

Bruikbaarheid van het Tobit model bij de analyse van Likertschalen in psychologisch onderzoek

Eindscriptie Bacheloropleiding Klinische Gezondheidspsychologie

Departement Psychologie en Gezondheid, Klinische Gezondheidspsychologie,
Universiteit van Tilburg

Auteur: Ben Koevoet

ANR: S389777

Begeleiders: A.A.J. Wismeijer & M.A.L.M. van Assen

Datum: 14 augustus 2009

Abstract

The utility of the Tobit model used to analyse Likert scales is investigated. Data from 113 psychology students were used. At two moments in time dysphoria, psychological quality of life and self-concealment were measured. The first session the dependent variable dysphoria was assessed using a 5-point Likert scale, while at the second session a 7-point Likert scale was used. The 2 supplementary categories were created by exaggerating the existing endpoints of the 5-point Likert scale. Using both Tobit and classical regression on the 5 and 7-point Likert scale it is investigated which model offers the best performance. This thesis contains evidence that supports the expectation that a Tobit analysis is more suitable for the analysis of a Likert scale than ordinary least squares (OLS) estimation.

Samenvatting

De bruikbaarheid van het Tobit model bij Likert schalen in psychologisch onderzoek is onderzocht. In een steekproef bestaande uit 113 studenten psychologie zijn op twee momenten dysforie, psychologische kwaliteit van leven en verbergen van het zelf gemeten. Tijdens de eerste meting is voor de afhankelijke variabele dysforie een 5-punts Likertschaal gebruikt en bij de tweede meting een 7-punts Likertschaal. Deze laatste is tot stand gekomen door extreme categorieën toe te voegen. In het onderzoek is getest of het Tobit model beter presteert dan klassieke regressie analyse. De verwachting dat het Tobit model beter past bij de analyse van een Likertschaal is door dit onderzoek gedeeltelijk bevestigend beantwoord.

Keywords: regression, Tobit, Likert scale, psychometrics, econometrics

Inleiding

In onderzoek wordt vaak aangenomen dat de waargenomen variabele normaal verdeeld is. In sommige gevallen kunnen niet alle waarden van de variabele correct gemeten worden. Een speciaal geval hiervan zijn gecensureerde waarnemingen. Een gedeelte van de waarnemingen wordt correct waargenomen maar een ander gedeelte wordt door een plafond- of vloereffect waargenomen op één punt.

Frances en Paap (2005) geven een voorbeeld hiervan; donaties aan goede doelen. Sommige mensen geven niets, anderen geven een beetje. De mensen die niet doneren zouden als ze de optie hadden wellicht een negatief bedrag hebben gegeven. Ze zouden geld willen ontvangen van het goede doel. Hier speelt een vloereffect, de vele waarnemingen met een gift van 0 zouden eigenlijk verspreid moeten liggen over de negatieve bedragen. Standaard regressie analyse zou onzuivere schattingen geven, de geobserveerde waarnemingen kloppen namelijk niet met het onderliggende verband.

Een manier om de onzuiverheid op te lossen is de extreme observaties (outliers) te verwijderen en alsnog een gewone regressieanalyse uitvoeren. Flood (1985) heeft echter aangetoond op basis van Monte-Carlosimulaties dat voor kansverdelingen onder normaliteitassumptie het Tobit model met maximum-likelihood schattingen beter presteert. Dit Tobit model zal nu besproken worden.

1.1. Het Tobit model

Om het bovenstaande vloereffect op te kunnen vangen heeft Tobin (1958) een statistisch model bedacht. Door Goldberger (1964) is dit model omgedoopt tot het Tobit model. Aan het klassieke regressie model wordt een latente variabele y_i^* toegevoegd:

$$y_i^* = \beta' x_i + \varepsilon_i,$$

$$\varepsilon_i | x_i \sim N(0, \sigma^2) \text{ iid},$$

$$y_i = y_i^* \quad \text{als } y_i^* > 0,$$

$$y_i = 0 \quad \text{als } y_i^* \leq 0.$$

In deze regressie is x een vector van onafhankelijke variabelen en β de bijbehorende vector met regressiecoëfficiënten. De fouten ε_i ($= y_i^* - \beta' x_i$) zijn onafhankelijke en identiek normaal verdeelde stochastische variabelen gegeven x_i . In het model bestaat de afhankelijke variabele y uit twee stukken. Er is een kansmassa op 0, terwijl voor waarden boven 0 een normale verdeling geldt zoals voor y_i^* . Met behulp van maximum-likelihood schatten kunnen de parameters β en σ geschat worden. De likelihood-functie is het product van de kansverdeling van de individuele waarnemingen. Deze is

