

Vereenvoudiging model waardering stroomonderbrekingen

NMa

B15236
Januari 2013

Copyright © 2013 Blauw Research bv

Alle rechten voorbehouden. Niets uit dit rapport mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen of enige andere manier, zonder voorafgaande schriftelijke toestemming van Blauw Research. Dit rapport is geleverd onder de leveringsvoorwaarden van de MOA.

All rights are reserved. Nothing from this report may be copied, saved in an authorised data bank or be made public in any form, whether it be electronically, mechanically or through photocopies without prior consent from Blauw Research. This report has been created following MOA conditions.

Inhoudsopgave

1 Kort & Krachtig	4
2 Totstandkoming model	6
3 Uitwerking model	9
4 Robuustheid model	14
5 Prijskaartjes	17
Bijlage Verantwoording	20
Bijlage Syntax opschonen dataset frequentie	21
Bijlage bepaling D0 en Fm	22

Hoofdstuk 1: Kort & Krachtig



Een eenvoudige, continue alternatieve waarderingfunctie

Blauw heeft eerder dit jaar voor de NMa een onderzoek uitgevoerd om het model voor de waardering van stroomonderbrekingen te actualiseren. De huidige waarderingfunctie is echter discontinu. Blauw heeft de opdracht gekregen te onderzoeken of een eenvoudigere, continue waarderingfunctie mogelijk is op basis van de onderzoeksdata die in 2012 verzameld is, en bij een positief resultaat deze aangepaste waarderingfunctie op te stellen.

Na een eerste beschouwing is Blauw tot de conclusie gekomen dat het mogelijk moet zijn een aangepaste waarderingfunctie op te stellen die aan beide genoemde voorwaarden (continu en eenvoudig) voldoet. Dit rapport beschrijft de totstandkoming van deze nieuwe waarderingfunctie, alsmede de uitwerking ervan en de resultaten.

De waarderingfunctie is samengevat in de volgende formules:

Algemeen:

$$C = (b1 * \ln(F - F_m + 1) + b2 * \ln(D - D_0 + 1)) * GSR$$

Huishoudens:

$$C = (15,71\% * \ln(F - 0,2 + 1) + 48,99\% * \ln(D - 0,083 + 1)) * €54$$

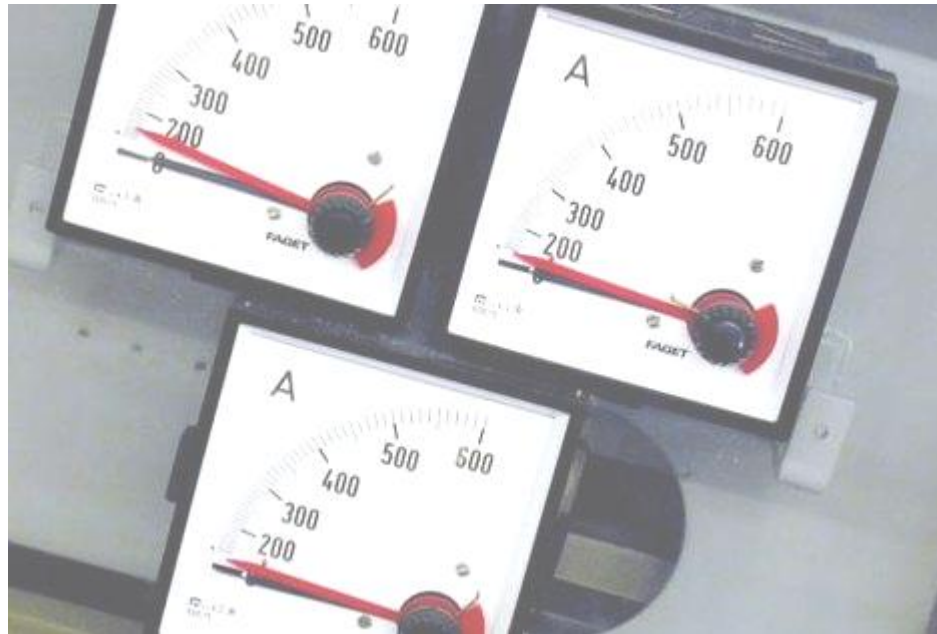
Bedrijven:

$$C = (21,93\% * \ln(F - 0,2 + 1) + 31,12\% * \ln(D - 0,083 + 1)) * €372$$

Hierbij staat C voor de gewenste compensatie in euro's
 F voor de frequentie in het aantal onderbrekingen per jaar
 D voor de gemiddelde duur van een onderbreking in uren
 GSR voor de gepercipieerde stroomrekening

Het eerste gedeelte van de formule beschrijft de gewenste compensatie in procenten. Door dit te vermenigvuldigen met de gemiddelde gepercipieerde stroomrekening kan het gewenste bedrag berekend worden.

