw arbitrair. Er is besloten de data vooral zelf te laten spreken.

Vervolgens werd al snel duidelijk dat het om hele heterogene data ging. Advertenties liepen uiteen van zaken als babykleertjes en theezakjes tot caravans en zeilboten. Adverteerders varieerden van de incidentele adverteerder tot partijen die dagelijks duizenden advertenties plaatsten. Ondanks dit gegeven is besloten karakteristieken en patronen van de gehele dataset te laten zien zonder dat altijd duidelijk werd welke krachten hier achter schuil gingen. De idee was dat deze enorme dataset niet binnengehaald was om uiteindelijk in te zoomen op de markt voor tweehandsbabykleertjes, omdat dat toevallig een schaal is waarop dingen nog begrepen kunnen worden. Om illustratieve redenen en om beter te kunnen begrijpen wat er achter de massa aan data schuilgaat is er hier en daar desalniettemin ook ingezoomd op onderliggende categorieën.

Het voorgaande geeft ook aan dat dit rapport geen uitputtende weergave kan zijn van wat er allemaal met de data van Marktplaats te doen is. Wat er in dit rapport is opgenomen is maar een fractie van wat er aan data is geproduceerd en dit is weer slechts een fractie van wat er aan data geproduceerd zou kunnen worden. Desalniettemin beoogt dit rapport een eerste inzicht te geven in de aard van de data zoals die door Marktplaats wordt bijgehouden en daarnaast in de mogelijkheden deze data te verrijken met data van het CBS zodat nieuwe, snellere of gedetailleerdere informatie ontstaat.

2. De dataset van Marktplaats

De centrale eenheid in de data van Marktplaats is de individuele advertentie. Alle kenmerken in de bestanden hebben in principe betrekking op de advertentie. Voor een groot aantal variabelen is dat ook logisch, zoals vraagprijs, hoogste bod, advertentiecategorie en datum en tijd van plaatsing. Voor andere kenmerken is dit iets minder natuurlijk zoals de postcode; een advertentie heeft immers geen woonplaats. In een aantal analyses wordt met name de postcode gehanteerd ('doorvertaald') als een kenmerk van de bij de advertentie behorende adverteerder en de bij het hoogste bod behorende bieder. Het is echter goed om in gedachten te houden dat de 'statistische eenheid' in de bestanden van Marktplaats de advertentie is en de meeste kenmerken worden vastgelegd vanuit het perspectief van de advertentie. Dit hoofdstuk is vooral bedoeld om inzicht te krijgen in algemene karakteristieken van de enorme hoeveelheid data in de dataset. Dit wordt overwegend vanuit het perspectief van de advertenties gedaan. In hoofdstuk 3 wordt specifiek ingegaan op kenmerken van de adverteerders en bidders op Marktplaats (de Marktplaatsgebruiker).

2.1 Datamanagement

De oorspronkelijk van Marktplaats ontvangen bestanden bestonden te samen uit meer dan 590 miljoen advertenties waarbij per advertentie (record) een aantal kenmerken is opgenomen (zie appendix 1). Het uitgangspunt was om niet al te veel aan de data te veranderen om de data zo veel als mogelijk voor zichzelf te laten spreken. De aantallen waren immers zo groot dat de invloed van een enkele 'rare' advertentie nooit groot kon zijn. Al werkende is echter een aantal malen besloten om toch in te grijpen in de dataset omwille van de helderheid van de te maken analyses.

Ten aanzien van de advertenties is besloten om advertenties in bepaalde advertentiecategorieën niet mee te nemen, omdat het hier geen (tweedehands) goederen betrof maar onder andere diensten, bijvoorbeeld iemand die zichzelf aanbiedt als glazenwasser. Daarnaast was er sprake van contactadvertenties, vacatures en het aanbieden van vastgoed. De aard van deze advertentiecategorieën en het gebruik hiervoor van Marktplaats, wijkt te veel af van de oorspronkelijke invalshoek van dit onderzoek, namelijk het beschrijven van aard en omvang van de handel in (tweedehands) goederen via Marktplaats. Een dienst bijvoorbeeld kan – in tegenstelling tot goederen – een oneindig aantal malen opnieuw en als nieuw worden verkocht. Daarnaast zijn advertenties waarbij een vraagprijs van boven de 99 999,- euro was vermeld ook weggelaten. Deze advertenties hadden ondanks 'de grote aantallen', toch een sterke invloed op de gemiddelde vraagprijs. In de meeste gevallen leek het ook niet aannemelijk dat de aangeboden goederen tegen een dergelijke hoge vraagprijs verkocht zouden worden (bijvoorbeeld een fiets voor 1 000 000 euro).

Tabel 2.1.1 geeft een overzicht van het totale aantal oorspronkelijke advertenties in de van Marktplaats verkregen dataset en het aantal advertenties dat uiteindelijk in de verdere analyses is betrokken. De uitsluiting op basis van de advertentiecategorie lag door de jaren rond de één procent van het oorspronkelijke aantal advertenties. Uitsluiting op basis van een vraagprijs van honderdduizend euro of meer betrof 'slechts' een paar duizend advertenties per jaar. Qua aantallen is de ingreep in het oorspronkelijk van Marktplaats verkregen aantal advertenties dus bescheiden. In appendix 1 is in detail toegelicht welke advertenties uit de dataset zijn weggelaten en welke gevolgen dit had niet alleen voor het aantal advertenties, maar ook voor de andere kenmerken van de dataset (vraagprijs, hoogste bod e.d.).

2.1.1 Advertenties: van Marktplaatsdata naar analysebestand

	2006		2007		2008		2009		2010		2011	
Advertenties Marktplaats (mln)	44,6	100%	66,5	100%	74,8	100%	84,7	100%	91,0	100%	103,1	100%
Uitgesloten op basis van:												
advertentiecategorie (duizenden)	471,5	1,1%	740,4	1,1%	1 353,1	1,8%	1 176,5	1,4%	911,4	1,0%	801,5	0,8%
vraagprijs > € 99.999,- (duizenden)	2,0	0,0%	2,5	0,0%	3,4	0,0%	3,8	0,0%	4,0	0,0%	4,6	0,0%
Advertenties analysebestand (mln)	44,1	98,9%	65,8	98,9%	73,4	98,2%	83,5	98,6%	90,1	99,0%	102,3	99,2%

Bron: CBS (Data: Marktplaats).

2.2 Advertenties en hun kenmerken

Tabel 2.2.1 geeft naast een overzicht van het totale aantal advertenties in het analysebestand, ook de gemiddelde waarde van een aantal kernvariabelen in de dataset. Het betreft hier gemiddelden van een enigszins opgeschoonde (zie paragraaf 2.1), maar nog immer buitengewoon heterogene dataset. Het aantal advertenties heeft namelijk betrekking op enerzijds zaken als postzegels en theezakjes en anderzijds op auto's en zeilboten. Zowel qua aantal advertenties als qua prijs ligt dit soort categorieën ver uit elkaar. Om de dataset 'te leren kennen' is toch een aantal robuuste kenmerken van de totale dataset samengesteld.

Het eerste dat opvalt is de substantiële toename in het gebruik van Marktplaats. Het aantal geplaatste advertenties is toegenomen van 44 miljoen in 2006 naar ruim 102 miljoen in 2011. Voor een aantal analyses is dit een probleem omdat deze autonome groei van het gebruik van Marktplaats algemene ontwikkelingen in de economie moeilijker zichtbaar maakt. Een voorbeeld hiervan is de mogelijke samenhang tussen de ontwikkeling van het aantal advertenties op Marktplaats en conjunctuurindicatoren zoals die door het CBS worden samengesteld. Om conjuncturele schommelingen in het gebruik van Marktplaats zichtbaar te maken moet op een of andere wijze gecorrigeerd worden voor de autonome groei van het aantal advertenties, uitsluitend voortvloeiende uit de toename van het internetgebruik in het algemeen en het gebruik van Marktplaats in het bijzonder (zie voor de samenhang tussen Marktplaatsdata en conjunctuur hoofdstuk 5).

Het jaarlijkse aantal unieke adverteerders dat een of meerdere advertenties op Marktplaats heeft geplaatst is in de periode 2006-2011 toegenomen van 2,5 tot 3,4 miljoen (+35 procent). Het jaarlijkse aantal geplaatste advertenties is in die periode meer dan verdubbeld. Dit komt tot uiting in een forse toename van het gemiddelde aantal advertenties per adverteerder; van 17 in 2006 tot 30 in 2011. De groei van het gebruik van Marktplaats uitgedrukt in het totale aantal geplaatste advertenties lijkt dus eerder veroorzaakt door een intensiever gebruik door (bestaande) adverteerders en in mindere mate door een toename van het aantal unieke adverteerders.

De gemiddelde vraagprijs van de aangeboden goederen varieert van 375 euro in 2010 en 2011 (laagste) tot 534 euro in 2008 (hoogste). Het gaat hier om een soort eenheidsprijs van alle advertenties waarbij een vraagprijs is vermeld. Aan deze gemiddelde prijs is dan ook geen heldere conclusie te verbinden. Deze prijs wordt beïnvloed door het achterliggende assortiment van advertentiecategorieën. Als er in het ene jaar meer auto's en zeilboten inzitten en het andere jaar relatief meer theezakjes en babykleertjes, dan heeft dit gevolgen voor de gemiddelde vraagprijs. Daarnaast is het aantal advertenties waarbij een vraagprijs vermeld is door de jaren heen niet stabiel, hoewel het nooit minder dan de helft van het aantal advertenties bedroeg. Een tweede maatstaf om de vraagprijs en de ontwikkeling hiervan in de tijd weer

te geven is de mediane vraagprijs. Dit is de vraagprijs waar de vraagprijs van de helft van de advertenties met een vraagprijs onder ligt (en de andere helft van de advertenties dus erboven). Dit is een nuttige aanvullende maatstaf voor verdelingen waarbij het gemiddelde niet in het midden ligt, maar zoals in het geval van Marktplaats wordt beïnvloed door een minderheid van advertenties met (veel) hogere vraagprijzen. Door de jaren bedroeg de mediane vraagprijs van de advertenties met een vraagprijs 25 euro en de laatste twee jaren iets minder. Deze mediane vraagprijs is betrekkelijk stabiel en fors lager dan de gemiddelde vraagprijs in de betreffende jaren.

2.2.1 Kenmerken analysebestand

	2006	2007	2008	2009	2010	2011
Aantal advertenties (mln)	44,1	65,8	73,4	83,5	90,1	102,3
Aantal adverteerders (mln)	2,5	2,9	3,0	3,2	3,3	3,4
Advertenties per adverteerder	17	23	25	26	27	30
Advertenties met vraagprijs > 0,1 €ct	51%	50%	51%	53%	55%	59%
Advertenties met een hoogste bod	39%	32%	27%	33%	24%	20%
Gemiddelde vraagprijs (euro)	466	438	534	423	375	375
Mediane vraagprijs (euro)	25	25	25	25	22	23
Gemiddelde hoogste bod (euro)	322	500	293	238	225	213
Mediane hoogste bod (euro)	20	20	20	18	20	20

Bron: CBS (Data: Marktplaats).

Een vergelijkbaar kenmerk is de gemiddelde waarde van het hoogste bod op een bepaalde advertentie. Ook hier geldt dat deze waarde wordt beïnvloed door (verschuivingen in) de achterliggende categorieën en door het aantal advertenties met een hoogste bod. Het aantal advertenties met een hoogste bod is kleiner dan het aantal advertenties met een vraagprijs. Daarnaast was het aantal advertenties met een hoogste bod in 2010 en 2011 beduidend lager dan in de voorgaande jaren. Ook voor het hoogste bod geldt dat de mediane waarde over de jaren heen stabiel is (20 euro) en net als bij de vraagprijs, vele malen lager is dan het gemiddelde hoogste bod.

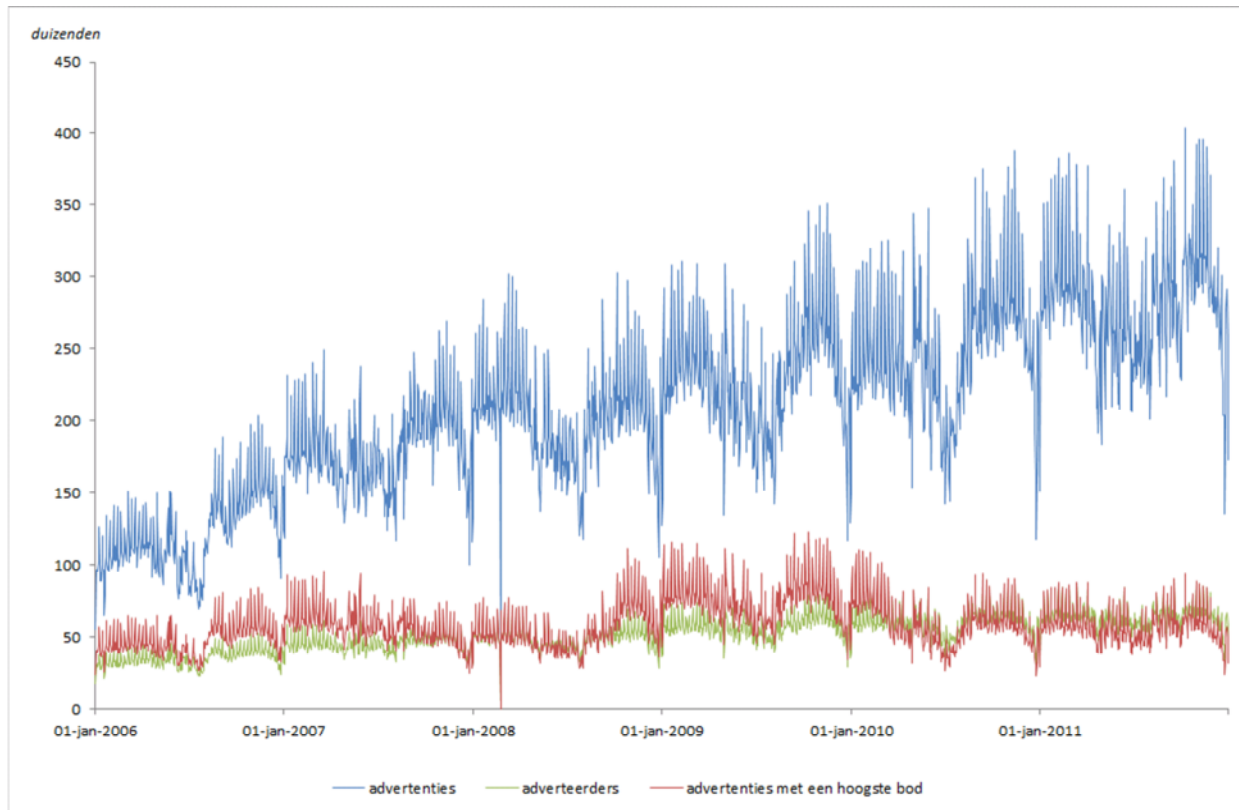
De onvergelijkbaarheid van de gemiddelde vraagprijs en het gemiddelde hoogste bod in de tijd maar ook met elkaar, wordt mooi geïllustreerd in 2007. Het gemiddelde hoogste bod van de advertenties waarbij een hoogste bod is uitgebracht ligt hoger dan de gemiddelde vraagprijs van de advertenties waarbij een vraagprijs vermeld is. De conclusie is hier niet per se dat bidders op grote schaal meer hebben geboden dan de vraagprijs, maar veel eerder dat het hier om twee grotendeels verschillende groepen van advertenties gaat. Immers, op advertenties met een vraagprijs zal vaak minder geboden (kunnen) worden. Advertenties waarbij voor het bepalen van de transactieprijs gekozen is voor de optie 'Bieden', kennen veelal geen vraagprijs. Een uitzondering vormt de categorie 'Bieden vanaf' waarbij de adverteerder een eerste (minimale) vraagprijs vermeld waarna potentiële kopers hier met een bod boven kunnen gaan zitten.

Dynamiek

In de tot nu toe gepresenteerde gegevens is de enorme hoeveelheid data van Marktplaats teruggebracht tot een paar statische kengetallen en kale aantallen per jaar. Figuur 2.2.2 geeft een indruk van de achterliggende dynamiek van het adverteren en bieden via Marktplaats. Het aantal advertenties, adverteerders en bidders zijn hier per dag weergegeven voor de gehele periode 2006-2011. Er zijn duidelijke (seizoens-)patronen zichtbaar zoals de 'dip' in de zomervakantie en de week van Kerst en

Oud en Nieuw. Ook komt de autonome groei van vooral het aantal advertenties op een indringender wijze over dan in de vorm van de kale gemiddelden zoals in de tabellen 2.1.1 en 2.2.1. In hoofdstuk 6 zal nader worden ingegaan op de patronen per jaar, per week, per etmaal en per categorie.

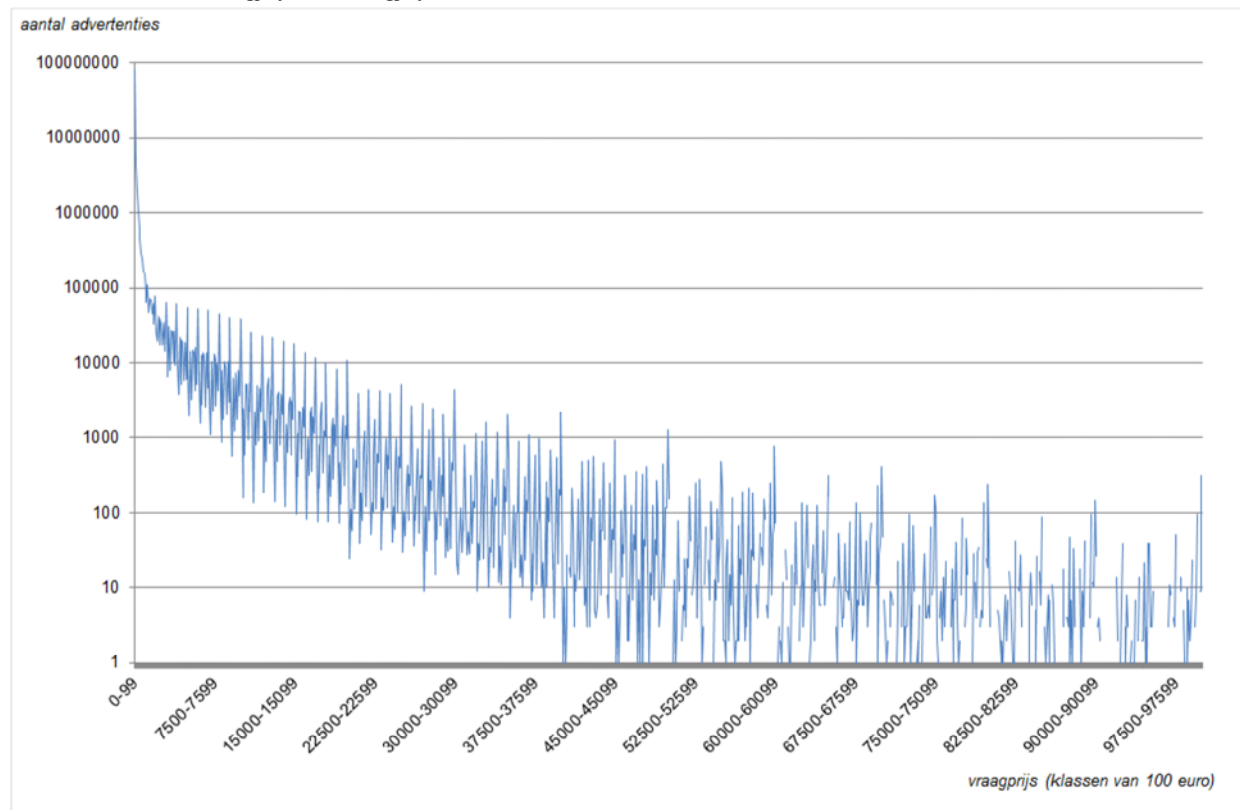
2.2.2 Advertenties, adverteerders en advertenties met een hoogste bod, per dag



Bron: CBS (Data: Marktplaats).

Ook voor twee andere kernvariabelen uit de dataset is voor 2010 de verdeling achter de in tabel 2.2.1 vermelde gemiddelden weergegeven. In figuur 2.2.3 zijn de advertenties met een vraagprijs verdeeld over de vraagprijs waarbij van begin tot eind van de verdeling een klasse-indeling is gekozen met een breedte van honderd euro. Dit leidt ertoe dat in de staart van de verdeling soms een score ontbreekt omdat in de betreffende klasse geen advertentie geplaatst was. De pieken en dalen in het aantal advertenties worden onder andere veroorzaakt doordat vooral bij de hogere bedragen er vaker ronde vraagprijzen worden ‘gekozen’ van vijfhonderd, duizend en tienduizend euro in plaats van 600, 1 100 of 10 100 euro. De figuur is niet bedoeld om keurig per klasse af te kunnen lezen hoeveel advertenties er zijn geplaatst. De figuur is bedoeld om de scheefheid van de verdeling te illustreren. In 2010 werd 87,5 procent van de advertenties met een vraagprijs aangeboden met een vraagprijs van minder dan honderd euro; 97,4 procent van de advertenties kende een vraagprijs van minder dan duizend euro. En, zoals uit tabel 2.2.1 al naar voren kwam lag de vraagprijs van de helft van de advertenties onder de 25 euro (de mediane vraagprijs).

2.2.3 Advertenties met vraagprijs naar vraagprijs, 2010



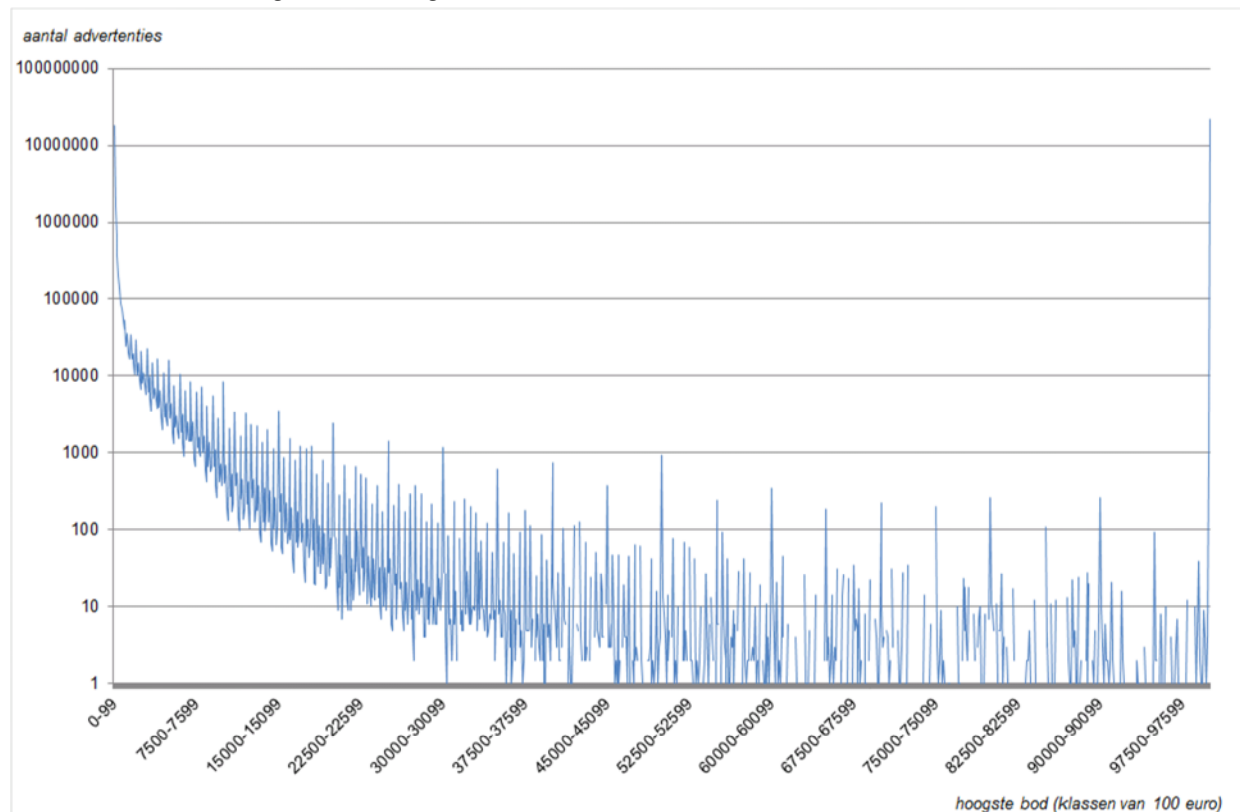
Bron: CBS (Data: Marktplaats).

Eenzelfde figuur is getekend voor het aantal advertenties met een hoogste bod in 2010. Hier is een vergelijkbare verdeling te zien van de advertenties met een hoogste bod over de klassen van dit hoogste bod. En ook hier zijn weer de pieken en dalen zichtbaar bij de meer gangbare respectievelijk de minder gangbare bedragen. Bij de advertenties met een hoogste bod geldt dat in 2010 bij 82,1 procent van de advertenties het hoogste bod onder de honderd euro lag; voor 96,7 procent van de advertenties lag het uiteindelijke hoogste bod onder de duizend euro. De mediaan van de verdeling van het hoogste bod lag in 2010 op 20 euro; de helft van het aantal hoogste biedingen lag dus onder de 20 euro.

Bij de biedingen zijn er relatief iets meer hogere biedingen dan bij de vraagprijs. Zo betrof 17,9 procent van de biedingen een bod van meer dan honderd euro; bij de vraagprijs gold dit maar voor 12,5 procent van de advertenties met een vraagprijs. Voor 3,3 procent van de biedingen betrof het een bod boven de duizend euro; ‘slechts’ 2,6 procent van de advertenties kende een vraagprijs van boven de duizend euro. In paragraaf 2.3 zal dan ook blijken dat het aantal advertenties met een hoogste bod in de duurdere advertentiecategorieën iets hoger ligt dan in de categorieën met goedkopere goederen (zie tabel 2.3.1). En, zoals al eerder opgemerkt betreft het ook hier een eerste ruwe ‘vergelijking’ tussen maar zeer ten dele dezelfde groep van advertenties.

Het doel van vooral de figuren 2.2.2., 2.2.3 en 2.2.4 is de beweging in de ‘zee van data’ te laten zien zonder dat direct duidelijk is door welke krachten deze beweging wordt veroorzaakt. Het is een saldo van wisselende aantallen advertenties, adverteerders, vraagprijzen en hoogste biedingen in de tijd en per categorie. Elke categorie heeft hierbij zijn eigen patronen en weegt de ene maand, week of dag, zwaarder mee dan de andere. In paragraaf 2.3 wordt geprobeerd verschillen in karakteristieken tussen de verschillende hoofdcategorieën in de dataset, enigszins zichtbaar te maken.

2.2.4 Advertenties met een hoogste bod naar hoogte bod, 2010



Bron: CBS (Data: Marktplaats).

2.3 Advertentiecategorieën

Marktplaats hanteert bij het rubriceren van de advertenties een bepaalde indeling in advertentiecategorieën. Deze bestaat uit een beperkt aantal hoofdcategorieën (30) en circa 2 200 onderliggende categorieën. Met name deze laatste indeling is door de jaren heen onderhevig aan wijzigingen. Zo zijn er categorieën samengevoegd, nieuwe bij gekomen en bestaande ‘bevroren’. Om de heterogeniteit van de complete dataset verder te illustreren zijn in tabel 2.3.1 de karakteristieken van de 30 hoofdcategorieën in 2010 weergegeven. In de categorie ‘Kinderen en Baby’s’ werden in dat jaar 15,5 miljoen advertenties geplaatst door 723 duizend verschillende adverteerders. Het is daarmee de hoofdcategorie met het hoogste aantal advertenties per jaar. Op het punt van het aantal verschillende adverteerders komt deze hoofdcategorie op de tweede plaats. De meeste unieke adverteerders zijn te vinden in de hoofdcategorie ‘Huis en Inrichting’. Intuïtief zijn dit soort verschillen wel te begrijpen. Onder de hoofdcategorie ‘Kinderen en Baby’s’ vallen goederen als babykleertjes, ledikantjes en commodes. Onder de hoofdcategorie ‘Huis en Inrichting’ vallen zaken als stoelen, eettafels en bankstellen. Dit zijn beide erkende categorieën van goederen op de markt voor tweedehands goederen. Babykleertjes en bijvoorbeeld ledikantjes zullen daarbij met een hogere frequentie worden aangeboden omdat de gebruiksduur van deze goederen korter is dan die van eettafels, stoelen en bankstellen. Daarnaast gaat het bij babykleding om grotere aantallen goederen die desgewenst in meerdere aparte advertenties kunnen worden aangeboden. Dit leidt dus tot meer advertenties. Daarentegen hebben bijna alle huishoudens meubels die ze vroeg of laat op Marktplaats aan zouden kunnen bieden en niet alle huishoudens hebben kleine kinderen. Voor de hoofdcategorie ‘Huis en Inrichting’ is het potentiële aantal adverteerders dus groter. Dit lijkt dan ook tot uiting te komen in het hogere aantal verschillende adverteerders.

2.3.1 Kenmerken analysebestand naar hoofdcategorie, 2010

Hoofdcategorie	Advertenties (duizend)	Adverteerders (duizend)	Advertenties per adverteerder	Vraagprijs (euro)		Hoogste bod (euro)		Advertenties met vraagprijs > 0.1 €ct	Advertenties met een hoogste bod
				gemiddeld	mediaan	gemiddeld	mediaan		
Antiek en Kunst	3 330	229	15	136	40	79	25	47%	19%
Audio, TV en Foto	1 988	483	4	144	50	151	45	60%	33%
Auto diversen	1 068	258	4	314	69	301	50	68%	27%
Auto-onderdelen	2 771	291	10	196	89	192	80	69%	18%
Auto's	1 842	335	6	7 257	4 900	4 433	2 500	91%	30%
Boeken	6 554	335	20	12	5	14	5	48%	13%
Caravans en Kamperen	448	155	3	3 284	80	1 476	51	60%	37%
Cd's en Dvd's	3 448	140	25	11	5	18	5	49%	16%
Computers en Software	1 721	348	5	134	50	180	50	67%	28%
Dieren en Toebehoren	2 549	387	7	162	35	169	28	57%	21%
Diversen	606	117	5	309	24	194	20	54%	19%
Doe-het-zelf en Verbouw	2 011	327	6	162	50	138	41	59%	23%
Fietsen en Brommers	2 741	571	5	202	95	221	55	69%	34%
Hobby en Vrije tijd	3 192	277	12	46	10	57	13	56%	20%
Huis en Inrichting	9 552	1 232	8	172	50	91	30	55%	25%
Kinderen en Baby's	15 549	723	22	26	10	26	9	47%	27%
Kleding Dames	9 425	430	22	28	15	26	10	53%	23%
Kleding Heren	1 819	217	8	38	25	40	15	59%	22%
Motoren	747	130	6	1 382	100	1 144	90	69%	24%
Muziek en Instrumenten	535	130	4	352	129	287	100	65%	37%
Postzegels en Munten	1 942	33	59	14	1	19	1	60%	25%
Sieraden, Tassen en Uiterlijk	2 675	236	11	82	18	89	18	64%	19%
Spelcomputers en Games	1 945	298	7	42	15	93	23	64%	28%
Sport en Fitness	1 524	404	4	152	35	71	25	63%	27%
Telecommunicatie	1 401	303	5	133	80	237	124	73%	37%
Tickets en Kaartjes	734	225	3	48	30	93	26	60%	27%
Tuin en Terras	1 624	353	5	184	50	127	35	59%	23%
Verzamelen	4 619	218	21	35	6	32	5	40%	21%
Watersport en Boten	568	119	5	1 431	75	1 002	75	60%	34%
Witgoed en Apparatuur	1 135	335	3	180	90	126	50	63%	36%
Totaal	90 065	9 639	9	375	22	225	20	55%	24%

N.B. Het aantal adverteerders en daardoor het gemiddelde aantal advertenties per adverteerder wijkt af van het totale aantal adverteerders omdat in deze tabel is gekeken naar het aantal adverteerders per hoofdcategorie. Er zijn echter adverteerders die in meerdere hoofdcategorieën actief zijn waardoor de som van het aantal unieke adverteerders per hoofdcategorie groter is dan het aantal unieke adverteerders voor alle hoofdcategorieën samen.

Bron: CBS (Data: Marktplaats).

Aan de andere kant van het spectrum bevinden zich hoofdcategorieën als 'Caravans en Kamperen', 'Muziek en Instrumenten' en 'Watersport en Boten'; goederen met veel minder advertenties en adverteerders per jaar. Een enigszins afwijkende hoofdcategorie is 'Postzegels en Munten'. In deze

categorie worden 1,9 miljoen advertenties geplaatst door slechts 33 duizend adverteerders. In deze categorie worden dan ook vaak individuele postzegels en munten aangeboden door een beperkt aantal adverteerders/verzamelaars die daar wel druk mee in de weer zijn.

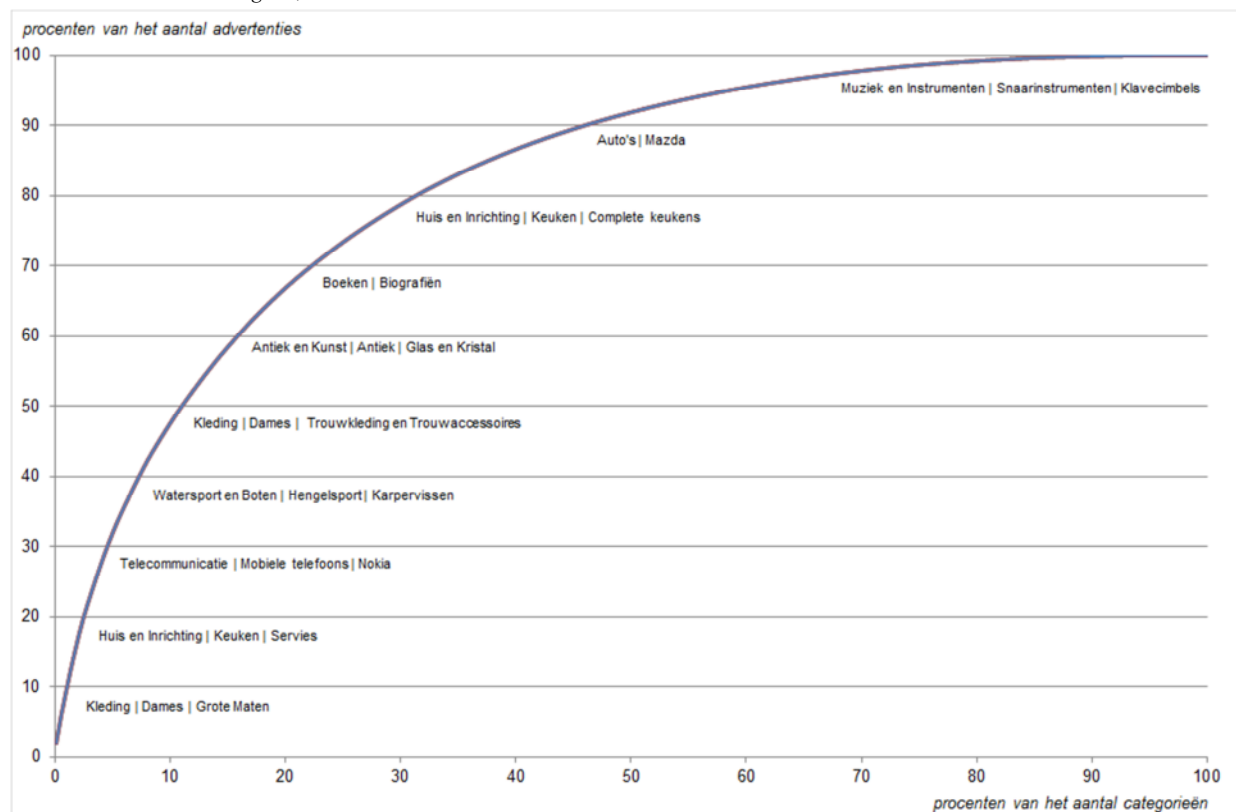
De gemiddelde vraagprijs voor de onderscheiden hoofdcategorieën lijkt in grote lijnen plausibel en is ook beter vergelijkbaar met de gemiddelde waarde van het hoogste bod voor de betreffende hoofdcategorie. Doordat ook de hoofdcategorieën nog betrekkelijk heterogeen van samenstelling zijn, komt het ook per hoofdcategorie echter nog voor dat het gemiddelde hoogste bod boven de gemiddelde vraagprijs ligt. Uiteindelijk worden er in de classificatie van Marktplaats circa 2 200 categorieën onderscheiden. En, zoals gezegd, de goederen in de hoofdcategorie ‘Kinderen en Baby’s’ variëren nog van rompertjes tot complete babykamers. Ook het aantal advertenties waarbij het betreffende goed met een vraagprijs wordt aangeboden varieert per hoofdcategorie. Bij ‘Auto’s’ is dit bijna regel (91 procent van de advertenties). Bij categorieën als ‘Boeken’, ‘Verzamelen’ en ‘Antiek en Kunst’ is het minder gangbaar om een uitgesproken vraagprijs op te nemen in de advertentie. Ook het aantal advertenties waarbij een hoogste bod tot stand komt loopt per hoofdcategorie uiteen van 13 procent (‘Boeken’) tot 37 procent (o.a. ‘Caravans en Kamperen’). Dit heeft uiteraard ook te maken met het aantal advertenties waarbij de mogelijkheid tot bieden wordt geboden. Dit lijkt bij de ene categorie gangbaarder dan bij de andere. Ook hier zijn voor de vraagprijs en het hoogste bod de mediane waarden bepaald. Deze liggen ook per hoofdcategorie stelselmatig fors onder respectievelijk de gemiddelde vraagprijs en het gemiddelde hoogste bod.