$$L(\theta; x, y) = \prod_{i=1}^n \left[\Phi\left(\frac{-\mu_i}{\sigma}\right)^{I[y_i=0]} \left(\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{y_i - \mu_i}{\sigma}\right)^2\right) \right)^{I[y_i>0]} \right]$$

met $\theta = (\beta, \sigma)$ en $\mu_i = \beta' x_i$. Het eerste gedeelte van de vergelijking

$$\Phi\left(\frac{-\mu_i}{\sigma}\right)$$

geldt voor observaties $y = 0$ met Φ de cumulatieve kansverdelingfunctie van de normale verdeling. Het tweede gedeelte

$$\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{y_i - \mu_i}{\sigma}\right)^2\right)$$

geldt voor observaties met $y_i > 0$, dit is de kansverdelingfunctie van de normale verdeling met gemiddelde μ_i en standaarddeviatie σ (Hendrickx, 2007). De schatter is nu de θ^{MLE} die deze functie maximaliseert (Bain & Engelhardt, 1991). Het Tobit model kan met enige

aanpassingen ook gebruikt worden voor censurering op een andere waarde dan 0 en zelfs voor een gecombineerd plafond- en vloereffect.

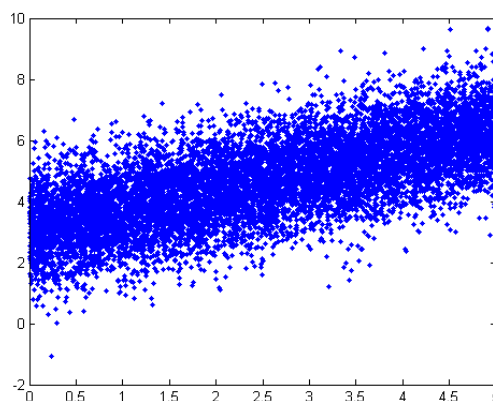
1.2. Het Tobit model toegepast op een Likertschaal

In psychologisch onderzoek wordt vaak gewerkt met een Likertschaal. Deze schaal bestaat uit de som van een aantal Likert items. De score op een Likert item komt tot stand doordat de proefpersoon een keuze maakt uit een set geordende antwoorden op een vraag van de onderzoeker. De set geordende antwoorden zijn gekoppeld aan een nummer (Likert, 1932). De antwoorden juist, gedeeltelijk juist, neutraal, gedeeltelijk onjuist en onjuist, welke gekoppeld zijn aan de nummers 1 tot en met 5 zijn hiervan een voorbeeld. Ze kunnen gebruikt worden als antwoord op de vraag: “Ik zie het leven positief in”.

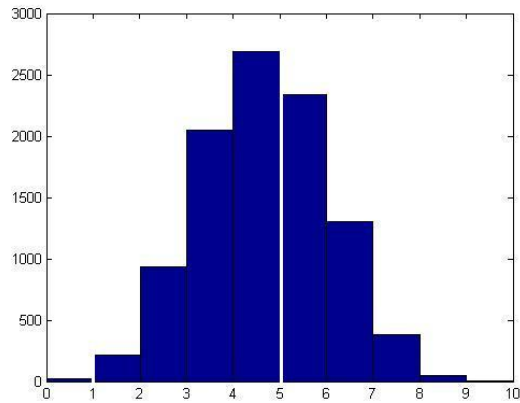
Stel nu dat in een populatie het volgende verband is tussen een afhankelijke en een onafhankelijke variabele

$$Y_i = 3 + (2/3) * X_i + e_i$$

De simulatie van het verband met $N = 10,000$ is te zien in figuur 1 en 2. Figuur 1 laat de puntenwolk zien van Y en X. Figuur 2 laat de verdeling van Y zien in een histogram.

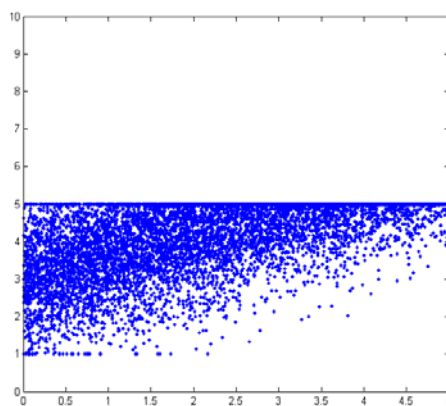


Figuur 1. Weergave gesimuleerde Y uitgezet tegen X.