Hoofdstuk 2: Totstandkoming model

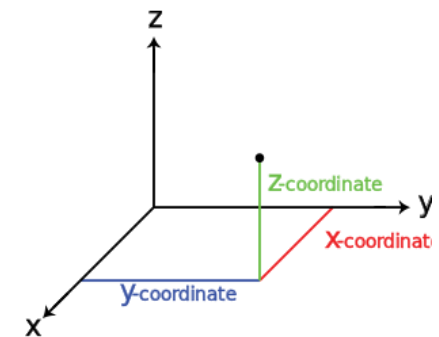


Een lineair additief model met logaritmische onderdelen

De eerste keuze die gemaakt dient te worden, is de vorm van het model. In overleg tussen Blauw en NMa is een waarderingsfunctie vastgesteld die gebaseerd is op één formule. Hierdoor verdwijnt de discontinuïteit zoals die optreedt als er meerdere functies (formules) worden gebruikt. De compensatie (C) wordt uitgedrukt in een functie afhankelijk van F (frequentie) en D (duur). In principe is de relatie lineair afgezien van twee transformaties:



- Van beide onafhankelijke variabelen (frequentie en duur) wordt eerst de logaritme genomen. Hiervoor zijn twee redenen: ten eerste sluit dit beter aan bij wat men intuïtief verwacht, te weten afnemende marginale dissatisfactie bij toenemende frequentie van het aantal onderbrekingen en de duur ervan. Ten tweede laat een vergelijking van de modellen (lineair, trapsgewijs en logaritmisch) zien dat een logaritmische transformatie een goede fit met de data geeft (zie voor verdere toelichting Blauw 2012).
- Bovendien wordt van beide variabelen eerst het in het onderzoek vastgestelde acceptabele niveau afgetrokken (zie bijlage voor methode vaststellen D_0 en F_m), en vervolgens het getal 1 opgeteld. Deze transformatie heeft tot gevolg dat in de situatie dat de frequentie acceptabel is en ook de duur het acceptabele niveau heeft, dat de gewenste compensatie dan per definitie gelijk aan nul is.



Een lineair additief model met logaritmische onderdelen

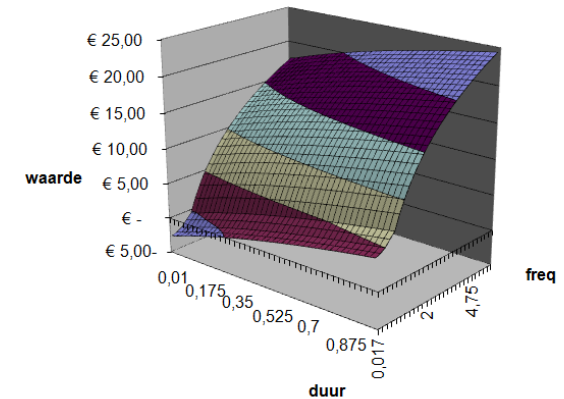
Voor de rest zijn alle verbanden lineair. In formulevorm komt het model er voor zowel huishoudens als bedrijven als volgt uit te zien:

$$C = (b1 * \ln (F-Fm+1) + b2 * \ln (D-D0+1)) * GSR$$

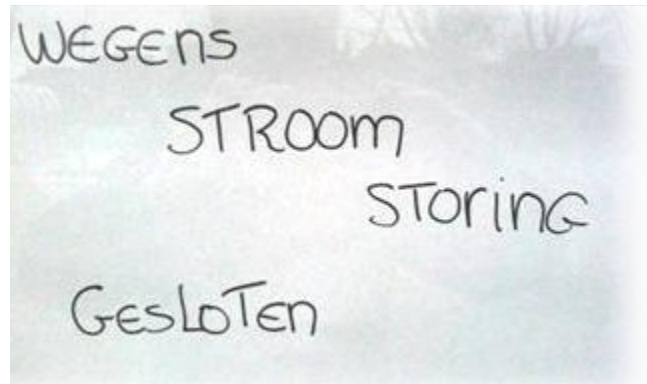
Waarbij :

- C staat voor compensatie (in euro's). Een negatieve uitkomst betekent dus dat de consument (of het bedrijf) bereid is extra te betalen.
- F is frequentie in aantal onderbrekingen per jaar
- D is de gemiddelde duur van een onderbreking in uren
- Fm en D0 zijn acceptabele waarden, waarbij geen compensatie gewenst is en men ook niet bereid is extra te betalen, deze waarden worden bepaald aan de hand van de Van Westendorpmethode. (zie rapport "Waardering stroomonderbrekingen 2012" voor verdere toelichting)
- b1 en b2 zijn parameters die de invloed van frequentie en duur weergeven. De parameters b1 en b2 worden bepaald aan de hand van de beide conjuncte onderdelen in de vragenlijst.
- GSR staat voor de (gemiddelde) gepercipieerde stroomrekening. Door met deze factor te vermenigvuldigen wordt het gemeten compensatiepercentage vertaald naar het bijbehorende absolute bedrag in euro's.