Verdeling over advertentiecategorieën

Om ook hier inzicht te geven in de achterliggende verdeling is in figuur 2.3.2 voor 2010 het aantal advertenties verdeeld over alle door Marktplaats onderscheiden categorieën. De advertentiecategorieën zijn hier op aflopend aantal advertenties gesorteerd. Rond ieder tiende percentiel van het aantal advertenties is de categorie vermeld die zich hier bevindt. Circa 70 procent van de advertenties werd in 2010 geplaatst in net iets meer dan de eerste 20 procent van het aantal advertentiecategorieën. De categorie ‘Biografieën’ bevond zich op de grens van dit percentiel. Het is aannemelijk dat deze indeling in categorieën vooral vanuit het perspectief van Marktplaats is vastgesteld bijvoorbeeld met het oogmerk om het zoeken naar bepaalde (tweedehands) goederen voor geïnteresseerden zo eenvoudig mogelijk te maken. Zo is kinderkleding soms onderverdeeld naar maat en auto’s en auto-onderdelen naar merk. Het is dus geen classificatie waarbij de goederen evenwichtig over de categorieën verdeeld zijn. En ook hier is het goed te bedenken dat deze verdeling per maand anders zal zijn dan per jaar. In de zomermaanden zullen zonnebrillen en cabrio’s beter vertegenwoordigd zijn en in de winter bijvoorbeeld schaatsen, met alle gevolgen van dien voor variabelen als de gemiddelde vraagprijs, het gemiddelde hoogste bod, maar ook het aantal advertenties met een vraagprijs en een hoogste bod.

De verdeling onderstreept ook dat wat in de ene categorie veel advertenties zijn, in de andere categorie maar als een gering aantal wordt gezien. Dit raakt aan de moeilijkheid om incidentele adverteerders te onderscheiden van ‘bedrijfsmatige’ adverteerders. Iemand die in de categorie ‘Postzegels en Munten’ 30 advertenties per jaar plaatst is een kleine adverteerder. Iemand die in de categorie ‘Witgoed en Apparatuur’ 30 advertenties per jaar plaatst is een grote adverteerder. Deze problematiek wordt verder uitgewerkt in paragraaf 2.4 waar advertenties en adverteerders op basis van verdelingskenmerken worden ingedeeld in profielen.

2.3.2 Advertenties naar categorie, 2010



Bron: CBS (Data: Marktplaats).

Om de gedachten verder te bepalen zijn in tabel 2.3.3 de tien categorieën met het hoogste en het laagste aantal advertenties in 2010 expliciet vermeld. De categorie met het hoogste aantal advertenties in 2010 was 'Kleding | Dames | Schoenen'; 1,7 miljoen of wel 1,9 procent van het totale aantal advertenties in 2010. De eerste tien categorieën waren samen goed voor 8,1 procent van het totale aantal advertenties in 2010. Deze categorieën worden gedomineerd door kinderkleding.

2.3.3 Tien categorieën met hoogste en laagste aantal advertenties, 2010

Bovenste 10 categorieën			Onderste 10 categorieën		
	aantal (duizend)			aantal	
Kleding Dames Schoenen	1 731	1,9%	Auto's Bentley	217	0,0%
Verzamelen Ansichtkaarten	765	0,8%	Snaarinstrumenten Mandolines	210	0,0%
Kinderkleding Maat 92	667	0,7%	Blaasinstrumenten Hobo's	203	0,0%
Kinderkleding Maat 104	660	0,7%	Caravans Alpenkreuzer	192	0,0%
Kinderkleding Maat 116	611	0,7%	Auto's Rolls-Royce	158	0,0%
Woonaccessoires Overige	611	0,7%	Caravans Biod	153	0,0%
Babykleding Maat 68	587	0,7%	Auto's Austin	132	0,0%
Merkkleding Truien en Vesten	565	0,6%	Auto's Aston Martin	125	0,0%
Babykleding Maat 86	553	0,6%	Caravans Solifer	109	0,0%
Kinderkleding Schoenen en Sokken	550	0,6%	Snaarinstrumenten Klavecimbels	104	0,0%

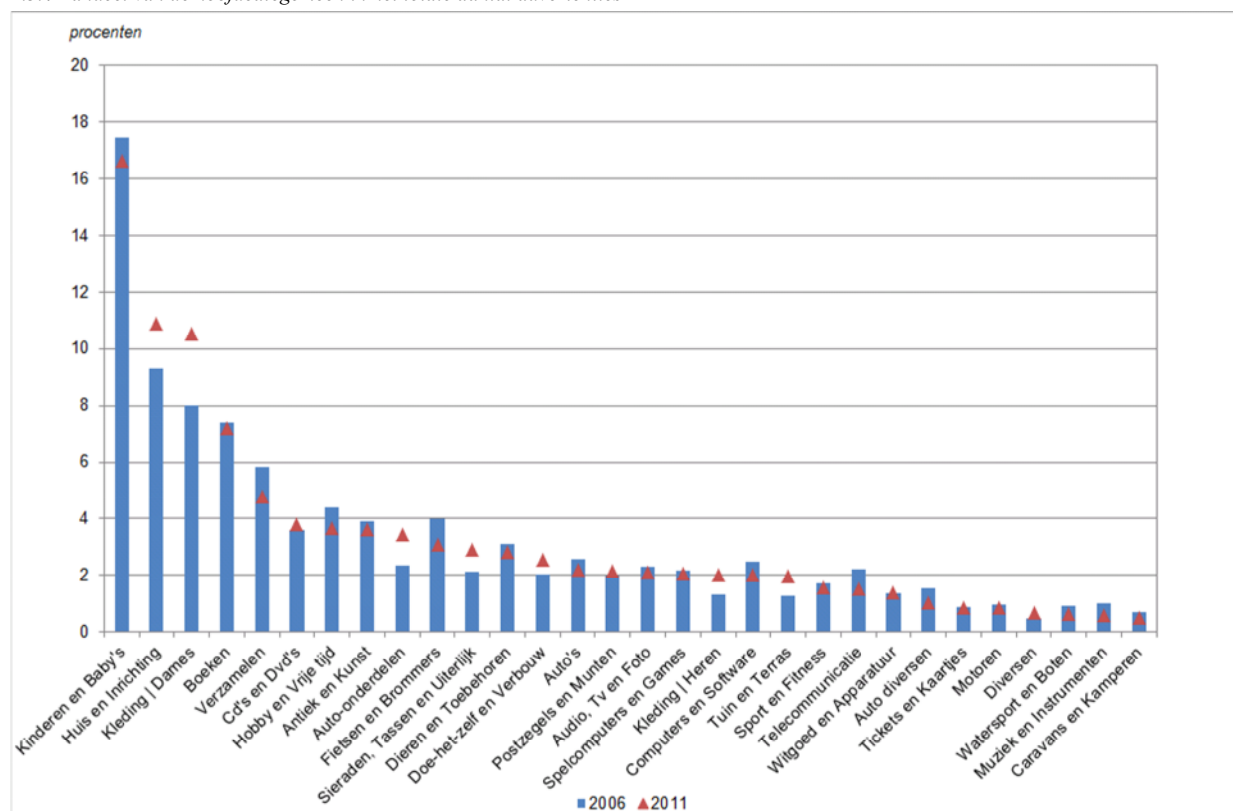
Bron: CBS (Data: Marktplaats).

De categorieën met het minste aantal advertenties in 2010 betreffen geheel andersoortige goederen. Bij deze onderste tien categorieën is nog een ondergrens van honderd advertenties gehanteerd om de exemplarische advertentie van die ene Dacia uit te sluiten. Deze categorieën worden gedomineerd door merkauto's en muziekinstrumenten en zijn samen goed voor 0,002 procent van het aantal advertenties in 2010.

Assortiment

Is het assortiment van de (tweedehands) goederen die op Marktplaats worden aangeboden in de hier beschouwde periode betrekkelijk stabiel gebleven of hebben er substantiële verschuivingen plaatsgevonden? In figuur 2.3.4 is het aandeel van de hoofdcategorieën in het totale aantal advertenties weergegeven voor 2006 en 2011. Zoals voor 2010 ook al uit tabel 2.3.1 naar voren kwam, worden ook in 2006 en 2011 in de hoofdcategorieën ‘Kinderen en Baby’s’ en ‘Huis en Inrichting’ de meeste advertenties geplaatst en de minste in de hoofdcategorieën ‘Caravans en Kamperen’ en ‘Muziek en Instrumenten’. Zoals uit figuur 2.3.4 blijkt is de volgorde van de categorieën praktisch ongewijzigd gebleven tussen 2006 en 2011. Dit suggereert dat de verdeling over de hoofdcategorieën betrekkelijk stabiel is.

2.3.4 Aandeel van de hoofdcategorieën in het totale aantal advertenties



Bron: CBS (Data: Marktplaats).

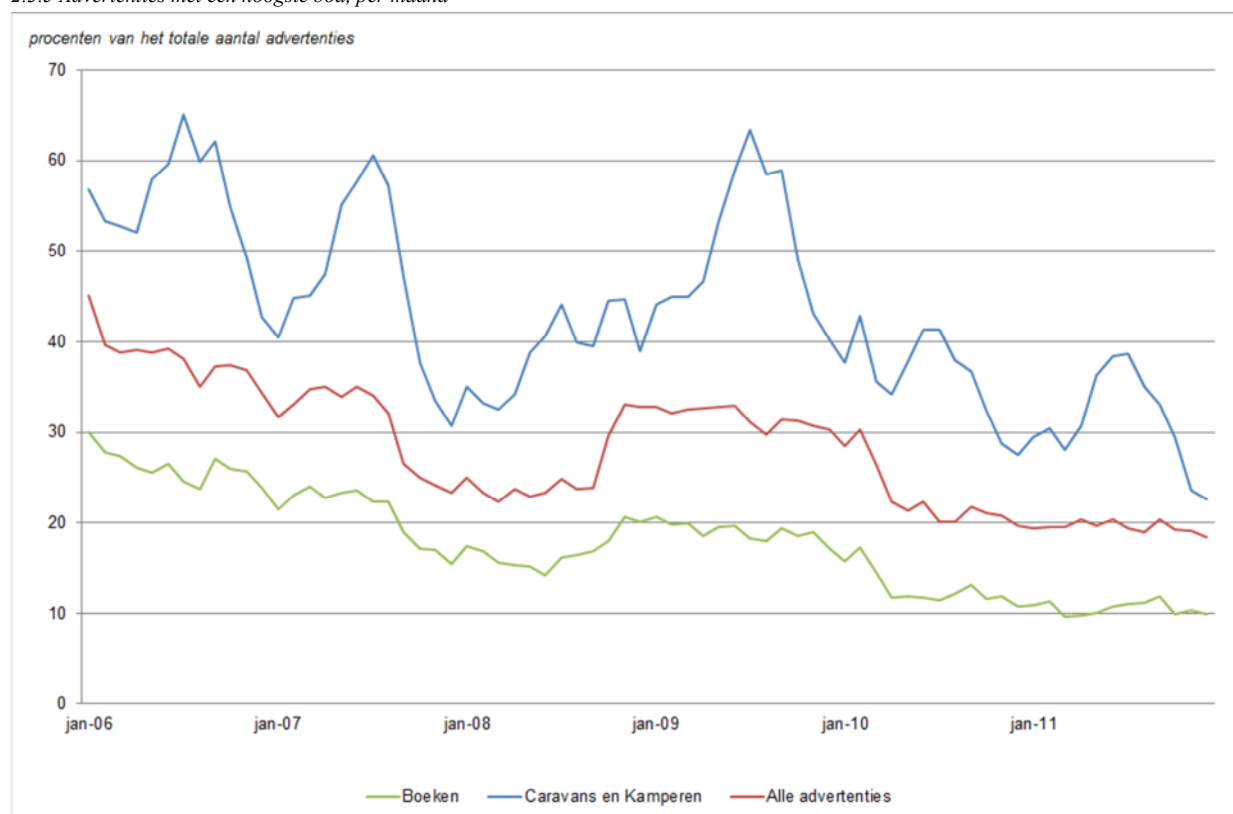
Het aandeel van de belangrijkste hoofdcategorieën is echter veel groter dan dat van de andere. De eerste drie hoofdcategorieën waren in 2006 goed voor 34,8 procent van het totale aantal advertenties; in 2011 was dit opgelopen tot 38,0 procent. Het lijkt dan ook moeilijk om ‘Kinderen en Baby’s’ van de eerste plaats te verdrijven. Toch is er wel sprake van verschuiving. Het aandeel van grotere hoofdcategorieën zoals ‘Kleding | Dames’ (+2,5 procentpunt) en ‘Huis en Inrichting’ (+1,6 procentpunt) nam toe. Ook kleinere hoofdcategorieën als ‘Tuin en Terras’ zagen hun aandeel relatief fors toenemen van 1,3 procent in 2006 tot 2,0 procent in 2011. Toch zijn dit geen grote verschuivingen die tot een totaal ander assortiment aan (tweedehands) goederen hebben geleid. Ook is het zo dat voor alle hoofdcategorieën geldt dat het absolute aantal advertenties in de periode 2006-2011 jaarlijks toenam; ook voor de hoofdcategorieën die hun aandeel zagen dalen. Ten slotte is het goed te bedenken dat binnen de hoofdcategorieën meer verandering in de achterliggende categorieën plaatsvindt. Zo bevat de categorie ‘Computers en Software’ in 2010 en 2011 in toenemende mate e-Readers en iPads en zijn dit ook aparte

categorieën geworden. In 2006 bestonden deze goederen nog praktisch niet en werden op Marktplaats dus ook niet aangeboden.

Bieden

In tabel 2.2.1. was een door de jaren heen afnemend aantal advertenties met een hoogste bod waarneembaar. Het betreft hier een afname van het aandeel advertenties met een hoogste bod. In figuur 2.3.5 is dit nogmaals geïllustreerd aan de hand van alle advertenties met een hoogste bod en de twee hoofdcategorieën met het hoogste (Caravan en Kamperen) en het laagste (Boeken) aandeel advertenties met een hoogste bod in 2011. In alle drie de gevallen is de dalende tendens zichtbaar. De patronen zijn echter zeer verschillend. Zo zijn de pieken in de zomermaanden bij de hoofdcategorie ‘Caravans en Kamperen’ duidelijk waarneembaar. De vraag (bieden) is in deze maanden dus groter dan in andere maanden van het jaar. Het bieden binnen de hoofdcategorie ‘Boeken’ kent daarentegen een gelijkmatiger verloop. Wat opvalt is dat 2009 wordt gekenmerkt door een hoger aandeel advertenties met een hoogste bod dan de jaren daarvoor en daarna. Hier is geen duidelijke verklaring voor gevonden. Wat figuur 2.3.5 beoogt te illustreren is dat de daling van het aandeel advertenties met een hoogste bod zoals waargenomen in tabel 2.2.1 en de ontwikkeling van het absolute aantal advertenties met een hoogste bod zoals weergegeven in figuur 2.2.2, het saldo zijn van de zeer uiteenlopende ontwikkelingen binnen de 2 200 onderscheiden categorieën van goederen. In absolute aantallen is het aantal advertenties met een hoogste bod overigens niet of nauwelijks gedaald zoals uit figuur 2.2.2 is op te maken.

2.3.5. *Advertenties met een hoogste bod, per maand*



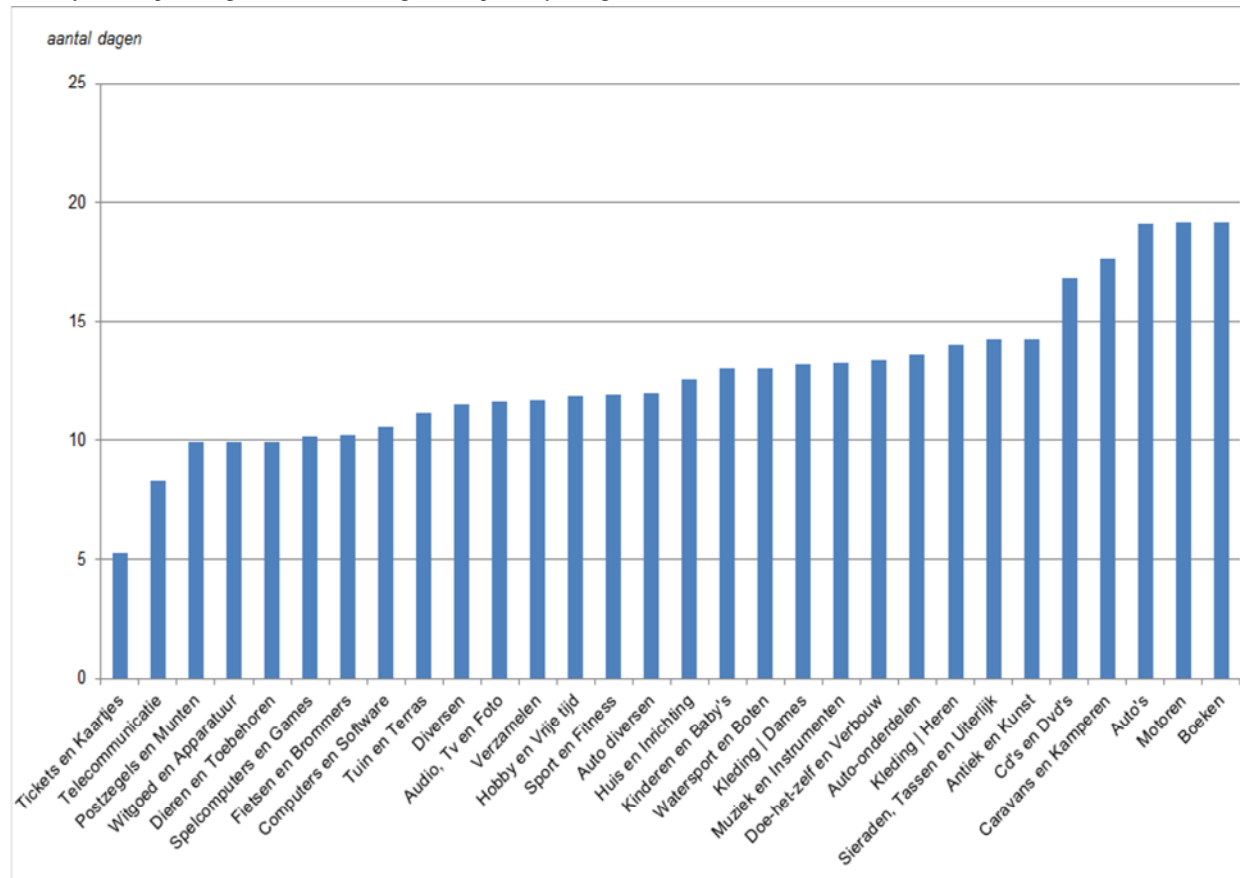
Bron: CBS (Data: Marktplaats).

Tijd tussen plaatsing advertentie en hoogste bod

Hoeveel tijd gaat er voorbij tussen de plaatsing van een advertentie en het moment dat er een hoogste bod wordt bereikt? In figuur 2.3.6 is dit per hoofdcategorie voor het jaar 2010 weergegeven. Bij de

hoofdcategorie ‘Tickets en Kaartjes’ is deze tijd het kortst. Dit is ook wel te begrijpen omdat hier sprake is van nog een extra ‘deadline’: namelijk de datum van het evenement zelf.

2.3.6 Tijd tussen plaatsing advertentie en hoogste bod, per hoofdcategorie, 2010



Bron: CBS (Data: Marktplaats).

Duurdere goederen en verzamelobjecten zoals ‘Boeken’ en ‘Antiek en Kunst’ lijken iets meer vertegenwoordigd in de tweede helft van de figuur: een groter aantal dagen tussen de plaatsing van de advertentie en de datum van het hoogste bod. Het aanbod is hier wellicht wat geringer en het product of de eisen zijn ook wat specifiekere. De potentiële koper is op zoek naar een bepaald merk auto met specifieke accessoires, een specifiek boek en een specifiek muziekinstrument; niet naar zomaar een auto, boek of muziekinstrument. Bij babykleertjes is dit minder het geval: een rompertje is een rompertje. Er is overigens een soort bovengrens voor de tijd dat een advertentie te vinden is op Marktplaats. Zonder aanvullende maatregelen verdwijnt een advertentie in de meeste categorieën na vier weken. Een hoogste bod kan of moet dus binnen vier weken gerealiseerd zijn. Voor alle advertenties met een hoogste bod gold in 2010 dat dit hoogste bod voor één op de drie advertenties binnen een dag was bereikt. Voor twee derde van de advertenties was het hoogste bod na tien dagen bereikt. Na vier weken – de standaard levensduur van veel advertenties – was voor 91 procent van de advertenties met een hoogste bod, dit bod bereikt. De andere negen procent van de advertenties incasseerden hun hoogste bod na genoemde vier weken omdat deze advertenties kennelijk langer mochten blijven staan bijvoorbeeld omdat de adverteerder hiervoor betaalde.

Views, vraagprijs en hoogste bod

Hoe vaak worden advertenties bekeken (*views*)? En heeft dit invloed op het hoogste bod? Zijn er advertenties die helemaal niet worden bekeken? Advertenties met een vraagprijs (49 keer) worden gemiddeld vaker bekeken dan advertenties zonder vraagprijs (36 keer). Advertenties waar geen bod op

is gedaan worden beduidend minder vaak bekeken (35 keer) dan advertenties waar wel een (hoogste) bod op is gedaan (83 keer). Kennelijk zijn sommige advertenties aantrekkelijker dan andere. Voor advertenties met een vraagprijs en een hoogste bod geldt dat de categorie waarbij het hoogste bod hoger is dan de vraagprijs gemiddeld ook het meeste aantal keren bekeken is. Het aantal views lijkt hiermee een indicatie voor de vraag mede gezien het ‘prijsoptimiserende’ effect ten opzichte van de initiële vraagprijs. Zo’n 7,7 miljoen advertenties of wel 8,6 procent is in 2010 helemaal niet bekeken.

2.3.7 Gemiddeld aantal keren dat een advertentie is bekeken, 2010

	Aantal keren bekeken	Aantal advertenties (mln)
<i>Vraagprijs</i>		
Advertenties met een vraagprijs < 0,1 €ct	36	40,3
Advertenties met een vraagprijs > 0,1 €ct	49	49,8
<i>Hoogste bod</i>		
Advertenties zonder bod	35	61,0
Advertenties met een hoogste bod	83	21,4
<i>Vraagprijs en hoogste bod</i>		
Advertenties met vraagprijs en een hoogste bod	116	8,5
hoogste bod is gelijk aan vraagprijs	61	1,9
hoogste bod is lager dan vraagprijs	124	4,6
hoogste bod is hoger dan vraagprijs	149	2,0
Niet bekeken advertenties	0	7,7
Totale aantal advertenties	43	90,1

Bron: CBS (Data: Marktplaats).

Ten slotte is ook nog een keer specifiek gekeken naar de verhouding tussen de vraagprijs van advertenties en het daarbij behorende hoogste bod en dan uitsluitend voor de groep advertenties met een vraagprijs én een hoogste bod. De vraag is of dit wél een consistent beeld geeft. Voor de 8,4 miljoen advertenties in 2010 met een vraagprijs en een hoogste bod is het gemiddelde hoogste bod (iets) lager dan de vraagprijs. Als de categorie ‘Bieden vanaf’ wordt weggelaten dan verandert het beeld enigszins. Ten eerste zijn de gemiddelde vraag- en biedprijzen hoger. Kennelijk wordt de optie ‘Bieden vanaf’ vaker gehanteerd in de mindere dure advertentiecategorieën. Ten tweede is het verschil tussen de vraagprijs en het hoogste bod wat groter geworden. Dit laatste is begrijpelijk. De categorie ‘Bieden vanaf’ begint met een soort minimum vraagprijs waarbij de verwachting is dat het uiteindelijke hoogste bod daarboven zal liggen. Toch zijn er nog immer (hoofd)categorieën waarbij ook na het weglaten van de categorie ‘Bieden vanaf’ het gemiddelde hoogste bod boven de gemiddelde vraagprijs lag. In incidentele gevallen het uiteraard goed mogelijk dat een bod hoger is dan de vraagprijs, maar gemiddeld genomen lijkt het een minder plausibele uitkomst. Dit kan veroorzaakt worden door een aantal extreme (‘niet serieuze biedingen’). Bij de advertenties zijn advertenties met een vraagprijs van meer dan 99 999,- euro weggelaten. Bij de biedingen is niet ingegrepen in de ontvangen data. Er blijven dus ongerijmdheden in de data zitten waar niet altijd direct een rationele verklaring voor te vinden is.

2.4 Ordening op basis van profielen

Een andere manier om ordening in de enorme hoeveelheid data aan te brengen is het detailleren van de data naar zogenaamde profielen. De data bevat namelijk enerzijds adverteerders die maar één of enkele advertenties per jaar plaatsen; de huis-tuin-en-keuken-adverteerder waar Marktplaats oorspronkelijk

voor bedoeld was of op zijn minst mee geassocieerd wordt. Anderzijds zijn er adverteerders die duizenden advertenties per dag plaatsen; de meer ‘bedrijfsmatige’ adverteerders van Marktplaats. Het is niet onwaarschijnlijk dat de karakteristieken van de advertenties en adverteerders voor deze groepen verschillend zijn.

Box 1: Advertentiekwartielen

Advertenties zijn voorzien van een kwartelaanduiding. Hiertoe is per jaar per adverteerder het aantal advertenties bepaald. Vervolgens zijn de adverteerders op grond van hun aantal advertenties van laag naar hoog gerangschikt en is het aantal advertenties verdeeld in kwartielen. De advertenties in het eerste kwartiel omvatten de eerste 25 procent van het totale aantal advertenties in een jaar geplaatst door adverteerders die het minst frequent adverteren op Marktplaats; de huis-tuin-en-keuken-advertenties. Het vierde advertentiekwartiel wordt gevormd door de 25 procent advertenties die geplaatst worden door de adverteerders die het frequentste een advertentie plaatsen op Marktplaats: de ‘bedrijfsmatige’ advertenties.

Het is hierbij belangrijk op te merken dat het bepalen van de advertentiekwartielen per categorie is gebeurd. Wat voor de ene categorie veel advertenties zijn is voor een andere categorie namelijk maar zeer gemiddeld. Een adverteerder die per jaar 30 postzegels aanbiedt is bijvoorbeeld niet erg actief binnen de categorie postzegels, terwijl een adverteerder die in één jaar 30 koelkasten aanbiedt wel een grote adverteerder is binnen de categorie koelkasten. Daarom zijn de advertentiekwartielen per categorie bepaald. Anders zou het vierde advertentiekwartiel in bovenstaand voorbeeld voornamelijk bestaan uit advertenties voor postzegels. Een bijkomstigheid van de gekozen benadering is dat de samenstelling van de advertentiekwartielen voor alle kwartielen gelijk is, namelijk 25 procent van alle onderscheiden categorieën.

Tabel 2.4.1 bevat dezelfde variabelen als tabel 2.2.1, maar dan uitgesplitst per advertentiekwartiel. In het eerste (vierde) kwartiel bevinden zich alle advertenties die geplaatst zijn door adverteerders met een laag (hoog) aantal advertenties binnen een willekeurige categorie. Elk kwartiel omvat 25 procent van het aantal geplaatste advertenties in het betreffende jaar. Dit volgt logischerwijze uit de wijze waarop de advertentiekwartielen zijn vastgesteld. Het aantal adverteerders in het eerste (vierde) kwartiel is het hoogst (laagst). Ook dit is logisch, omdat het eerste (vierde) kwartiel wordt gevormd door de minst (meest) frequent adverterende adverteerders. Daar zijn dus veel (weinig) adverteerders voor nodig om tot 25 procent van het totale aantal advertenties te komen. De verdeling lijkt door de jaren heen overigens iets schever geworden. In 2006 was 51 procent van de adverteerders nodig om te komen tot de ‘eerste’ 25 procent van de advertenties (het eerste kwartiel); in 2011 was dit opgelopen tot 57 procent van de adverteerders. In dezelfde periode nam het aantal adverteerders behorende bij de advertenties in het vierde kwartiel af van 2,5 procent in 2006 tot 1,7 procent in 2011. Het grote verschil tussen het gemiddelde aantal advertenties per adverteerder tussen het eerste en het vierde kwartiel vloeit voort uit het voorgaande en spreekt voor zich.

Interessanter zijn de verschillen in andere kenmerken tussen de onderscheiden advertentiekwartielen. Bij huis-tuin-en-keuken-advertenties op Marktplaats (eerste kwartiel) is het minst vaak sprake van een vraagprijs. Dit aandeel loopt per kwartiel op en dit patroon is over de jaren heen stabiel. Het omgekeerde geldt voor het aantal advertenties met een hoogste bod. Dit is juist het hoogste voor de advertenties in het eerste kwartiel en neemt per volgend kwartiel af. Ook dit patroon lijkt over de jaren heen stabiel. De gemiddelde vraagprijs van advertenties in het vierde kwartiel ligt voor alle jaren

beduidend hoger dan de vraagprijs in de andere drie kwartielen. Hoewel niet voor alle jaren, lijkt de gemiddelde vraagprijs van advertenties in het eerste kwartiel het laagst. Ook hier doet zich bij het gemiddelde hoogste bod het omgekeerde voor. Dit lijkt voor advertenties in het eerste kwartiel over het algemeen het hoogst te zijn.

2.4.1. Kenmerken van advertenties per advertentiekwartiel

	advertentiekwartiel	2006	2007	2008	2009	2010	2011
Aantal advertenties (mln)	1	11,0	16,4	18,4	20,9	22,5	25,6
Aantal advertenties (mln)	2	11,0	16,4	18,4	20,9	22,5	25,6
Aantal advertenties (mln)	3	11,0	16,4	18,4	20,9	22,5	25,6
Aantal advertenties (mln)	4	11,0	16,4	18,4	20,9	22,5	25,6
Aantal advertenties (mln)	Totaal	44,1	65,8	73,4	83,5	90,1	102,3
Aantal adverteerders (mln)	1	2,2	2,6	2,7	2,9	3,0	3,2
Aantal adverteerders (mln)	2	1,4	1,4	1,4	1,5	1,7	1,7
Aantal adverteerders (mln)	3	0,6	0,5	0,5	0,6	0,6	0,6
Aantal adverteerders (mln)	4	0,1	0,1	0,1	0,1	0,1	0,1
Aantal adverteerders (mln)	Totaal	4,2	4,5	4,7	5,1	5,5	5,5
Advertenties per adverteerder	1	5	6	7	7	7	8
Advertenties per adverteerder	2	8	12	13	14	13	15
Advertenties per adverteerder	3	19	33	38	37	35	43
Advertenties per adverteerder	4	106	234	260	224	213	266
Advertenties per adverteerder	Totaal	11	14	16	16	16	19
Advertenties met vraagprijs > 0,1 €ct	1	42%	40%	40%	43%	46%	48%
Advertenties met vraagprijs > 0,1 €ct	2	45%	45%	46%	48%	51%	54%
Advertenties met vraagprijs > 0,1 €ct	3	52%	53%	54%	55%	58%	61%
Advertenties met vraagprijs > 0,1 €ct	4	63%	63%	64%	65%	67%	71%
Advertenties met vraagprijs > 0,1 €ct	Totaal	51%	50%	51%	53%	55%	59%
Advertenties met een hoogste bod	1	52%	46%	41%	48%	36%	33%
Advertenties met een hoogste bod	2	47%	38%	32%	40%	29%	24%
Advertenties met een hoogste bod	3	37%	28%	22%	28%	20%	16%
Advertenties met een hoogste bod	4	21%	14%	11%	15%	11%	8%
Advertenties met een hoogste bod	Totaal	39%	32%	27%	33%	24%	20%
Gemiddelde vraagprijs (euro)	1	445	410	451	349	290	286
Gemiddelde vraagprijs (euro)	2	431	393	462	369	329	328
Gemiddelde vraagprijs (euro)	3	432	407	522	414	374	375
Gemiddelde vraagprijs (euro)	4	533	512	647	520	470	471
Gemiddelde vraagprijs (euro)	Totaal	466	438	534	423	375	375
Gemiddelde hoogste bod (euro)	1	309	525	311	270	262	248
Gemiddelde hoogste bod (euro)	2	322	509	300	229	207	184
Gemiddelde hoogste bod (euro)	3	329	475	252	200	187	178
Gemiddelde hoogste bod (euro)	4	343	438	287	228	221	226
Gemiddelde hoogste bod (euro)	Totaal	322	500	293	238	225	213

N.B. Het aantal adverteerders en daardoor het gemiddelde aantal advertenties per adverteerder wijkt af van het totale aantal adverteerders omdat in deze tabel is gekeken naar het aantal adverteerders per advertentiekwartiel. Er zijn adverteerders die in meerdere advertentiekwartielen voorkomen waardoor de som van het aantal unieke adverteerders per advertentiekwartiel groter is dan het aantal unieke adverteerders voor alle kwartielen samen.

Bron: CBS (Data: Marktplaats).

Tabel 2.4.2 laat zien voor welke categorieën het relatieve verschil tussen het aantal advertenties per adverteerder in het eerste en het vierde advertentiekwartiel het grootst respectievelijk het kleinst is. Dit geeft een idee van de scheefheid van de verdeling van advertenties over de adverteerders per categorie.

In het extreme geval plaatsen de adverteerders in het vierde kwartiel gemiddeld evenveel advertenties per jaar als de adverteerders in het eerste kwartiel. In dat geval is de intensiteit van gebruik van Marktplaats gelijk voor alle adverteerders binnen de categorie. Uit de tabel blijkt dat deze situatie het best benaderd wordt door de categorie ‘Caravans | Alpenkreuzer’. Het andere uiterste wordt gevormd door de categorie ‘Accessoires | Wioldoppen’. Hier plaatsen (enkele) adverteerders in het vierde kwartiel vierduizend keer zoveel advertenties als adverteerders in het eerste kwartiel. Dit duidt op een groot aantal in bulk geplaatste ‘bedrijfsmatige’ advertenties op Marktplaats binnen deze categorie. De advertenties binnen de categorie ‘Caravans | Alpenkreuzer’ lijken daarentegen voornamelijk uit huis-tuin-en-keuken advertenties te bestaan.

Het nadeel van de gekozen benadering is dat elke categorie een vierde kwartiel kent, ook als er zoals in het geval van ‘Caravans | Alpenkreuzer’, geen sprake lijkt te zijn van in grote aantallen geplaatste bedrijfsmatige advertenties. Er is echter niet gekozen voor een absoluut minimum aantal advertenties om tot een vierde kwartiel ‘toegelaten’ te worden. Aan de andere kant zou het ontbreken van dit verschil tot uiting moeten komen in vergelijkbare karakteristieken van beide kwartielen.

2.4.2 Advertenties per adverteerder per advertentiekwartiel, per categorie, 2010

	Advertentiekwartiel			
	1	2	3	4
	<i>aantal</i>			
<i>Categorieën met grootste verschil</i>				
Accessoires Wioldoppen	1,3	13,5	403,8	4 038,0
Slaapkamer Bedden	1,0	3,3	358,3	1 954,0
Automaten Gokkasten en Fruitautomaten	1,6	16,9	367,9	2 575,0
Slaapkamer Matrassen en Bedbodems	1,4	38,8	657,6	1 845,4
Slaapkamer Beddengoed	1,8	22,5	310,8	2 240,9
Auto's Boeken	3,7	130,9	856,3	4 403,6
Optische apparatuur Microscopen	1,2	9,4	146,1	1 315,0
Beren en Cherished Teddies	2,4	28,0	251,3	2 462,2
Speelgoed Knuffels	1,9	15,1	185,3	1 667,8
Accessoires Speakers en Versterkers	1,4	5,3	69,5	1 209,4
<i>Categorieën met kleinste verschil</i>				
Caravans Alpenkreuzer	1,0	1,0	1,0	1,3
Tv's Breedbeeld	1,0	1,0	1,0	1,4
Overige Motoren	1,0	1,0	1,0	1,5
Overige Caravans en Kamperen	1,0	1,0	1,0	1,5
Home Cinema-sets	1,0	1,0	1,0	1,8
Overige Muziek en Instrumenten	1,0	1,0	1,0	2,0
Caravans Biod	1,0	1,0	1,0	2,1
Snaarinstrumenten Klavecimbels	1,0	1,0	1,0	2,2
Open zeilboten	1,0	1,0	1,0	2,2
Brommers Crossbrommers	1,0	1,0	1,0	2,2

N.B. De categorieën zijn gesorteerd op het aantal keren dat het aantal advertenties per adverteerder in het vierde advertentiekwartiel groter is dan in het eerste kwartiel.

Bron: CBS (Data: Marktplaats).

Box 2: Adverteerderskwartielen

Naast de advertenties zijn ook de adverteerders ingedeeld in kwartielen. Adverteerders worden hierbij ingedeeld conform het hoogste advertentiekwartiel dat ze in een bepaalde categorie toegekend hebben gekregen. Het vierde adverteerderskwartiel bestaat dus uit adverteerders die in minimaal een categorie advertenties hebben geplaatst die tot het vierde advertentiekwartiel behoorden. Omgekeerd bestaat het eerste kwartiel uit adverteerders die in het betreffende jaar nooit advertenties hebben geplaatst anders dan in een eerste advertentiekwartiel. Omdat een adverteerder – in tegenstelling tot een advertentie – zo tot meerdere kwartielen kan behoren, is er voor gekozen de adverteerder het hoogste kwartiel toe te kennen dat aan hem of haar is toegekend binnen een bepaalde categorie. Een adverteerder kan dus op basis van het aantal advertenties binnen de categorie ‘Postzegels’ in het vierde kwartiel zijn ingedeeld, terwijl hij of zij in de categorie ‘Caravans’ in kwartiel een is ingedeeld. Deze adverteerder wordt gezien als een adverteerder in het vierde kwartiel inclusief al zijn of haar advertenties.