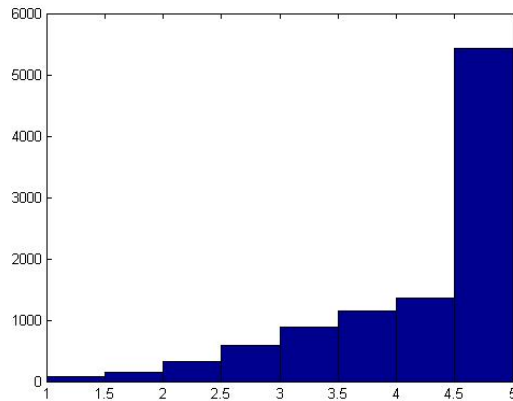


Figuur 2. Histogram van gesimuleerde werkelijke Y. De verticale strepen bij $X = 1$ en $X = 5$ geven de censurering weer.

Wanneer dit verband met een Likertschaal gemeten wordt, zijn waarnemingen onder 1 en boven de 5 slechts waarneembaar als 1 en 5. Deze censurering is weergegeven met verticale strepen in figuur 2. De gesimuleerde resultaten zijn analoog aan de voorgaande figuren weergegeven in figuur 3 en 4, maar nu na censurering.



Figuur 3. Gesimuleerde waargenomen Y uitgezet tegen X.



Figuur 4. Histogram van gesimuleerde gecensureerde Y.

In tabel 1 zijn de geschatte coëfficiënten weergegeven wanneer op bovenstaande data, inclusief gecensureerde waarnemingen, een klassieke regressieanalyse wordt uitgevoerd.

Tabel 1

Regressie coëfficiënten klassieke regressieanalyse gecensureerde waarnemingen.

Variabele	Coëfficiënt	SE	z-Waarde	p-Waarde
Constante	3.24*	0.01	227.47	0.00
Predictor X	0.40*	0.01	81.99	0.00

* $p < 0.05$

Note. SE: standaardfout.

Het effect van X op Y was 0.67, maar is nu verkeerd geschat als 0.40. In het algemeen zal deze onderschatting optreden bij normaal verdeelde variabelen. De regressiecoëfficiënt bij een OLS schatting zal vergeleken met de Tobit schatting in absolute waarde kleiner zijn

$$|\beta^{\text{OLS}}| < |\beta^{\text{Tobit}}|$$

Deze onzuiverheid kan ook invloed hebben op schattingen van effecten van andere variabelen.

Een klassieke regressieanalyse op de ongecensureerde data geven de in tabel 2 getoonde correcte schattingen.

Tabel 2

Regressie coëfficiënten klassieke regressieanalyse op ongecensureerde data.

Variabele	Coëfficiënt	SE	z-Waarde	p-Waarde
Constante	2.98*	0.02	148.60	0.00
Predictor X	0.68*	0.01	97.78	0.00

* $p < 0.05$

Note. SE: standaardfout.

Het Tobit model houdt in tegenstelling tot het klassieke regressie model rekening met het feit dat de waarnemingen gecensureerd zijn. Hierbij wordt verondersteld dat proefpersonen bij het beantwoorden van een Likert item categorieën misten. Daar waar deze personen “dit is extreem juist” gekozen zouden hebben, was er slechts de mogelijkheid juist te kiezen. Zij hadden dus een 5 ingevuld terwijl zij eigenlijk 6 of misschien wel hoger zouden kiezen. Het Tobit model geeft de in tabel 3 weergegeven correcte schattingen.

Tabel 3

Tobit analyse op gecensureerde data.

Variabele	Coëfficiënt	SE	z-Waarde	p-Waarde
Constante	2.98*	0.02	139.65	0.00
Predictor X	0.67*	0.01	77.22	0.00

* $p < 0.05$

Note. SE: standaardfout.

1.3. Het gebruik van het Tobit model in verschillende wetenschappen

Het Tobit model wordt bij economisch onderzoek al veelvuldig gebruikt in bijvoorbeeld de onderzoeken van Fang en Rizzo (2009), Farrel en Walker (1999), Barrera (1990) en vele anderen. Het van oorsprong econometrische model wordt momenteel onderzocht door andere wetenschappen op praktische bruikbaarheid. Psychologen Smithson en Verkuilen (2006) hebben verschillende methoden vergeleken om gecensureerde variabelen te analyseren. Zij komen bij hun onderzoek met een dataset over een leestest tot de conclusie dat de door hen gebruikte bèta kansverdeling beter presteert dan het Tobit model. Met hun methode wordt niet onderzocht wat er zou gebeuren wanneer de meetschaal zou worden uitgebreid, in hun analyse wordt slechts gebruik gemaakt van een maat voor de fit van de data.

Twisk en Rijmen (2009) komen tot de conclusie dat Tobit modellen bij longitudinale studies een betere fit geven dan lineaire modellen en dat deze theoretisch beter zijn uitgerust voor gecensureerde variabelen. Echter wanneer slechts een kleine proportie van de waarnemingen gecensureerd is, is de onzuiverheid van de lineaire schatters te verwaarlozen.