2 Totstandkoming model

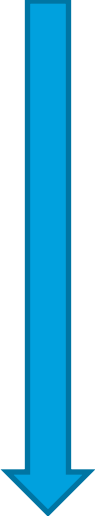


Hoofdstuk 3: Uitwerking model



Stappen voor uitwerking en invullen model

Voor verdere uitwerking van het model, dat wil zeggen het schatten van parameters b_1 en b_2 voor zowel bedrijven als huishoudens, zijn de volgende stappen genomen:

- 
1. Het opnieuw (strenger) opschonen van de data (verwijderen onbetrouwbare en inconsistente cases) om de parameters nauwkeuriger te kunnen schatten dan in het vorige onderzoek (Blauw 2012). De mogelijkheid om strenger te selecteren, bestaat omdat er minder parameters geschat worden.
 2. Het transformeren van de variabelen zodat deze overeenkomen met het nieuwe model
 3. Het uitrekenen van de bèta-gewichten van de duur en de bijbehorende compensatie en de bèta-gewichten van de frequentie en bijbehorende compensatie
 4. Het bepalen van de verhouding tussen deze bèta-gewichten

Deze stappen worden in het huidige hoofdstuk verder toegelicht.

Stap 1 en 2: opschonen en transformeren data

1. Het opschonen van de data

Omdat de data voor de duur en frequentie op een verschillende manier vergaard zijn, is ook het opschonen van de dataset op twee verschillende manieren gebeurd.

Duur (Choice Based Conjoint)

De data van de Choice Based Conjoint (CBC) zijn op twee verschillende manieren opgeschoond.

- Ten eerste zijn alleen de respondenten meegenomen die de vragenlijst volledig hebben ingevuld, en die in de vragenlijst geen inconsistente antwoorden hebben gegeven (zoals laag stroomverbruik en zeer hoge kosten).
- Ten tweede is de CBC-dataset nog opgeschoond op basis van de RLH (Root Likelihood), een maat voor de interne consistentie van de antwoorden van de respondent. Alle respondenten met een RLH lager dan 450 zijn uit de dataset verwijderd.

Frequentie (vignet-analyse op basis van rapportcijfers)

- Ook voor de analyse op de frequentie-dataset zijn alleen de respondenten meegenomen die de vragenlijst volledig hebben ingevuld, en die in de hele vragenlijst geen inconsistente antwoorden hebben gegeven.
- Daarnaast zijn respondenten verwijderd indien de antwoorden op het vignet-gedeelte niet plausibel waren (zie bijlage met syntax voor verdere toelichting)
- Tot slot zijn respondenten die alle voorgelegde situaties exact hetzelfde rapportcijfer hebben geven uitgesloten.

2. Het transformeren van de variabelen zodat deze overeenkomen met het nieuwe model

In het nieuwe model worden de bèta-gewichten vermenigvuldigd met het logaritme van de frequentie of duur, gecompenseerd voor de acceptabele waarde. Om het model goed te kunnen schatten was het nodig om de variabelen in de dataset overeenkomstig te transformeren. De duur is als volgt omgerekend: $D(\text{nieuw}) = \ln(D - D_0 + 1)$, en de frequentie op dezelfde wijze: $F(\text{nieuw}) = \ln(F - F_m + 1)$.

Stap 3 en 4: berekenen gewichten en verhoudingen– duur

3. Het berekenen van de bèta-gewichten van de duur en de bijbehorende compensatie en de bèta-gewichten van de frequentie en bijbehorende compensatie

De volgende stap is het bepalen van de bèta-gewichten op basis van de getransformeerde variabelen. Voor de duur (en de bijbehorende compensatie) is dit gedaan door middel van een Hierarchical Bayes analyse waarbij de onafhankelijke variabele, duur, getransformeerd is (door middel van de volgende transformatie: $\ln(D-D_0+1)$). In deze analyse zijn ook de variabelen dag waarop de storing valt, het seizoen, het dagdeel en of er al dan niet gewaarschuwd is meegenomen om het model zo betrouwbaar mogelijk te schatten. De resulterende scores zijn vervolgens gemiddeld voor alle respondenten.

Dit leidde tot de volgende resultaten:

	Huishoudens	Bedrijven
Bèta duur	-3,9674	-2,9499
Bèta compensatie	0,08098	0,09478
Verhouding	48,99	31,12

De verhouding van de bèta-gewichten representeert de relatieve invloed van duur op compensatie: hoeveel compensatie moet worden gegeven als de duur toeneemt om het resultaat (in dit geval de kans op keuze van het vignet, ofwel het gemiddelde nut), en daarmee de waardering, gelijk te houden.