Er is niet voor gekozen om adverteerders in te delen op grond van hun totale aantal advertenties in een jaar. Omdat advertenties voor postzegels nu eenmaal in grotere aantallen worden geplaatst dan bijvoorbeeld advertenties voor auto's, zouden de adverteerders van postzegels zo immers oververtegenwoordigd zijn in het vierde adverteerderskwartiel.

Adverteerders in het eerste adverteerderskwartiel – de incidentele adverteerders – nemen door de jaren heen circa vijf procent van het aantal advertenties voor hun rekening. Deze groep adverteerders omvat 45 tot 50 procent van het totale aantal adverteerders. Dat dit eerste adverteerderskwartiel niet overeenkomt met 25 procent van het aantal adverteerders, zoals bij de advertenties wel het geval is, komt doordat bij het toekennen van adverteerderskwartielen gebruik is gemaakt van de advertentiekwartielen (zie box 2). Het gemiddelde aantal advertenties per jaar is voor adverteerders in het vierde kwartiel vele malen groter dan in het eerste adverteerderskwartiel. Dit verschil is groter dan bij de advertentiekwartielen (zie tabel 2.4.1) omdat ook advertenties die niet behoren tot een vierde advertentiekwartiel in het vierde adverteerderskwartiel terecht kunnen komen omdat de bijbehorende adverteerder nu eenmaal in het vierde kwartiel is ingedeeld. Uit de cijfers blijkt dat dit regelmatig voorkomt: het aantal advertenties in het vierde advertentiekwartiel in 2011 bedroeg 25,6 miljoen, het aantal advertenties behorende bij de adverteerders in het vierde adverteerderskwartiel bedroeg in datzelfde jaar 48,0 miljoen. Veel-adverteerders zorgen voor de plaatsing van een aanzienlijk deel van de advertenties; in 2011 werd 46,9 procent van de advertenties geplaatst door 2,8 procent van de adverteerders die tenminste in één categorie advertenties heeft geplaatst die tot het vierde advertentiekwartiel behoorden.

Adverteerders in het vierde kwartiel vermelden het vaakst een vraagprijs in hun advertentie. Adverteerders in het eerste kwartiel ontvangen daarentegen het vaakst een bod op hun advertentie. De gemiddelde vraagprijs van de incidentele adverteerder ligt in alle jaren hoger dan die van de andere adverteerderskwartielen. Ditzelfde geldt voor het gemiddelde hoogste bod. Dit wijkt af van het patroon zoals dat bij de advertentiekwartielen was te zien. Het aandeel duurdere goederen in het assortiment advertenties van het eerste adverteerderskwartiel is hoger dan in het eerste advertentiekwartiel. Kennelijk vindt er een verschuiving plaats van (goedkopere) advertenties die wel tot een eerste advertentiekwartiel behoren maar niet op het conto van een adverteerder in een eerste adverteerderskwartiel worden geschreven. Dit levert per saldo een groter aandeel duurdere goederen op in het eerste adverteerderskwartiel. Een verklaring waarom dit zo is ontbreekt op dit moment.

2.4.3 Kenmerken van advertenties per adverteerderskwartiel

	adverteerderskwartiel	2006	2007	2008	2009	2010	2011
Aantal advertenties (mln)	1	2,2	3,5	4,2	4,6	4,4	5,1
Aantal advertenties (mln)	2	6,4	10,9	12,5	13,8	14,6	17,3
Aantal advertenties (mln)	3	13,5	20,5	22,8	26,2	28,5	32,0
Aantal advertenties (mln)	4	22,1	30,7	33,9	38,9	42,5	48,0
Aantal advertenties (mln)	Totaal	44,1	65,8	73,4	83,5	90,1	102,3
Aantal adverteerders (mln)	1	1,1	1,3	1,4	1,6	1,5	1,6
Aantal adverteerders (mln)	2	0,9	1,0	1,0	1,1	1,1	1,2
Aantal adverteerders (mln)	3	0,5	0,4	0,4	0,5	0,6	0,5
Aantal adverteerders (mln)	4	0,1	0,1	0,1	0,1	0,1	0,1
Aantal adverteerders (mln)	Totaal	2,5	2,9	3,0	3,2	3,3	3,4
Advertenties per adverteerder	1	2	3	3	3	3	3
Advertenties per adverteerder	2	7	11	12	13	13	15
Advertenties per adverteerder	3	28	46	52	53	51	61
Advertenties per adverteerder	4	213	438	480	418	403	499
Totaal	Totaal	17	23	24	26	27	30
Advertenties met vraagprijs > 0,1 €ct	1	49%	46%	45%	50%	54%	55%
Advertenties met vraagprijs > 0,1 €ct	2	46%	44%	43%	46%	49%	51%
Advertenties met vraagprijs > 0,1 €ct	3	45%	44%	45%	47%	50%	54%
Advertenties met vraagprijs > 0,1 €ct	4	56%	57%	59%	59%	61%	65%
Advertenties met vraagprijs > 0,1 €ct	Totaal	51%	50%	51%	53%	55%	59%
Advertenties met een hoogste bod	1	57%	52%	47%	54%	41%	38%
Advertenties met een hoogste bod	2	55%	48%	42%	50%	37%	33%
Advertenties met een hoogste bod	3	49%	39%	33%	40%	29%	25%
Advertenties met een hoogste bod	4	27%	18%	14%	19%	14%	11%
Advertenties met een hoogste bod	Totaal	39%	32%	27%	33%	24%	20%
Gemiddelde vraagprijs (euro)	1	819	743	807	645	525	497
Gemiddelde vraagprijs (euro)	2	566	477	508	375	314	283
Gemiddelde vraagprijs (euro)	3	410	358	413	321	279	287
Gemiddelde vraagprijs (euro)	4	439	440	578	469	430	438
Gemiddelde vraagprijs (euro)	Totaal	466	438	534	423	375	375
Gemiddelde hoogste bod (euro)	1	595	935	555	517	518	479
Gemiddelde hoogste bod (euro)	2	407	608	351	282	260	223
Gemiddelde hoogste bod (euro)	3	288	429	222	174	161	152
Gemiddelde hoogste bod (euro)	4	255	361	234	194	194	195
Gemiddelde hoogste bod (euro)	Totaal	322	500	293	238	225	213

Bron: CBS (Data: Marktplaats).

Bieders naar adverteerderskwartiel

Zoals in hoofdstuk 3 zal blijken is het voor een bepaalde groep bidders mogelijk deze te koppelen aan het uniek identificerende nummer van een adverteerder (SELLER_ID). Dit geldt namelijk voor die bidders die bij het uitbrengen van het (hoogste) bod gebruik hebben gemaakt van hun account die ze ook bij het plaatsen van hun advertentie(s) hebben gehanteerd. Dit kan dus alleen voor bidders die ook geadvertiseerd hebben. Het is aldus mogelijk deze beperkte groep bidders te voorzien van adverteerderskenmerken. In tabel 2.4.4 zijn deze unieke bidders en hun hoogste biedingen ingedeeld naar adverteerderskwartielen. De achterliggende gedachte is om te zien welke groep adverteerders verantwoordelijk is voor het grootste deel van de hoogste biedingen en bijbehorende unieke bidders (vraagzijde) en vooral of dit gelijk is aan de verdeling van advertenties en adverteerders (aanbodzijde)

binnen deze groep. De analyse heeft maar een beperkte geldigheid omdat bij 82 procent van de hoogste biedingen en 42 procent van de unieke hoogste bidders geen SELLER_ID (en dus geen adverteerderskwartiel) gevonden kon worden.

Van de hoogste biedingen in 2010 kwam 23 procent van de incidentele adverteerders en 11 procent uit de groep van veel-adverteerders. Incidentele adverteerders brengen dus relatief vaker een hoogste bod uit dan veel-adverteerders. Dit lijkt op het eerste gezicht niet onlogisch, immers veel-adverteerders willen vooral verkopen. De verdeling van de bijbehorende unieke bidders laat echter een iets ander beeld zien. Vergeleken met de adverteerders zijn de uitbrengers van een hoogste bod relatief vaker te vinden in de hogere adverteerderskwartielen. Dit suggereert dat regelmatige adverteerders ook regelmatige bidders zijn; gebruikers die zowel als adverteerder en bidder meer dan gemiddeld actief zijn. Maar, nogmaals, vooral van de bidders is voor het overgrote deel geen adverteerderskenmerk gevonden omdat ze nooit een advertentie hebben geplaatst of daarbij een andere account hebben gebruikt.

2.4.4 Aantal biedingen en unieke bidders, naar adverteerderskwartiel, 2010

	Advertenties	Hoogste bod	Adverteerders	Unieke bidders
	<i>procenten</i>			
<i>Adverteerderskwartiel</i>				
1	5	23	46	36
2	16	34	34	37
3	32	32	17	22
4	47	11	3	4
Totaal	100	100	100	100
Onbekend	-	82	-	42

N.B. Betreft alleen bidders die ook een of meerder advertenties hebben geplaatst en bij hun bod dezelfde account hebben gebruikt als bij het plaatsen van een of meerder van deze advertenties.

Bron: CBS (Data: Marktplaats).

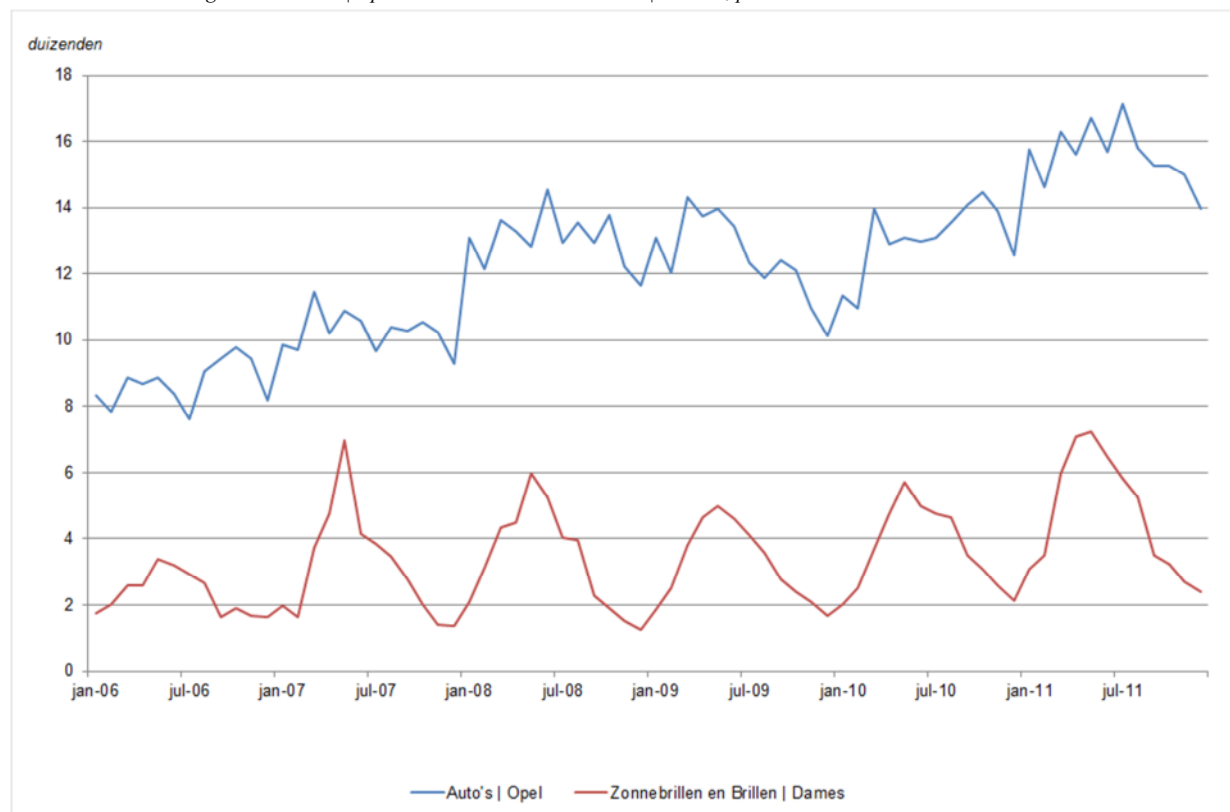
2.5 Zonnebrillen en Opels

In deze paragraaf worden de kenmerken van twee willekeurig gekozen categorieën goederen behandeld. In 2010 waren deze categorieën verantwoordelijk voor respectievelijk 0,17 procent (Opels) en 0,05 procent (Zonnebrillen) van het totale aantal advertenties. De voornaamste reden is om een concreet beeld te schetsen voor twee homogene categorieën opdat beter begrepen wordt wat zich in de totale dataset afspeelt. De beweging in de totale dataset is zoals gezegd het resultaat van de bewegingen en patronen in de honderden individuele categorieën.

Auto's | Opels

Het aantal advertenties van Opels is in de periode 2006-2011 toegenomen. Over het algemeen worden er in de zomermaanden en in de maand december minder Opels aangeboden en in het vroege voorjaar meer. Het jaar 2011 wijkt iets af van dit patroon. In 2010 was het aantal advertenties met een vraagprijs 91 procent. De gemiddelde vraagprijs van de Opels in deze advertenties bedroeg 6 244 euro. Het aantal advertenties met een hoogste bod in 2010 was 30 procent; het gemiddelde hoogste bod bedroeg 3 098 euro. Het gemiddelde aantal advertenties per adverteerder bedroeg vier. Het gemiddelde aantal

2.5.1 Advertenties categorieën 'Auto's | Opel' en 'Zonnebrillen en Brillen | Dames', per maand



Bron: CBS (Data: Marktplaats).

2.5.2 Kenmerken categorieën 'Auto's | Opel' en 'Zonnebrillen en Brillen | Dames', 2010

	Aantal advertenties (duizenden)	Aantal adverteerders (duizenden)	Advertenties per adverteerder	Advertenties met vraagprijs > 0,1 €ct	Advertenties met een hoogste bod	Gemiddelde vraagprijs (euro)	Gemiddelde hoogste bod (euro)
<i>Auto's Opel</i>							
advertentiekwartiel 1	39,2	34,6	1	77%	67%	6 963	3 179
advertentiekwartiel 2	39,2	4,6	9	91%	28%	4 831	2 784
advertentiekwartiel 3	39,2	0,8	47	97%	18%	5 473	3 109
advertentiekwartiel 4	39,2	0,2	166	100%	10%	7 743	3 421
Totaal	156,9	40,2	4	91%	30%	6 244	3 098
<i>Zonnebrillen en Brillen Dames</i>							
advertentiekwartiel 1	11,1	8,8	1	56%	37%	55	41
advertentiekwartiel 2	11,1	2,4	5	59%	29%	51	46
advertentiekwartiel 3	11,1	0,6	17	69%	17%	45	60
advertentiekwartiel 4	11,1	0,1	123	86%	6%	54	35
Totaal	44,4	12,0	4	68%	22%	52	46

Bron: CBS (Data: Marktplaats).

advertenties per adverteerder loopt op van één in het eerste advertentiekwartiel tot 166 in het vierde kwartiel. Dit duidt op een aantal veel-adverteerders binnen deze categorie.

Zonnebrillen en Brillen | Dames

Het aantal advertenties in deze categorie is beduidend kleiner dan in de categorie ‘Auto’s | Opel’ en is de loop van de jaren ook praktisch niet toegenomen. In de zomermaanden schommelt het aantal rond de zesduizend; in de wintermaanden rond de tweeduizend. Het seizoenpatroon is sterker en consistentere dan bij de Opels. Het aantal advertenties met een vraagprijs ligt met 68 procent ruim boven het gemiddelde van alle advertenties (24 procent) maar beduidend lager dan bij Opels. Het aantal advertenties met een hoogste bod in 2010 was 20 procent met een gemiddelde hoogste bod van 46 euro. Ook hier geldt dat het aantal advertenties per adverteerder in het vierde advertentiekwartiel zoveel malen groter is dan dat in het eerste dat sprake moet zijn van een aantal veel-adverteerders.

2.6 Conclusies

In dit hoofdstuk is geprobeerd ‘grip’ te krijgen op de dataset van Marktplaats. Het beeld dat hier uit naar voren komt is dat van een heel grote, heel heterogene en heel scheef verdeelde dataset waar eenvoudige maatstaven als een gemiddelde waarde van variabelen, niet afdoende zijn om de data tot leven te brengen, laat staan te doorgronden. Een gemiddelde vraagprijs in een advertentie is de gemiddelde vraagprijs van honderdduizenden babykleertjes en duizenden auto’s. Het gemiddelde aantal advertenties per adverteerder is het gemiddelde van de eenmalige adverteerder en de bedrijfsmatige adverteerder die dagelijks vele advertenties plaatst. Dit hoofdstuk schetst in eerste instantie voornamelijk een beeld van:

- Groei:

- 44 miljoen advertenties in 2006, 102 miljoen advertenties in 2011;
- 2,5 miljoen adverteerders in 2006; 3,4 miljoen in 2011.

en

- Verdeeldheid (cijfers 2010):

- 15,5 miljoen advertenties in ‘Kinderen en Baby’s; 535 duizend in de hoofdcategorie ‘Muziek en Instrumenten’;
- 3,2 procent van de adverteerders plaatst 47,2 procent van alle advertenties;
- een gemiddelde vraagprijs van 375 euro; een mediane vraagprijs van 22 euro;
- een gemiddelde hoogste bod van 225 euro; een mediane hoogste bod van 20 euro.

Toch zijn ook patronen zichtbaar (gemaakt):

- het aandeel advertenties met een vraagprijs schommelt tussen de 50 en de 60 procent;
- het aandeel advertenties met een bod is in de laatste jaren wat afgenomen;
- ‘Auto’s’ hebben bijna altijd een vraagprijs (2010: 91 procent);
- advertenties van incidentele adverteerders hebben het vaakst een (hoogste) bod.

3. Data van Marktplaats: koppeling met regionale CBS-data

Eén van de geanticiperde gebruiksmogelijkheden van de data van Marktplaats was het op regionaal niveau koppelen van deze gegevens aan gegevens van het CBS. Op deze wijze ontstaat nieuwe informatie over bijvoorbeeld kenmerken van regio's waar meer of minder gebruik wordt gemaakt van Marktplaats. Ook biedt het de mogelijkheid om de variabele afstand te betrekken in de analyses en onderzoeksvragen te beantwoorden als: Wat is de gemiddelde afstand tussen adverteerder en bidder voor de verschillende categorieën goederen? En wordt deze beïnvloed door de prijs van de betreffende goederen? Met andere woorden: wat is letterlijk het bereik van de markt voor tweedehands goederen? Naast deze wat verdergaande analyses worden er ook figuren gepresenteerd waar de gegevens van Marktplaats gewoonweg geregionaliseerd worden weergegeven. Zo ontstaat een beeld van mogelijke verschillen in aard en omvang van het gebruik van Marktplaats binnen Nederland.

In dit hoofdstuk wordt ook ingegaan op achtergrondkenmerken van gebruikers van Marktplaats (leeftijd, inkomen e.d.). Een gebruiker is hier een persoon die in een bepaalde tijdsperiode een of meerdere advertenties heeft geplaatst op Marktplaats. Bidders die in de analyse worden betrokken zijn een deelverzameling van deze adverteerders, namelijk die adverteerders die ook een bod hebben gedaan op een of meerder advertenties en hierbij dezelfde account hebben gebruikt als bij het plaatsen van hun advertentie(s). Bidders die nooit geadverteerd hebben op Marktplaats, of onder een andere account dan die gebruikt bij hun advertenties, vallen buiten de analyses omdat hier geen postcode van beschikbaar is. Deze analyses worden in dit hoofdstuk behandeld omdat het feitelijk niet de kenmerken zijn van de persoon die daadwerkelijk gebruik heeft gemaakt van Marktplaats. Deze gegevens ontbreken namelijk in de dataset van Marktplaats. De gebruiker van Marktplaats wordt gekarakteriseerd op basis van de kenmerken van het 6-positie postcodegebied dat bij de advertentie is vermeld (en op de adverteerder van toepassing wordt verklaard).

3.1 Koppeling op postcode

De variabele die gehanteerd is om de gegevens van Marktplaats en het CBS op regionaal niveau aan elkaar te koppelen is de postcode. Voor het overgrote deel van de advertenties en dus ook voor de bijbehorende adverteerder, is in het bestand van Marktplaats de postcode bekend. Dit komt onder andere omdat adverteerders er zelf belang bij hebben dat hun advertenties getoond worden bij regionale zoekopdrachten door potentiële kopers. In appendix 2 wordt in detail toegelicht hoe deze koppeling op postcode tot stand is gekomen en hoe ook bidders op Marktplaatsadvertenties aan een postcode zijn gekoppeld. Aan deze laatste groep wordt door Marktplaats namelijk geen postcode gevraagd. Deze groep gebruikers heeft er zelf – in tegenstelling tot de adverteerders – ook geen belang bij om hun postcode te verstrekken.

Ten behoeve van de analyses in dit hoofdstuk zijn de gegevens van Marktplaats gekoppeld aan de CBS database Kerncijfers Postcodegebieden 2008-2010.¹ Dit biedt zoals gezegd de mogelijkheid om de gebruikers van Marktplaats te beschrijven aan de hand van de karakteristieken van de postcodegebieden

¹ De database Kerncijfers Postcodegebieden 2008-2010 bevat informatie over het aantal inkomensontvangers, het fiscaal maandinkomen en de percentages zelfstandigen en uitkeringsontvangers in 6-positie postcodegebieden in december 2008. De database is aangevuld met demografische gegevens, de woningvoorraad en de gemiddelde woningwaarde per 6-positie postcodegebied in 2010. De gegevens in de databases hebben dus betrekking op twee verschillende peilmomenten.

van waaruit ze actief zijn. Omdat ook voor een bepaalde groep bidders de postcode kan worden vastgesteld kan er naast de afstand tussen adverteerder en bidder ook iets gezegd worden over verschillen in karakteristieken tussen adverteerder en bidder. Dit kan bijvoorbeeld het verschil in inkomen betreffen tussen de adverteerder en bidder voor de verschillende categorieën goederen. Op deze wijze kunnen er goederenstromen worden gedetecteerd die van ‘armere’ naar ‘rijkere’ postcodegebieden gaan of omgekeerd.

Dit hoofdstuk sluit af met een eenvoudige regressieanalyse om te zien of de kenmerken van de 6-positie postcodegebieden tot op zekere hoogte kunnen verklaren waarom in het ene gebied juist meer en in het andere gebied juist minder Marktplaatsgebruikers zijn.

3.2 Advertenties en adverteerders naar regio

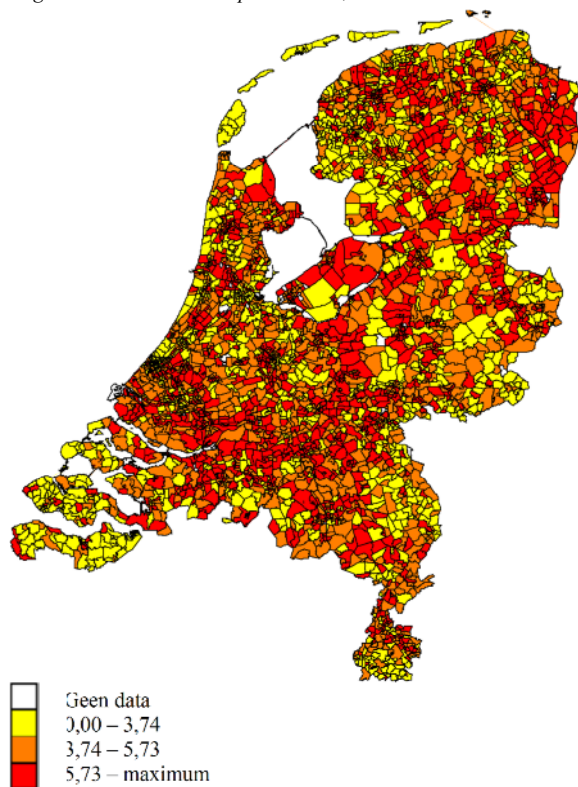
Het aantal advertenties op Marktplaats hangt af van het aantal unieke adverteerders en het aantal advertenties per adverteerder. Omdat de analyse wordt uitgevoerd op postcodeniveau is het aantal advertenties per inwoner en het aantal unieke adverteerders per inwoner berekend. Het nadeel van het aantal advertenties per inwoner is dat de bedrijfsmatige adverteerders op Marktplaats het beeld erg kunnen vertekenen door vanaf één of een beperkt aantal postcodegebieden, veel advertenties te plaatsen. Voor het aantal unieke adverteerders telt een bedrijfsmatige adverteerder daarentegen net zo zwaar mee als een huis-tuin-en-keuken adverteerder.

Figuur 3.2.1 laat de regionale verdeling van het aantal advertenties per inwoner over 4-positie postcodegebieden zien voor 2010. Het valt voornamelijk op dat de advertentiedichtheid laag is in een groot gedeelte van Zeeland en op de Waddeneilanden. Een verklaring hiervoor kan liggen in het feit dat een intermediair als Marktplaats meer toegevoegde waarde heeft wanneer de potentiële afzetmarkt groter is. In relatief perifere en (dus) minder dichtbevolkte gebieden is er minder vraag vanuit de nabije omgeving. Dit verkleint de kans dat een advertentie tot een transactie leidt en waarschijnlijk zal er daardoor ook minder geadverteerd worden. Gebieden waar nauwelijks ‘geel’ voorkomt zijn Zuid-Holland, grote delen van Utrecht en Noord-Brabant én het noordoosten van de provincies Drenthe en Groningen.

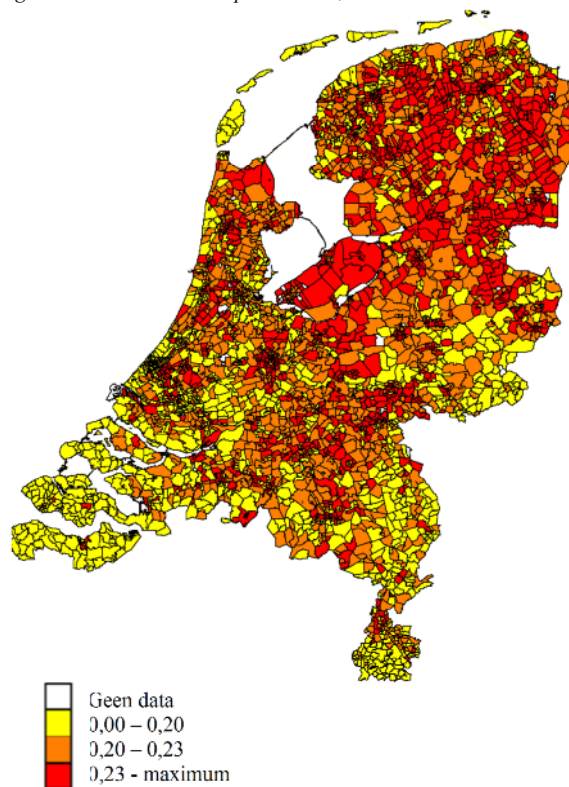
Figuur 3.2.2 laat de regionale verdeling van het aantal unieke adverteerders per inwoner per 4-positie postcodegebied zien voor 2010. Het valt opnieuw op dat de gebruikersdichtheid laag is in Zeeland en op de Waddeneilanden. Echter, hier komen ook grote delen van Limburg en het oosten van Gelderland naar voren als regio's met een minder intensief Marktplaatsgebruik. Over het algemeen wordt zichtbaar dat het gebruik van Marktplaats in het noordelijke gedeelte van Nederland wijder verspreid is dan in het zuiden van Nederland.

Een eerste conclusie is dat in 2010 vanuit alle 4-positie postcodegebieden in Nederland advertenties worden geplaatst op Marktplaats en zich in alle 4-positie postcodegebieden (dus) ook adverteerders bevinden. Daarnaast is er geen duidelijk patroon van grotere aaneengesloten gebieden die nauwelijks advertenties plaatsen. Overal in Nederland zijn ‘gele’ postcodegebieden te zien. Op het punt van het aantal adverteerders is er wel sprake van een zeker patroon. In het noorden van Nederland zijn er relatief veel meer adverteerders dan in het zuiden. Een directe verklaring hiervoor ontbreekt.

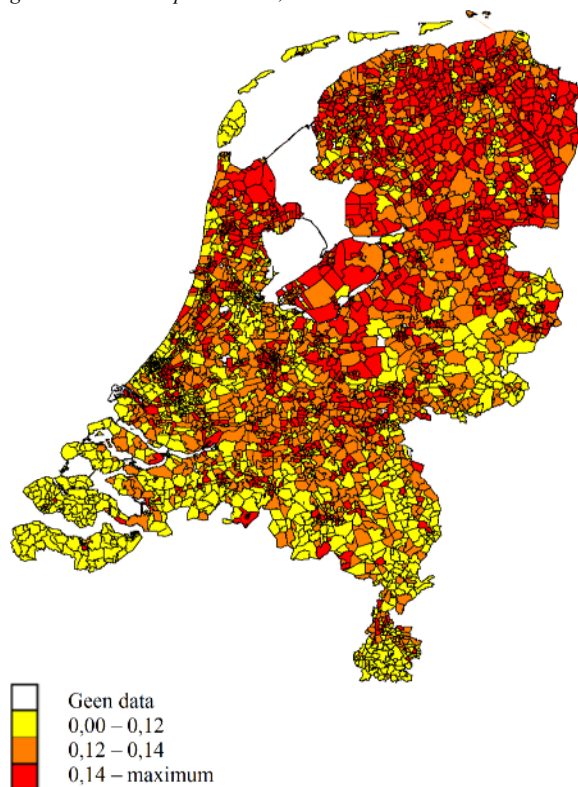
Figuur 3.2.1 Advertenties per inwoner, 2010



Figuur 3.2.2 Adverteerders per inwoner, 2010



Figuur 3.2.3 Bieders per inwoner, 2010



Bron: CBS (Data Marktplaats).

De tweedeling tussen Noord- en Zuid-Nederland wat betreft het aantal adverteerders op Marktplaats komt in nog sterkere mate terug bij eenzelfde weergave van het aantal bieders per inwoner naar regio. Noord-Nederland is hier overwegend rood en Zuid-Nederland is hier overwegend geel. Over het algemeen gaan achter het aantal adverteerders en bieders gewoon mensen van ‘vlees en bloed’ schuil. Een opvallende conclusie is dan ook dat het aantal mensen dat als adverteerder en bieder gebruik maakt van Marktplaats in Noord-Nederland relatief groter is dan in Zuid-Nederland. Ten overvloede zij nogmaals vermeld dat het in figuur 3.2.3 bieders betreft die ook een of meerdere advertenties hebben geplaatst op Marktplaats en dus niet iedereen die een bod heeft gedaan op Marktplaatsadvertenties. Daarnaast gaat het om de bieders van het uiteindelijke hoogste bod.

3.3 Karakteristieken Marktplaatsgebruiker

Door de 6-posities postcode behorende bij een advertentie te koppelen aan de regionale CBS-data kan iets gezegd worden over de karakteristieken van de adverteerders en advertenties op Marktplaats. De idee hierbij is dat de kenmerken van een 6-posities postcodegebied van toepassing zijn op de adverteerders en advertenties die vanuit het betreffende postcodegebied zijn geplaatst.² Beschikbare variabelen zijn onder andere het fiscale maandinkomen, de leeftijdsverdeling, het percentage niet-westerse allochtonen en het aantal adressen per vierkante kilometer (de omgevingsadressendichtheid (OAD)).³ Opgebouwd vanuit de 6-posities postcodegebieden wordt eerst de gemiddelde kenmerken voor de totale Nederlandse bevolking berekend (als referentie). Hierbij wordt het fiscale maandinkomen vermenigvuldigd met het aantal inkomensontvangers en gedeeld door het totale aantal inwoners om te komen tot een gemiddeld inkomen per inwoner. Het percentage niet-westerse-allochtonen, mannen, vrouwen en de leeftijdscategorieën wordt gewogen met het aantal inwoners en de OAD met het aantal huishoudens (als benadering voor het aantal adressen). Vanuit de 6-posities postcodegebieden worden zo door weging de kenmerken van de ‘gemiddelde Nederlander’ opgebouwd. Vervolgens worden de waarden voor deze variabelen opnieuw berekend waarbij als weging het aantal adverteerders wordt gebruikt in plaats van het totale aantal inwoners. Op vergelijkbare wijze als voor de ‘gemiddelde Nederlander’ worden zo de kenmerken van de ‘gemiddelde adverteerder’ opgebouwd.

Voorbeeld: Twee postcodegebieden met allebei 100 inwoners wegen even zwaar bij het bepalen van het gemiddelde percentage mannen in Nederland. Stel nu dat het eerste postcodegebied 30 adverteerders heeft, terwijl er in het tweede gebied 60 adverteerders op Marktplaats zijn. In dat geval weegt het tweede postcodegebied twee keer zo zwaar mee als het eerste bij de berekening van de kenmerken van de gemiddelde adverteerder op Marktplaats; in dit voorbeeld het percentage mannen.

De verschillen tussen de vanuit de postcodegebieden opgebouwde populatie Marktplaatsgebruikers en de totale Nederlandse bevolking zijn niet voor alle kenmerken even groot. Bij deze vergelijking is het aandeel van de verschillende groepen in de totale bevolking als referentie op 100 gesteld. De interpretatie van tabel 3.3.1 luidt als volgt: het vanuit de postcodegebieden opgebouwde aandeel 75-jarigen en ouder in de populatie Marktplaatsgebruikers is 76,6; dit is 23,4 procent lager dan het aandeel van deze groep in de totale bevolking. Postcodegebieden met relatief veel 75-jarigen en ouder zijn dus ondervertegenwoordigd onder de Marktplaatsgebruikers.

² Merk op dat het werken met gemiddelden mogelijke verschillen tussen Marktplaatsgebruikers kan onderdrukken. Bij een analyse op 4-cijferig postcodeniveau zijn Marktplaatsgebruikers nog nauwelijks te onderscheiden van de gemiddelde Nederlander.

³ De inkomensgegevens hebben betrekking op 2008; de demografische gegevens op 2010.

3.3.1 Karakteristieken Marktplaatsgebruikers, 2010

	Marktplaatsgebruikers	Marktplaatsgebruikers gewogen met aantal advertenties
	<i>index (totale bevolking = 100)</i>	
<i>Kenmerken populatie</i>		
0-14-jarigen	101,3	100,2
15-24-jarigen	102,8	99,3
25-44-jarigen	106,0	105,0
45-64-jarigen	99,1	100,9
65-74-jarigen	89,9	92,0
75-jarigen en ouder	76,6	81,4
Niet-westerse-allochtonen	87,2	79,6
Mannen	101,1	101,3
Vrouwen	98,9	98,7
Omgevingsadressendichtheid	98,4	90,6
Fiscaal maandinkomen per inwoner	100,6	98,5

N.B. Kenmerken per postcodegebied gewogen met totale aantal inwoners (totale bevolking), adverteerders op Marktplaats (Marktplaatsgebruikers) en met het aantal adverteerders gewogen met het aantal advertenties (Marktplaatsgebruikers gewogen met aantal advertenties).

Bron: CBS (Data: Marktplaats).

Het grootste verschil is de ondervertegenwoordiging van ouderen in de postcodegebieden met relatief veel Marktplaatsgebruikers (leeftijdscategorieën 65-74-jarigen en 75-jarigen en ouder). Ook het aandeel niet-westerse-allochtonen is in postcodegebieden met veel Marktplaatsgebruikers kleiner dan op grond van hun aandeel in de totale bevolking verwacht zou mogen worden. Postcodegebieden met veel 25-44-jarigen zijn daarentegen oververtegenwoordigd onder de Marktplaatsgebruikers.

Als de kenmerken van postcodegebieden van Marktplaatsgebruikers ook nog eens gewogen worden met het aantal advertenties – kenmerken van postcodegebieden van waaruit relatief veel advertenties worden geplaatst tellen dan zwaarder mee – dan verandert het beeld hier en daar. Het aandeel van de ouderen bijvoorbeeld neemt dan toe. In postcodegebieden met veel ouderen ligt het aantal advertenties per adverteerder blijkbaar iets hoger dan gemiddeld. Het aandeel van de postcodegebieden met veel niet-westerse-allochtonen neemt daarentegen verder af. Vanuit deze gebieden worden dus relatief weinig advertenties geplaatst. Een laatste opvallende verschuiving is de verdere afname van de omgevingsadressendichtheid in postcodegebieden van Marktplaatsgebruikers als gewogen wordt met het aantal advertenties. Er komen dus relatief veel advertenties uit gebieden met een lagere omgevingsadressendichtheid. Dit zou deels veroorzaakt kunnen worden door bedrijfsmatige adverteerders die buiten de bevolkingscentra zijn gevestigd.

In tabel 3.3.2 zijn de gewogen kenmerken van de postcodegebieden van Marktplaatsgebruikers naar adverteerderskwartiel opgenomen en vergeleken met de kenmerken van de postcodegebieden van de gemiddelde Marktplaatsgebruiker. Adverteerders in kwartiel één zijn de adverteerders die incidenteel adverteren op Marktplaats, adverteerders in kwartiel vier zijn adverteerders die frequent en in grote aantallen adverteren (zie box 2 in hoofdstuk 2).