In *Social Science & Medicine* wordt het Tobit model al wat vaker gebruikt bijvoorbeeld door Curtis, Dooley en Phipps (2004) en Leigh en Fries (1993). Verscheidene takken van de wetenschap gebruiken dus het Tobit model al ter analyse van data. In deze thesis wordt het nut van de Tobit analyse voor psychologische data getoetst.

De operationalisatie van de verwachting dat een Tobit analyse beter past bij de analyse van een Likertschaal dan klassieke regressie analyse is niet eenvoudig. Respondenten vullen vragen met 7 antwoorden misschien anders in dan dezelfde vragen met 5 antwoorden. Dat wil zeggen, het kan voorkomen dat de verdeling van y^* verandert wanneer meer antwoordcategorieën toegevoegd worden. Als nu de Tobit en klassieke regressie analyse met

5-punts Likertschaal enerzijds met de klassieke regressie analyse op basis van een 7-punts Likertschaal anderzijds vergeleken worden is dus niet duidelijk of het verschil tussen Tobit en klassieke regressie analyse ligt in de verandering van y^* als gevolg van een verandering in het aantal antwoorden of in het tekort komen van categorieën.

Als daarentegen Tobit met 5 categorieën vergeleken wordt met Tobit met 7 categorieën zit het effect van het ontbreken van extreme categorieën niet in het verschil tussen deze twee procedures, want het Tobit model corrigeert hiervoor zowel bij 5 als 7 antwoordcategorieën. Alleen de verandering in y^* komt in het verschil tussen de laatstgenoemde 2 modellen voor. De klassieke regressie modellen op basis van een 5-punts en 7-punts Likertschaal kunnen in het geval van het ontbreken van extreme categorieën wel verschillen, bij de 5-punts schaal ontbreken deze extra categorieën. Bij de vergelijking tussen de klassieke regressie modellen komen dus zowel de verandering van y^* als het ontbreken van extreme categorieën naar voren.

Kort gezegd komen bij de vergelijking tussen de klassieke regressie analyse modellen wel de effecten van het ontbreken extremere schalen naar voren, terwijl deze bij de vergelijking tussen de Tobit modellen niet naar voren komen. De verwachting is dat het Tobit model beter past bij de analyse van een Likertschaal dan een klassieke regressieanalyse. In dat geval is het verschil tussen de coëfficiënten bij een Tobit analyse bij een 5 en 7-punts Likertschaal ongelijk aan het verschil tussen de coëfficiënten bij een klassieke regressieanalyse op een 5 en 7-punts Likertschaal.

Methode

2.1. Deelnemers aan het onderzoek en procedure

De steekproef bestond uit 113 psychologiestudenten (20 mannen, 93 vrouwen) van de universiteit van Tilburg, faculteit sociale wetenschappen. De gemiddelde leeftijd van de

steekproef was 20.10 jaar ($SD = 2.88$). 7 maanden na de eerste meting is de tweede meting uitgevoerd, op beide momenten zijn vragenlijsten afgenomen. Er werd gebruik gemaakt van de vragenlijsten die gebruikt zijn voor het onderzoek “Geheimen – piekeren – symptomen” van de universiteit van Tilburg. Uit deze vragenlijsten is een selectie van vragen gebruikt om de schalen dysforie, psychologische kwaliteit van leven en verbergen van het zelf te meten.

De vragenlijsten zijn per computer afgenomen. Tegen de studenten werd gezegd dat zichtbaar zou zijn wanneer ze de vragenlijsten niet serieus zouden invullen en dat zij dan proefpersoonuren niet bijgeschreven zouden krijgen. Aangezien de studenten proefpersoonuren moeten verzamelen ter behalen van hun diploma is dit een motivatie om de vragenlijst eerlijk in te vullen.

2.2. Variabelen

2.2.1 Dysforie

De DS14 vragenlijst meet type-D persoonlijkheid, deze bestaat uit twee factoren: negatieve affectiviteit en sociale inhibitie. Emons, Meijer en Denollet (2007) concluderen dat de DS14 deze twee factoren adequaat meten. Daarnaast komt Denollet (2005) tot de conclusie dat de DS14 een korte, goede vragenlijst is die toegepast kan worden in de praktijk. Dysforie is een subschaal van negatieve affectiviteit. De afhankelijke variabele in dit onderzoek is dysforie. Deze werd gemeten aan de hand van 3 items (“4. Ik voel me vaak ongelukkig”, “7. Ik zie de zaken somber in”, “13. Ik zit vaak in de put”). Tijdens de eerste meting werd een keuze gemaakt uit een 5-punts Likertschaal (laagste score 1 betekent ‘onjuist’, middelste score 3 betekent ‘neutraal’, hoogste score 5 betekent ‘juist’). Tijdens de tweede meting werd een keuze gemaakt uit een 7-punts Likertschaal (laagste score 1 betekent ‘geheel onjuist’, middelste score 4 betekent ‘neutraal’, hoogste score 7 betekent ‘geheel juist’). Alle items zijn

positief geformuleerd en zijn gesommeerd om tot de totaal score te komen. Cronbach's alpha bij de eerste meting was 0.89 en bij de tweede meting 0.87.