Stap 3 en 4: berekenen gewichten en verhoudingen– frequentie

Om de bèta-gewichten van de frequentie en de bijbehorende compensatie te berekenen is een regressie-analyse uitgevoerd met het rapportcijfer als afhankelijke variabele, en de getransformeerde frequentie en de compensatie als onafhankelijke variabelen. Het gebruik van de getransformeerde frequentie leidde tot een hogere verklaarde variantie (R^2) bij zowel huishoudens (0,075, eerst 0,052) als bij bedrijven (0,068, eerst 0,043). Dit betekent dat het nieuwe model een grotere voorspellende waarde heeft dan wanneer de ongetransformeerde frequentie wordt gebruikt .

Dit leidde tot de volgende resultaten:

	Huishoudens	Bedrijven
Bèta frequentie	-0,424	-0,443
Bèta compensatie	0,027	0,0202
Verhouding	15,71	21,92

Voor duur en frequentie samen leidt dat tot de volgende formules:

$$\text{Huishoudens: } C = (15,71\% * \ln (F- 0,2 +1) + 48,99\% * \ln (D-0,083+1)) * €54$$

$$\text{Bedrijven: } C = (21,92\% * \ln (F- 0,2 +1) + 31,17\% * \ln (D-0,083+1)) * €372$$

Hoofdstuk 4: Robuustheid model



Robuustheid

De volgende stap was het bepalen van de robuustheid van het nieuwe model. Dit betekent in dit geval dat de compensatie niet sterk verandert bij kleine wijzigingen in de achterliggende parameters. Hiertoe zijn de bèta-gewichten met 10% en 50% gevarieerd (zowel in negatieve als positieve richting). Ook zijn de gevolgen berekend van het veranderen van de waarden van D_0 en F_m .

Robuustheid model bèta-gewichten

Het bepalen van de robuustheid van de bèta-gewichten is uitgevoerd door bij de bètagewichten 10% en 50% op te tellen, en er 10% en 50% vanaf te trekken, en door dit bij beide bèta-gewichten gelijktijdig te doen.

Uit de uitgevoerde analyse kan geconcludeerd worden dat het model robuust is wat betreft deze parameters. Wanneer één van beide bèta-gewichten 10% of 50% groter of kleiner wordt, verandert de compensatie met een lager percentage. Indien zowel b_1 als b_2 10% toe- of afnemen, neemt de compensatie ook precies 10% toe. Hetzelfde geldt (uiteraard) voor een toename van 50%, of ieder ander percentage.

Robuustheid model ten aanzien van verschuivingen acceptabele waarden

Om de robuustheid van het nieuwe model voor veranderingen van D_0 en F_m te bepalen, is het model opnieuw geschat voor de D_0 en F_m die gevonden zijn in het SEO-onderzoek uit 2004 ($D_0 = 0,35$ en $F_m = 0,12$ voor huishoudens en $D_0 = 0,24$ en $F_m = 0,08$ voor bedrijven). Dit leverde een verschuiving van de prijskaartjes op van minder dan 15%, terwijl de acceptabele waarden meer dan 15% afwijken. Dit leidt tot de conclusie dat het nieuwe model ook robuust is voor verschuivingen van de acceptabele waarden.

Randvoorwaarden

Er zijn situaties te bedenken waarin het model geen oplossing geeft. Dit is het geval als F is kleiner dan $F_m - 1$, of D is kleiner dan $D_0 - 1$. Zolang F_m en D_0 zelf tussen 0 en 1 liggen, zoals in de huidige situatie, geeft dit geen probleem. Echter als méér dan 1 onderbreking per jaar acceptabel zou zijn, of een onderbreking van meer dan 1 uur, dan is het theoretisch mogelijk dat $F - F_m + 1$ kleiner dan nul is, waardoor de formule geen oplossing geeft. Het is niet voor de hand liggend dat de acceptabele waarden deze kant op zullen verschuiven, aangezien de consument de neiging heeft steeds kritischer te worden, en de afhankelijkheid van elektriciteit in de nabije toekomst niet zal afnemen.