Tabel 3.3.2 Karakteristieken Marktplaatsgebruikers, naar adverteerderskwartiel, 2010

	Adverteerderskwartiel			
	1	2	3	4
	<i>index</i>			
<i>Kenmerken populatie</i>				
0-14-jarigen	99,3	100,7	100,8	98,1
15-24-jarigen	101,3	99,2	98,5	98,6
25-44-jarigen	99,6	100,4	100,6	98,9
45-64-jarigen	100,4	99,7	99,4	101,1
65-74-jarigen	100,3	99,3	99,9	103,6
75 jaar en ouder	100,0	98,9	100,4	107,9
Niet-westerse-allochtonen	104,2	96,5	96,9	98,2
Mannen	100,0	100,0	100,0	100,2
Vrouwen	100,0	100,0	100,0	99,8
Omgevingsadressendichtheid	103,1	98,3	96,6	94,8
Fiscaal maandinkomen per inwoner	100,4	100,8	99,7	100,2

N.B. Kenmerken per postcodegebied gewogen met het aantal Marktplaatsgebruikers, totaal en per adverteerderskwartiel (zie ook tabel 3.3.1).

Bron: CBS (Data: Marktplaats).

Vergeleken met de kenmerken van de totale populatie Marktplaatsgebruikers zijn postcodegebieden met relatief veel niet-westerse-allochtonen onder de incidentele adverteerders oververtegenwoordigd. De incidentele adverteerders zitten dus relatief vaak in postcodegebieden met veel relatief veel niet-westerse-allochtonen. Ook komt de relatieve oververtegenwoordiging van postcodegebieden met veel ouderen in de groep veel-adverteerders weer naar voren (kwartiel 4). Postcodegebieden met veel ouderen zijn in de populatie Marktplaatsgebruikers dus ondervertegenwoordigd, maar vanuit diezelfde postcodegebieden wordt wel bovengemiddeld geadverteerd. Ook het verschijnsel dat veel-adverteerders vaker vanuit gebieden met een lagere omgevingsadressendichtheid adverteren, terwijl de incidentele adverteerders in gebieden met een hoger adressendichtheid wonen, is weer zichtbaar. Dit lijkt ook weer te duiden op de tweedeling in de data: enerzijds de huis-tuin-en-keuken-adverteerder van Marktplaats en anderzijds de bedrijfsmatige adverteerder.

3.4 Adverteerder versus bieder

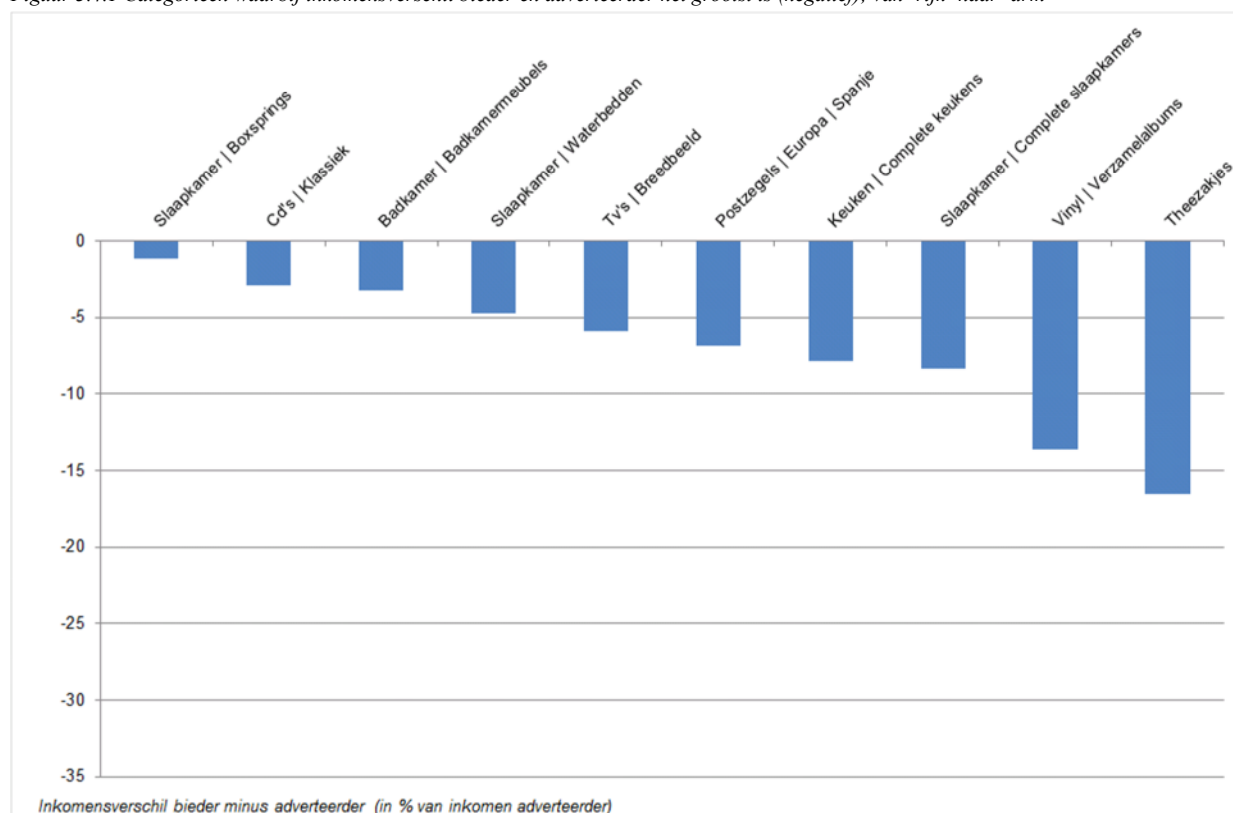
In 2010 was op ongeveer 20 procent van alle geplaatste advertenties een hoogste bod gedaan. Een bieder dient bij het uitbrengen van een bod een account aan te maken of een bestaande account te gebruiken. Een deel van de bidders plaats ook advertenties op Marktplaats en heeft uit hoofde daarvan reeds een of meerdere accounts. Een deel van deze bidders hanteert bij het bieden een van deze accounts. Zo kan een deel van de bidders gekoppeld worden aan de SELLER_ID en zo aan een postcode en vervolgens aan kenmerken van dat postcodegebied (zie ook appendix 2). Met deze gegevens kunnen dus karakteristieken van postcodegebieden van adverteerders én bidders worden vergeleken. In deze paragraaf wordt dit voor twee kenmerken uitgewerkt: inkomen en afstand.

Inkomen

Een invalshoek om mogelijke patronen in de data zichtbaar te maken is de vraag in hoeverre bepaalde goederen betrekkelijk structureel van ‘arm naar rijk’ lijken te gaan en andersom. Figuur 3.4.1 laat voor het jaar 2010 de tien categorieën zien waarvoor het relatieve inkomensverschil tussen de hoogste bieder en de adverteerder in negatieve zin, het grootst is. Het betreft hier dus goederen die van relatief rijke naar relatief arme postcodegebieden gaan. De goederencategorieën kunnen ruwweg in twee groepen worden ingedeeld. De eerste groep lijkt te bestaan uit weinig ‘prestigieuze’ verzamelgoederen

(theezakjes, postzegels). De tweede groep bestaat vooral uit gebruiksgoederen. Deze tweede groep wordt gedomineerd door goederen die te maken hebben met de inrichting van het huis. Een achterliggende verklaring zou kunnen zijn dat personen die relatief minder te besteden hebben eerder geneigd zijn om hun huis in te richten met tweedehands spullen. Het inrichten van een huis kan toch gezien worden als een noodzakelijk goed en via Marktplaats kan er op een eenvoudige manier bespaard worden op de inrichtingskosten. Het inkomensverschil tussen bidder en adverteerder wordt overigens al snel klein. Voor de categorie ‘Slaapkamer | Boxsprings’ gaat het nog maar om iets meer dan een procent.

Figuur 3.4.1 Categorieën waarbij inkomensverschil bidder en adverteerder het grootst is (negatief); van ‘rijk’ naar ‘arm’

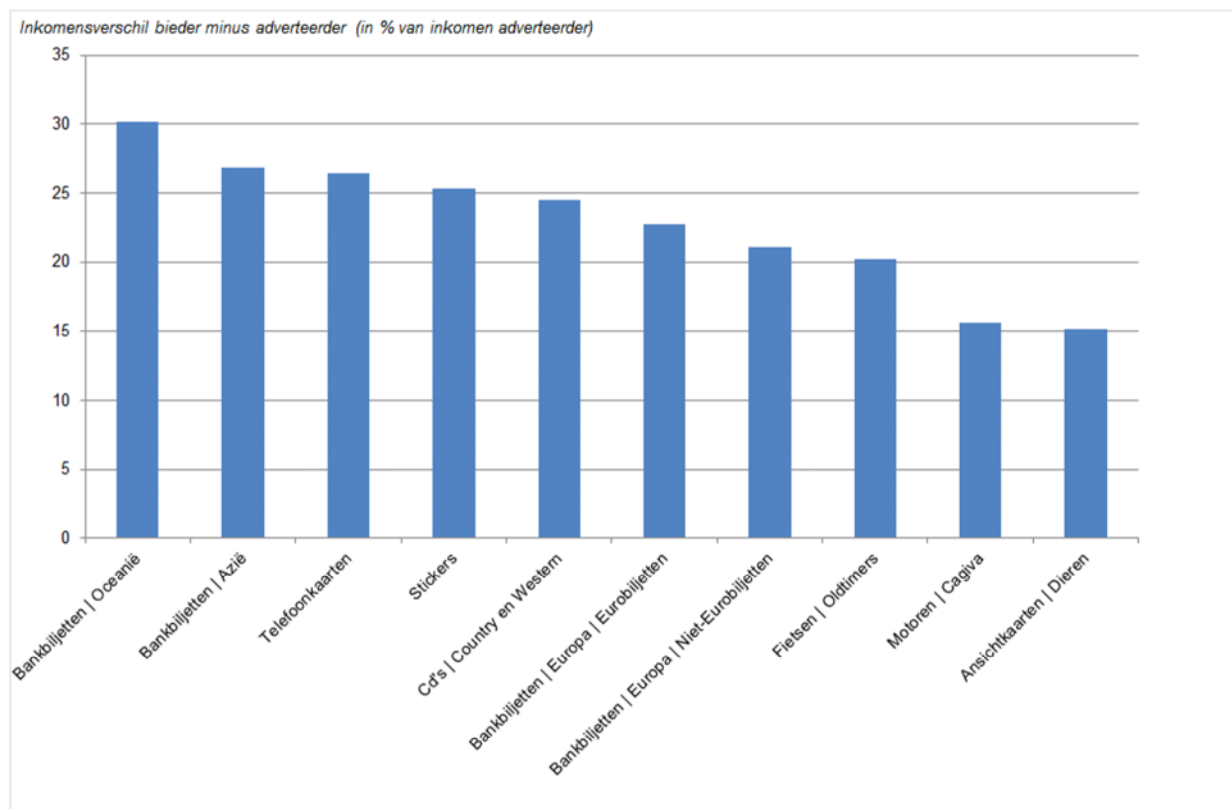


Bron: CBS (Data: Marktplaats).

Figuur 3.4.2 laat het omgekeerde plaatje zien namelijk die goederencategorieën die van relatief armere postcodegebieden naar relatief rijkere postcodegebieden gaan. Het betreft hierbij bijna uitsluitend verzamelobjecten, zoals bankbiljetten en ansichtkaarten. Het verschil met de vorige figuur is typerend. Van de tien categorieën waarbij het inkomensverschil tussen de bidder en adverteerder in positieve zin het grootst is, heeft er geen enkele betrekking op gebruiksgoederen.

De aanduidingen arm en rijk dienen hier overigens als zeer relatief te worden geïnterpreteerd. Ten eerste gaat het ook om relatieve inkomensverschillen tussen de postcodegebieden van bidder en adverteerder en dus niet per se tussen de in absolute zin hoogste inkomens- en laagste inkomensgroepen in de samenleving. Daarnaast worden de relatieve inkomensverschillen al gauw klein. Dit geldt vooral voor de goederencategorieën die van ‘rijk’ naar ‘arm’ gaan. Dit geldt opvallend genoeg in mindere mate voor de goederencategorieën die van ‘arm’ naar ‘rijk’ gaan. Hier geldt dat het inkomensverschil tussen het postcodegebied van bidder en adverteerder voor alle tien geselecteerde categorieën 15 procent of meer bedraagt.

Figuur 3.4.2 Categorieën waarbij inkomensverschil bidder en adverteerder het grootst is (positief); van ‘arm’ naar ‘rijk’



Bron: CBS (Data: Marktplaats).

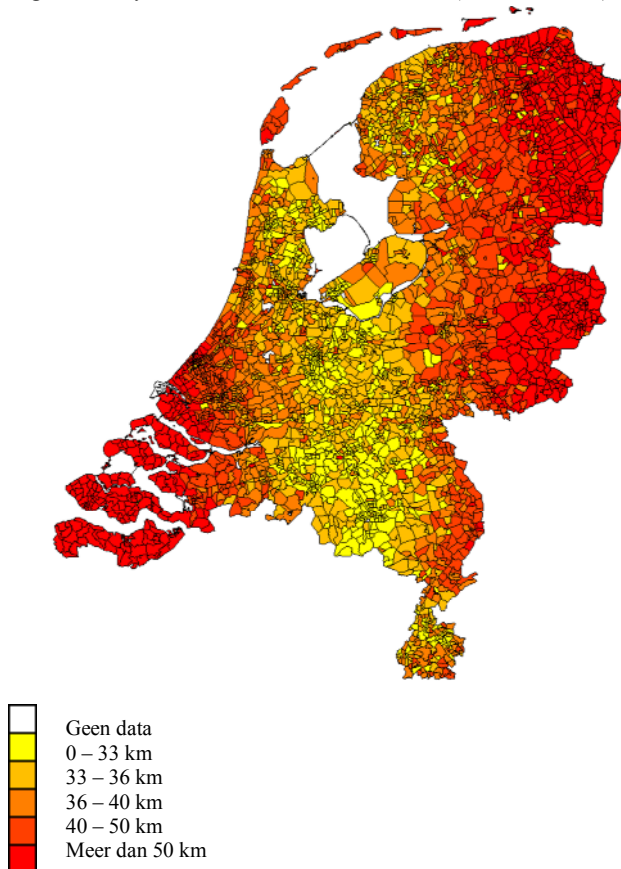
Afstand

Voor dezelfde groep adverteerders en bidders waarvan op grond van de postcode de inkomensgegevens achterhaald konden worden, kan de afstand tussen adverteerder en bieder worden berekend. Deze informatie zou iets kunnen zeggen over de reisbereidheid van potentiële kopers. De idee is dat kopers bereid zijn verder te reizen voor producten met een hogere prijs, omdat de transactiekosten (reiskosten) dan in verhouding niet zo hoog zijn. Deze relatie kan wel beïnvloed worden door allerlei aspecten die samenhangen met de categorie goederen waar het om gaat. Bepaalde goederen kunnen eenvoudigweg per post verstuurd worden (bankbiljetten, postzegels), anderen niet (fietsen, kasten). De transactiekosten voor een bankbiljet zijn onafhankelijk van de afstand tussen adverteerder en bieder (pakketje versturen binnen Nederland). Transactiekosten voor een fiets zijn wel afhankelijk van de afstand tussen adverteerder en bieder (zelf ophalen). Vertrouwen in de adverteerder en/of het aangeboden goed kan ook een rol spelen. Misschien wil men een postzegel wel eerst met eigen ogen zien voordat men tot aankoop besluit. In dat geval zou er toch ook voor postzegels een relatie kunnen zijn tussen afstand en prijs. Ook aanbod en vraag zelf spelen een rol. Het aanbod van schoenen en babykleertjes is bijvoorbeeld betrekkelijk groot en is het niet aannemelijk dat een potentiële koper hier ver voor hoeft te reizen. Dat ene verzamelobject of die ene saxofoon is wel uniek zodat een potentiële koper hiervoor wel bereid is verder te reizen.

Figuur 3.4.3 laat de gemiddelde afstand tussen bieder en adverteerder zien voor alle advertenties op 4-posities postcodeniveau. Het is goed te zien dat de gemiddelde afstand tussen bieder en adverteerder hoger is in de oostelijke grensgebieden, op de Waddeneilanden, in Zeeland en het zuidwesten van Zuid-Holland. Het ligt voor de hand dat dit te maken heeft met het aanbod van advertenties in de directe omgeving. Voor potentiële kopers in bovengenoemde gebieden is het moeilijker om een advertentie te vinden, vanwege de restricties die opgelegd worden door de aanwezigheid van de zee of landsgrenzen.

Bieders die woonachtig zijn in meer centraal gelegen gebieden vinden gemiddeld gezien hun goederen op een meer nabijgelegen locatie; ze kunnen letterlijk alle kanten op.

Figuur 3.4.3 Afstand tussen bidder en adverteerder (alle advertenties), 2010



Bron: CBS (Data: Marktplaats).

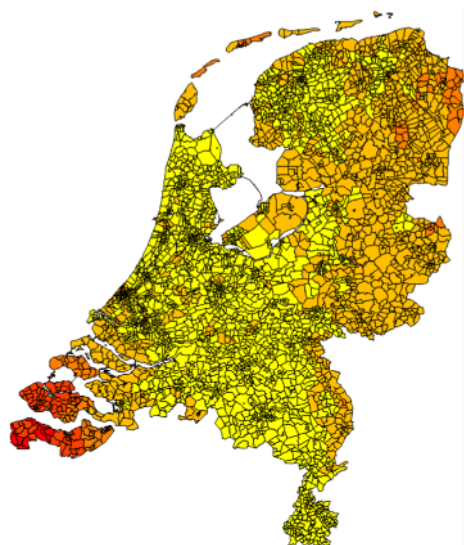
De bevindingen uit figuur 3.4.3 zijn plausibel, maar niet erg verrassend. Immers, als men op een meer centraal gelegen locatie woont is de verwachte afstand naar een willekeurig gebied kleiner dan wanneer men in een uithoek van het land woont.

Om die reden is het interessant om te kijken naar de gemiddelde afstanden voor een aantal specifieke categorieën. Onderstaande serie van vier kaartjes zoomt daarom in op vier verschillende hoofdcategorieën: 'Huis en Inrichting', 'Tuin en Terras', 'Kinderen en Baby's' en 'Postzegels en Munten'. Voor deze vier hoofdcategorieën is de gemiddelde afstand tussen bidder en adverteerder op 3-positie postcodeniveau weergegeven in de figuren 3.4.4a tot en met 3.4.4d. Hierbij zijn bewust dezelfde afstandscategorieën onderscheiden (legenda).

Bieders zijn minder bereid om ver te reizen voor goederen uit de categorieën 'Huis en Inrichting' en 'Tuin en Terras'. Een mogelijke verklaring kan zijn dat de bidder voor deze categorieën zelf moet reizen om het product bij de adverteerder op te halen.

Figuur 3.4.4 Afstand tussen bieder en adverteerder, vier hoofdcategorieën, 2010

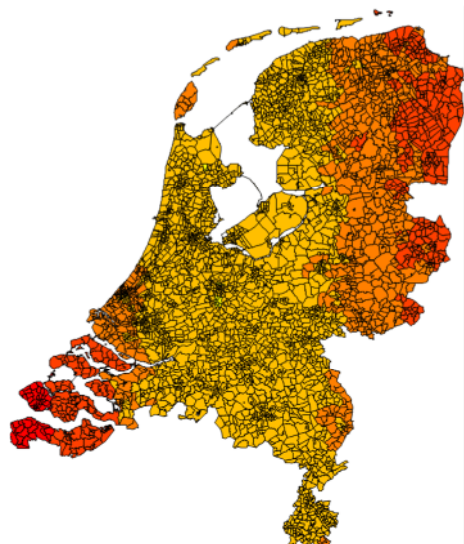
a. Huis en Inrichting



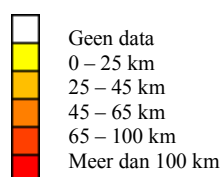
b. Tuin en Terras



c. Kinderen en Baby's



d. Postzegels en Munten



Bron: CBS (Data: Marktplaats).

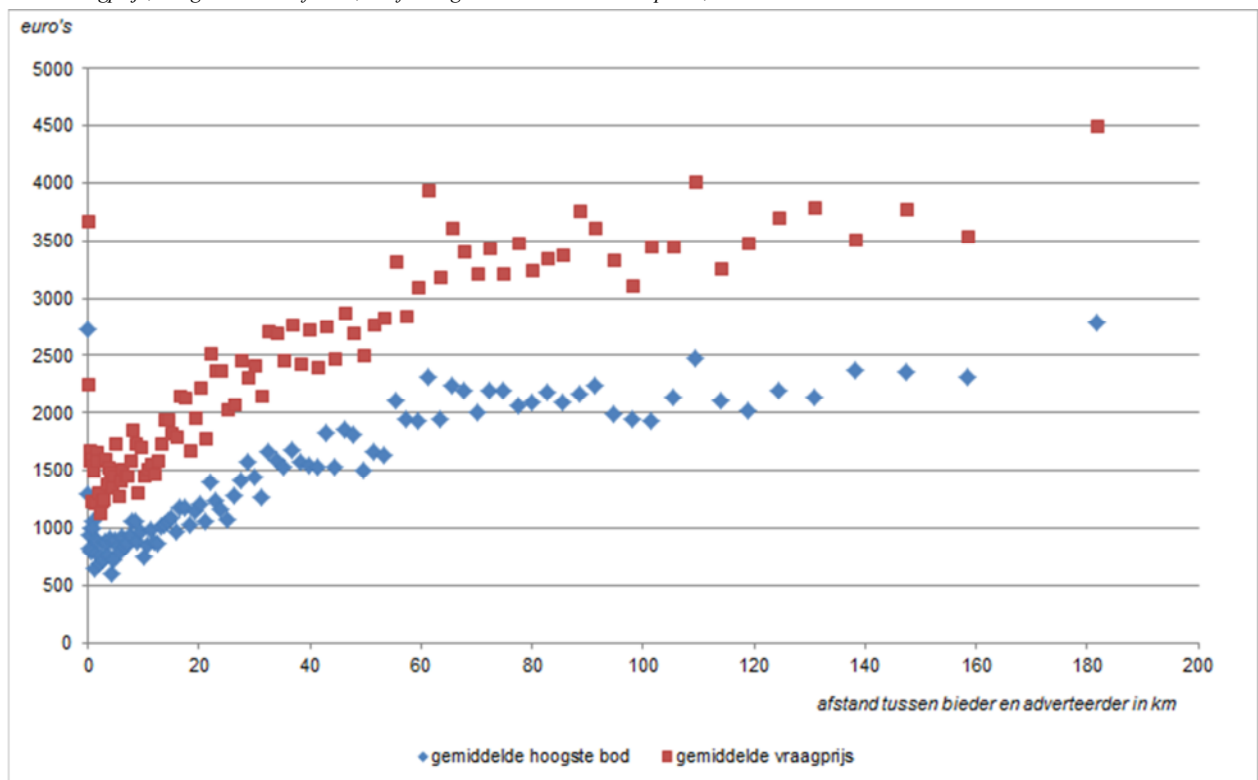
Het gaat om producten van groter formaat, die niet per post opgestuurd kunnen worden. Het blijft overigens ook hier zichtbaar dat bidders in de meer afgelegen gebieden van Nederland hoe dan ook een grotere afstand afleggen dan de bidders in de centrale gedeelten van het land. De afstanden tussen bieder en adverteerder zijn een stuk groter voor de categorieën 'Kinderen en Baby's' en 'Postzegels en Munten'. Dit wordt mogelijk verklaard door het feit dat deze goederen gemakkelijk per post verzonden kunnen worden. In dat geval speelt afstand een minder belangrijke rol doordat de totale transactiekosten (de kosten om het product van de adverteerder naar de bieder te transporteren) niet afstandsafhankelijk

zijn. Wederom is de gemiddelde afstand tussen bidder en adverteerder het grootst in de uithoeken van het land.

Moelijker te vervoeren goederen lijken dus letterlijk een kleinere (tweedehands) markt te hebben dan goederen die bijvoorbeeld per post kunnen worden bezorgd. De lokale component van de markt in tweedehands goederen verschilt dus per goederencategorie.

Een andere manier om nog eens in te gaan op de relatie vraagprijs, hoogste bod en afstand is uitgewerkt in de figuren 3.4.5 tot en met 3.4.7. In deze figuren zijn voor een drietal hoofdcategorieën de advertenties met een vraagprijs én een hoogste bod verdeeld over advertentiepercentielen gerangschikt op de afstand tussen adverteerder en bidder. De afstandsklassen tussen bidder en adverteerder zijn dus zo gekozen dat in elke klasse telkens één procent van de advertenties valt. Bij de laatste percentielen zijn deze klassebreedtes afstandsklassen dus ‘breder’. De afstandsklasse waar de laatste één procent van de advertenties in valt is vaak enkele kilometers ‘breed’, terwijl de één procent advertenties met de geringste afstand tussen bidder en adverteerder vaak al binnen een bereik van enkele honderden meters liggen. Voor elk van deze percentielen is vervolgens de gemiddelde vraagprijs en het gemiddelde hoogste bod bepaald.

3.4.5 Vraagprijs, hoogste bod en afstand, hoofdcategorie Caravans en Kamperen, 2010



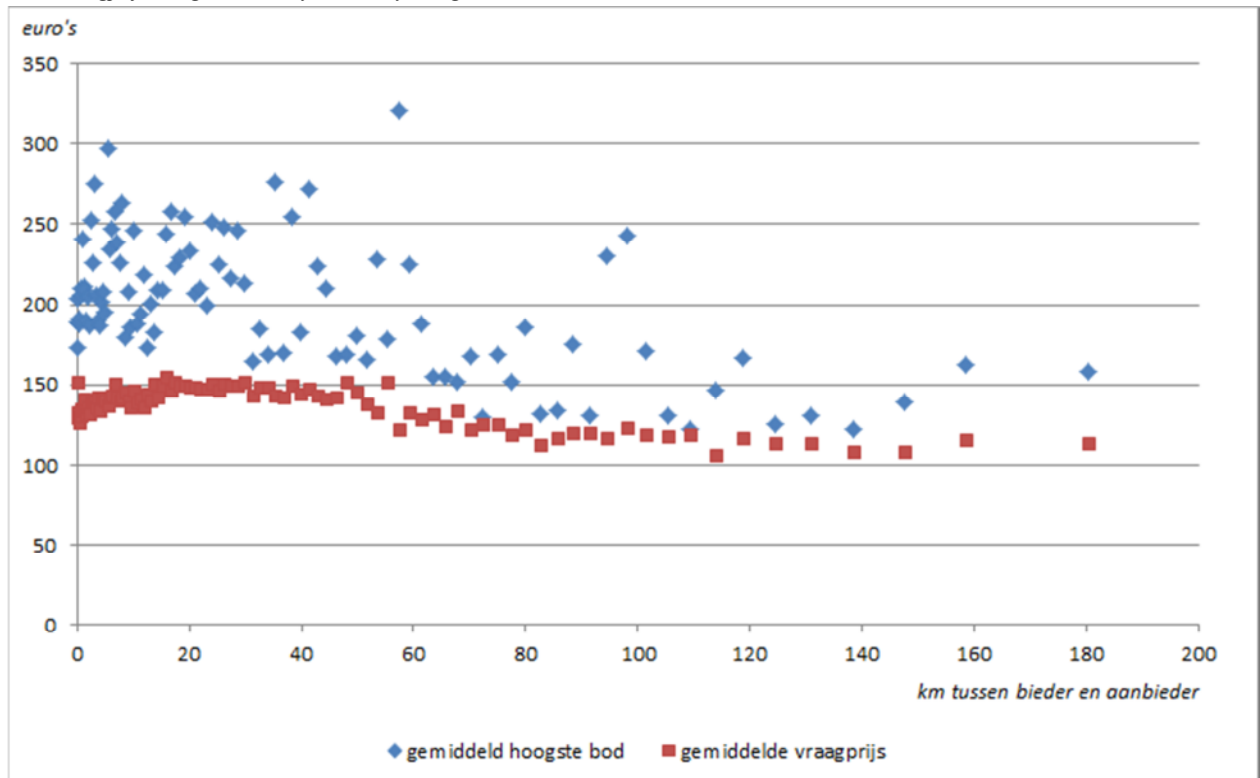
N.B. De indeling van de afstand tussen bidder en adverteerder is zo gekozen dat er telkens één procent van de advertenties in het betreffende afstandpercentiel valt.

Bron: CBS (Data: Marktplaats).

Voor de hoofdcategorie ‘Caravans en Kamperen’ (figuur 3.4.6) geeft dit het volgende beeld: een bidder biedt op goederen met een relatief hoge vraagprijs als het op steenworpafstand van zijn woning is (‘de caravan van de burens mag wat kosten’). Wordt eenmaal een bepaalde afstand overschreden dan biedt een bidder op goederen met een hogere vraagprijs naarmate de afstand groter is. Voor duurdere goederen in deze hoofdcategorie is men bereid wat verder te reizen. De verhouding tussen de

transactiekosten en de prijs van het eventueel te kopen goed zijn in dit geval ook lager. Een afstandspersentiel bevat hier 1 442 advertenties.

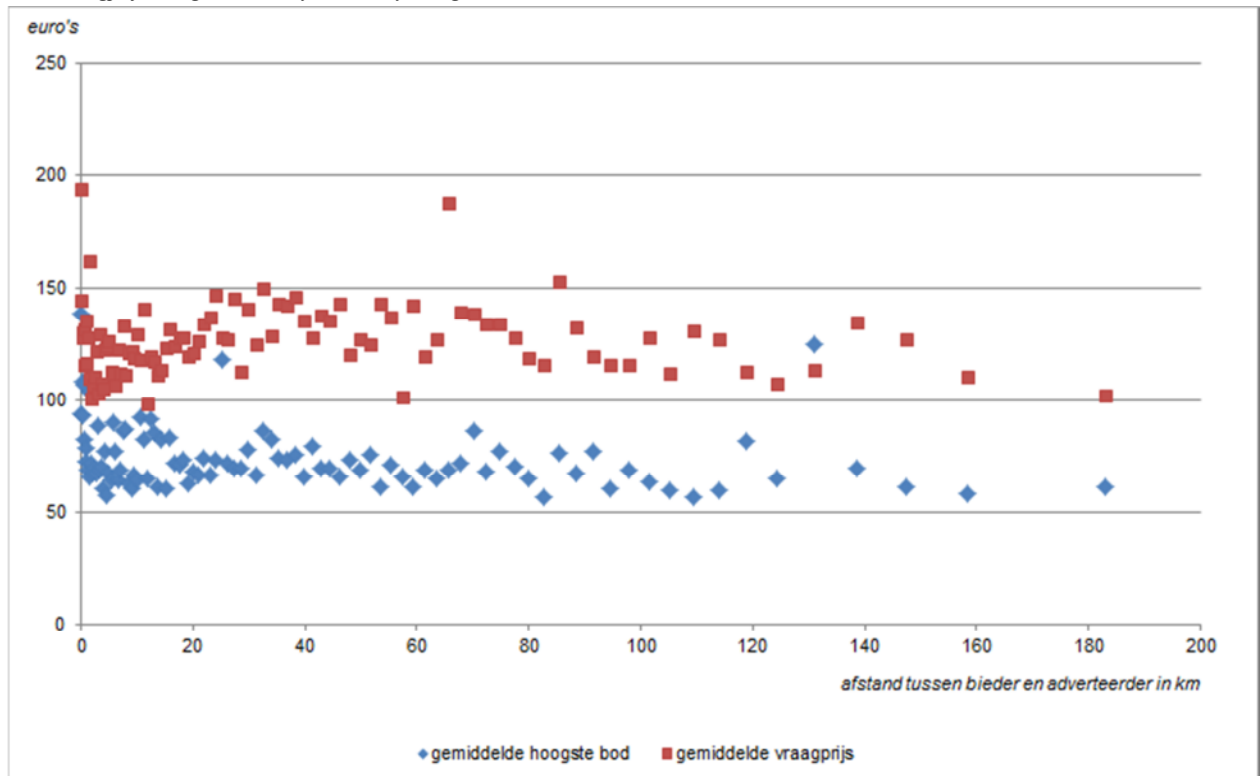
3.4.6 Vraagprijs, hoogste bod en afstand, hoofdcategorie Telecommunicatie, 2010



N.B. De deling van de afstand tussen bieder en adverteerder is zo gekozen dat er telkens één procent van de advertenties in het betreffende afstandspersentiel valt.

Bron: CBS (Data: Marktplaats).

3.4.7 Vraagprijs, hoogste bod en afstand, hoofdcategorie Antiek en Kunst, 2010



N.B. De indeling van de afstand tussen bidder en adverteerder is zo gekozen dat er telkens één procent van de advertenties in het betreffende afstandspersentiel valt.

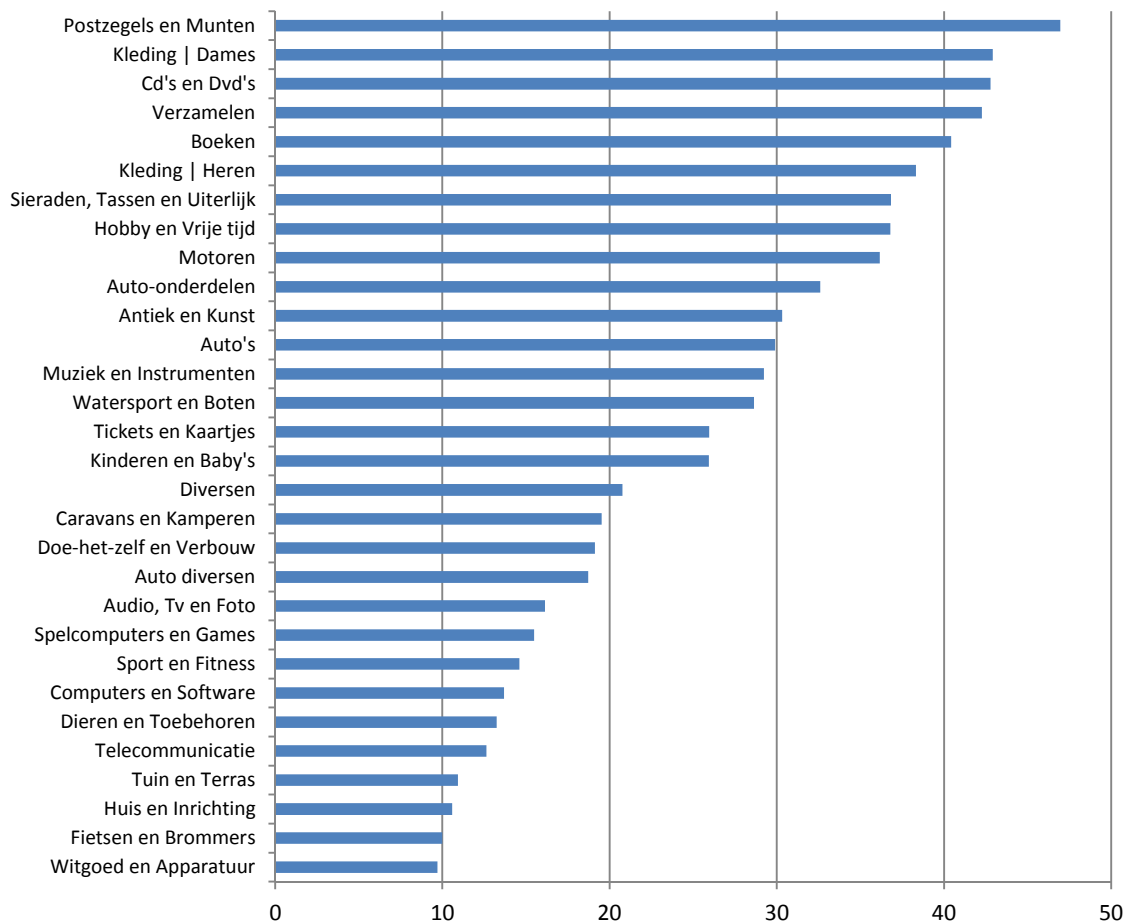
Bron: CBS (Data: Marktplaats).

In figuur 3.4.6 is hetzelfde gedaan, maar dan voor de hoofdcategorie ‘Telecommunicatie’. Hier lijkt iets anders aan de hand te zijn dan bij de hoofdcategorie ‘Caravans en Kamperen’. Het aanbod van goederen in deze categorie is groter; een advertentiepersentiel bevat hier 3 505 advertenties. Voor advertenties binnen een straal van circa 40 kilometer van de bidder lijkt er nauwelijks sprake van een verband tussen vraagprijs, bod en afstand. Na deze 40 kilometer lijkt er echter sprake van een soort omslagpunt: de gemiddelde vraagprijs gaat omlaag naarmate de afstand tussen bidder en adverteerder toeneemt. Kennelijk wordt een deel van de reiskosten dan opeens verdisconteerd met de vraagprijs (‘Als ik verder moet reizen, dan moet het wel goedkoper zijn’).