2.2.2 Psychologisch welbevinden

Er is een kortere versie ontwikkeld van de World Health Organization Quality of Life 100 (WHOQOL-100) vragenlijst die de kwaliteit van leven vaststelt, de WHOQOL-bref. Uit deze vragenlijst is de subschaal psychologische kwaliteit van leven geselecteerd als onafhankelijke variabele. De WHOQOL-bref heeft goede psychometrische eigenschappen (The WHOQOL Group, 1998). Het construct psychologische kwaliteit van leven werd gemeten aan de hand van 6 items. Tijdens zowel de eerste als de tweede meting werd bij 3 van de items een keuze gemaakt uit een 5-punts Likertschaal (laagste score 1 betekent 'helemaal niet', middelste score 3 betekent 'middelmatic', hoogste score 5 betekent 'een extreme hoeveelheid'), bij 2 items uit een alternatieve 5-punts Likertschaal (laagste score 1 betekent 'helemaal niet', middelste score 3 betekent 'gemiddeld', hoogste score 5 betekent 'helemaal' en laagste score 1 betekent 'erg ontevreden', middelste score 3 betekent 'ontevreden noch tevreden', hoogste score 5 betekent 'helemaal'). Het overige item was negatief geformuleerd op een 5-punts Likertschaal (laagste score 1 betekent 'nooit', middelste score 3 betekent 'zo nu en dan', hoogste score 5 betekent 'altijd'). De score op psychologische kwaliteit van leven is de som van de eerste 5 vragen met de omgescoorde waarde van het laatste negatief geformuleerde item hierbij opgeteld. Cronbach's alpha voor de eerste meting was 0.69 en voor de tweede meting 0.74.

2.2.3. Verbergen van het zelf

Op basis van de Tilburg Secrecy Scale-25 (Wismeijer et al, 2009) is de schaal "self-concealment" gemeten en vertaald met "verbergen van het zelf". Deze subschaal van de

Tilburg Secrecy Scale 25 bestaat uit 5 items gemeten op een 5-punts Likertschaal (laagste score 1 betekent ‘dit is geheel niet voor mij van toepassing’, middelste score 3 betekent ‘dit geldt enigszins voor mij’, hoogste score 5 betekent ‘dit is zeer voor mij van toepassing’). Doordat alle items positief zijn geformuleerd, kunnen deze opgeteld worden om tot een totale score te komen. Cronbach’s alpha voor de eerste meting was 0.88 en voor de tweede meting 0.82.

2.3. Analyse

De vragenlijsten zijn twee maal afgenomen; de eerste keer dat de vragenlijst is afgenomen is een 5-punts Likertschaal gebruikt voor de afhankelijke variabele, terwijl de tweede meting plaats vond op basis van een 7-punts-Likertschaal. Deze 7-punts-Likertschaal is geen standaarduitbreiding waardoor meer tussencategorieën ontstaan ten opzichte van een 5-punts-Likertschaal, maar een uitbreiding waarbij extreme categorieën worden toegevoegd. Naast de categorieën juist, eerder juist, neutraal, eerder onjuist en onjuist worden nu ook geheel juist en geheel onjuist als keuzemogelijkheid gebruikt.

Door de verschillen tussen Tobit analyse en klassieke regressie analyse kan niet zomaar de fit vergeleken worden: Het Tobit model geeft een likelihood-ratio, terwijl een OLS procedure een schatting van R^2 oplevert. In de literatuur staan nog andere methoden beschreven om de modellen te vergelijken zoals de Akaike information criterion (Akaike, 1974) of de Bayesian information criterion (Schwarz, 1978). In deze thesis wordt voor een alternatieve oplossing gekozen om de modellen te vergelijken.

De 7-punts Likertschaal van de 2^{de} meting wordt gezien als een betere representatie van het gemeten construct, doordat de proefpersonen met extremere responses deze nu ook kunnen geven. De coëfficiënten voor het intercept, de parameter die de samenhang tussen de score op de DS14 en psychologische kwaliteit van leven geeft en de parameter die de

samenhang tussen de score op de DS14 en het verbergen van het zelf weergeeft worden geschat.