Hoofdstuk 5: Prijskaartjes



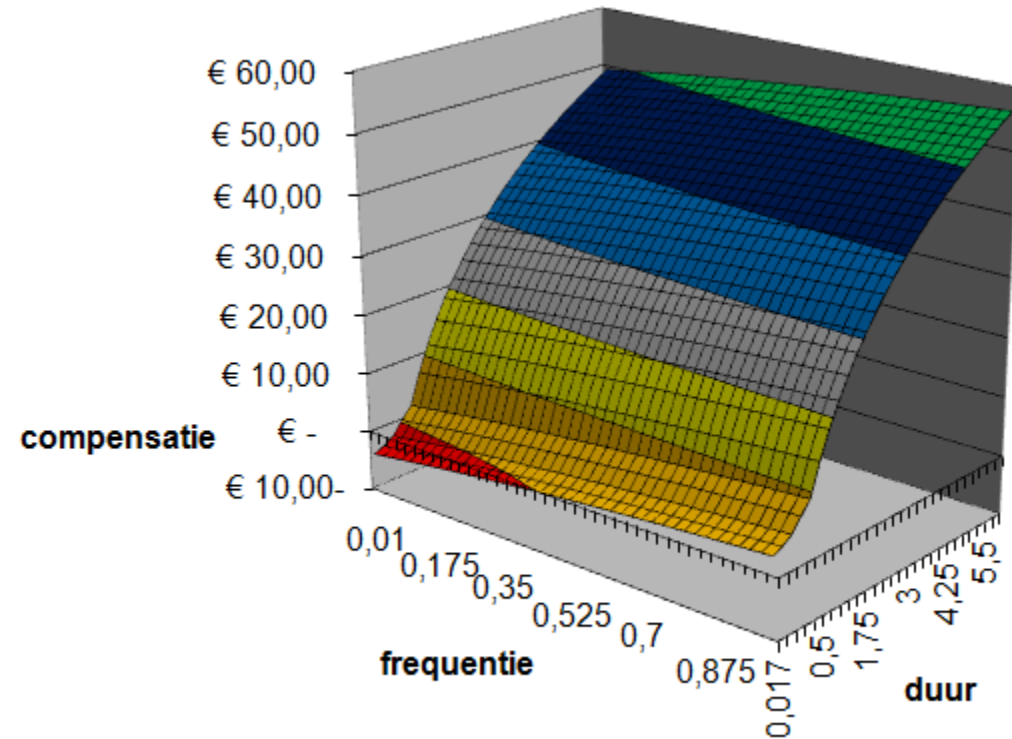
Prijkaartjes huishoudens

Door de vastgestelde functie in te vullen voor een reëel bereik van de frequentie en duur, kan een visuele weergave van het model gemaakt worden. Uit deze driedimensionale grafiek kunnen de 'prijkaartjes' en het verloop van de compensatie ten opzichte van de frequentie en duur van stroomonderbrekingen worden afgelezen.

De gewenste compensatie is €0 op de lijn die loopt over het punt (D0,Fm), dus bij een onderbreking van 5 minuten eens in de vijf jaar. De gewenste compensatie laat vanuit deze lijn een logaritmisches dalend en stijgend verband zien, waarbij de snelheid van de toe- of afname van de gewenste compensatie sterker is dicht bij de nul-lijn.

De gebruikte range loopt van een frequentie van 0,01 tot 1 stroomonderbrekingen per jaar en een duur van 1 minuut tot 6 uur. De voor huishoudens gewenste compensatie is het hoogst bij een frequentie van 1 stroomonderbreking per jaar en een duur van 6 uur, namelijk €56,15. De compensatie is het laagst bij een onderbreking van 1 minuut eens in de 100 jaar, namelijk -€3,61.