In figuur 3.4.7 tenslotte is voor de hoofdcategorie ‘Antiek en Kunst’ eenzelfde figuur getekend. Ook hier geldt net als bij ‘Caravans en Kamperen’ dat goederen in de directe nabijheid een streepje voor hebben; bidders bieden op advertenties met een hoge vraagprijs. Eenmaal over deze grens van directe nabijheid heen, is er geen relatie tussen vraagprijs, hoogste bod en afstand te zien. Het betreft hier dan ook een categorie goederen die betrekkelijk heterogeen is (‘unica’) en een deel van de bidders het object gewoonweg willen hebben ongeacht de afstand. In deze hoofdcategorie bestaat een advertentiepersentiel uit 5 148 advertenties.

Wat opvalt is dat voor al deze drie hoofdcategorieën de afstand tussen bidder en adverteerder ‘doorloopt’ tot bijna 200 kilometer. Zowel in de categorie ‘Caravans en Kamperen’ als in de hoofdcategorie ‘Telecommunicatie’ zijn er dus bidders die een hoogste bod uitbrengen op een advertentie waarbij het gewenste artikel bijna 200 kilometer verder weg ligt. Voor deze drie geselecteerde hoofdcategorieën lijkt het bereik van de tweedehands markt dus even groot. Dit blijkt voor alle hoofdcategorieën het geval; de maximum afstand tussen bidder en adverteerder is dus geen goede maat voor de geografische omvang van de markt. Om toch een beeld te vormen van hoe groot de markt voor iedere advertentiecategorie is staat in figuur 3.4.8 de afstand tussen bidder en adverteerder van de middelste advertentie. Zo is te zien dat vijftig procent van de advertenties in de categorie ‘Witgoed en apparatuur’ in 2010 van een bidder binnen een straal van 9,7 kilometer een hoogste bod ontvangt, voor boeken is dit 40,4 kilometer.

Figuur 3.4.8: Mediane afstand tussen bieder en adverteerder per hoofdcategorie in kilometer. 2010.



3.5 Regressiemodel

Met behulp van een regressiemodel zal geprobeerd worden een indicatie te krijgen welke kenmerken van de 6-positie postcodegebieden er nu toe doen bij het verklaren van verschillen in het aantal adverteerders en het aantal advertenties tussen 6-positie postgebieden. Het model dat zal worden geschat is geformuleerd in vergelijking (1).

$$\log(y_i) = a + bX_i + e_i \quad (1)$$

In vergelijking (1) geeft y_i het aantal adverteerders in postcodegebied i , a is een constante, X_i is een vector met J verklarende variabelen ($x_1, x_2, x_3, \dots, x_j$) en e_i is de storingsterm. De verklarende variabelen die in het model zijn opgenomen, zijn:

- leeftijdscategorieën (0-14-jarigen, 15-24-jarigen, 25-44-jarigen, 45-64-jarigen, 65-74-jarigen en 75-jarigen en ouder);
- geslacht (mannen, vrouwen);
- huishoudtype (huishoudens met kinderen, overige huishoudens);
- bevolkingsgroep (niet-westerse-allochtonen, overige inwoners);
- het gemiddelde fiscale maandinkomen per inwoner;
- het aantal adressen per vierkante kilometer (OAD);
- het absolute aantal inwoners.

Deze regressievergelijking in (1) wordt voor het jaar 2010 geschat. De reden hiervoor is dat de meeste variabelen per postcodegebied betrekking hebben op het jaar 2010.

Naast het aantal adverteerders wordt ook het aantal advertenties geschat. In termen van vergelijking (1) is y_i in dit geval het absolute aantal advertenties per 6-positie postcodegebied. De verklarende variabelen blijven hierbij dezelfde.

Data

De analyse is gebaseerd op alle advertenties en bijbehorende adverteerders waar in 2010 een 6-positie postcode van bekend was die ook gekoppeld kon worden aan een van de 6-positie postcodes zoals die voor dat jaar in de bestanden van het CBS voorkwamen. Uiteindelijk betrof dit 69,0 miljoen advertenties uit het in 2010 gebruikte analysebestand met 90,1 miljoen advertenties; of wel bijna 77 procent van alle advertenties. Het grootste aantal advertenties – 15,3 miljoen of wel 17 procent – ging verloren doordat advertenties waarvan slechts de 4-positie postcode bekend was, niet zijn meegenomen in de regressieanalyse. De advertenties waar alleen de 4-positie postcode van bekend was, waren niet normaal verdeeld over Nederland. Dit hoeft niet direct te duiden op andere sociaaleconomische of demografische kenmerken, maar het waren in ieder geval wel specifieke regio's. Het aantal 6-positie postcodegebieden in de uiteindelijke regressieanalyse was ruim 389 duizend. Dit zijn de 6-positie postcodegebieden uit het bestand van Marktplaats die gekoppeld konden worden aan de bestanden van het CBS en waarvan de verklarende variabelen ook beschikbaar waren.

Beschrijvende statistiek

Tabel 3.5.1 geeft een overzicht van het gemiddelde en de standaardafwijking van de afhankelijke en onafhankelijke variabelen. Het gemiddelde aantal adverteerders per 6-positie postcodegebied in 2010 bedroeg zeven en liep uiteen van één tot 661. Het aantal advertenties varieerde van één tot bijna 500 duizend. De invloed van dergelijke uitschieters op de regressieanalyse wordt verminderd door als te verklaren variabele de logaritme van respectievelijk het aantal adverteerders en het aantal advertenties te nemen.

3.5.1 Variabelen in de regressievergelijking per 6-positie postcodegebied, 2010

Variabele/kenmerk	Gemiddelde	Standaarddeviatie	Minimum	Maximum
Aantal adverteerders	7	6	1	661
Aantal advertenties	177	1 076	1	495 533
Fractie 0-14-jarigen	0,17	0,11	0	0,89
Fractie 15-24-jarigen	0,12	0,10	0	1
Fractie 25-44-jarigen	0,30	0,15	0	1
Fractie 45-64-jarigen	0,27	0,14	0	1
Fractie 65-74-jarigen	0,08	0,10	0	1
Fractie 75-jarigen en ouder	0,06	0,11	0	1
Fractie mannen	0,50	0,08	0	1
Fractie huishoudens met kind(eren)	0,39	0,23	0	1
Fractie niet-westerse-allochtonen	0,09	0,16	0	1
Fiscaal maandinkomen per inwoner	1 813	3 675	0	1 717 117
Adressendichtheid	1 698	1 731	0	12 342
Aantal inwoners	40	28	1	958

N.B. Aantal 6-positie postcodegebieden: 389 495; Aantal advertenties: 69,0 mln.

Bron: CBS (Data: Marktplaats).

Ook voor de verklarende variabelen is de variatie tussen de 6-positie postcodegebieden groot. Zo variëren de fracties van de verschillende leeftijdscategorieën van nul tot één; zijn er postcodegebieden waar praktisch geen kinderen wonen en postcodegebieden waar alleen maar kinderen wonen en postcodegebieden met een omgevingsadressendichtheid van nul en van meer dan 12 duizend. Juist omdat er verschil is in karakteristieken tussen de postcodegebieden kan dit een bijdrage leveren aan het verklaren van verschillen in aantallen adverteerders en advertenties per postcodegebied.

Schattingen

De resultaten van de meervoudige regressie staan in tabel 3.5.2. In de eerste specificatie wordt het aantal adverteerders geschat. Op het punt van leeftijd geldt dat een hogere fractie 15-24-jarigen of 25-44-jarigen in een postcodegebied, gegeven de andere variabelen, leidt tot een hoger aantal adverteerders. Het tegenovergestelde geldt voor oudere inwoners. De referentiegroep is hier de fractie 0-14-jarigen. Een hoger percentage mannen in een postcodegebied leidt gemiddeld tot meer adverteerders in dat gebied. Een groter aandeel huishoudens met kinderen en een hoger percentage niet-westerse-allochtonen leiden juist tot minder adverteerders in een postcodegebied. Dit beeld komt overeen met het beeld dat in tabel 3.3.1 naar voren kwam waar de kenmerken van postcodegebieden van de populatie adverteerders is afgezet tegen die van de totale bevolking. Ook hier waren jongere mensen oververtegenwoordigd in de populatie adverteerders en niet-westerse-allochtonen juist ondervertegenwoordigd.

Van het inkomen per inwoner, de omgevingsadressendichtheid en het aantal inwoners is de logaritme opgenomen in de regressievergelijking, ook met het oogmerk om uitschieters te 'neutraliseren'. Alle drie de variabelen hebben een positieve invloed op het aantal adverteerders. Hoe hoger het inkomen, de omgevingsadressendichtheid of het aantal inwoners in een postcodegebied, hoe meer adverteerders.

3.5.2 Gestandaardiseerde regressiecoëfficiënten

	log(adverteerders)	log(advertenties)
Fractie 15-24-jarigen	0,041	0,017
Fractie 25-44-jarigen	0,050	0,083
Fractie 45-64-jarigen	-0,085	-0,029
Fractie 65-74-jarigen	-0,107	-0,054
Fractie 75-jarigen en ouder	-0,177	-0,117
Fractie mannen	0,022	0,010
Fractie huishoudens met kind(eren)	-0,034	0,012
Fractie niet-westerse-allochtonen	-0,117	-0,127
log(fiscaal maandinkomen per inwoner)	0,232	0,177
log(adressendichtheid)	0,015	0,027
log(inwoners)	0,720	0,995

N.B. Alle regressiecoëfficiënten zijn significant met $p < 0,01$. $R^2 \log(\text{adverteerders}) = 0,505$; $R^2 \log(\text{advertenties}) = 0,252$.

Bron: CBS (Data: Marktplaats).

Verandert dit beeld als het aantal advertenties als te verklaren variabele wordt ingevoerd? De positieve invloed van 15-24-jarigen op het aantal advertenties is geringer in vergelijking met het verklaren van het aantal adverteerders in een 6-positie postcodegebied. En, de drukkende invloed van de ouderen wordt minder. Dit is ook consistent met het beeld dat naar voren kwam in tabel 3.3.1. Postcodegebieden met veel jongere inwoners adverteren meer dan gemiddeld, maar plaatsen niet zo heel veel advertenties. Inwoners van gebieden met veel ouderen adverteren beneden gemiddeld, maar de inwoners die adverteren, adverteren relatief veel. De aanwezigheid van huishoudens met kinderen heeft een drukkend effect op het aantal adverteerders, maar een positief effect op het aantal advertenties. Een hoger

percentage niet-westerse allochtonen in een postcodegebied leidt zowel bij de adverteerders als bij het aantal advertenties tot minder gebruik van Marktplaats. Inkomen, adressendichtheid en aantal inwoners hebben evenals bij het aantal adverteerders, een positief effect op het aantal advertenties in een postcodegebied.

3.6 Conclusies

In dit hoofdstuk zijn de Marktplaatsdata gekoppeld aan – kenmerken van – 6-positie postcodegebieden zoals het CBS die beschikbaar heeft. Vanuit heel Nederland worden advertenties geplaatst op Marktplaats. Het aantal verschillende personen dat gebruik maakt van Marktplaats (adverteerders, bidders) lijkt in het noorden van Nederland echter relatief groter dan in het zuiden. Kenmerken van 6-positie postcodegebieden die positief gecorreleerd zijn met het aantal adverteerders is het aandeel jongere mensen (15-24-jarigen, 25-44-jarigen). Een groot aandeel niet-westerse-allochtonen draagt juist niet bij aan het aantal adverteerders. Adverteerders die bovengemiddeld vaak adverteren zijn oververtegenwoordigd in postcodegebieden met een lagere omgevingsadressendichtheid. Twee kenmerken zijn nader bekeken: inkomen en afstand. De categorieën goederen waarbij het relatieve inkomensverschil tussen het postcodegebied van de bidder en de adverteerder in negatieve zin het grootst is – goederen die van rijkere naar armere postcodegebieden gaan – worden gedomineerd door gebruiksgoederen. Categorieën die van armere naar rijkere postcodegebieden gaan worden daarentegen gedomineerd door ‘verzamelobjecten’. De afstand tussen bidder en adverteerder is voor de personen woonachtig in het midden van Nederland (logischerwijze) het kleinst. Voor een drietal categorieën is nagegaan of er een zekere relatie is tussen vraagprijs, bod en afstand tussen adverteerder en bidder. Dit gaf een wisselend beeld. Voor goederen in de categorie ‘Caravans en Kamperen’ leek de bidder bereid verder te reizen voor duurdere goederen. In de categorie ‘Telecommunicatie’ was dit echter juist niet zo. In de hoofdcategorie ‘Antiek en Kunst’ lijkt afstand geen rol te spelen. Ook hier lijkt het dus moeilijk om een eenduidige relatie vast te stellen tussen vraagprijs, bod en afstand. Dit lijkt sterk afhankelijk van de categorie goederen waar het omgaat. Daarnaast is er een aantal categorieën waarbij het artikel verstuurd zou kunnen worden; afstand speelt dan nauwelijks een rol.

4. Data van Marktplaats; omzet C2C-handel

Een schatting van de omzet van de zogenaamde consumer-to-consumer-handel (C2C) is om tal van redenen interessant. Niet omdat het een nieuw verschijnsel is; de briefjes op het prikbord van de lokale supermarkt en de advertenties in het plaatselijke huis-aan-huis blad bestonden al veel eerder. Het is vooral de enorme schaalvergroting die er heeft plaatsgevonden doordat vraag en aanbod van tweedehands goederen via internet bij elkaar kunnen worden gebracht. De moeite om tweedehands goederen aan te bieden is niet bijzonder groot. Het bereik van de advertentie – en daarmee de verkoopkans – is echter vele malen groter dan de eerder genoemde lokale supermarkt of het huis-aan-huis blad. Dit wordt bevestigd door het succes of de omvang van websites die de handel in tweedehands goederen faciliteren. In economische zin is er bij de handel in tweedehands goederen tussen particulieren sprake van een herverdeling van goederen binnen de sector huishoudens. Deze handel draagt niet bij tot de economische groei in termen van het bruto binnenlands product. Dit wil echter niet zeggen dat het een ‘waardeloos’ verschijnsel is. Alleen al het feit dat het gebeurt, illustreert dat het in een behoefte voorziet en daarmee waardevol is. Ook draagt het op grote schaal verkopen van voor het ene huishouden niet meer nuttige goederen aan een ander huishouden, bij aan de gebruiksduur van consumptiegoederen en daarmee aan de duurzaamheid van de economie (‘spullen gaan langer mee’). Deze transacties beïnvloeden op een of andere manier ook de omzet van de Detailhandel, met andere woorden: het is niet een verschijnsel zonder belang voor de ‘echte’ economie.

4.1 Kan een schatting van de C2C-handel worden gemaakt?

De omzet van de consumer-to-consumer markt die via Marktplaats.nl gefaciliteerd wordt, is een grootte die niet zonder meer te bepalen is op basis van de van Marktplaats verkregen gegevens. Hieronder zal worden beschreven wat de problemen hierbij zijn. De C2C-omzet in jaar t is de waarde van alle via Marktplaats tussen consumenten verhandelde artikelen bij elkaar opgeteld. Deze rekensom luidt als volgt:

$$\text{C2C-omzet}_t = \sum_{i=1}^N \text{verkoopprijs}_{it}$$

De bouwstenen voor deze rekensom zijn dus het aantal verkochte tweedehands goederen (N) en de daarbij behorende verkoopprijs. In de vergelijking wordt bewust gesproken over tweedehands goederen. Nieuwe goederen kunnen per definitie niet door consumenten worden verkocht. Ook een TV die ‘niet uit de doos is geweest’ maar door een consument wordt aangeboden, is tweedehands, want in economische zin al een keer ‘geconsumeerd’. In dit hoofdstuk zal stap-voor-stap aangegeven worden wat de (onoverkomelijke) moeilijkheden zijn om tot een enigszins betrouwbare schatting van de C2C-omzet te kunnen komen.

Vraag 1: Is het aantal aangeboden tweedehands goederen bekend?

Adverteerders kunnen in het veld CONDITIE bij de advertentie aangeven of het aangeboden artikel ‘Gebruikt’, ‘Z.g.a.n.’ (‘Zo goed als nieuw’) of ‘Nieuw’ is. Voor alle in 2010 geplaatste advertenties was deze conditie voor 61 procent van de advertenties bekend. Voor een schatting van het aantal advertenties van tweedehands goederen zou de categorie ‘Nieuw’ worden weggelaten en de categorieën ‘Gebruikt’ en ‘Z.g.a.n.’ worden samengenomen. Wat rest is de 39 procent van de advertenties waarbij in het midden is gelaten of het hier nieuwe of tweedehands goederen betreft (zie tabel 4.1.1).

4.1.1 Vulling variabelen benodigd voor schatting C2C-omzet Marktplaats, 2010

	Aantal advertenties (duizend)	Aandeel advertenties
<i>Conditie aangeboden goederen</i>		
Niet ingevuld	35 006	39%
Gebruikt	14 664	16%
Z.g.a.n. ('Zo goed als nieuw')	22 146	25%
Nieuw	18 250	20%
<i>Reden verwijdering advertentie</i>		
Verkocht	11 627	13%
Levens einde	27 676	31%
Geen selectie	36 511	41%
Overige redenen	14 251	16%
Totaal	90 065	100%

N.B. Verkocht (verkocht op Marktplaats, elders verkocht); Levens einde ((gratis) advertentieperiode verstrekken).

Bron: CBS (Data: Marktplaats).

Vraag 2: Is het aantal verkochte tweedehands goederen bekend?

Een advertentie op marktplaats kent een aantal mogelijke 'levensindes'. Een advertentie kan verlopen. Er is namelijk een maximum aan de periode dat een advertentie (gratis) op Marktplaats mag staan. Een advertentie kan om welke reden dan ook door Marktplaats zelf verwijderd worden. Ook kan de adverteerder zelf zijn of haar advertentie verwijderen. Bij dit verwijderen kan de adverteerder aangeven waarom hij of zij dit doet. Eén van de keuzemogelijkheden is dat de adverteerder aangeeft dat het artikel verkocht is; via Marktplaats of anderszins. Het verwijderen van een advertentie en het opgeven van de reden is een vrijwillig veld voor de adverteerder en doet dus een beroep op de discipline van adverteerders om dit ook 'braaf' te doen. Bij een deel van de adverteerders lijkt deze discipline te ontbreken. Andere adverteerders hebben er ook helemaal geen belang bij hun advertenties van Marktplaats te verwijderen, bijvoorbeeld de bedrijfsmatige adverteerders die min of meer continu zichtbaar willen zijn op Marktplaats. Tabel 4.1.1 illustreert de geringe informatie die uit de opgave van de reden van verwijdering van de advertentie (AD_RMVL_RSN_ID) is te halen. In 2010 was slechts voor 13 procent van de advertenties min of meer uitsluitend over het feit of de advertentie tot een transactie heeft geleid.

Vraag 3: Is de verkoopprijs bekend?

Een eerste aanname zou kunnen zijn dat ieder artikel dat is aangeboden en een hoogste bod heeft ontvangen, verkocht is tegen dit hoogste bod. Op deze manier zou voor 2010 van circa 24 procent van de advertenties een verkoopprijs bekend zijn. Het hanteren van de vraagprijs voor de waarde van een artikel is ook mogelijkheid. In 2010 was voor 55 procent van alle advertenties een vraagprijs bekend. Afgezien van de overlap tussen beide groepen is van meer dan de helft van de advertenties dus op een of andere wijze een indicatie voor de prijs van de aangeboden goederen voor handen. Hierbij is niet de plausibiliteit van de vraagprijs en het hoogste bod betrokken. Zoals in hoofdstuk 2 aangegeven zijn advertenties met een vraagprijs van meer dan 99 999,- euro uit het analysebestand weggelaten. Verder is er niet 'gaaf gemaakt' op rare vraagprijzen en/of ongeloofwaardige hoogste biedingen. Ongeloofwaardige vraagprijzen en/of hoogste biedingen zijn echter wel 'aan te wijzen' in de dataset. Ten slotte is bekend dat de totstandkoming van een aantal transacties zich buiten het gezichtsveld van Marktplaats (en dus de dataset) afspeelt. Zo had in 2011 'slechts' 63 procent van de advertenties waarbij

aangegeven was dat ze verkocht zijn via Marktplaats, een hoogste bod ontvangen. In een aantal gevallen is de prijs dus anderszins tot stand gekomen; bijvoorbeeld doordat de (potentiële) koper zelf contact heeft gezocht met de adverteerder (telefoon, e-mail) en de transactie verder langs deze weg heeft afgehandeld.

De ‘witte vlekken’ in de dataset om tot een sluitende berekening van de C2C-omzet via Marktplaats te komen is bij de primaire bouwstenen van deze rekensom direct al groot. Vooral het bepalen hoeveel advertenties nu daadwerkelijk tot een transactie hebben geleid, is zeer moeilijk. Maar ook het bepalen van de verkoopprijs is zonder aanvullende (arbitraire) veronderstellingen niet te doen. En zelfs het totale aanbod van tweedehands goederen is nog met onzekerheden omkleed (zie hieronder).

Andere complicaties

Los van de kale beschikbaarheid van een variabele is er op een aantal punten ook nog sprake van ‘ruis’ in de variabele: meet de variabele wel wat het suggereert te meten? Een voorbeeld hiervan is het feit dat een advertenties lang niet altijd een uniek artikel betreft. Zo kan men bij het aanklikken van ‘Fietsen’ terecht komen bij advertenties waar een heel assortiment fietsen wordt aangeboden. Hierbij is het onduidelijk welk product nu precies wordt aangeboden en als er een transactie plaatsvindt, wat er dan precies verkocht is. Het omgekeerde geldt ook. Sommige artikelen worden vele malen achtereenvolgens aangeboden. Er zitten dus dubbele advertenties in van bijvoorbeeld commodes, zonnebanken e.d. Als een zonnebank bij de vierde keer adverteren uiteindelijk wordt verkocht is uit de dataset niet af te leiden dat dezelfde zonnebank al drie keer eerder, zonder succes, is aangeboden. Dit maakt het ook moeilijk om tot verantwoorde aannames te komen hoeveel unieke goederen er nu uiteindelijk via Marktplaats worden aangeboden en verkocht. In 2011 plaatste driekwart van alle zonnebankaanbieders slechts één enkele advertentie voor een zonnebank. Hier kunnen dus geen dubbele advertenties tussen zitten. Echter, 25 procent van de aanbieders plaatste meerdere keren een advertentie voor een zonnebank; de kans dat hier dubbele advertenties inzitten, lijkt reëel.⁴

4.1.3 In 2011 op Marktplaats aangeboden zonnebanken

Aangeboden zonnebanken per adverteerder	Percentage van adverteerders
1	74,3
2 tot 5	22,7
5 of meer	3,0

Bron: (CBS (Data: Marktplaats).

Een andere kwestie is de vraag tot wanneer een adverteerder en koper nog een particulier zijn en niet een bedrijfsmatige handelaar (koper en verkoper) in tweedehands goederen? Of een artikel daadwerkelijk van consument naar consument verhandeld wordt is slechts naar te gissen. Wel is er op basis van de eerder besproken kwartielenaanpak een inschatting te maken van veel-gebruikers; maar of een veel-gebruiker een actieve particulier is of een bedrijf, is niet zonder meer te bepalen. Daarbij zijn bedrijven op Marktplaats niet per se direct te herkennen als bedrijfsmatig; een architectenbureau die een tweedehands bureaustoel en een nietmachine aanbiedt is anders dan een auto-onderdelenhandel die dagelijks twaalf complete sets winterbanden aanbiedt; beide zijn bedrijven, maar een ander profiel.

⁴ In de dataset kon hier niet op ‘gezocht worden’ omdat de advertentietekst geen onderdeel was van de gevraagde dataset.

Iets soortgelijks geldt voor de bieder. Van de hoogste bieder is nog net iets minder bekend dan van de adverteerder, maar intuïtief is wel aan te geven dat een bieder nauwelijks nog een particulier te noemen is als men in één jaar bij meer dan 100 auto's, 50 motoren en 230 boten het hoogste bod heeft voor een bedrag van circa 4,5 miljoen euro. Maar waar de grens te leggen is, net zoals bij een aantal voorgaande kwesties, arbitrair.

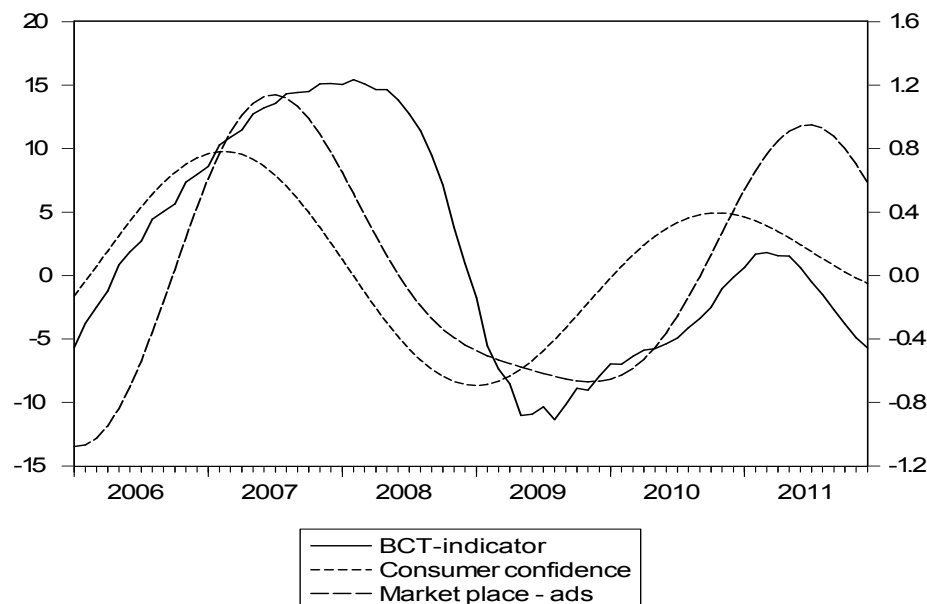
Omdat alle eerder genoemde punten te veel haken en ogen opleveren bij het vaststellen van de waarde, het volume en de aard van de bij de transactie betrokken partijen, wordt afgezien van het maken van een schatting van de C2C-handel via Marktplaats. De invloed van de aannames zou te groot zijn en de aannames zelf, te arbitrair. Zonder dat het in geld is uit te drukken, lijkt de conclusie dat de omvang van de handel in tweedehands goederen gezien het aantal betrokken adverteerders, het aantal advertenties en de spreiding van het verschijnsel over heel Nederland, substantieel is, gerechtvaardigd. Tot slot dient hierbij opgemerkt dat het hier weliswaar om een grote speler gaat op het terrein van het faciliteren van de handel in tweedehands goederen, maar zeker niet de enige. Er zijn meerdere algemene intermediairs die deze dienst bieden. Daarnaast zijn er speciaal op niches gerichte sites. Een volledig beeld van aard en omvang van de online handel in tweedehands goederen vergt dus een grotere mobilisatie van bronnen.

5. Data van Marktplaats; conjunctuur⁵

Samenvatting

Data van online activiteiten zijn potentieel een aantrekkelijke bron voor het construeren van economische indicatoren. De populatie is vaak erg breed, er is veel data en de actualiteit is potentieel zeer hoog. Zelden is het echter zo dat de data direct bruikbaar zijn, over het algemeen zijn het bijproducten van specifieke online activiteiten. Er moet dus getest worden welke relevante informatie er eigenlijk in de data zit, en wat voor filtering er noodzakelijk is. In dit Engelstalige hoofdstuk wordt verslag gedaan van diverse verkennende analyses die zijn uitgevoerd om te testen in hoeverre de Marktplaatsdata nuttige macro-economische indicatoren kunnen opleveren. De focus lag daarbij op het gebruik bij het signaleren van korte-termijn economische ontwikkelingen, omdat de actualiteit en frequentie waarmee de Marktplaatsdata gegenereerd worden ze daar potentieel geschikt voor maken. In eerste instantie is getest in hoeverre data van Marktplaats zuivere conjunctuurindicatoren zouden kunnen opleveren. Daarna is de relatie tussen ontwikkelingen op Marktplaats en de detailhandel onderzocht. Op Marktplaats worden zowel nieuwe als tweedehands goederen verhandeld, met name consumptiegoederen. Vandaar dat een analyse van de invloed op en samenhang met ontwikkelingen in de detailhandel voor de hand ligt. De Marktplaatsdata zijn in dit onderdeel op twee manieren geanalyseerd; in de vorm van het aantal advertenties en als het aantal advertenties dat een hoogste bod heeft ontvangen. De eerste benadering is een aanbod maat, de tweede geeft een indicatie van de daadwerkelijke vraag omdat een succesvol bod hoogst waarschijnlijk resulteert in een daadwerkelijke transactie. Deze data geven geen inzicht in de waarde of hoeveelheid van de producten die aangeboden of verhandeld worden, het zijn meer volume- of activiteitenindicatoren.

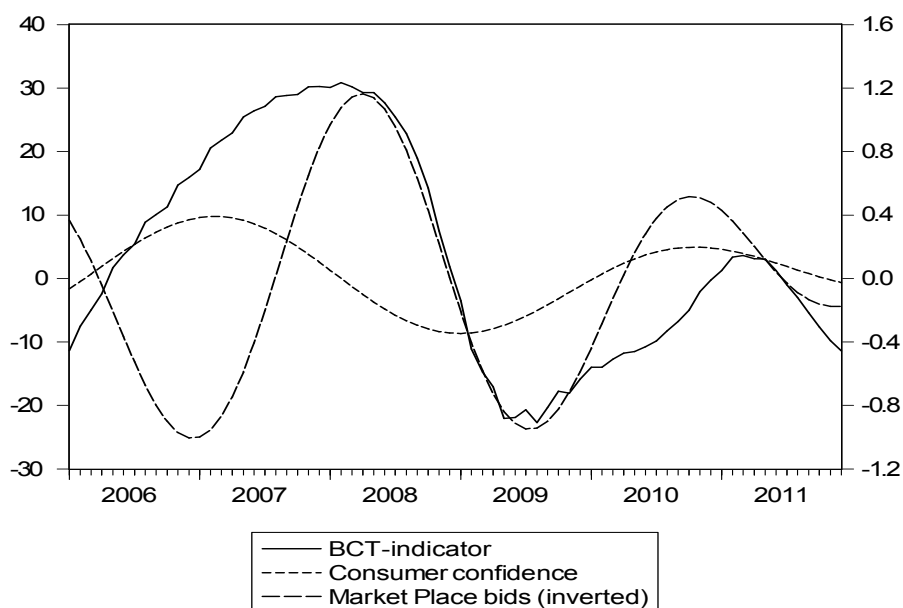
1. Totaal volume Marktplaatsadvertenties [Christiano-Fitzgerald band pass filter, 2-5 jaar], vergeleken met de Nederlandse conjunctuur weergegeven door de conjunctuurklokindicator (BCT-indicator) en het consumentenvertrouwen (Consumer-confidence).



⁵ Dit hoofdstuk is verzorgd door Floris van Ruth. Het betreft een Engelstalige bijdrage. Daarom is er voor gekozen deze bijdrage vooraf te laten gaan door een Nederlandstalige samenvatting. Floris van Ruth is werkzaam bij de sector Nationale Rekeningen.

Uit de data van zowel het totale aantal aangeboden consumptiegoederen als het aantal succesvolle biedingen op consumptiegoederen werd de cyclus gefilterd. Beide cycli vertoonden een duidelijke samenhang met de Nederlandse conjunctuurcyclus. Het patroon was in belangrijke mate hetzelfde, alleen de overeenkomst op het gebied van datering van omslagpunten varieerde. De activiteit op Marktplaats, zowel aan de aanbod als aan de vraag kant, wordt dus beïnvloed door de conjunctuur. Dit betekent dat Marktplaats geen eigen leven leidt, maar onderdeel is van de Nederlandse economie. Dit werd verder bevestigd door de nog sterkere samenhang tussen het consumentenvertrouwen en de cycli van de Marktplaatsdata. Opvallend was dat het aanbod (advertenties) op Marktplaats positief samenhangt met de conjunctuur, maar de vraag (biedingen) negatief. In slechtere tijden koopt men meer op Marktplaats, afgaand op de biedingen ten minste.

2. Totaal volume biedingen [Christiano-Fitzgerald band pass filter, 2-5 jaar], vergeleken met de Nederlandse conjunctuur weergegeven door de conjunctuurklokindicator (BCT-indicator) en het consumentenvertrouwen (Consumer-confidence).



Deze negatieve samenhang is een interessant resultaat, dat bevestigd werd in verdere analyses. Ook gespiegelde gedrag van aanbod en biedingen komt terug. Via tijdreeksmodellen is het verband tussen ontwikkelingen op Marktplaats en in de detailhandel formeel getest. Aan de ene kant geeft dit een indicatie of Marktplaats en de “gewone” economie elkaar beïnvloeden, en aan de andere kant is het een test of Marktplaatsdata informatie bevat die nuttig kan zijn bij het volgen van economische ontwikkelingen. De analyses zijn uitgevoerd zowel voor de totale detailhandel en het totale Marktplaats aanbod van consumptiegoederen, als voor de detailhandel in consumentenelektronica en een speciaal geconstrueerd Marktplaats consumenten elektronica-aggregaat. Deze laatste twee hebben qua samenstelling een betere overeenkomst dan de totale aggregaten. Dat bleek ook uit een betere fit in de modellen voor deze combinatie. Het opvallendste resultaat was dat de biedingen meer informatie over ontwikkelingen in de detailhandel blijken te bevatten dan het aantal advertenties. Het voorspellend vermogen was groter. Ook in de modellen bleek het verband tussen ontwikkelingen in de detailhandel en het aantal biedingen negatief; een hogere groei van het aantal biedingen op Marktplaats ging gepaard met een lagere groei van de detailhandel. Maar ook het aantal advertenties bleek bij de specifieke analyse voorspellende waarde te hebben, met een positief verband. De algemene conclusie is dat er een duidelijk verband is tussen Marktplaats activiteit en economische ontwikkelingen, en dat

Marktplaatsdata belangrijke (voorspellende) informatie bevatten. Vooral het aantal biedingen lijkt potentieel interessant als conjunctuurindicator. Een afsluitend voorbehoud; de periode waarop deze analyse is gebaseerd, is aan de korte kant en was economisch gezien nogal woelig. Het betreffen hier dan ook eerste tentatieve conclusies.

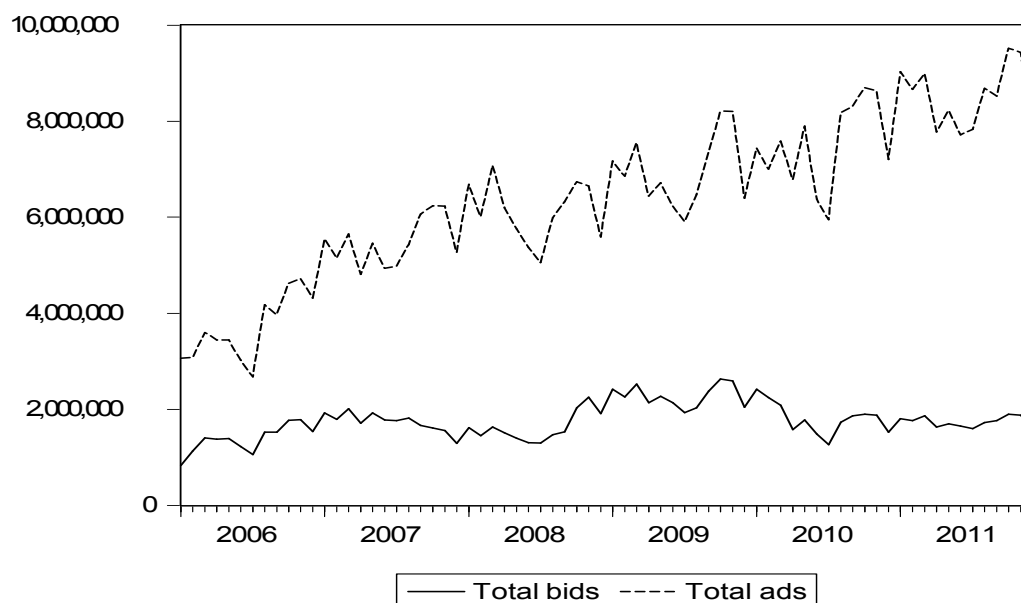
5.1 Introduction and general analysis

More and more business to consumer retail activity is shifting to the internet. The Dutch classified advertising site “Marktplaats.nl” (since 2004 a subsidiary of eBay) started out in 1999 as solely a marketplace for transactions of second hand goods between consumers on which an increasing number of professional sellers of both second hand and new goods have become active.

Statistics Netherlands has been given access to the data archive of Marktplaats, to analyse their significance for social and economic statistics. It concerns micro data of the individual advertisements. This part of the research project tests whether the monthly aggregates of these data can be useful as macro-economic indicators. This could be in the form of early indicators of general economic developments, or in forecasts for the retail sector.

For this purpose the categories most relevant to consumption, i.e. concerning consumer goods, were selected. This means that non-consumer related categories such as job vacancies, business services, contact ads, real estate etc., were filtered out. A total aggregate of these categories was computed and used in the first analysis testing a possible relationship between Marktplaats activity and general economic activity and retail sales. In the second phase the same relationships are tested again but now for a selected number of categories representing a specific and more homogeneous group of consumer goods: consumer electronics. This group of consumer goods coincides more with part of the retail sales and might also be more sensible to general economic trends (see paragraph 5.2). A detailed description of the dataset used is given in appendix 3.