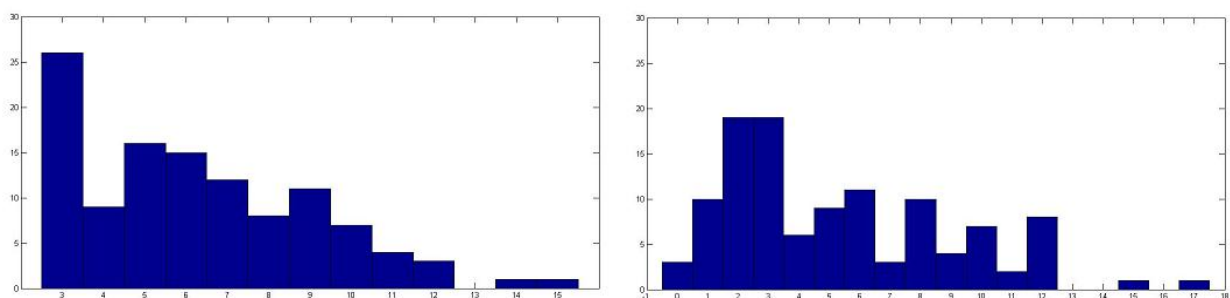
De afgenomen 5 en 7-punts Likertschalen worden onderworpen aan zowel een standaard regressie analyse als een gecensureerde regressieanalyse (Tobit). Deze analyses geven de schattingen voor de bovengenoemde parameters. Vervolgens wordt gezien of het verschil tussen de coëfficiënten verkregen met de Tobit analyse anders is dan het verschil tussen de coëfficiënten verkregen met de klassieke regressie analyses.

$$\beta_7^{Tobit} - \beta_5^{Tobit} \neq \beta_7^{OLS} - \beta_5^{OLS}$$

De software programma's SPSS en Eviews worden gebruikt. SPSS bevat namelijk niet de mogelijkheid een Tobit analyse uit te voeren.

Resultaten

Er waren 26 (23%) respondenten die op de eerste meting de laagst mogelijke score op dysforie aangaven, dat wil zeggen een score die op de linkergrens ligt. Er was 1 (0.88%) respondent van wie de score op de rechtergrens lag, deze persoon gaf de hoogst haalbare score op dysforie aan. Tijdens de tweede meting waren er nog slechts 3 (2.6%) scores die op de linkergrens lagen en lagen geen scores op de rechtergrens. Figuur 5 geeft in een histogram weer welke score de respondenten behaalden voor de afhankelijke variabele dysforie tijdens de eerste meting en de tweede meting.



Figuur 5. Histogram van dysforie; links op 5-punts Likertschaal, rechts op 7-punts Likertschaal.

De extra mogelijkheid “geheel onjuist” is dusdanig gebruikt dat nu de scores 0,1 en 2 ook gehaald zijn, tegelijkertijd is de frequentie van score 3 afgenomen. De vorm van de linkerstaart van het histogram benadert in de tweede meting meer de normale verdeling.

3.1. Afstand tussen de coëfficiënten binnen de gebruikte modellen

Tabel 4 geeft de schattingen weer voor de parameters van de regressievergelijking met dysforie als afhankelijke variabele en psychologische kwaliteit van leven (PKvL) samen met verbergen van het zelf (VZ) als onafhankelijke variabelen. Het intercept is hierbij ook opgenomen.

Tabel 4

Schattingen coëfficiënten volgens OLS en Tobit op basis van 5 en 7-punts Likertschalen.

Variabele	OLS5		OLS7		Tobit5		Tobit7	
	Coëf	SE	Coëf	SE	Coëf	SE	Coëf	SE
Intercept	17.88*	1.37	20.31*	1.97	20.00*	1.71	20.76*	1.99
PKvL	-0.62*	0.06	-0.79*	0.07	-0.76*	0.07	-0.81*	0.08
VZ	0.13*	0.04	0.13*	0.06	0.17*	0.05	0.13*	0.06

* $p < 0.05$

Note. Coëf: Coëfficiënt, SE: standaardfout, PKvL: Psychologische kwaliteit van leven, VZ: Verbergen van het zelf.

Duidelijk is te zien dat bij OLS schattingen van het intercept een groter verschil is tussen de metingen met 5 en 7-punts Likertschaal dan bij het Tobit model. De OLS schattingen, 17.88 en 20.31, verschillen 2.43, terwijl de Tobit schattingen 20.00 en 20.76, 0.76 verschillen. Dit wijst op een betere prestatie van het Tobit model.

De schattingen van de psychologische kwaliteit van leven verschillen 0.17 bij de klassieke regressie analyse, terwijl deze bij de Tobit analyse 0.05 verschillen. Wederom is de afstand tussen de coëfficiënten bij het Tobit model kleiner dan de afstand tussen de coëfficiënten bij het klassieke regressie model. Dit wijst ook op een betere prestatie van het Tobit model.