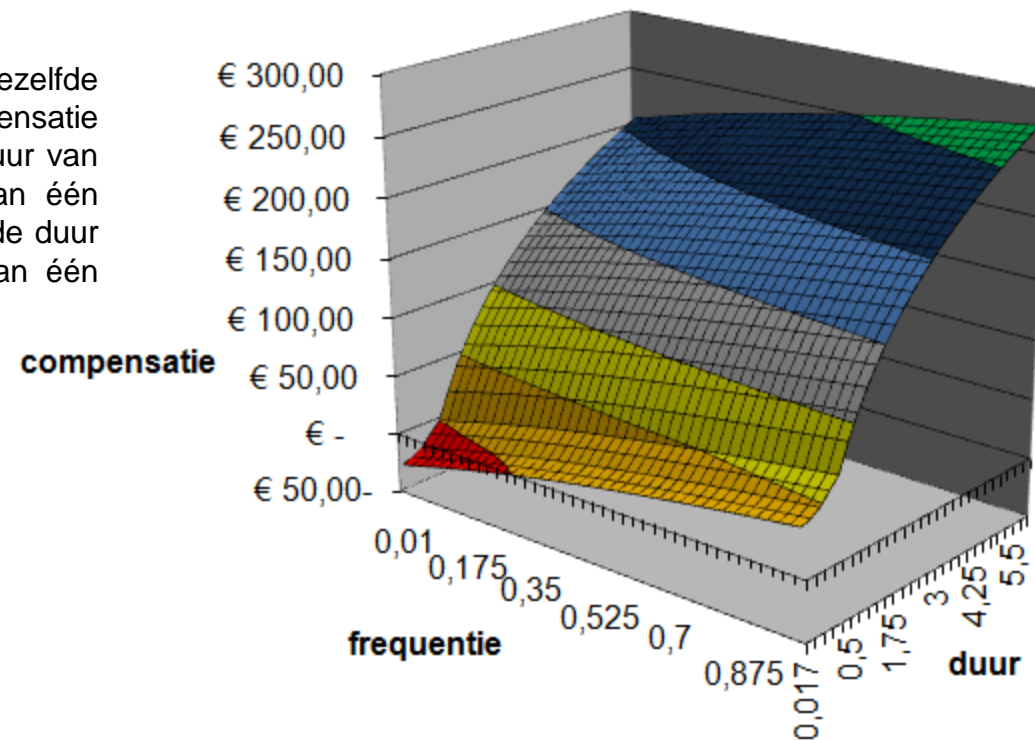
5 Prijkaartjes



Prijskaartjes bedrijven

Ook voor de bedrijven is het model uitgewerkt voor een reëel bereik van de frequentie en duur van stroomonderbrekingen, met nevenstaande grafiek als resultaat. De 'prijskaartjes' en de verhouding tussen de gewenste compensatie en de duur en frequentie van stroomonderbrekingen kunnen uit de grafiek worden afgelezen.

Voor het bepalen van de prijskaartjes van de bedrijven is dezelfde range gekozen als bij de huishoudens. De gewenste compensatie loopt bij de bedrijven van -€25,18 voor een gemiddelde duur van een onderbreking van één minuut en een frequentie van één onderbreking per 100 jaar, tot €271,87 voor een gemiddelde duur van een onderbreking van zes uur en een frequentie van één onderbreking per jaar.



Bijlage Verantwoording

Keuze voor gepercipieerde stroomrekening

Er is gekozen om de gepercipieerde stroomrekening te gebruiken (de bedragen die de respondenten hebben ingevuld) in plaats van de werkelijke gemiddelde stroomrekening. De eerste reden hiervoor is van praktische aard: de daadwerkelijke gemiddelde stroomrekening van het MKB is niet bekend. De tweede reden is onderzoekstechnisch: respondenten hebben de bedragen die zij zelf hebben ingevuld getoond gekregen bij het invullen van de vragenlijst, dus is het logisch dat deze bedragen worden gebruikt als basis voor het model.

Keuze geen gestandaardiseerde rapportcijfers

In het onderzoek van SEO uit 2004 is ervoor gekozen om de rapportcijfers uit de vignettenanalyse op individueel niveau te compenseren voor het gemiddelde rapportcijfer. Deze stap is in het huidige onderzoek echter overgeslagen omdat er dit maal (in tegenstelling tot in 2004) geen referentievignet is getoond aan alle respondenten. Het is dus mogelijk dat alle getoonde vignetten vrij negatief of vrij positief beoordeeld worden. Dit maakt het onnodig en zelfs onjuist om te compenseren voor het gemiddelde gegeven rapportcijfer.

Keuze model

Alle drie de meest voor de hand liggende modellen (lineair, stapsgewijs en logaritmisch) zijn nagerekend. Net als in 2004 verklaren elk van de modellen de uitkomsten ongeveer even goed. De keuze is gemaakt voor het logaritmische model om de volgende redenen:

- Zelfde keuze als in 2004
- Intuïtief het meest logisch
- Weinig verschil in verklarende kracht tussen de modellen
- Robuuste uitkomsten

(voor verdere toelichting, zie rapport “Waardering stroomonderbreking 2012”)

Resultaten opschonen data

Het opschonen van de data leidt tot respondentenaantallen die nog steeds zeer betrouwbare en robuuste resultaten waarborgen.

Duur	Huishoudens	Bedrijven
Aantal respondenten	4183	1145
Na verwijderen RLH < 450	3669	1117

Frequentie	Huishoudens	Bedrijven
Aantal respondenten	4183	1145
Na plausibiliteitscontrole	3525	987

Bijlage Syntax opschonen dataset frequentie

*opschonen op basis van plausibiliteit:

*stap 1: niet plausible beoordelingen verwijderen:.

compute plausibel1=0.

if (((E2>6) and (frequentie>1) and (korting<0)) or ((E2<4)

and (frequentie<0.25) and (korting>-5))) plausibel1=1.

select if plausibel1 = 0.

*stap 2:

E2_nu staat voor het aantal plausible waarnemingen per respondent.

E2_sd staat voor variantie in rapportcijfers.

Verwijderen respondenten met meer dan 1 niet plausibel antwoord en

verwijderen respondenten zonder variantie in antwoorden.

DATASET ACTIVATE DataSet2.

AGGREGATE

/OUTFILE=* MODE=ADDVARIABLES

/BREAK=CODERESP

/E2_nu=N

/E2_sd=SD(E2).

select if (E2_nu > 3).

select if (E2_sd >0).

EXECUTE.

Bijlage Bepaling D0 en Fm

De D0 en de Fm, de waarden die worden gebruikt als acceptabele waarden om het model te ijken, worden bepaald via de Van Westendorpmethode. Dit wordt gedaan door eerst de lijn te berekenen die het verschil tussen acceptabel en onacceptabel aangeeft (zie volgende pagina voor voorbeeld). Het punt waar deze lijn het sterkst stijgt, en waar minstens de helft van de respondenten het niveau acceptabel vindt, is de plek waar het ijkpunt zich bevindt. Het punt waar de stijging het sterkst is wordt bepaald door de maximale waarde van de tweede afgeleide te berekenen.