5.1.1 Monthly volume of newly placed consumer goods related advertisements and highest bids on Marktplaats.



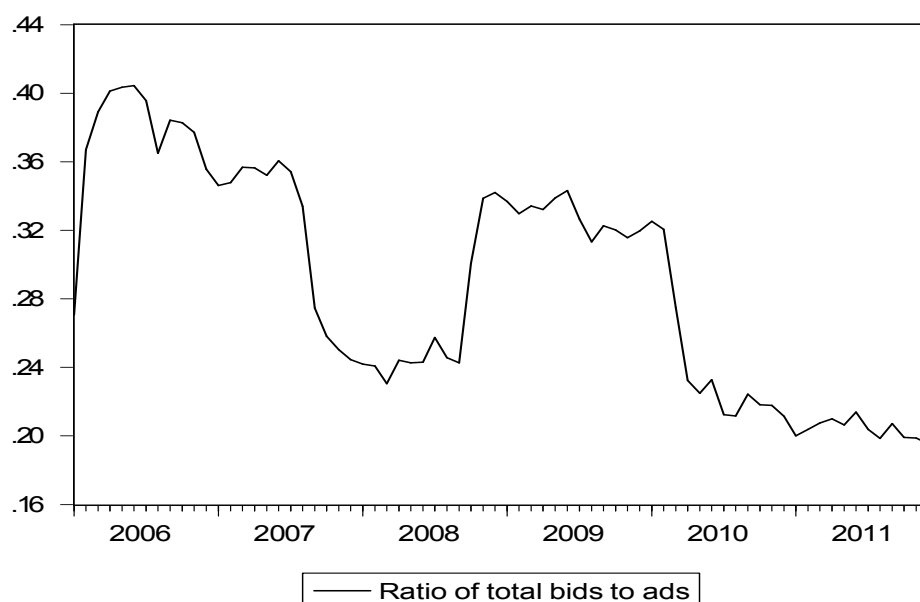
Marktplaats offers the option of holding an online auction, and part of the items offered are thus sold. This means that there are two types of data available relating to Marktplaats offerings. For a start, there

is the number of advertisements, but there is also the number of auctions concluded with a highest bid. In the latter case, it is highly likely that a transaction has occurred. Unfortunately, only part of the transactions is conducted by auction. An issue with the advertisements is that these are in principle unbound; it is both possible to put up a large amount of advertisements relating to a single item and to put up one advertisement for a batch of goods.

Both the aggregates on advertisements (“ads”) and highest bids (“bids”) are pure volume indicators, i.e. they are just a count, no information on the number or value of items concerned is taken into account.⁶ For each type of data, indicators were constructed by counting per month the new occurrences. Data were available from January 2006 to December 2011. Graph 5.1.1 shows the resulting series.

There is a marked difference in development between the number of advertisements and the number of successful auctions. Growth of the latter has been much slower, and has even levelled off. On the other hand, the number of advertisements possesses a strong trend. This could partly be due to a growth of multiple advertisements relating to single items. On the other hand selling by auction is only one of the possibilities to conduct a transaction. The number of advertisements which offer this option is not necessarily stable over time. It is thus instructive to consider the ratio of the total number of bids to the total number of advertisements (see graph 5.1.2).

5.1.2 Monthly ratio of new highest bids to new advertisements of consumer-goods on Marktplaats.



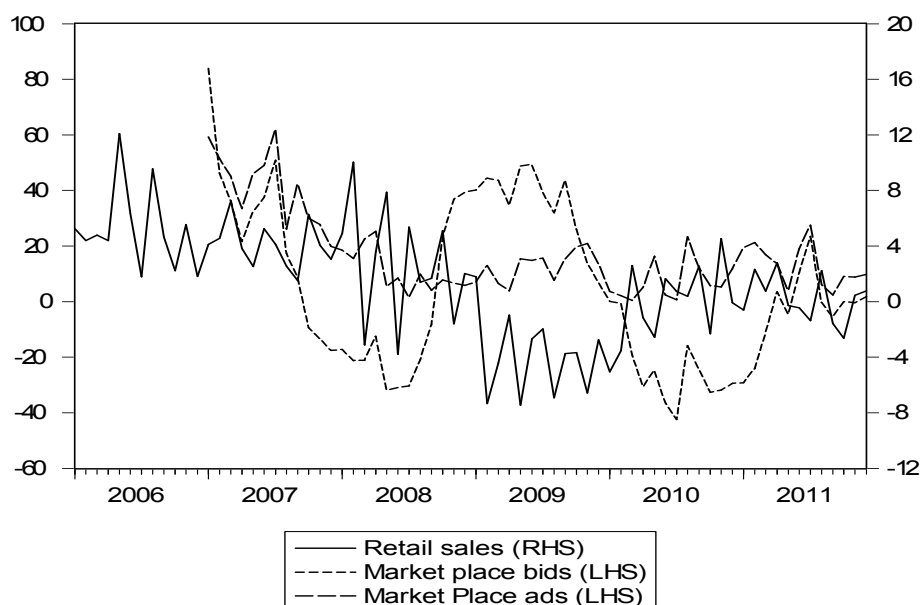
This ratio exhibits a clear downward trend, apart from an aberrant period in 2009. All in all, Marktplaats has shown remarkable growth, from 3 million advertisements put up in January 2006 to 9 million in January 2011 (because of seasonal effects, it is best to make comparisons to corresponding periods in the year before). Growth in these years was most explosive in 2006 and 2007, as is also visible in graph 5.1.3 which shows the year on year growth rate in the number of Marktplaats advertisements and bids. This explosive uptake of Marktplaats is exactly what makes it difficult to determine its usability as a business cycle indicator. For this phase of its existence, the trend in online activity on Marktplaats is a mixture of the autonomous trend of Marktplaats diffusion, a general trend in the uptake of internet usage

⁶ Values like highest bid, ask price etc. are available on a per item level, but these are not used in the analysis.

and online shopping, and a trend due to general economic conditions. The first trend component is likely to disappear as the website further matures; efforts to econometrically separate the two have proven problematic. For now, this mixture of dynamics makes analysis difficult. Graph 5.1.3 shows the growth of Marktplaats activity together with the growth rate of the volume of retail sales in The Netherlands.

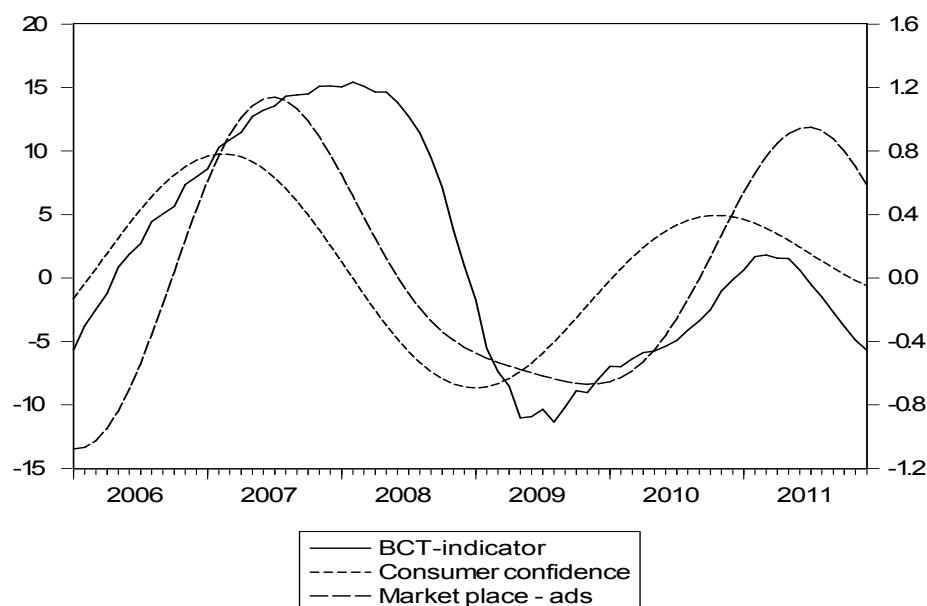
Overall, the visual match is not very strong, and as a complicating factor, the retail and advertisements series are rather noisy. There is however a very interesting difference in behaviour between the advertisements series and the bids series. Whilst the growth rate of ads seems to have a very rough similarity to the development of the growth rate of retail sales (though with a much larger amplitude), the growth rate of the number of bids seems to be inversely related to the growth rate of retail sales. This is reflected in the respective correlations with the growth rate of total retail sales, 0.46 and with a lead of seven months for ads, and -0.28 at a lead of three months for bids. Both correlations are rather low, but the different signs are a first confirmation of the difference in relationship with retail sales. It is also interesting and potentially useful that Marktplaats activity seems to be leading.

5.1.3 Relative year-on-year growth rate of the total volume of newly placed consumer goods related advertisements and bids on Marktplaats (left hand side axis) compared to the growth rate of the volume of retail sales (right hand side axis).



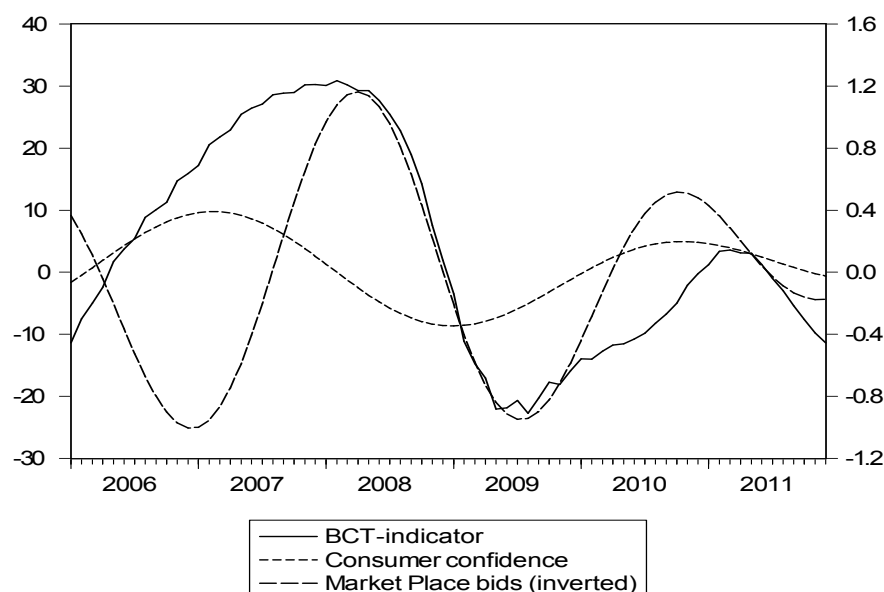
As volatility is a potential obstacle to the use of total Marktplaats activity as a short-term economic indicator, a different approach might be more fruitful. It is common practice in business cycle analysis to use so-called band-pass filters, which remove unwanted short-term effects (noise, seasonality), and also abstract from long-term trend developments. Here, the Christiano-Fitzgerald band-pass filter has been applied to the indices of total Marktplaats advertisements and bids. The lower limit was set at 2 years, the upper at 5, as the time series is not long enough to allow for a higher limit. The results are shown in graphs 5.1.4 and 5.1.5, together with the cycle of consumer confidence and the aggregated indicator from the Statistics Netherlands Business Cycle Tracer, which represents the Dutch business cycle.

5.1.4 Cycle of total volume of Marktplaats advertisements (Christiano-Fitzgerald band pass filter, 2-5 years), compared to the Dutch business cycle as represented by the Statistics Netherlands Business Cycle Tracer indicator and filtered cycle of consumer confidence.



The difference in behaviour between advertisements and bids is present again, the cycle of bids having to be inverted to achieve an optimal match with the two business cycle indicators. Overall, the cycles from both Marktplaats indicators, and especially for advertisements, are very similar to the Dutch business cycle as represented by the Business Cycle Tracer indicator. The cyclical pattern in advertisements in the period 2006-2011, consisting of a peak, a trough and a peak exactly match the cyclical chronology of the Dutch business cycle. There is little or no evidence of unrelated idiosyncratic cycles, or uninformative mini-cycles in the Marktplaats indicators.

5.1.5 Cycle of total volume of Marktplaats bids (Christiano-Fitzgerald band pass filter, 2-5 years), compared to the Dutch business cycle as represented by the Statistics Netherlands Business Cycle Tracer indicator and filtered cycle of consumer confidence.



This shows that there is a clear and relevant business cycle component in Marktplaats activity. Unfortunately, the location of cyclical maximums and minimums shows no clear lead profile when compared to those of the Business Cycle Tracer indicator. Marktplaats advertisements are leading in one turning point (compared to the BCT-indicator) and lagging in two others, whilst the bids cycle is once leading, once coincident and once lagging. A possible explanation for this varying behaviour is that although Marktplaats activity is linked to economic activity, this is not a fixed or firm link, as for example for production and consumption.

The link with consumer confidence is stronger than with the business cycle, another interesting result, and consumer confidence leads Marktplaats activity. The results of this analysis are confirmed and summarized in the correlation analysis presented in table 5.1.6.

5.1.6 Correlations of cycles (Christiano-Fitzgerald 2-5 years) of Marktplaats total ads and total bids with Dutch business cycle (Statistics Netherlands Business cycle tracer indicator) and consumer confidence (cycle). Leads and lags between parentheses.

	BCT-indicator	Consumer confidence
Marktplaats - total ads	0.605 (0)	0.827 (+6)
Marktplaats - total bids	-0.478 (+3)	-0.557 (+10)

Leads and lags in months, + means Marktplaats indicator is lagging, - means leading.

The analysis so far has been somewhat qualitative, and seems to indicate that there is a link present between Marktplaats activity and retail sales, but that the connection is not all that strong and somewhat ambiguous. By using more formal time series techniques, the nature of the connection can be further investigated. The focus is on whether Marktplaats data contain relevant information for developments in retail trade.

5.1.7 Results of tests for presence of unit roots in levels of Marktplaats total ads and bids data and volume of retail sales.

Levels	Augmented Dickey-Fuller	Phillips-Perron	KPSS-test	ERS point optimal
Retail; total sales	-1.67 (0.440)*	-7.29 (0.000)	0.206	677.6*
Marktplaats - total ads	-2.63 (0.092)*	-2.19 (0.210)*	1.072*	1623.1*
Marktplaats - total bids	-3.57 (0.001)	-3.45 (0.012)	0.337	9.07*

*Indicates presence of unit root at the 5% level.

To start with, it is necessary to test whether the indicators considered here are stationary or contain a unit root, i.e. a stochastic trend. Table 5.1.7 presents the outcomes of various unit root tests. The conclusion is that Marktplaats advertisements contain an unit root, Marktplaats bids do not, and on balance total retail sales are integrated as well.

These outcomes have a number of consequences for the modelling strategy. As both the total volume of retail sales and Marktplaats advertisements contain a unit root, it is possible that these series are co-integrated, meaning that estimating a vector error correction model would be the most efficient modelling strategy in that case. The results presented in table 5.1.8 indicate that on balance, this is not the case. Total retail sales and Marktplaats advertisements do not contain the same stochastic trend.

5.1.8 Co-integration tests for Marktplaats total ads, and total volume of retail sales.

Data Trend:	None	None	Linear	Linear	Quadratic
<i>Selected (0.05 level*) Number of Co-integrating Relations by Model</i>					
Test Type	No Intercept	Intercept	Intercept	Intercept	Intercept
	No Trend	No Trend	No Trend	Trend	Trend
Trace	0	1	0	1	2
Max-Eig	0	0	0	0	2
<i>Log Likelihood by Rank (rows) and Model (columns)</i>					
0	347.7322	347.7322	343.6094	343.6094	343.1599
1	343.5486	341.4810	337.6551	334.6543	334.4786
2	342.7929	337.5653	337.5653	330.5915	330.5915
<i>Akaike Information Criteria by Rank (rows) and Model (columns)</i>					
0	13.41465	13.41465	13.34269	13.34269	13.39525
1	13.40843	13.37224	13.27644	13.20862*	13.23656
2	13.51840	13.40899	13.40899	13.24039	13.24039
<i>Schwarz Criteria by Rank (rows) and Model (columns)</i>					
0	15.10485	15.10485	15.10332*	15.10332*	15.22630
1	15.23948	15.23850	15.17792	15.14531	15.20846
2	15.49030	15.45132	15.45132	15.35314	15.35314

Lag length 12 determined by LR, AIC, SC lag length criteria

In this case, a useful type of model to estimate is a Vector Auto-regression (VAR). First, the data are rendered stationary by computing relative year on year growth rates. This differencing operation removes seasonal effects, as well as trends. The results for the VAR of retail sales and Marktplaats advertisements are presented in tables 5.1.9 and 5.1.10. Lag length criteria indicated that a lag length of three was required. The fit statistics in table 5.1.14 do not indicate an especially good fit. Confirming this are the outcomes of the Granger causality tests in table 5.1.10. These show that in the context of this VAR, there is no significance information on retail sales in Marktplaats ads and vice versa. Further analysis of this model was therefore not worthwhile.

5.1.9 Estimation results VAR model Marktplaats total ads and total volume of retail sales (year on year growth rates)

	Retail sales	Marktplaats total ads
R-squared	0.576759	0.568036
Adj. R-squared	0.525970	0.516201
Log likelihood	-132.1040	-201.5719
Akaike AIC	4.880841	7.318311
Schwarz SC	5.131742	7.569212
Log likelihood		-331.9567
Akaike information criterion		12.13883
Schwarz criterion		12.64063

The situation was rather different in the VAR of retail sales and Marktplaats bids (both again in year on year growth rates). Lag length was 10, and as table 5.1.11 shows the fit was much better. According to the results from the Granger causality tests (table 5.1.12), Marktplaats bids contain information on the development of retail sales, but much less so the other way round.

5.1.10 Granger causality/block exogeneity Wald tests VAR Marktplaats total ads and total volume of retail sales (year on year growth rates)

Dependent variable: Retail sales			
Excluded	Chi-sq	df	Prob.
Market place-ads	2.509160	3	0.4736

Dependent variable: Market place-ads

Excluded	Chi-sq	df	Prob.
Retail sales	1.178330	3	0.7582

5.1.11 Estimation results VAR model Marktplaats total-bids and total volume of retail sales (year on year growth rates)

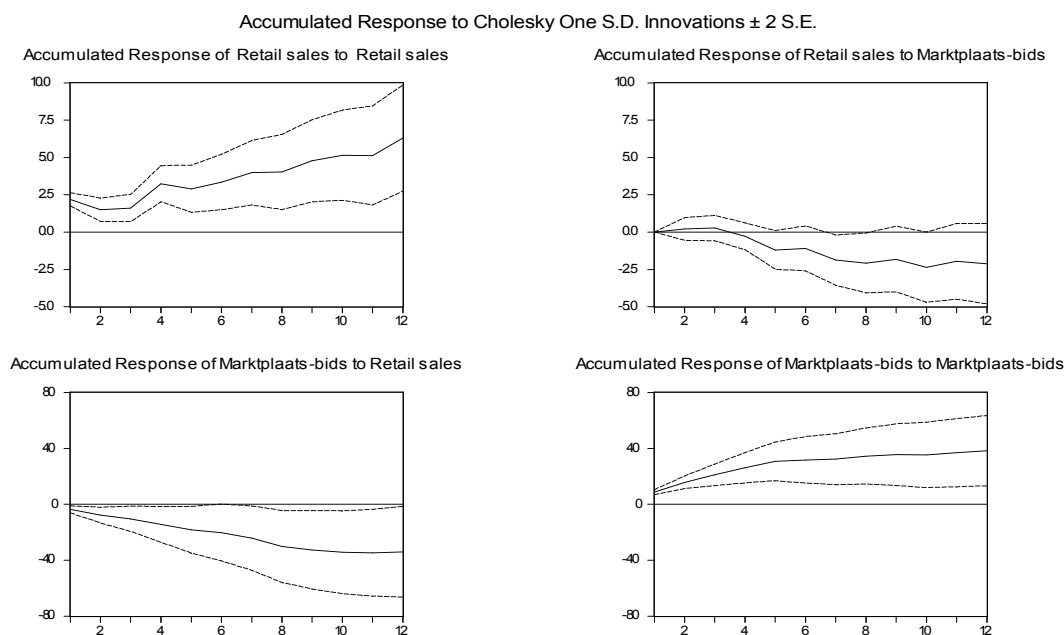
	Retail sales	Marktplaats total bids
R-squared	0.801445	0.929726
Adj. R-squared	0.664511	0.881261
Log likelihood	-96.69161	-169.4952
Akaike AIC	4.707665	7.619806
Schwarz SC	5.510714	8.422856
Log likelihood		363.8930
Akaike information criterion		122.4136
Schwarz criterion		-262.0790

5.1.12 Granger causality/block exogeneity Wald tests VAR Marktplaats total bids and total volume of retail sales (year on year growth rates)

Dependent variable: Retail sales			
Excluded	Chi-sq	df	Prob.
Market place-bids	28.66463	10	0.0014
Dependent variable: Market place-bids			
Excluded	Chi-sq	df	Prob.
Retail sales	15.78821	10	0.1059

These results are strong enough to warrant the performance of an impulse-response analysis. This summarizes the response of one series to a change (impulse) in the other one, in the context of this VAR. Graph 5.1.13 shows the cumulative effect over time (y-axis) of a standardized change in one of the series on itself and the other series. The impulse-response graphs show that there is a clear and consistent response in both cases, i.e. a real influence. Even more interesting, the results from the VAR confirm the inverse relationship observed earlier; an increase in the growth rate of Marktplaats bids results in a decrease in the growth rate of retail sales. This could be a symptom of crowding out, or just reflect switches between new and second hand purchases as economic conditions change.

5.1.13 Impulse-response analysis from VAR of Marktplaats total bids and total volume of retail sales. Upper-right; response of retail sales to impulse of Marktplaats, lower-left; response of Marktplaats to impulse of retail sales.



Another modelling strategy to investigate the links between indicators is to test for predictive content using ARMA-models. Forecasting is not only interesting in itself, but also provides another angle from which to analyze the relationship between two variables. If one variable contains information useful in forecasting another, this indicates that a causal relationship could be present, especially when the explanatory variable possesses a clear lead. In this sense, the analysis performed here is a variant on Granger causality. At the least, having forecasting power shows that a series contains information on the target indicator. The retail sales series was chosen as target indicator as this is of general interest for sectoral and business cycle analysis. Also, the Marktplaats indicators are potentially available on a much higher frequency and much more timely, making them the logical exogenous variables. Significance in a time series model does not mean all that much by itself; the performance of an exogenous variable needs to be compared to alternatives to assess its explanatory power.

For the retail series, a univariate ARMA model was estimated as the baseline model, (using year-on-year growth rates). Then, it was tested whether a variable of interest, for example a Marktplaats indicator, was significant in the model. The strongest test is whether an indicator improves the forecasting performance of the ARMA model, measured here as the root mean square forecasting error from one step ahead forecasts. One month ahead forecasts were extracted from a rolling regression framework over the period 2011:01-2011:12. This exercise was performed for both Marktplaats ads and bids, and consumer confidence and economic expectations were tested as predictors as well to supply some context for the Marktplaats results. The outcomes are summarized in table 5.1.14.

5.1.14 Results ARMA modeling and forecasting (rolling regression over 2011:01 to 2011:12) total volume of retail sales (relative year on year growth rate)

Exogenous variable	AR-terms	MA-terms	Coef. exogenous (sig.)	Lag*	R ²	AIC	Jarque-Berra (prob.)	Q-stat (lag 12)	Forecast RMSE
-	3, 5	1, 12	-	-	0.75	4.35	27.85 (0.000)	13.31 (0.102)	1.99
Marktplaats – ads (yoy)			n.s.						
Marktplaats – bids (yoy)	3, 5	1, 12	-0.069 (0.0000)	-3	0.79	4.19	0.37 (0.370)	13.79 (0.087)	1.75
Consumer confidence [†]	3, 5	1, 12	0.165 (0.0009)	-6	0.77	4.22	0.653 (0.721)	13.17 (0.106)	1.94
Future economic situation [†]	3	12	0.053 (0.0224)	-10	0.73	4.37	0.974 (0.615)	22.16 (0.012)	1.89

[†]=levels, *=lag in months, - means lead, RMSE=root mean square error.

The results from the ARMA modeling confirm the earlier conclusions. The development of Marktplaats advertisements does not contain useful information on the development of retail sales. The bids data on the other hand are highly significant, and improves the forecasting performance of the model. The information in Marktplaats bids is even more relevant than that in the consumer confidence and economic expectations series. In agreement with earlier results, the negative coefficient of Marktplaats bids indicates an inverse relationship.

The overall conclusion from this general analysis is that there are real links between Marktplaats activity and economic developments, though they are not easy to summarize. Complicating factor is the fact that there are two sides to marketplace activity; offering goods and buying them. Offers take the form of advertisements, and there is no fixed link between the number of advertisements and the number of sales, or even the number of items on offer. The number of Marktplaats advertisements exhibits a strong rising trend, but this series seems to have no relationship, either positive or negative with total retail sales. But the cyclical component, when filtered from the original data, shows a clear similarity to the Dutch business cycle. This means that Marktplaats advertisement activity is linked to general economic activity.

The other side of Marktplaats activity is on the buying side, where there is the option to sell items via an auction. In this case, a registered highest bid is assumed to represent a transaction. Therefore, these can be considered to represent actual sales. But as auctions are an option, and even auctions can be concluded by phone or email only part of Marktplaats transactions will be registered in this way, and it is unknown how large a part. The behaviour of the Marktplaats total bids series is markedly different from the advertisements series. The latter exhibits a clear rising trend, while the number of bids seems to be roughly stationary, from 2007 onwards. Again, analysis of the cyclical component shows a real relationship with the Dutch business cycle, but less strong and anti-cyclical. This means that cyclical bidding activity picks up when the economy slows down and vice versa. This inverse relationship between Marktplaats bidding activity and economic developments is confirmed when confronting the Marktplaats bidding data with the volume of retail sales in a formal modelling exercise. Both in a VAR model and ARMA model, the relationship between Marktplaats bids and retail sales is a negative one; i.e. an increasing growth in successful Marktplaats bids means a decreasing growth of retail sales.

The relationship is very strong, there is clear predictive information in Marktplaats bids concerning the development of retail sales. A very strong disclaimer is necessary regarding all these conclusions; the analysis is based on only five years of data, a period which contains the continuing uptake of online shopping and a major economic crisis. All conclusions are tentative. Furthermore, there is the issue that the Marktplaats series and total retail sales are rather heterogeneous in composition, which results in only partly overlapping aggregates. A more focussed analysis might yield more robust results. Therefore, in the next section the exercise will be repeated for a selected number of categories, where Marktplaats and retail aggregates have a relatively good match; consumer electronics.

5.2 Single category analysis: Consumer Electronics

Total Marktplaats advertisements is a rather heterogeneous aggregate, even after filtering out non-consumer related categories such as job vacancies, business services, contact ads etc., as was done here. Therefore, the link with retail sales is not necessarily very strong. It seems probable that a stronger link between certain more specific advertisement categories and corresponding retail sales sub-aggregates is present. In this section, this will be tested for a specially constructed Marktplaats aggregate containing data on advertisements in the consumer electronics category. This category contains among others audio and TV equipment. A basic question is what the relationship between online-advertisements of consumer goods and retail sales of these same consumer goods (SBI 474) should be.

Four main hypotheses have been put forward:

- (1) People put up an add to sell their old appliances when buying a new one. This means that Marktplaats activity should be coincident or slightly lagging retail sales (positive relationship).
- (2) Marktplaats is just an alternate selling channel for retailers, and should therefore exhibit the same dynamics as retail sales (positive relationship, possibly with co-integration).
- (3). Marktplaats is a competitor of “normal” retail sales, and should therefore crowd out retail sales.
- (4) Second hand-goods available via Marktplaats are bought instead of new, store bought goods.

Both (3) and (4) suggest a negative relationship, but will be hard to distinguish with the data available here. Of course, it is probable that a mixture of all these effects is presents, which will make it difficult to draw any hard conclusions. A complicating factor is that Marktplaats, like eBay, is used by professional sellers of new goods, in addition to consumers selling and buying second hand goods. Using several modelling techniques, it will be tested if any conclusions can be drawn on the relation between Marktplaats activity and retail sales.

As all these series are more or less trending, not much can be learned from graphs of the level series. As table 5.2.1 shows, the Marktplaats advertisements and retail sales series contain an unit root.

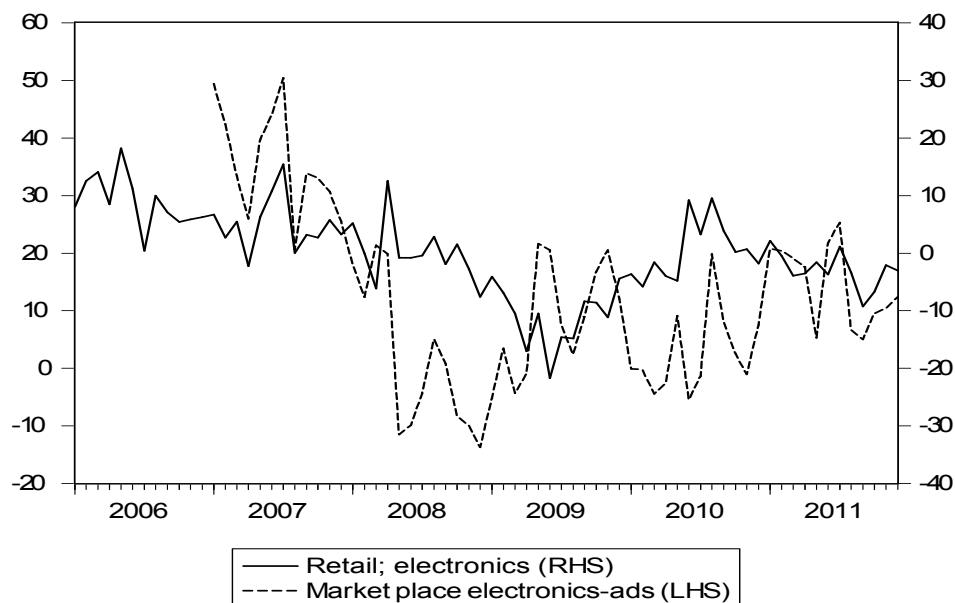
5.2.1 Results of tests for presence of unit roots in levels of Marktplaats category data and retail sales sub-aggregates (volume).

Levels	Augmented Dickey-Fuller	Phillips-Perron	KPSS-test	ERS point optimal
Retail; consumer electronics	-2.58 (0.102)*	-4.97 (0.001)	1.082*	1118.8*
Marktplaats electronics ads	-1.90 (0.327)*	-2.29 (0.177)*	0.946*	23.23*
Marktplaats electronics bids	-3.15 (0.03)	-3.19 (0.02)	0.129	1.59

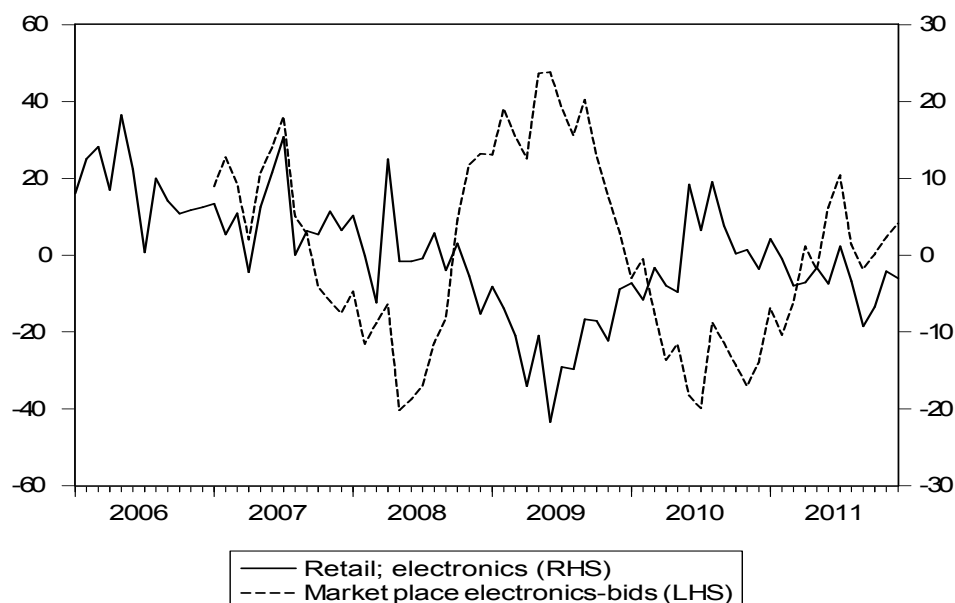
*Indicates presence of unit root at the 5% level

Therefore, a transformation is required to render them stationary. Year-on-year growth rates are the standard reporting format for the short-term retail sales statistics of Statistics Netherlands, and this transformation has the added advantage of removing seasonality. Graphs 5.2.2 and 5.2.3 show the series transformed to relative year-on-year growth rates.

5.2.2 Relative year-on-year growth rate of the volume of consumer electronics advertisements on Marktplaats compared to the growth rate of the volume of consumer electronics retail sales.



5.2.3 Relative year-on-year growth rate of the volume of consumer electronics bids on Marktplaats compared to the growth rate of the volume of consumer electronics retail sales.



The graphs suggest the presence of a relationship between retail sales of the consumer electronics category and the Marktplaats series, though markedly different for advertisements than for bids. The fit is far from perfect though, for a start the Marktplaats data possess a much higher average growth rate. The relatively weak match is reflected in the relatively low correlation of 0.484 (lead of 6 months) for

ads and -0.51 (lead of 2 months) for bids, suggesting a negative relationship. For more firm and specific conclusions on these relationships, formal modelling is required.

Given the results of the unit root tests, the integrated level series of Marktplaats advertisements and retail sales of consumer electronics were tested for co-integration. The results can be found in table 5.2.4.

5.2.4 Co-integration tests for Marktplaats consumer electronics ads, and retail sales of consumer electronics.

Data Trend:	None	None	Linear	Linear	Quadratic
<i>Selected (0.05 level*) Number of Cointegrating Relations by Model</i>					
Test Type	No Intercept No Trend	Intercept No Trend	Intercept No Trend	Intercept Trend	Intercept Trend
Trace	0	1	1	1	2
Max-Eig	0	1	1	0	0
<i>Log Likelihood by Rank (rows) and Model (columns)</i>					
0	863.6033	863.6033	861.8672	861.8672	861.3695
1	860.4355	855.1587	853.7424	853.3423	852.8447
2	859.3336	852.2027	852.2027	848.3382	848.3382
<i>Akaike Information Criteria by Rank (rows) and Model (columns)</i>					
0	29.62634	29.62634	29.63499	29.63499	29.68424
1	29.65362	29.51340	29.49975*	29.51942	29.53589
2	29.74864	29.58042	29.58042	29.51929	29.51929
<i>Schwarz Criteria by Rank (rows) and Model (columns)</i>					
0	31.01052*	31.01052*	31.08838	31.08838	31.20684
1	31.17622	31.07060	31.09156	31.14583	31.19691
2	31.40966	31.31064	31.31064	31.31872	31.31872

Lag length 10 determined by LR, AIC, SC lag length criteria

There are some indications of the presence of a common trend. However, the Schwarz criteria selects no co-integration and the AIC selects a relatively complex model. Given the relatively short period of data available, co-integration testing is bound to be difficult. No co-integration is the simpler hypothesis, therefore, the presence of a common trend is rejected here. Instead of a vector error correction model, a regular vector auto-regression will be estimated. Lag length criteria indicated the use of 13 lags in the VAR, and the fit results are presented in table 5.2.5. The R^2 values seem respectable, although that is not surprising given the inclusion of 14 lags in the equations. The results from the Granger causality/exogeneity test give a somewhat different picture. The conclusion is that retail sales do contain significant information on Marktplaats developments, but not the other way round.

5.2.5 Estimation results VAR model Marktplaats consumer electronics ads and retail sales of consumer electronics

	Retail CE	Marktplaats CE - ads
R-squared	0.936415	0.873328
Adj. R-squared	0.831686	0.664691
Log likelihood	-113.3443	-103.8633
Akaike AIC	6.188884	5.776667
Schwarz SC	7.341723	6.929506
Log likelihood		-216.3463
Akaike information criterion		11.92810
Schwarz criterion		14.23378

Lag length 14 determined by LR, AIC, SC lag length criteria

5.2.6 Granger causality test Marktplaats consumer electronics ads and retail sales consumer electronics (year on year growth rates)

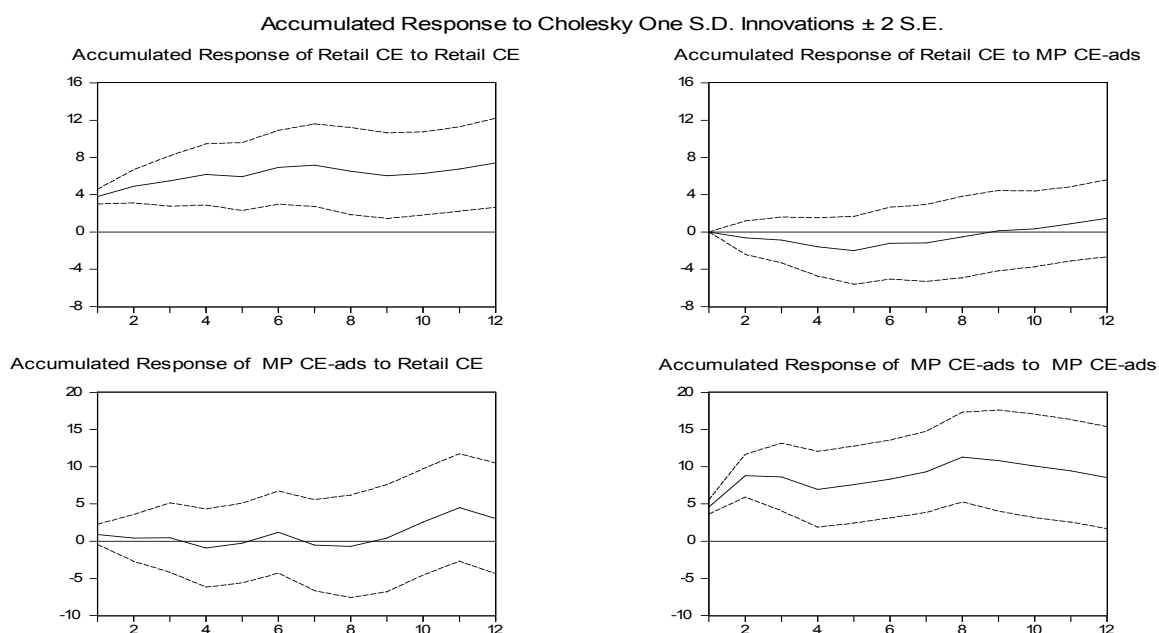
Pairwise Granger Causality Tests

Lags: 14

Null Hypothesis:	Obs	F-Statistic	Prob.
Marktplaats CE - ads does not Granger Cause retail sales CE	46	1.43483	0.2374
Retail sales CE does not Granger Cause Marktplaats CE - ads		4.73572	0.0016

These conclusions are confirmed by the impulse-response analysis, which show that the responses include zero in the 95% confidence band. The relationship can roughly be characterized as weak but positive.