De laatste variabele VZ vertoont een tegengesteld effect, de OLS schattingen liggen dichter bij elkaar dan de Tobit schattingen, het gaat hier om een klein verschil. Het klassieke regressie model geeft schattingen met een verschil van 0.005 en de Tobit schattingen verschillen 0.04. De afstanden tussen de coëfficiënten binnen de Tobit en OLS schattingen verschillen, ook dit laatste resultaat ondersteunt de verwachting.

Discussie

In deze bachelor thesis is de bruikbaarheid van het Tobit model bij Likert schalen in psychologisch onderzoek onderzocht. De verwachting dat het Tobit model beter past bij de analyse van een Likert schaal is door dit onderzoek bevestigend beantwoord. De toevoeging van 2 extreme categorieën zorgt ervoor dat het vloereffect nagenoeg verdwijnt, wat een sterke aanwijzing is voor het terechte gebruik van de Tobit procedure wegens censurering. De regressiecoëfficiënten verkregen middels de Tobit procedure geven sterkere verbanden weer tussen de afhankelijke en onafhankelijke variabelen dan de klassieke regressie analyse. Laatstgenoemde algemeen geaccepteerde methode in de psychologie onderschat dus de

verbanden tussen variabelen. Meer onderzoek is nodig om het nut van het Tobit model nog overtuigender duidelijk te maken.

Tijdens dit onderzoek is ook gekeken naar het effect van de Tobit analyse op complete vragenlijsten, deze analyses zijn niet opgenomen in de resultaten. Zowel de Tobit als de OLS schatters gaven dezelfde waarden, omdat er op het totaal van 14 items in de DS14 geen vloer of plafond effect speelde. Belangrijk is op te merken dat Tobit in ieder geval niet slechter presteerde in gevallen dat het vloereffect niet speelde. Echter, zoals aangetoond geldt, op subschalen die slechts uit enkele items bestaan, dat het effect kan optreden dat OLS schattingen niet correct zijn. Wel is voorzichtigheid geboden voor de gevallen dat in een onderzoek de geobserveerde grensscores van een afhankelijke variabele bijdragen aan het werkelijke verband tussen variabelen. Zo ontstaat een schijn vloer- of plafondeffect waarbij geen censurering plaatsvindt. In dat geval zou eerder gedacht moeten worden aan een analyse aan de hand van een bèta kansverdeling zoals Smithson en Verkuilen (2006) opperen.

Men kan zich afvragen waarom niet gewoon de extreme categorieën toegevoegd moeten worden aan de 5-punts Likertschaal. Drenth en Sijtsma (2006) beargumenteren dat zowel een 5 als 7-punts Likertschaal zinvol gebruikt kunnen worden in psychologisch onderzoek, waarbij nog meer categorieën toevoegen geen voordeel meer oplevert. Het idee is dan dat de Tobit analyse overbodig is omdat nu ook de extreme meningen gegeven en gemeten kunnen worden. Het punt blijft dan echter dat zelfs na die schaaluitbreiding naar 7 items de censurering aanwezig blijft, zij het in mindere mate. De Tobit analyse is dan nog steeds zinvol.

Een aantal zaken zijn nog te verbeteren voor vervolgonderzoek. Er is in dit onderzoek gekozen voor de schaaluitbreidingen van onjuist naar geheel onjuist en van juist naar geheel juist. Wellicht dat de extreme categorieën met een ander label zoals “dit definieert mij als persoon” betere uitkomsten geven. Deze opmerking leidt ook tot een andere vraag; namelijk

of de categorieën wel goed geïnterpreteerd zijn. Door een regressieanalyse uit te voeren wordt aangenomen dat de onderlinge afstanden tussen de categorieën even groot zijn. Het verschil tussen geheel juist en juist, wordt dus net zo belangrijk geacht als het verschil tussen juist en eerder juist. Wellicht dat deze categorieën niet even ver uit elkaar liggen, waardoor een OLS schatting sowieso zinloos is. De vraag in deze thesis was echter of het Tobit model het beter zou doen dan de OLS schattingen. Gegeven dat er een OLS schatting gedaan wordt, kan men beter het Tobit model gebruiken aangezien het beter of gelijk presteert. Een mogelijk nadeel van het Tobit model is echter de complexiteit van de analyse vergeleken met de klassieke regressie analyse.

Dit voorgaande argument gaat ook op wanneer kritisch gekeken wordt naar de verdeling van de afhankelijke variabele. Bij zowel Tobit analyse als een OLS schatting wordt uitgegaan van een normale verdeling van de afhankelijke variabele als respectievelijk onderliggend model of werkelijke veronderstelling. In dit onderzoek is de data niet geheel normaal verdeeld, maar gegeven dat er gekozen wordt voor klassieke regressieanalyse, is het beter om een Tobit analyse uit te voeren.