Rechts op deze pagina is de bijbehorende data weergegeven voor de D0 van huishoudens. De oranje cijfers geven aan dat minder dan 50% van de respondenten de situatie op het gegeven punt acceptabel vindt. De hoogste waarde van de tweede afgeleide is geel gemarkeerd.

Resultaten huishoudens - duur

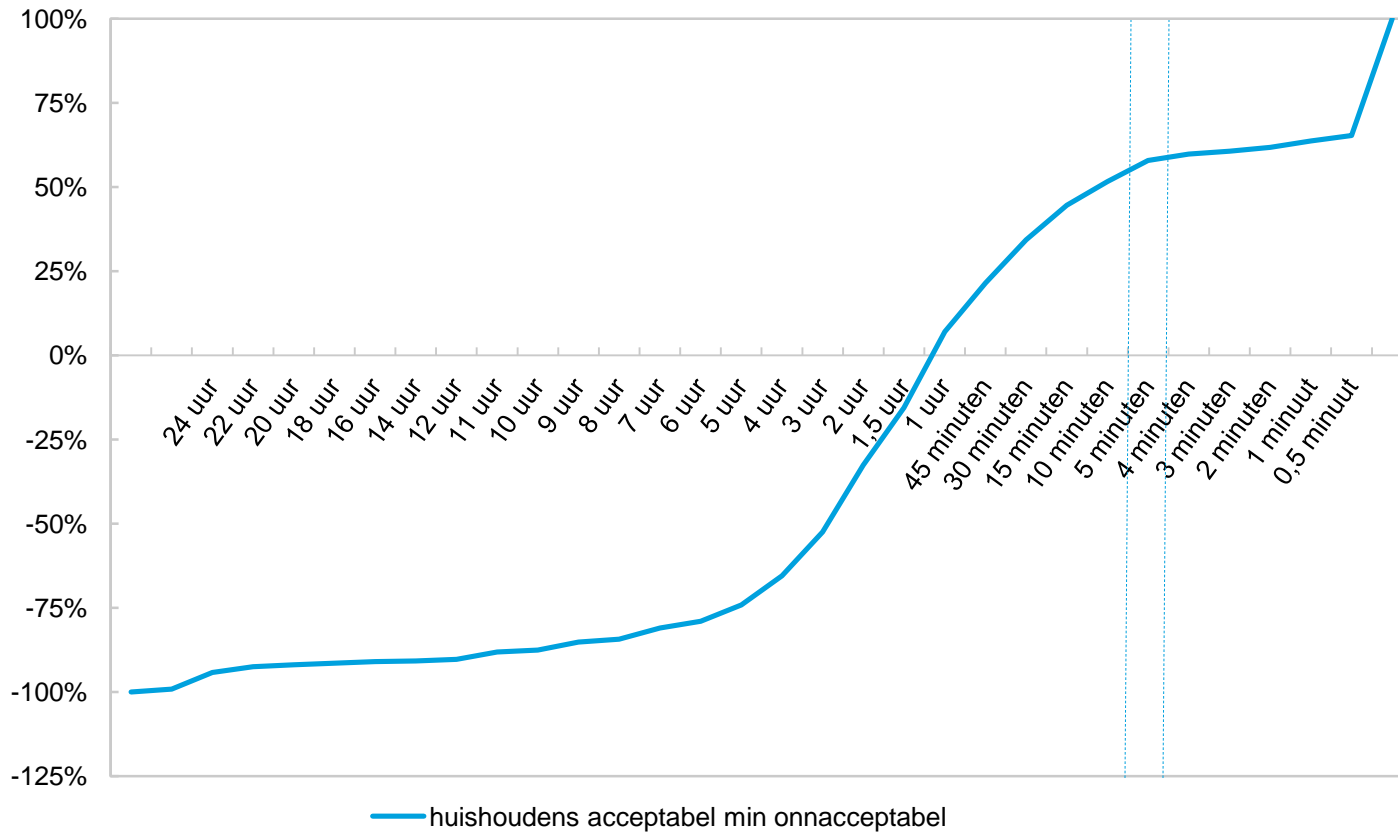
	afgeleide	tweede afgeleide
2 uur	0,17	0,03
1,5 uur	0,23	-0,06
1 uur	0,14	0,08
45 minuten	0,13	0,02
30 minuten	0,10	0,03
15 minuten	0,07	0,03
10 minuten	0,06	0,01
5 minuten	0,02	0,04
4 minuten	0,01	0,01
3 minuten	0,01	0,00
2 minuten	0,02	-0,01
1 minuut	0,02	0,00
0,5 minuut	0,35	-0,33