5.2.7 Impulse-response analysis from VAR model of Marktplaats consumer electronics ads and retail sales of consumer electronics (year on year growth rates). Upper-right; response of retail sales to impulse of Marktplaats., lower-left; response of Marktplaats to impulse of retail sales



Marktplaats bids in the consumer electronics category did not contain a stochastic trend, therefore a regular VAR is the model of choice. Both retail data and Marktplaats data were entered as year on year growth rates to assure comparability. Lag length criteria indicated the use of 15 lags, resulting in rather high R^2 values. However, in contrast with the results for Marktplaats advertisements, in this case the Granger causality/exogeneity test shows that both series contain significant information for the other one, see table 5.2.8.

5.2.8 Estimation results VAR model Marktplaats consumer electronics bids and retail sales of consumer electronics

	Retail CE	Marktplaats CE-bids
R-squared	0.932818	0.987205
Adj. R-squared	0.788857	0.959786
Log likelihood	-87.75209	-114.0657
Akaike AIC	5.277871	6.447364
Schwarz SC	6.522460	7.691954
Log likelihood		-199.8641
Akaike information criterion		11.63841
Schwarz criterion		14.12759

Lag length 15 determined by LR, AIC, SC lag length criteria

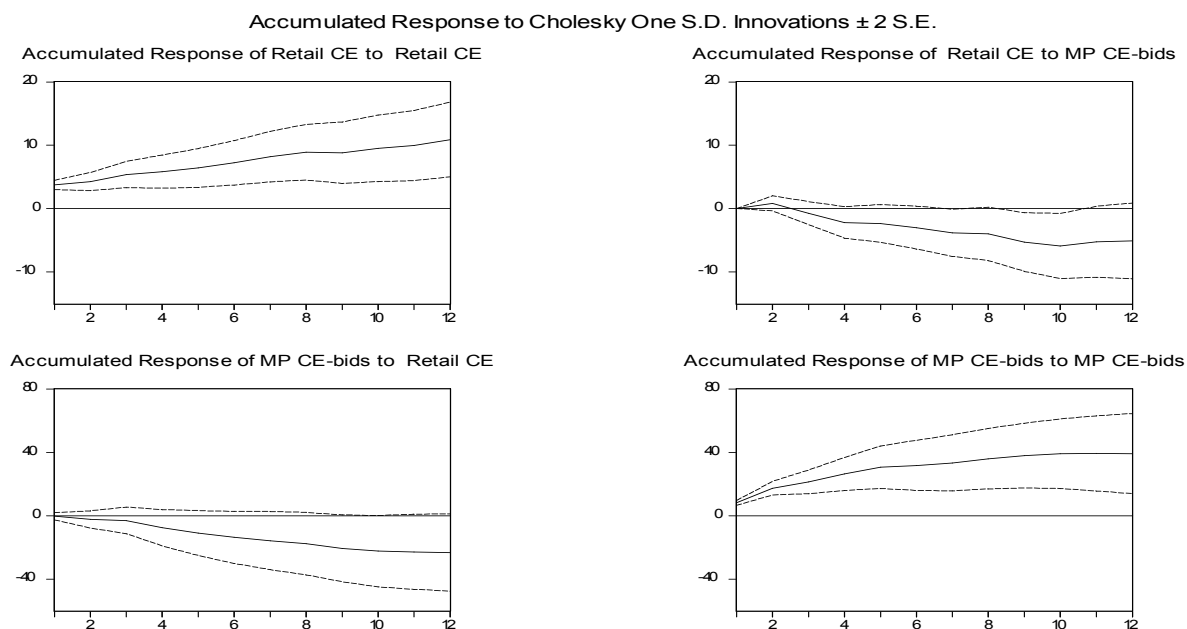
5.2.9 Granger causality test Marktplaats consumer electronics bids and retail sales consumer electronics (year on year growth rates)

Pairwise Granger Causality Tests

Lags: 15			
Null Hypothesis:	Obs	F-Statistic	Prob.
Marktplaats CE -bids does not Granger Cause retail sales CE	45	2.78318	0.0315
Retail sales CE does not Granger Cause Marktplaats CE -bids		5.32745	0.0016

This interesting result is confirmed by the impulse-response analysis (graph 10), which shows for both series a significant response to a change in the other series. Just as for the total aggregates, for consumer electronics the estimated relationship between Marktplaats bids and retail sales is a negative one, an increase in the growth rate of Marktplaats bids resulting in a decrease of the growth of retail sales.

5.2.10 Impulse-response analysis from VAR model of Marktplaats consumer electronics bids and retail sales of consumer electronics (year on year growth rates). Upper-right; response of retail sales to impulse of Marktplaats, lower-left; response of Marktplaats to impulse of retail sales.



Finally, to further analyze the relation between retail sales of consumer electronics and Marktplaats activity from a forecasting perspective, an ARMA modeling exercise was performed. Just as for the total aggregates, it was tested whether the Marktplaats data can enhance the forecasting performance of

an ARMA-model. Again, the model was based on the year-on-year growth rates and one month ahead forecasts were extracted from a rolling regression framework over the period 2011:01-2011:12. The results are presented in table 5.2.11.

5.2.11 Results ARMA modeling and forecasting retail sales of consumer electronics (volume, relative year on year growth rate)

Exogenous variable	AR-terms	MA-terms	Coef. exogenous (sig.)	Lag*	R ²	AIC	Jarque-Berra (prob.)	Q-stat (lag 12)	Forecast RMSE
-	1, 3	12	-	-	0.76	5.57	4.218 (0.121)	11.15 (0.266)	3.75
Marktplaats – ads (yoy)	1, 3	12	0.158 (0.0302)	0	0.80	5.30	22.60 (0.000)	16.42 (0.059)	2.89
Marktplaats – bids (yoy)	1, 3	12	-0.181 (0.0000)	-2	0.85	5.02	15.15 (0.000)	9.18 (0.421)	2.96
Consumer confidence [†]	1, 3	12	0.311 (0.0003)	-5	0.80	5.30	10.20 (0.006)	6.80 (0.658)	3.11
Future economic situation [†]	1, 3	12	0.107 (0.0112)	-9	0.77	5.42	0.434 (0.805)	8.19 (0.515)	3.74

[†]=levels, * =lag in months, - means lead, RMSE=root mean square error.

These estimation results show that the Marktplaats online advertising data are significant in explaining retail sales, and that they have considerable forecasting power. The Marktplaats indicators improve forecasting performance not only relative to the baseline model, but also compared to the models containing consumer sentiment indicators. In the case of sales of consumer electronics, the coefficient of Marktplaats bids is negative with a small lead of two months. This indicates that higher growth of online activity signals lower growth of retail sales. This could mean crowding-out or substitution of retail sales by Marktplaats. In the case of Marktplaats advertisements, the Marktplaats indicator has a positive coefficient without a lead. This indicates that Marktplaats advertisements activity and retail sales move more in sync.

Conclusions

Online markets like Marktplaats have become an important channel for the second hand trade, and also for the sale of new goods. The question this study tried to address is in how far online activity is connected to (general) economic activity, and whether the data from Marktplaats could yield useful macro-economic indicators. The potential links with retail sales and consumption are obvious, but an interesting possibility is a use as business cycle indicators, signaling general economic developments. Online market place data are potentially interesting in this respect because of their high frequency and timeliness. This study analyzed the Marktplaats data from two different angles; data relating to the number of advertisements vs. the number of items bought by bidding (bids), and development of total Marktplaats activity vs. that of a selected category of consumer goods. The number of advertisements is a pure supply indicator; there is no direct connection to the number of actual transactions. A highest bid in an auction indicates an actual transaction, but only a part of online offerings uses the auction option. The range of items on offer on Marktplaats, even after filtering out the consumer goods-related advertisements, is staggering, and not always relevant for the analysis of economic conditions. They range from advertisements for antiques, via boats to new kitchens. For example, stamps are a very large sub-category. A more specific category, as consumer electronics used in this study, might therefore exhibit a stronger link with economic developments than the total.

The data from the online marketplace Marktplaats are somewhat difficult to analyze. As this study shows, it takes some effort to access the useful information contained in the data. The time series available are relatively short, and the development of the Marktplaats indicators is, heavily influenced by the autonomous trend originating in the diffusion of Marktplaats as an online marketplace. On top of this, the timespan considered here, January 2006 to December 2011, contains a major economic crisis.

The first step was to test whether Marktplaats activity has a link to general economic conditions. The cycles extracted from total consumer goods related advertisements and bids were shown to have a clear resemblance to the general Dutch business cycle, as represented by the Statistics Netherlands Business Cycle Tracer indicator, though with somewhat different timing of turning points. This shows that Marktplaats activity is influenced by the business cycle. The relationship between the Marktplaats cycles and filtered consumer confidence was even stronger, though consumer confidence leads the Marktplaats indicators. But again, this indicates a real relation between Marktplaats activity and macro-economic developments.

Further quantitative tests yielded a number of interesting results. The most striking outcome is the difference between advertisements and bids. The number of advertisements has grown much faster than the number of bids. It is the development of the number of bids which was shown to possess the most valuable information on economic developments, more specifically on the growth of retail sales. The link between retail sales and advertisements was weak, but bids contained significant information. The difference is probably due to the fact that bids relate more to actual sales. Marktplaats bids were shown to have clear predictive power both for total retail sales and for the sales of consumer electronics. The relationship was stronger for the better matching consumer electronics aggregates, as expected. Most interesting, the relation between the development of Marktplaats bids and the development of retail sales was a negative one; higher growth of Marktplaats activity signals lower growth of retail sales. Whether this is caused by switching to Marktplaats for the purchase of new items, or a switch to used goods instead of new items bought in retail stores is not clear. Marktplaats data have shown to contain real and important information on economic developments. It remains the case however that “raw” Marktplaats indicators are relatively poor economic indicators. Perhaps more filtering of the underlying data could improve their performance. As a final note it should be mentioned that the results reported in this paper should be treated carefully, as they are based on observations from a relatively limited period of time. Therefore, all conclusions are tentative.

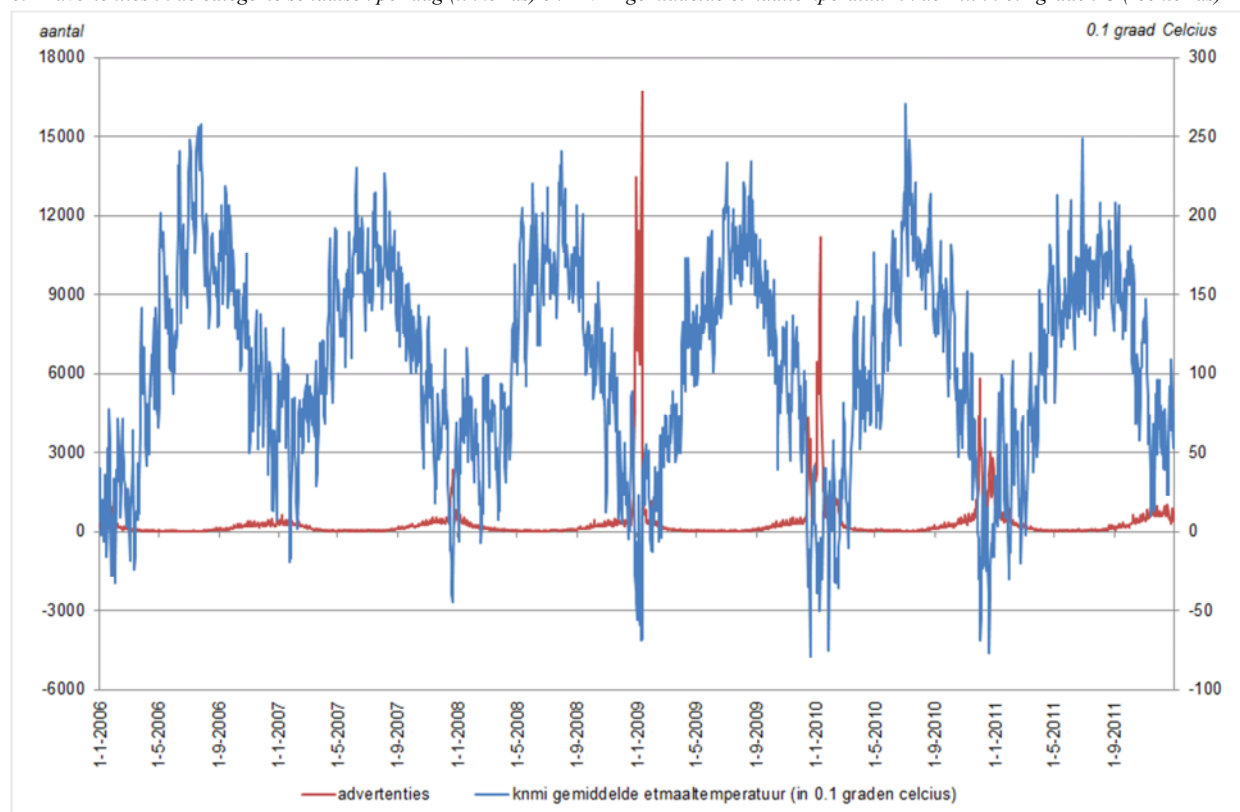
6. Wetenswaardigheden

Gedurende het werken met de van Marktplaats verkregen dataset zijn talloze zaken de revue gepasseerd. Vele figuren zijn samengesteld om ideeën en hypothesen te toetsen. Een belangrijk doel hiervan was om inzicht in en grip op de data te krijgen. Hoe zit de dataset in elkaar? Zijn de uitkomsten plausibel? Zijn ze te interpreteren? Een aantal van deze figuren wordt in dit hoofdstuk getoond ter illustratie van wat er allemaal uit de dataset gehaald kan worden, al dan niet in combinatie met andere gegevens. Het zijn niet allemaal schoolvoorbeelden van figuren die ten grondslag lagen aan het verzoek om de dataset van Marktplaats te mogen hebben, maar meer een bloemlezing uit hetgeen allemaal is ‘uitgedraaid’ aan data.

Specifieke categorieën

Sommige advertentiescategorieën zijn temperatuursafhankelijk, zoals schaatsen. Niet in alle winters worden evenveel schaatsen aangeboden en gevraagd. Het moet ook echt een koude winter zijn (2009, 2010, 2011). Het verschil tussen wel en geen (enige tijd) ijs in de sloten is extreem. In de strengere winters van 2009 tot en met 2011 werden duizenden schaatsen aangeboden. In minder strenge winters is er wel sprake van een opleving, maar lang niet in de orde van grootte als in de strenge winters. Kennelijk wordt er pas massaal geadverteerd als er ook echt ijs ligt of een langere vorstperiode is voorzien.

6.1 Advertenties in de categorie schaatsen per dag (linker as) en KNMI gemiddelde etmaaltemperatuur in de Bilt in 0.1 graden C (rechter as)

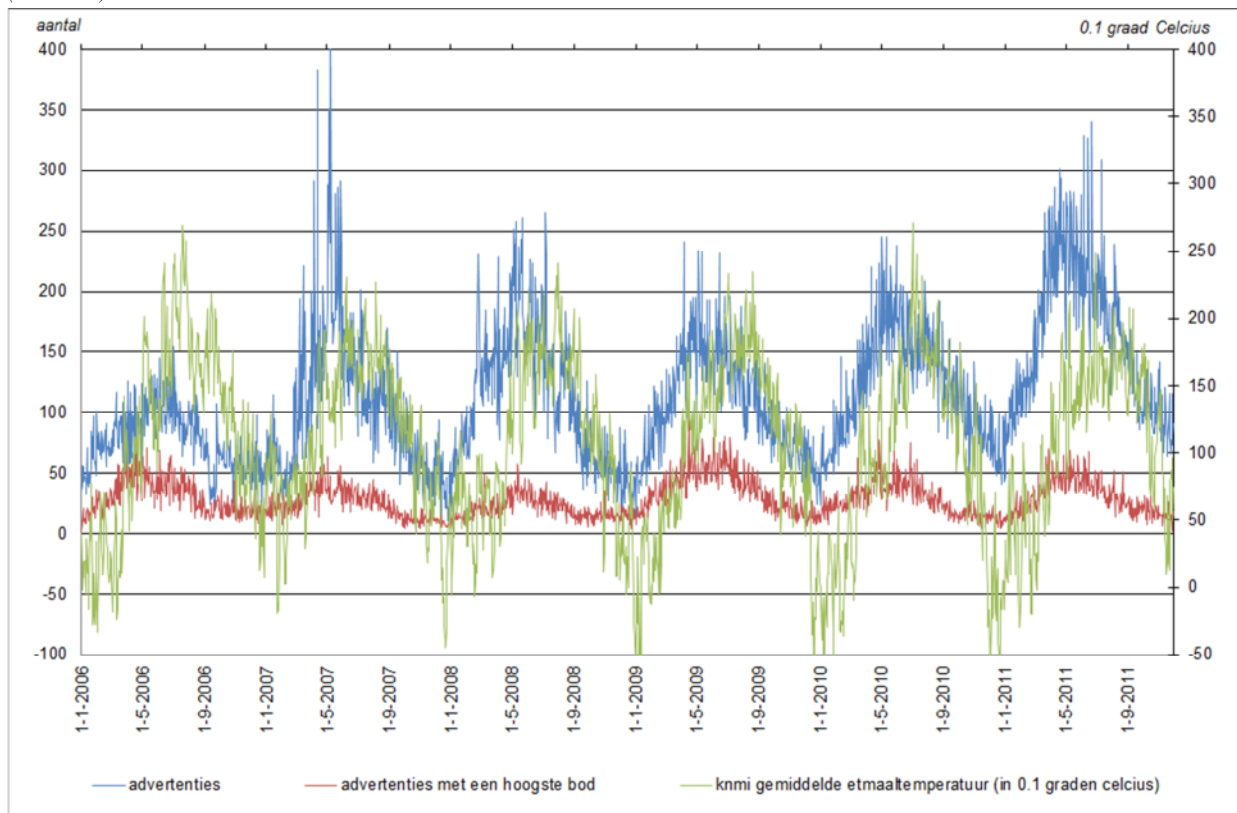


Bron: CBS (Data: Marktplaats, KNMI).

Andere advertentiecategorieën zijn meer seizoensafhankelijk zoals dameszonnebrillen en reageren niet zo extreem op temperatuur als de categorie schaatsen. In deze categorie wordt stelselmatig in het voorjaar veel geadverteerd en veel geboden. Dit soort figuren laten zien wat verwacht mag worden. Niet zozeer bijzonder verrassend, maar de dataset lijkt consistent. Het helpt om de verschillende variabelen in de dataset te begrijpen én mochten er onverhoopt in de zomer veel schaatsen worden aangeboden,

dan zou dat een aanwijzing zijn dat variabelen in de dataset niet weergeven wat werd aangenomen en dus verkeerd werden geïnterpreteerd.

6.2 Advertenties in de subcategorie dames zonnebrillen per dag (linker as) en KNMI gemiddelde etmaaltemperatuur in de Bilt in 0.1 graden C (rechter as)



Bron: CBS (Data: Marktplaats, KNMI).

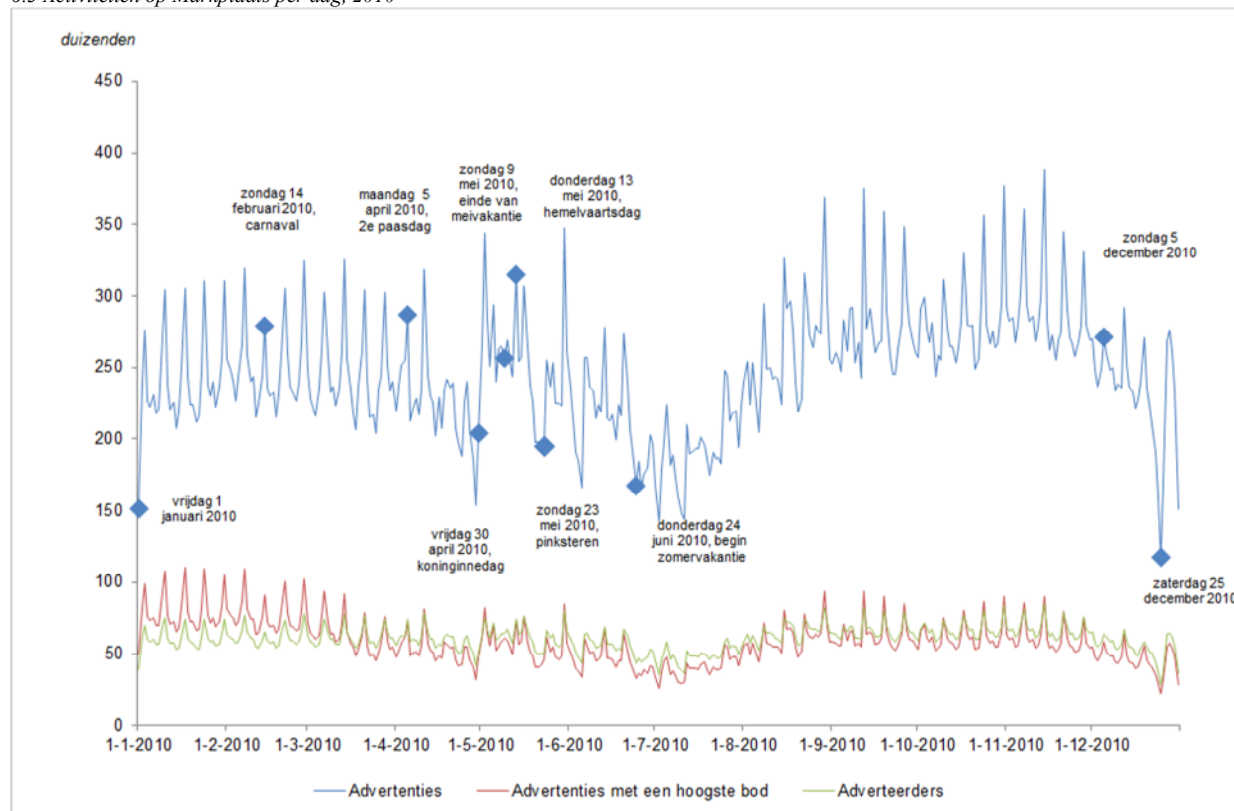
Patronen in de tijd

Ook om grip op de data te krijgen zijn gedragspatronen van adverteerders en bidders in de tijd vastgelegd, achtereenvolgens per jaar, week en etmaal.

Figuur 6.3 toont de activiteiten van adverteerders en bidders per dag voor het jaar 2010. De vakantieperiodes zijn duidelijk zichtbaar en ook de pieken en dalen op de verschillende feestdagen. Eerste Kerstdag, Nieuwsjaardag, Koninginnedag, Eerste Pinksterdag en Sinterklaas lijken hierbij feestdagen waarbij het 'not done' is om op Marktplaats te zitten; Hemelvaartsdag en tweede Paasdag juist wel. De wekelijkse pieken in de grafiek zijn de zondagen. Dit is de dag dat men het meest actief is op Marktplaats.

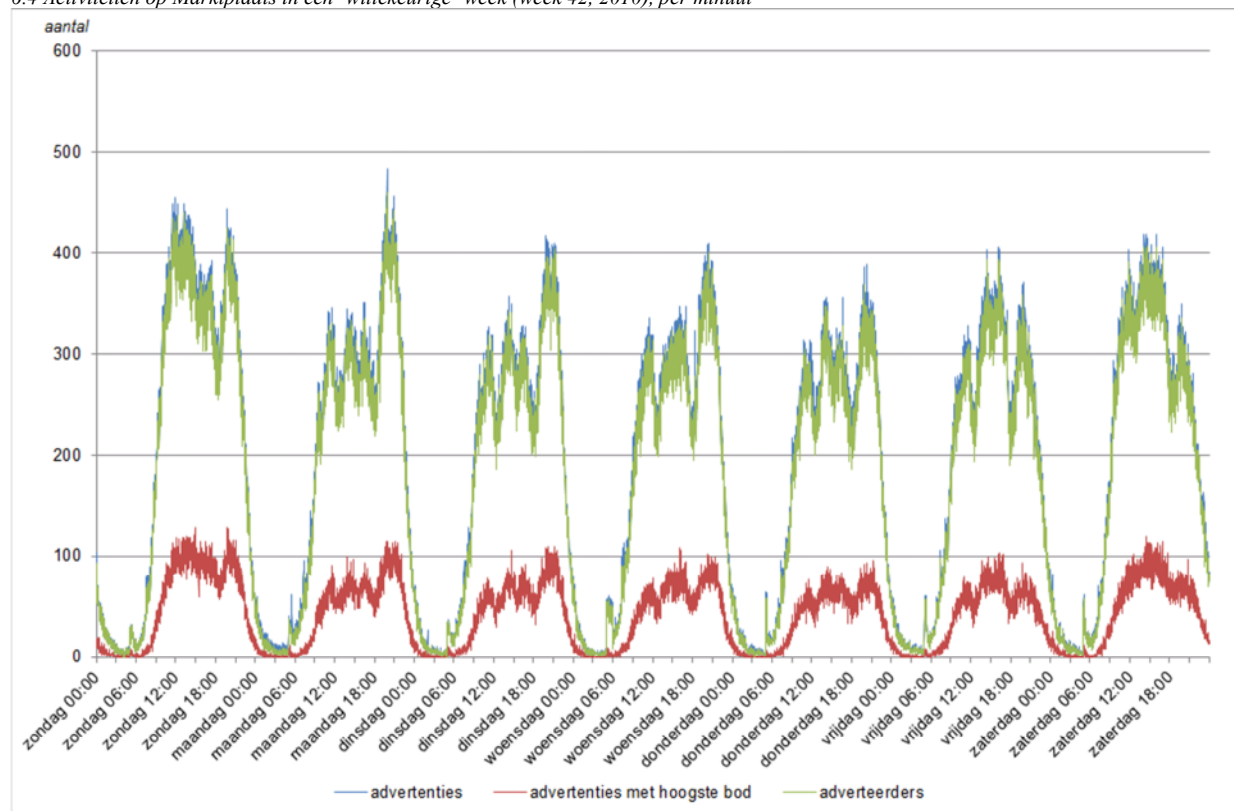
De meer dan gemiddelde activiteit op zondagen komt in figuur 6.4 terug. In figuur 6.4 zijn de activiteiten op Marktplaats weergegeven voor een 'willekeurige' week in 2010. Weergegeven is week 42, een week zonder feestdagen en ook niet in een vakantieperiode. De activiteit is weergegeven per minuut. Dit brengt met zich mee dat zondag niet direct met kop en schouders boven de andere dagen uitsteekt, maar gedurende de dag is de activiteit op zondag op elk willekeurig tijdstip groter dan op elke andere dag. Gecumuleerd over de dag levert dat een verschil op zoals te zien was in figuur 6.3.

6.3 Activiteiten op Marktplaats per dag, 2010



Bron: CBS (Data: Marktplaats).

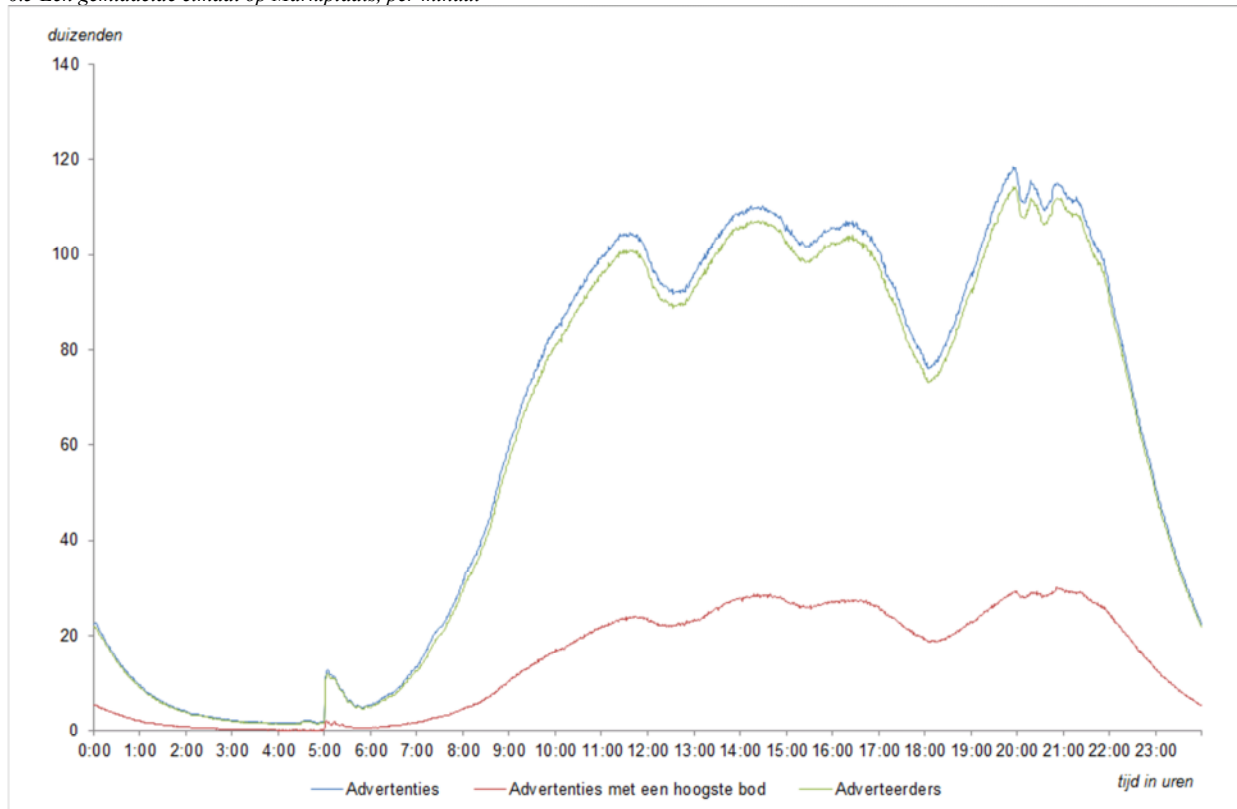
6.4 Activiteiten op Marktplaats in een 'willekeurige' week (week 42, 2010), per minuut



Bron: CBS (Data: Marktplaats).

Figuur 6.5 ten slotte toont de activiteiten op Marktplaats gedurende een gemiddeld etmaal. Momenten als ‘de lunch’(12.00 uur), ‘het avondeten’ (18.00 uur) en ‘de kinderen van school halen’ (15.00 uur) zijn duidelijk zichtbaar in de figuur. De bovengemiddelde activiteit van adverteerders en hun advertenties om 5.00 uur ’s ochtends kan duiden op bedrijfsmatige adverteerders die hun advertenties automatisch uploaden. ’s Avonds tussen 21.00 en 23.00 uur lijkt ook een periode van meer dan gemiddelde activiteit op Marktplaats. Afgezien van de piek op 5.00 uur ’s ochtends lijkt de figuur redelijk synchroon te lopen met de dagindeling van een gemiddeld gezin in Nederland. Dit geldt overigens ook voor de patronen in de ander twee figuren. De bedrijfsmatige adverteerders blijven bij dit soort patronen kennelijk onder de oppervlakte.

6.5 Een gemiddelde etmaal op Marktplaats, per minuut



Bron: CBS (Data: Marktplaats).

7. Slotbeschouwing

Dit project is voor het CBS een van de eerste ervaringen met zogenaamde ‘big data’; een grote hoeveelheid betrekkelijk ongestructureerde data die niet op voorhand met een specifiek doel is verzameld. Een eerste reflex was dan ook om de data te structureren en gaaf te maken zodat alles plausibel en begrijpelijk werd. Uiteindelijk is dit maar in zeer bescheiden mate gedaan omdat het in zekere zin in tegenspraak lijkt met het fenomeen ‘big data’. In het extreme geval ‘gooi je zo veel weg’ dat er nauwelijks nog sprake is van big data.

Een tweede ‘probleem’ dat op de achtergrond permanent aanwezig was, was het begrip representativiteit. Het CBS is gewoon statistische gegevens te verzamelen die representatief zijn voor heel Nederland, een bedrijfstak of een bepaald verschijnsel. Waar waren deze gegevens representatief voor? Misschien wel niet meer dan alleen voor zichzelf en moeten de gegevens ook als zodanig beschouwd en gepresenteerd worden?

Uiteindelijk is een aantal zaken uitgewerkt. Zo is in eerste instantie in hoofdstuk 2 de van Marktplaats verkregen data gepresenteerd zonder dat daar andere gegevens aan zijn toegevoegd. Men kan hier veel verder in gaan dan in hoofdstuk 2 is gebeurd en ook met veel meer tot de verbeelding sprekende ‘visualisaties’. De vraag die hier achterblijft is in hoeverre dit tot de kerntaken van het CBS behoort?

In hoofdstuk 3 is de Marktplaatsdata verrijkt met gegevens van het CBS. Op deze wijze ontstaat nieuwe statistische informatie. Zo is een beeld geschetst van de verspreiding van het gebruik van Marktplaats in Nederland en is aan de hand van kenmerken van 6-posities postcodegebieden geschetst welke karakteristieken bijdragen tot het gebruik van Marktplaats en welke juist niet.

In hoofdstuk 4 was er een concrete doelstelling geformuleerd. Kan op basis van de Marktplaatsdata een schatting worden gemaakt van in eerste instantie de omzet in tweedehands goederen tussen particulieren via Marktplaats? Hier is van afgezien omdat er te veel cruciale gegevens ontbraken en er derhalve te veel (arbitraire) aannames gemaakt moesten worden om tot een dergelijke schatting te kunnen komen. Daarnaast speelde hierbij de overweging dat een dergelijk ruwe schatting toch een ‘eigen leven’ zou kunnen gaan leiden.

In hoofdstuk 5 is onderzocht in hoeverre activiteiten op Marktplaats overeenkomst vertoonden met conjunctuurindicatoren zoals het CBS die samenstelt. Het ‘conjuncturele’ verloop in het aantal Marktplaatsadvertenties en het aantal biedingen kende overeenkomsten met de algemene conjunctuur en met name ook het consumentenvertrouwen. Opvallend was dat het aanbod (advertenties) positief samenhangt met de conjunctuur, maar de vraag (biedingen) negatief. Dit beeld werd bevestigd bij de samenhang tussen activiteiten op Marktplaats en de ontwikkeling van de detailhandelsomzet. Deze samenhang kwam in algemene zin uit de verf maar beter nog voor een speciaal geconstrueerde groep consumptiegoederen: de consumentenelektronica. De vraag die hier bovenkomt drijven is in hoeverre en op welke wijze het CBS gebruik zou kunnen maken van de hier geconstateerde samenhang?

Bestandsbeheer en datamanagement

Het van Marktplaats ontvangen bestand bestond uit alle in de jaren 2006 tot en met 2011 door ‘particulieren’ op Marktplaats geplaatste advertenties. Erkende zakelijke adverteerders en advertenties waar niet zozeer een product werd aangeboden, maar meer de aandacht werd gevestigd op een bepaalde site, zijn buiten beschouwing gelaten (admarkt). Dit neemt niet weg dat ook in de verkregen dataset ‘particulieren’ zaten die feitelijk een bedrijf waren of op zijn minst een bedrijfsmatige gebruiker van Marktplaats; in ieder geval niet de huis-tuin-en-keuken Marktplaatsgebruiker.

De levering bestond uit twee bestanden per jaar:

- I. Een bestand met advertenties mét een hoogste bod, 2006 t/m 2011.
- II. Een bestand met advertenties waar niet op geboden is, 2006 t/m 2011.

In tabel 1 zijn de per advertentie (record) meegeleverde variabelen vermeld.