Men zou zich naast bovenstaande punten kunnen afvragen of de 7 maanden die tussen de eerste en tweede meting zat nog invloed heeft gehad op de resultaten. Dit zou betekenen dat een verband, bijvoorbeeld tussen dysforie en psychologisch welbevinden sterker of zwakker zou zijn geworden. In dat geval zouden de coëfficiënten van de tweede meting niet hetzelfde verband schatten als de coëfficiënten van de eerste meting. Er is echter geen theoretische basis om aan te nemen dat het verband tussen de onderzochte variabelen in 7 maanden zou zijn veranderd.

Referenties

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716–723.
- Bain, L.J., & Engelhardt, M. (1991). *Introduction to probability and mathematical statistics*. Californië: Duxbury press.
- Barrera, A. (1990). The interactive effects of mother's schooling and unsupplemented breastfeeding on child health. *Journal of Development Economics*, 34, 81-98.
- Curtis, L. J., Dooley, M.D., & Phipps, S. A. (2004). Child well-being and neighbourhood quality: Evidence from the Canadian National Longitudinal Survey of Child and Youth. *Social Science & Medicine*, 58, 1917-1927.
- Denollet, J. (2005). DS14: Standard assessment of negative affectivity, social inhibition, and type D personality. *Psychosomatic Medicine*, 67, 89-97.
- Drenth, P. J. D. & Sijtsma, K. (2006). *Testtheorie: inleiding in de theorie van de psychologische test en zijn toepassingen*. Houten: Bohn Stafleu van Loghum.
- Emons, W. H. M., Meijer, R. R., & Denollet, J. (2007). Negative affectivity and social inhibition in cardiovascular disease: Evaluating type-D personality and its assessment using item response theory. *Journal of Psychosomatic Research*, 63, 27-39.
- Fang, H., & Rizzo, A. J. (2009). Competition and physician-enabled demand: The role of managed care. *Journal of Economic Behavior & Organization*. Ontleend: 21 juli, 2009, aan http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6V8F4WBT422-2-2&_cdi=5869&_user=522558&_orig=search&_coverDate=05%2F22%2F2009&_sk=999999999&view=c&wchp=dGLbVIW-zSkzV&md5=d3d9535c2cfdd5a5adeb9ef8bfc81bfl&ie=/sdarticle.pdf
- Farrell, L., & Walker, I. (1999). The welfare effects of lotto: Evidence from the UK. *Journal of Public Economics*, 72, 99-120.

- Flood, L. (1985). A Monte Carlo comparison of the maximum likelihood and corrected OLS estimators for Tobit models. *Economics Letters*, 19, 155-163.
- Franses, P H., & Paap, R. (2005). *Quantitative models in marketing research*. Cambridge: Cambridge University Press.
- Goldberger, A. S. (1964). *Econometric Theory*. New York: Wiley.
- Hendrickx, R. (2007). Orientation ME/EM course notes. Tilburg: Department of econometrics and operations research faculty of economics and business administration; Tilburg University.
- Leigh, J. P., & Fries J. F. (1993). Tobit, fixed effects, and cohort analyses of the relationship between severity and duration of rheumatoid arthritis. *Social Science & Medicine*, 36, 1495-1502.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 22, 55.
- Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6, 461-464.
- Smithson, M., & Verkuilen, J. (2006). A better lemon squeezer? Maximum-Likelihood regression with beta-distributed dependend variables. *Psychological Methods*, 11, 54-71.
- Tobin, J. (1958). Estimation of relationships for limited dependend variables. *Econometrica*, 26, 24-36.
- Twisk, J., & Rijmen, F. (2009). Longitudinal Tobit regression: A new approach to analyze outcome variables with floor or ceiling effects. *Journal of Clinical Epidemiology*.
- Ontleend: 17 juli, 2009, aan http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6T84-4VK6939-1-7&_cdi=5076&_user=522558&_orig=search&_cover

Date=02%2F10%2F2009&_sk=999999999&view=c&wchp=dGLbVlb-zSkzV&md5
=a174e767f9260dd40e08580ac3b6f0cd&ie=/sdarticle.pdf

The WHOQOL Group (1998). Development of the World Health Organization WHOQOL-

bref. Quality of life assessment instrument. *Psychological Medicine*, 28, 551–558.

Wismeijer, A. A. J., van Assen, M. A. L. M., Sijtsma, K., & Vingerhoets, A. J. J. M. (2009).

Development and dimensionality testing of the Tilburg Secrecy Scale-25 (TSS25).

Ingestuurd voor publicatie.