Resultaten huishoudens - frequentie

	afgeleide	tweede afgeleide
1/2jaar	0,19	0,122
1/3jaar	0,07	0,042
1/4jaar	0,02	-0,030
1/5jaar	0,05	0,029
1/6jaar	0,02	0,020
1/7jaar	0,00	0,002
1/8jaar	0,00	0,001
1/9jaar	0,00	-0,044
1/10jaar	0,05	0,026
1/11jaar	0,02	0,016

Bijlage Bepaling D0 en Fm

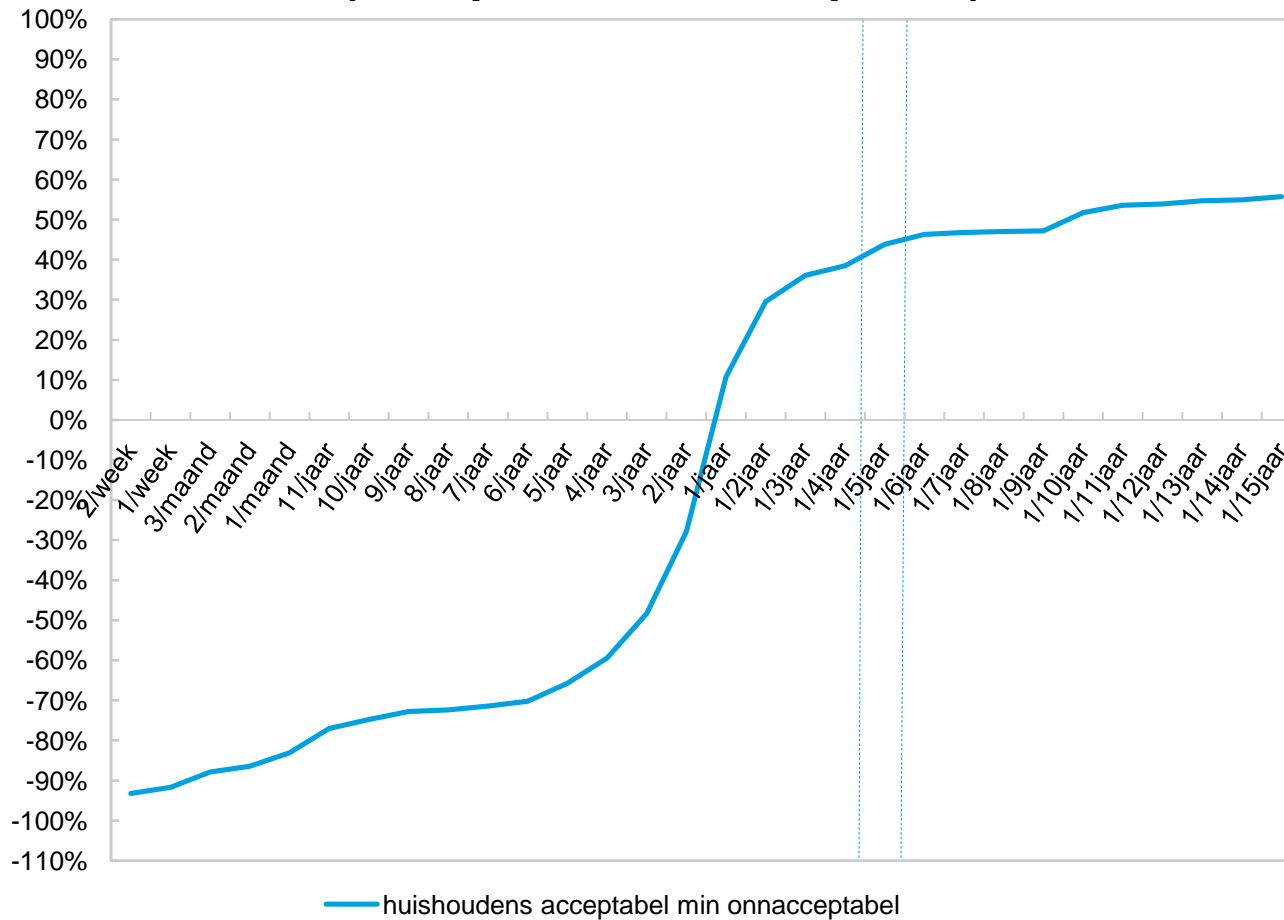
Acceptatie duur huishoudens (Acceptabel - Onacceptabel)



De knik die de maximale stijging aangeeft ligt op 5 minuten. Deze is wiskundig bepaald door de tweede afgeleide (de stijging van de stijging) te berekenen.

Bijlage Bepaling D0 en Fm

Acceptatie frequentie huishoudens (Acceptabel - Onacceptabel)



De knik die de maximale stijging aangeeft ligt op één onderbreking per vijf jaar. Deze is wiskundig bepaald door de tweede afgeleide (de stijging van de stijging) te berekenen.