1. Naam en omschrijving variabelen in bestand(en) Marktplaats

Naam	Omschrijving
AD-ID	Uniek identificerend nummer advertentie
SELLER_ID	Uniek identificerend nummer adverteerder
AD_ORGNL_START_DT_TM	Begindatum / -tijd van advertentie
AD_ACTL_END_DT_TM	Einddatum / -tijd van advertentie
CONDITIE	Conditie van aangeboden artikel (nieuw / z.g.a.n. / gebruikt)
ASK_PRICE_LC	Vraagprijs
ASK_PRICE_TYPE_ID	Prijstype (t.e.a.b. / vrij bieden / gratis / n.o.t.k. / op aanvraag / bieden vanaf / ruilen)
AD_STATUS_ID	Status advertentie (actief / verlopen / verwijderd door gebruiker)
L1_ID	Indeling advertenties in hoofdcategorieën van artikelen (ca. 30)
L2_ID	Indeling advertenties in categorieën van artikelen (ca. 2 200)
VIEW_AD_COUNT	Aantal keren dat advertentie is bekeken
CLSFD_AD_CNTCT_PSTL_CD	Postcode adverteerder
AD_RMVL_RSN_ID	Reden van verwijdering advertentie (verkocht via MP / elders verkocht / etc.)
XCOORD	x-coördinaat rijksdriehoeksstelsel
YCOORD	y-coördinaat rijksdriehoeksstelsel
BIDDER_ID	Uniek identificerend nummer bidder
CLSFD_BID_STATUS_ID	Status van bod
CLSFD_BID_AMT	Hoogte van het hoogste bod
CLSFD_BID_DT	Datum van het hoogste bod
CLSFD_BID_RMVL_RSN_ID	Reden verwijdering bod
CLSFD_BID_STATUS_ID2	Duplicaat van CLSFD_BID_STATUS_ID
SELLER_GENDER_ID	Geslacht adverteerder

Bron: CBS (Data: Marktplaats).

De bestanden zijn gezippt en versleuteld aangeleverd op een harddisk. De bestanden waren in ASCII-formaat en met een vaste breedte. De AD-ID's waren in wetenschappelijke notatie weergegeven. Om ruimte te besparen en de bestanden geschikt te maken voor plaatsing op het netwerk zijn deze met dank aan het Innovatielab met een Python-script omgezet naar tab-gescheiden bestanden, waarbij de losse

datum- en tijdvelden zijn samengevoegd tot één datum-tijd-veld. Getallen in wetenschappelijke notatie zijn teruggezet naar een normale numerieke notatie.

Deze stap leverde een reductie van data van 431 bytes per record naar ca. 147 bytes per record en bij de bestanden met advertenties plus bod, van 660 bytes per record naar ca. 109 bytes per record.

2. Bestandsbeheer: van dataset Marktplaats tot SQL-database (1)

	Aantal records	Bytes per record	Bestands grootte geleverde bestand					
			in bytes	in Gb	na hercoderen, in Gb	in % van origineel	na zippen, in Gb	in % van origineel
MP2006	1,8E+07	431	7,5E+09	7,0	2,3	33%	0,7	10%
NBMP2006	4,5E+07	660	2,9E+10	27,4	4,5	16%	1,2	5%
MP2007	2,1E+07	431	9,0E+09	8,4	2,6	31%	0,8	10%
NBMP2007	6,7E+07	660	4,4E+10	40,9	6,8	17%	1,9	5%
MP2008	2,0E+07	431	8,5E+09	7,9	2,5	31%	0,8	10%
NBMP2008	7,5E+07	660	4,9E+10	46,0	7,7	17%	2,1	5%
MP2009	2,8E+07	431	1,2E+10	11,0	3,5	32%	1,1	10%
NBMP2009	8,5E+07	660	5,6E+10	52,1	8,7	17%	2,5	5%
MP2010	2,2E+07	431	9,3E+09	8,7	3	34%	1,0	12%
NBMP2010	9,1E+07	660	6,0E+10	55,9	10,2	18%	3,3	6%
MP2011	2,1E+07	431	9,0E+09	8,4	2,9	34%	1,0	12%
NBMP2011	1,0E+08	660	6,8E+10	63,4	11,6	18%	3,8	6%
TOTAAL	5,9E+08	611	4,0E+11	337,0	66,2	20%	20,2	6%

Bron: CBS (Data: Marktplaats).

Na plaatsing op het netwerk zijn de bestanden ingelezen op een SQL-server, waarna de bestanden met advertenties met en zonder bod gekoppeld zijn.

3. Bestandsbeheer: van dataset Marktplaats tot SQL-database (2)

	SQL-ruimte-gebruik, in Gb	in % van originele bestand	SQL-ruimte-gebruik ná conversie, in Gb	in % van originele bestand	Tabelnaam	SQL-ruimte-gebruik ná samenvoeging bestanden, in Gb	in % van originele bestand
MP2006	2,9	41%	1,4	20%	Compleet 2006	3,3	10%
NBMP2006	5,9	21%	3,3	12%			
MP2007	3,2	39%	1,5	18%	Compleet 2007	4,9	10%
NBMP2007	8,8	22%	4,9	12%			
MP2008	3,1	39%	1,5	18%	Compleet 2008	5,5	10%
NBMP2008	10	22%	5,5	12%			
MP2009	4,3	39%	2,0	18%	Compleet 2009	6,3	10%
NBMP2009	11,3	22%	6,3	12%			
MP2010	3,6	42%	1,6	18%	Compleet 2010	6,7	10%
NBMP2010	13,1	23%	6,7	12%			
MP2011	4,7	56%	1,5	18%	Compleet 2011	7,6	11%
NBMP2011	14,8	23%	7,6	12%			
TOTAAL	85,8	25%	44,0	13%		34,4	10%

Bron: CBS (Data: Marktplaats).

Filtering van advertenties

Bij een eerste analyse van de records in het bestand zoals dat op de SQL-server stond, bleek al vrij snel dat er enkele 'vreemde' dingen aan de hand waren. Zo was de vraagprijs soms wel erg hoog. Ook waren er 'advertentiecategorieën' die niet direct tot het doel van dit onderzoek behoorden, namelijk het

schetsen van aard en omvang van de handel in (tweedehands) goederen op basis van de van Marktplaats ontvangen data. Uiteindelijk is besloten een aantal advertentiecategorieën niet mee te nemen in de verschillende analyses (Diensten, Vacatures, Contactadvertenties en Vastgoed). Daarnaast is besloten advertenties met een vraagprijs van meer dan 99 999,- euro ook niet op te nemen in het analysebestand. In tabel 4 is stap voor stap aangegeven om hoeveel advertenties het gaat en hoe de filtering de kenmerken van de oorspronkelijk van Marktplaats ontvangen dataset heeft beïnvloed.

4. Filtering van advertenties

	2006	2007	2008	2009	2010	2011
<i>Advertenties Marktplaats</i>						
Aantal advertenties (mln)	44,6	66,5	74,8	84,7	91,0	103,1
Advertenties met vraagprijs > 0,1 €ct (mln)	22,6	33,3	37,8	44,8	50,1	60,2
Advertenties met een hoogste bod (mln)	17,5	20,8	19,6	27,5	21,6	20,9
Som alle vraagprijzen (mld €)	25,0	46,8	55,1	121,2	31,7	34,8
Som alle hoogste biedingen (mld €)	5,9	10,8	6,0	6,7	5,0	4,6
Gemiddelde hoogste bod (in €)	338	520	304	245	232	219
Gemiddelde vraagprijs (in €)	1 105	1 407	1 457	2 706	632	577
<i>Verwijderd op basis van categorie</i>						
Aantal advertenties (mln)	0,5	0,7	1,4	1,2	0,9	0,8
Advertenties met vraagprijs > 0,1 €ct (mln)	0,3	0,3	0,4	0,6	0,3	0,3
Advertenties met een hoogste bod (mln)	0,1	0,1	0,1	0,1	0,0	0,0
Som alle vraagprijzen (mld €)	11,4	27,9	31,6	93,4	4,1	3,3
Som alle hoogste biedingen (mld €)	0,3	0,5	0,2	0,2	0,1	0,1
Gemiddelde hoogste bod (in €)	2 309	3 736	3 202	3 325	3 163	3 312
Gemiddelde vraagprijs (in €)	43 332	80 506	77 415	148 134	13 343	11 063
<i>Na filtering op categorieën</i>						
Aantal advertenties (mln)	44,1	65,8	73,4	83,5	90,1	102,3
Advertenties met vraagprijs > 0,1 €ct (mln)	22,3	32,9	37,4	44,2	49,8	59,9
Advertenties met een hoogste bod (mln)	17,3	20,7	19,6	27,4	21,5	20,8
Som alle vraagprijzen (mld €)	13,5	19,0	23,5	27,8	27,5	31,5
Som alle hoogste biedingen (mld €)	5,6	10,4	5,8	6,5	4,9	4,4
Gemiddelde hoogste bod (in €)	323	500	294	238	226	213
Gemiddelde vraagprijs (in €)	606	575	629	629	554	525
<i>Advertenties met vraagprijs > 99.999,-€</i>						
Aantal advertenties (duizenden)	2,0	2,5	3,4	3,8	4,0	4,6
Advertenties met vraagprijs > 0,1 €ct (duizenden)	2,0	2,5	3,4	3,8	4,0	4,6
Advertenties met een hoogste bod (duizenden)	0,5	0,6	0,6	0,6	0,4	0,4
Som alle vraagprijzen (mld €)	3,1	4,5	3,6	9,1	8,9	9,0
Som alle hoogste biedingen (mld €)	12,5	14,5	14,5	14,9	8,1	7,7
Gemiddelde hoogste bod (in €)	26 400	24 300	25 200	27 000	20 900	19 400
Gemiddelde vraagprijs (in €)	1 590 000	1 850 000	1 060 000	2 370 000	2 230 000	1 980 000

Bron: CBS (Data: Marktplaats).

De filtering op advertentiecategorieën heeft tot gevolg dat er 0,5 (2006) tot 1,4 (2008) miljoen advertenties uit het bestand worden verwijderd. De gemiddelde vraagprijs van de overblijvende advertenties in de dataset neemt fors af. Ditzelfde geldt voor het gemiddelde hoogste bod. Dit is grotendeels te verklaren door het schrappen van de advertentiecategorie ‘Vastgoed’.

Bij nadere beschouwing bleek een zeer klein aantal advertenties voor een aanzienlijk deel van de totale vraagprijs te zorgen. In 2010 zorgde 0,004 procent van de advertenties na filtering op advertentiecategorieën voor meer dan 32 procent van de som van alle vraagprijzen van de advertenties met een vraagprijs. Deze advertenties lijken grotendeels niet serieus te nemen; van de advertenties met een vraagprijs van meer dan 99 999,- euro heeft een aanzienlijk deel ‘mooie’ ronde vraagprijzen, zoals 1000 000 € voor een fiets. Hoewel het in een aantal gevallen een reële vraagprijs kan zijn is toch besloten alle advertenties met een vraagprijs van meer dan 99 999,- euro niet op te nemen.

Het aantal advertenties en hun kenmerken die in het uiteindelijke analysebestand terecht zijn gekomen varieerde van 44,1 miljoen in 2006 tot 102,3 miljoen in 2011. Dit zijn ook de advertenties en hun kenmerken zoals vermeld in tabel 2.1.1 in hoofdstuk 2.

Vulling van de dataset

In tabel 5 is de vulling van een aantal op het eerste gezicht interessante variabelen in het uiteindelijke analysebestand vermeld. Identificerende variabelen zoals de AD_ID, SELLER_ID en ook CLSFD_AD_CNTCT_PSTL_CD (postcode adverteerder (zie ook Appendix 2)) zijn praktisch altijd bekend. Ook de aanduiding van de advertentie(hoofd)categorieën is altijd beschikbaar evenals de datum- en tijdvelden van het plaatsen van een advertentie of een (hoogste) bod. Deze laatste worden automatisch op de advertentie ‘gestempeld’.

De meer analytische of inhoudelijke variabelen zijn echter minder consequent ingevuld. De CONDITIE waarin het goed wordt aangeboden (nieuw, zo goed als nieuw (z.g.a.n.) of gebruikt) was in 2010 voor 61 procent van de advertenties ingevuld. Als de categorie ‘z.g.a.n.’ ook wordt beschouwd als ‘gebruikt’ dan bestaat twee derde van de advertenties waar het kenmerk CONDITIE voor bekend is, uit tweedehands goederen en een derde wordt als ‘nieuw’ aangeboden. De vraag of deze verhouding ook geldt voor de 39 procent van de advertenties waarvoor deze variabele niet bekend is, is moeilijk te beantwoorden.

De variabele ASK_PRICE_TYPE_ID (vraagprijs) is wel altijd ingevuld. Hier is echter een betrekkelijk groot aantal mogelijkheden. In 2010 had 44 procent van de advertenties een uitgesproken vraagprijs (*‘User enters amount’*). Voor 22 procent van de advertenties werd gekozen voor prijsvorming via de categorie ‘Vrij bieden’. De derde categorie is een mengvorm tussen een vraagprijs en bieden, namelijk ‘Bieden vanaf’. ‘Tegen elk aannemelijk bod (t.e.a.b.) kwam bij 11 procent van de advertenties voor. De andere categorieën komen minder vaak voor.

De derde variabele kent de meest teleurstellende vulling. De reden waarom een advertentie is verwijderd (AD_RMVL_RSN_ID) is slechts voor een minderheid van het aantal advertenties bekend. Voor 31 procent van de verwijderde advertenties is de reden ‘Levens einde advertentie’. Dit laatste betekent dat de advertentie automatisch na het verstrijken van de advertentietermijn is verwijderd. Een veronderstelling zou kunnen zijn dat deze advertentie niet tot een transactie hebben geleid, maar dat is gezien de ‘discipline’ waarmee dit veld wordt ingevuld door adverteerders allerminst zeker. Het aantal advertenties waarvan bekend is dat deze wel tot een transactie heeft geleid is slechts 13 procent; voor 12 procent van de verwijderde advertenties is aangegeven dat het artikel is verkocht via Marktplaats, voor één procent van de advertenties is aangegeven dat het artikel anderszins is verkocht. Voor een groot deel van de advertenties is het dus onduidelijk hoe het ‘afgelopen’ is. Dit bemoeilijkt ook het schatten van de omzet van de handel in (tweedehands) goederen via Marktplaats (zie ook hoofdstuk 4).

5. Vulling kernvariabelen in dataset Marktplaats, 2010

5. Vallende kernvariabelen in dataset Marktprijzen, 2010

Naam	Omschrijving	Advertenties	
		aantal (duizend)	percentage
CONDITIE			
NULL	Niet ingevuld	35 006	39
101	Gebruikt	14 664	16
102	Z.g.a.n	22 146	25
103	Nieuw	18 250	20
Totaal		90 065	100
ASK_PRICE_TYPE_ID			
-99	Invalid	6	0
0	User enters amount	39 267	44
2	T.e.a.b.	9 537	11
3	Vrij bieden	19 920	22
4	N.o.t.k	1 419	2
5	Ruilen	173	0
7	Gereserveerd	6 200	7
8	Zie omschrijving	1 887	2
9	Op aanvraag	522	1
10	Bieden vanaf	11 136	12
Totaal		90 065	100
AD_RMVL_RSN_ID			
NULL	Leeg	158	0
1	Verkocht op MP	6 344	7
2	Elders verkocht	975	1
3	Iets anders	12 636	14
8	Levens einde advertentie	27 676	31
9	Handmatige verwijdering	1	0
11	Verkocht op MP -Post	1 971	2
12	Verkocht op MP - opgehaald	2 223	2
13	Verkocht op MP - Anders	114	0
100	Geen selectie	36 511	41
101	API removal by API user	736	1
102	API removal by cronjob	721	1
Totaal		90 065	100

Bron: CBS (Data: Marktplaats).

Koppelen CBS- en Marktplaatsdata op basis van postcode

Een adverteerder op Marktplaats kan kiezen welke (aanvullende) informatie hij of zij wil verstrekken over de advertentie en/of zichzelf. Om de drempel om een advertentie te plaatsen zo laag mogelijk te houden zijn er echter weinig velden verplicht om in te vullen. Dit resulteert in een zeer wisselende vulling van de variabelen in de dataset van Marktplaats (zie appendix 1). Variabelen waar de adverteerder een belang bij heeft om goed in te vullen zijn over het algemeen ook goed gevuld. De postcode is een voorbeeld van een zeer goed gevuld veld aangezien adverteerders er belang bij hebben om in regionale zoekopdrachten van potentiële kopers gevonden te worden. Zo was in 2011 bij bijna 99 procent van de advertenties de postcode ingevuld. In 83 procent van de gevallen betrof dit de 6-posities postcode en in 17 procent van de gevallen de 4-posities postcode. Dit is dan ook goed als aanknopingspunt te gebruiken om de Marktplaatsdata te verrijken met informatie die het CBS op postcodeniveau beschikbaar heeft. Het CBS hanteert in haar reguliere statistische publicaties overigens nauwelijks het regionale aggregatieniveau postcode, maar maakt eerder gebruik van de – statistisch gezien relevantere – indeling in wijken en buurten. Aangezien het – niet beschikbare – huisnummer nodig is om te bepalen welk deel van een 6-posities postcodegebied in welke buurt valt, is het niet mogelijk om zonder verlies van gegevens de Marktplaatsadvertenties in de betreffende wijken en buurten te positioneren. Om toch gegevens te kunnen koppelen is er voor gekozen het maatwerkbestand⁷ van het Centrum voor Beleidsstatistiek te gebruiken. Dit bestand bevat statistische informatie per 6-posities postcodegebied (zie tabel 4). Daarnaast is er gebruik gemaakt van de jaarlijkse plaatsbepaling van postcodegebieden door de GIS-afdeling van het CBS om de afstand tussen adverteerders en bidders waarvan de postcode bekend is, te kunnen berekenen.

Van advertentie naar adverteerder

De postcodes van de advertenties in de Marktplaatsdata zijn in twee stappen gekoppeld aan de postcodes van het CBS. In de eerste stap zijn alle geldige 6-posities postcodes gekoppeld. In de tweede stap zijn de ongeldige 6-posities postcodes ingekort tot 4-posities postcodes, waarna opnieuw een koppeling wordt gemaakt op 4-posities postcode. De advertenties waarvoor sowieso alleen maar een 4-posities postcode bekend was, zijn in deze stap ook mee gekoppeld. De advertenties waar geen postcode bij vermeld was – in 2011 dus net iets meer dan een procent van de advertenties – blijven buiten de analyses die gerelateerd zijn aan de postcode.

De postcode is in eerste aanleg een kenmerk van een advertentie en niet van een adverteerder. Een adverteerder heeft echter een unieke identiteit (gebruikers-id) die gekoppeld is aan zijn of haar advertentie. De postcode van de advertentie wordt dan ook van toepassing verklaard op de bijbehorende adverteerder. Het staat een adverteerder overigens vrij om per advertentie telkens een nieuwe account aan te maken en zo (telkens) onder een ander gebruikers-id te adverteren. Omgekeerd kan een adverteerder onder meerdere postcodes adverteren bijvoorbeeld vanuit huis en vanaf het werk, maar ook

⁷ <http://www.cbs.nl/nl-NL/menu/themas/dossiers/nederland-regionaal/cijfers/incidenteel/maatwerk/default.htm>.

gewoon omdat hij of zij daadwerkelijk verhuisd is. Hier wordt in principe niet voor gecorrigeerd. De postcode van de advertentie wordt voor waar aangenomen en doorgegeven aan de bijbehorende adverteerder. Tabel 1 geeft een extreem voorbeeld van een adverteerder die in 2011 bijna 14 duizend 6-positie postcodes hanteerde voor zijn of haar advertenties. Dit soort adverteerders wijkt vaak ook op andere punten af van het gedrag dat van een ‘normale’ Marktplaatsgebruiker mag worden verwacht. In hoofdstuk 2 en appendix 1 van dit rapport is uitvoerig toegelicht welke gegevens buiten de analyses met de Marktplaatsdata zijn gehouden, om welke redenen en op welke punten dit de uiteindelijk gebruikte dataset heeft beïnvloed. Merk op dat ook na opschoning van het bestand er een aantal adverteerders overblijft dat in 2011 nog meer dan honderd postcodes hanteerde om zijn of haar advertenties onder te plaatsen.

1. Top 10 aantal postcodes per adverteerder, 2011

Top 10 aantal postcodes per adverteerder (SELLER_ID) (alle advertenties)	Top 10 aantal postcodes per adverteerder (SELLER_ID) (advertenties in analysebestand)
<i>aantal</i>	
13 902	178
1 095	136
231	119
179	118
173	117
149	113
136	107
119	106
118	103
118	103

Bron: CBS (Data: Marktplaats).

Van adverteerder naar bidder

Om ook aan de vraagzijde gebruikers van Marktplaats van een postcode te voorzien is als volgt te werk gegaan. Personen die een bod doen op een advertentie moeten zich via een account bekend maken. Omdat de gebruikers-id van een account uniek is, is de gebruikers-id van bidders gerelateerd aan die van de adverteerders. Van deze laatste is zoals uit het voorgaande is gebleken, praktisch altijd de postcode bekend. Op deze wijze is ook een deel van de bidders voorzien van een postcode. Aan bidder wordt dus alleen een postcode toegewezen als ze ooit een advertentie hebben geplaatst en daarbij hetzelfde account hebben gebruikt als bij het bieden. Daarnaast geldt dat het hier niet alle bidders betreft maar alleen degene die uiteindelijk het hoogste bod hebben uitgebracht.

De hiervoor beschreven koppeling lijkt eenvoudig, maar aangezien de postcode in deze dataset een advertentiekenmerk is en geen persoonskenmerk zijn er voor veel gebruikers meerdere postcodes te vinden. Om in die gevallen zo dicht mogelijk bij de meest waarschijnlijke postcode te komen is er per gebruikers-id voor de hele periode van 2006-2011 een lijstje gemaakt met alle gebruikte postcodes met daarbij de datum dat deze het eerst en de datum dat deze het laatst gebruikt is. Vervolgens is gekeken naar de datum waarop met de betreffende gebruikers-id een bod is uitgebracht. In het gunstigste geval valt deze datum tussen twee plaatsingen van advertenties onder dezelfde gebruikers-id en met dezelfde postcode. Deze postcode wordt dan ook van toepassing verklaard op de bidder. Het blijkt echter dat er bij vijf procent van de bidders op een koppelmoment tussen twee advertenties, meer dan één postcode te vinden is. In de meeste gevallen zal dit een zeer legitieme reden hebben, zoals dat men zowel op het

thuis- als het werkadres artikelen aanbiedt, maar in een klein aantal gevallen kan het een kwestie van opzet zijn om in zoveel mogelijk regio's 'gevonden' te worden. Zo zijn er adverteerders die op één en hetzelfde moment enkele honderden postcodes verspreid door het hele land, actief gebruikten. In het geval dat er meerdere mogelijke postcodes zijn bij het hoogste bod zoals dat op een bepaalde datum is uitgebracht, is er gekozen om de laatst gebruikte van deze postcodes te nemen. Mochten er meer gevonden zijn die op hetzelfde moment als laatst gebruikt zijn, dan is gekozen voor de postcode waar de meeste advertenties bij hoorden.

2. Hoogste biedingen/bieders die aan een postcode zijn gekoppeld

	2006	2007	2008	2009	2010	2011
<i>aantal (duizend)</i>						
Totaal aantal hoogste biedingen	13 248	17 228	16 487	23 845	18 697	17 863
Gekoppelde hoogste biedingen	9 009	13 896	13 482	19 871	15 062	11 327
w.v. hoogste biedingen met meer dan 1 koppelmogelijkheid	461	737	651	856	499	235
w.v. hoogste biedingen met meer dan 3 koppelmogelijkheden	10	22	20	25	15	7
Gekoppelde hoogste biedingen	68%	81%	82%	83%	81%	63%
w.v. gekoppelde hoogste biedingen met meer dan 1 mogelijkheid	5%	5%	5%	4%	3%	2%
w.v. gekoppelde hoogste biedingen met meer dan 3 mogelijkheden	0%	0%	0%	0%	0%	0%

Bron: CBS (Data: Marktplaats).

Kwaliteit koppeling

De beste kwaliteit koppeling (situatie 1 in tabel 3) komt voor bij gebruikers die een bod uitbrengen op een datum die tussen twee advertenties valt waarbij zij dezelfde postcode hebben ingevoerd (voorbeeld 1).

Voorbeeld 1: gebruiker voert postcode 1234AA in op 10 januari 2009 bij plaatsing eerste advertentie, gebruiker voert postcode 1234AA in op 18 januari 2009 bij tweede advertentie, gebruiker heeft op 13 januari 2009 een hoogste bod uitgebracht. In dit geval is de hoogste bieder/hoogste bod gekoppeld aan postcode 1234AA.

De kwaliteit van de koppeling wordt minder wanneer dit niet het geval is. Wanneer er maar één referentiepunt is in de tijd dan wordt het bod aan een postcode gekoppeld als het bod minder dan een jaar voor de plaatsing van de eerste advertentie ligt (situatie 2), of het bod minder dan een jaar na plaatsing van de laatste advertentie is gedaan (situatie 3). Wanneer een bod langer dan een jaar na de plaatsing van een laatste advertentie volgt (situatie 4) wordt de kans dat de gebruiker verhuisd is steeds groter, maar gezien het hier nog slechts een tijdsbestek van maximaal vier jaar betreft is deze koppeling toch uitgevoerd. Bij voortzetting van het onderzoek in de tijd dient een beslissing genomen te worden of er op een bepaald moment een afkappunt bereikt wordt (of niet).

Voorbeeld 2: gebruiker uit voorbeeld 1 brengt op 12 mei 2008 een bod uit; dit is minder dan 1 jaar voor zijn eerste advertentie van 10 januari 2009; dus deze wordt gekoppeld. Gebruiker brengt ook een hoogste bod uit op 23 oktober 2006; deze ligt meer dan een jaar voor de eerste plaatsing van een advertentie, en deze wordt dus niet gekoppeld.

Het is verder mogelijk om te koppelen op biedingen waarvan de datum meer dan een jaar voor de eerste plaatsing van een advertentie valt (situatie 5). Dit is echter uit pragmatisch oogpunt niet handig. Ieder jaar dat er nieuwe informatie bijkomt zouden oudere jaren gereviseerd moeten worden. Om dit te limiteren wordt er niet op advertenties gekoppeld die meer dan een jaar na een bod van dezelfde

gebruikers-id zijn geplaatst. De onderste rij in tabel 3 is dan ook als niet gekoppeld te beschouwen en beschouwd.

3. Gekoppelde biedingen per onderscheiden situatie

Gekoppelde hoogste biedingen naar plaatsingsjaar van de advertentie waar het betreffende hoogste bod op uit is gebracht en koppelingssituatie	2006	2007	2008	2009	2010	2011
<u>Situatie 1:</u> Datum van bod valt tussen twee data van advertenties in waarbij twee keer dezelfde postcode is gehanteerd	62%	76%	78%	81%	80%	63%
<u>Situatie 2:</u> Datum van bod ligt tot een jaar voor de plaatsing van de eerste advertentie	25%	12%	10%	8%	8%	5%
<u>Situatie 3:</u> Datum van bod ligt tot een jaar na de datum van plaatsing laatste advertentie	4%	5%	5%	5%	7%	24%
<u>Situatie 4:</u> Datum van bod ligt langer dan een jaar na de datum van plaatsing laatste advertentie	0%	1%	2%	3%	4%	8%
<u>Situatie 5:</u> Datum van bod ligt meer dan een jaar voor datum plaatsing eerste advertentie(deze koppeling wordt niet in verdere analyses gebruikt)	9%	6%	4%	3%	1%	

N.B. De linker- en rechterrand van bovenstaande tabel zijn enigszins afwijkend. Dit is te wijten aan het ontbreken van eerdere en latere jaren. In de kolom van 2011 zou met de levering van de data over 2012 een verschuiving optreden van rij 3 naar rij 1 en 2 en van rij 4 naar rij 1. In kolom 2006 zou er met het toevoegen van een eerder jaar een verschuiving optreden van rij 5 naar rij 3 en 1 en van rij 2 naar rij 1.

Bron: CBS (Data: Marktplaats).

Toepassing op postcode gekoppelde bestanden

De koppeling van Marktplaatsadvertenties, -adverteerders, -biedingen en -bieders aan 6- en 4-positie postcodes is gemaakt om er aanvullende gegevens van het CBS aan te kunnen koppelen. De gegevens die het CBS op postcodeniveau beschikbaar heeft zijn vermeld in tabel 4. Met name in hoofdstuk 3 van dit rapport is veelvuldig gebruik gemaakt van deze gekoppelde datasets van enerzijds Marktplaats en anderzijds het CBS.

Afstand

Als van een advertentie en het hoogste bod de postcode bekend zijn dan kan – zoals reeds eerder gememoreerd – de afstand tussen adverteerder en bieder berekend worden en zo onderdeel van de verschillende analyses worden. In tabel 5 zijn het aantal 6- en 4-positie postcodes vermeld waarvoor de coördinaten van de rijksdriehoeksmeting bekend zijn. De coördinaten van de rijksdriehoeksmeting maken het mogelijk om de afstand tussen twee postcodes te berekenen met behulp van de stelling van Pythagoras (in meters). Daarnaast geeft tabel 5 een indruk van met name het grote aantal 6-positie postcodes dat hiermee een betrekkelijk fijnmazige geografische eenheid lijkt te zijn.

5. Aantal 6-positie en 4-positie postcodegebieden

	6-positie postcodes	4-positie postcodes
<i>aantal</i>		
2006	445 185	4 043
2007	447 434	4 046
2008	449 104	4 046
2009	450 994	4 048
2010	452 620	4 050
2011	453 852	4 051

Bron: CBS.

Selectie goederengroepen vergelijking Marktplaats en Detailhandel

Voor het hoofdstuk Marktplaats en conjunctuur is een tijdreeks van categorieën advertenties samengesteld die zo goed mogelijk aansluit bij de bedrijfsgroep Detailhandel in consumentenelektronica in de classificatie van het CBS (SBI-groep 47.4). Zoals eerder vermeld zijn de (hoofd-)categorieën in de onderscheiden advertentierubrieken van Marktplaats ingegeven vanuit het perspectief een website op een handzame manier in te delen, en niet op basis van consequent doorgevoerde uitgangspunten van een ‘officiële’ statistische classificatie. Dit zorgt er voor dat sommige categorieën te vinden zijn in hoofdcategorieën die vanuit de Marktplaatsgebruiker intuïtief logisch zijn, maar voor een ‘officiële’ classificatie ten behoeve van het samenstellen van statistieken, niet. Daarnaast is de koppeling tussen een winkel waar een goed verkocht wordt en een categorie goederen soms lastig. In de SBI-groep 47.4 *‘Winkels in consumentenelektronica’* vallen bijvoorbeeld ook de *‘Winkels in een algemeen assortiment van wit- en bruingoed’* (SBI 47.43.2), terwijl de *‘Winkels in witgoed’* buiten de 47.4 vallen. De Marktplaatscategorie witgoed valt hiermee dus feitelijk deels binnen en deels buiten de *‘Winkels in consumentenelektronica’*. Omdat het deel dat er binnen valt niet is te bepalen is de gehele categorie witgoed niet meegenomen in de analyse.

Een vergelijkbare situatie treedt op bij digitale camera’s; waar vroeger de handel in fotocamera’s voornamelijk via de *‘Winkels in fotografische artikelen’* liep is er met de komst van de digitale camera meer en meer ruimte voor de winkels in consumentenelektronica gekomen. Hier is op Marktplaats iets anders aan de hand dan met witgoed; de voormalige hoofdcategorie “Fotografie” is in 2008 samengevoegd met “Audio en TV”, maar het onderliggende onderscheid tussen digitale en analoge fotografie is hierbij in stand gebleven. Daarnaast is er een onderscheid gemaakt tussen professionele en amateur fotografie. Deze onderscheiden maken het mogelijk om de amateur digitale fotografie aan de consumentenelektronica toe te schrijven, terwijl de professionele en analoge fotografie hier buiten gehouden wordt.

1. Relevante onderdelen van de Detailhandel in Standaard Bedrijfsindeling 2008

- 47.4 Winkels in consumentenelektronica
 - 47.41 Winkels in computers, randapparatuur en software
 - 47.42 Winkels in telecommunicatieapparatuur
 - 47.43 Winkels in audio- en videoapparatuur of in een algemeen assortiment van wit- en bruingoed
 - 47.43.1 Winkels in audio- en videoapparatuur
 - 47.43.2 Winkels in een algemeen assortiment van wit- en bruingoed
 - 47.54 Winkels in elektrische huishoudelijke apparatuur en onderdelen daarvoor
 - 47.54.1 Winkels in witgoed
 - 47.54.2 Winkels in naai- en breimachines
 - 47.54.3 Winkels in overige elektrische huishoudelijke apparatuur
 - 47.54.4 Winkels in onderdelen voor elektrische huishoudelijke apparatuur
 - 47.78.1 Winkels in fotografische artikelen
 - 47.78.2 Winkels in optische artikelen

Bron: CBS.

Buiten de categorie “Audio, Tv en Foto” zijn ook de categorieën “Computers en Software” en “Telecommunicatie” toegewezen aan de consumentenelektronica. Daarnaast zijn er nog stukken van de

categorieën “Fotografie”, “Muziek en instrumenten” en “Video en Camera’s” meegenomen in de conjunctuuranalyse.

2. Geselecteerde categorieën voor het samenstellen van de tijdreeks consumentenelektronica

(Hoofd-)categorie	(Hoofd-)categorie	(Hoofd-)categorie
Audio, Tv en Foto Accu's en Batterijen Afstandsbedieningen" Bandrecorders Beamers Blu-ray-spelers en Blu-ray-recorders Buizenversterkers Cassettedecks Cassettedecks Cd-spelers Dvd-spelers en Dvd-recorders Fotografie Accu's en Batterijen Fotografie Camera's Digitaal Fotografie Digitale fotolijstjes Fotografie Flitsers Fotografie Fotoprinters en Fotopapier Fotografie Geheugenkaarten Harddiskrecorders Home Cinema-sets Kabels Koptelefoons en Headsets Luidsprekers Mediaspelers Minidiskspelers Mp3-spelers Accessoires Apple iPod Mp3-spelers Accessoires Overige merken Mp3-spelers Apple iPod Mp3-spelers Overige merken Mp4-spelers Opladers Overige Audio, Tv en Foto Platenspelers en Pick-ups Professionele Audio-, Tv- en Video-apparatuur Radio's Receivers Stereo-sets Tuners Tv-decoders en Schotelantennes Tv's Accessoires Tv's Breedbeeld Tv's Overige Tv's Tv's Plasma, Lcd en Led Versterkers en Receivers Videocamera's Digitaal Videospelers en Videorecorders Walkmans, Discmans en Minidiscspelers Weerstations en Barometers	Computers en Software(vervolg) Laptops Accu's en Batterijen Laptops Mini Laptops en Netbooks Laptops Onderdelen en Toebehoren Laptops Tassen en Hoezen Laptops en Notebooks Modems, ISDN en Faxen Moederborden Monitoren en Beeldschermen Mp3 spelers Muizen en Joysticks Netwerkkarten, Routers en Switches Overige Computers en Software Printerbenodigdheden Printers Processors Scanners Servers Software Antivirus en Beveiliging Software Apple Macintosh Software Besturingssoftware en Drivers Software Educatie en Cursussen Software Grafisch, Foto en Film Software Kinderen en Jeugd Software Muziek en Audio Software Navigatie en Geografie Software Overige Sun, Sgi en Unix Sun, Unix en Sgi Tablet-pc's Toetsenborden USB Sticks en Toebehoren Videokaarten Webcams ZX Spectrum Fotografie Digitale camera's Canon Digitale camera's Fuji Digitale camera's HP Digitale camera's Kodak Digitale camera's Konica Digitale camera's Leica Digitale camera's Minolta Digitale camera's Nikon Digitale camera's Olympus Digitale camera's Pentax Digitale camera's Samsung Muziek en Instrumenten Harddiskrecorders Mengpanelen Microfoons	Telecommunicatie Midi-apparatuur Speakers Antennes en Masten Antwoordapparaten Carkits Carkits Motorola Carkits Nokia Datacommunicatie en VoIP Faxen ISDN en ADSL Meetapparatuur Mobiele telefoons Apple iPhone Mobiele telefoons Batterijen en Accu's Mobiele telefoons Blackberry Mobiele telefoons Bosch Mobiele telefoons Ericsson Mobiele telefoons Headsets Mobiele telefoons Hoesjes en Frontjes Apple iPhone Mobiele telefoons Hoesjes en Frontjes Blackberry Mobiele telefoons Hoesjes en Frontjes HTC Mobiele telefoons Hoesjes en Frontjes Nokia Mobiele telefoons Hoesjes en Frontjes Overige merken Mobiele telefoons Hoesjes en Frontjes Samsung Mobiele telefoons HP Mobiele telefoons HTC en Qtek Mobiele telefoons KPN/PTT Mobiele telefoons LG Mobiele telefoons Mitsubishi Mobiele telefoons Motorola Mobiele telefoons Nec Mobiele telefoons Nokia Mobiele telefoons Opladers en Autoladers Mobiele telefoons Overige merken Mobiele telefoons Panasonic Mobiele telefoons Philips Mobiele telefoons Sagem Mobiele telefoons Samsung Mobiele telefoons Siemens Mobiele telefoons Software Mobiele telefoons Sony Ericsson Mobiele telefoons Toebehoren en Onderdelen Overige Telecommunicatie Pda's Pda's Toebehoren Portofoons en Walkie-talkies Prepaidkaarten en Simkaarten Satellietontvangers en Schotelantennes Scanners Semafoons Telefooncentrales Vaste telefoons Handsets en Draadloos Vaste telefoons Niet Draadloos Video en Camera's Wap Motorola Wap Nokia Zenders en Ontvangers Accessoires Digitale videocameras Dvd-spelers SVHS Videocamera's SVHS Videorecorders Videocamera's Videoprojectoren Videorecorders en -spelers

Bron: Marktplaats.