

Gewonnen, maar ook gezond?

Het effect van loterijprijzen op gezondheidsgedrag

Simone Reij

Bachelorscriptie Economie & Bedrijfseconomie

Erasmus Universiteit Rotterdam

Juli 2012



Gewonnen, maar ook gezond?

Het effect van loterijprijzen op gezondheidsgedrag

Bachelorscriptie Economie & Bedrijfseconomie

Auteur	Simone Reij
Studentnummer	323566
Opleiding	Economie & Bedrijfseconomie
Datum	juli 2012
Plaats	Maassluis
Begeleider	Dr. J.L.W. (Hans) van Kippersluis
Universiteit	Erasmus Universiteit Rotterdam
	Burgemeester Oudlaan 50
	3062 PA Rotterdam

Dit onderzoek gebruikt Britse paneldata om het exogene inkomenseffect op een aantal gezondheidsgedragingen te testen: sporten, vrije tijd en roken. Loterijprijzen worden gebruikt als exogene inkomensverandering. De positieve inkomensschokken hebben een positief significant effect op sporten, vrije tijd en roken. Enerzijds zal een positieve inkomensverandering leiden tot positief gezondheidsgedrag, meer sporten en meer uitgaven aan vrije tijd. Anderzijds zal een positieve inkomensverandering leiden tot negatief gezondheidsgedrag, meer roken

Kernwoorden: inkomen; gezondheidsgedrag; loterijprijzen; roken; sporten; vrije tijd.

Inhoudsopgave

1. Inleiding	4
1.1 Relevantie	4
1.2 Onderzoeken	5
1.3 Aanpak	6
2. Theoretisch kader	8
3. Data	11
3.1 Gezondheidsgedrag	11
3.1.1 Positief gezondheidsgedrag	11
3.1.2 Negatief gezondheidsgedrag	12
3.2 Loterijprijzen	12
4. Methode	14
4.1 Gebruik van de loterijdata	14
4.1.1. Winnaars tegenover niet-winnaars	14
4.1.2 Winnaars tegenover onsuccesvolle spelers	15
4.1.3 Kleine winnaars tegenover grote winnaars	15
4.2 Tijdsaspect	15
4.3 Regressie	16
5. Resultaten	17
5.1 Inkomenseffect op positief gezondheidsgedrag	17
5.2 Inkomenseffect op negatief gezondheidsgedrag	18
5.3 Regressiemodel exclusief controlevariabelen	19
6. Interpretatie	20
7. Discussie	22
7.1 Toevoeging variabelen	22
7.2 Paneldata	22
7.3 Variabele Bedragvrijetijd gezondheidsbevorderend?	23
7.4 Loterijdata	23
8. Referenties	25
9. Appendix	26

1. Inleiding

1.1 Relevantie

De relatie tussen individueel inkomen en gezondheid is een onderwerp wat al aan vele studies is onderworpen. Een hoger inkomen is geassocieerd met een betere gezondheid, is de algemene bevinding. Deze relatie is gevonden in een groot aantal landen en voor een grote variatie van gezondheidsvariabelen (Ettner, 1996, Lindahl, 2005, Meer et al., 2003).

Hoewel gezondheidsgedrag invloed heeft op de gezondheid, hebben maar enkele onderzoeken gezondheidsgedrag in hun analyse meegenomen.

Wanneer er vanuit wordt gegaan dat individuen hun nutsfunctie maximaliseren, waarbij de nutsfunctie wordt gedefinieerd over gezondheid en andere goederen die onderworpen zijn aan begroting en tijdsdruk, zal een positieve schok in inkomen leiden tot een verbreding van de budgetrestrictie en dus leiden tot een betere gezondheid. Hier wordt er vanuit gegaan dat gezondheid een normaal goed is. Gezondheid kan men echter niet kopen, tevens is gezondheid afhankelijk van andere elementen in de nutsfunctie.¹ Zo is de consumptie van sigaretten, alcohol en drugs negatief gecorreleerd met de gezondheid. Als er wordt uitgegaan dat sigaretten, alcohol en drugs normale goederen zijn, is het nut positief gecorreleerd met de consumptie van deze goederen. Is gezondheidsgedrag dan wèl van belang om mee te nemen in de analyse van inkomen op gezondheid?

Evenzeer staat de positieve relatie tussen inkomen en gezondheid voor volwassenen open voor een aantal interpretaties: de causaliteit is van inkomen op gezondheid, hoe hoger het inkomen hoe meer men in gezondheid kan investeren; van gezondheid op inkomen, gezond zijn stelt je in staat te werken en dus inkomen te genereren; of beide zijn beïnvloed door verborgen gemeenschappelijke factoren. Een gemeenschappelijke factor is bijvoorbeeld opleiding; opleiding heeft een causaal effect op zowel gezondheid als inkomen (Card, 1999, Arendt 2005). Het testen van een causale relatie van inkomen op gezondheid vereist exogene bewegingen in inkomen; de verandering in inkomen mag niet afhangen van andere factoren en moet dus aselekt voorkomen.

Dit onderzoek zal zich niet richten op het inkomenseffect op de gezondheid zelf, maar op het inkomenseffect op gezondheidsgedrag. Dit zal worden benaderd door middel van exogene inkomensveranderingen.

¹ Zie theoretisch kader

1.2 Onderzoeken

Hieronder worden een aantal onderzoeken besproken waarin het exogene inkomenseffect op gezondheid is onderzocht.

Lindahl (2005) test het inkomenseffect op gezondheid en sterfte met Zweedse longitudinale gegevens. Loterijprijzen worden hier gebruikt als exogene variatie in het inkomen. Een positieve causale relatie is gevonden tussen inkomen en gezondheid. Echter zijn in de metingen van de gezondheid geen gezondheidsgedragingen meegenomen.

Meer et al. (2003) maken een analyse over de invloed van welvaart op de gezondheid. Om het probleem omgekeerde causaliteit op te lossen, gebruiken ze erfenissen als exogene variabele. Men concludeert dat de connectie tussen welvaart en gezondheid niet wordt gedreven door een verandering van de welvaart op korte termijn. Tevens heeft men in dit onderzoek het gezondheidsgedrag niet meegenomen.

Onderzoeken die wel het gezondheidsgedrag in hun analyse hebben meegenomen:²

Ettner (1996) schat het effect van inkomen op gezondheid met een Amerikaanse dataset. Ze maakt gebruik van verschillende metingen voor gezondheid, zo ook het gezondheidsgedrag met betrekking tot alcoholmisbruik. Ze test op het gemiddelde dagelijkse alcoholgebruik tijdens de afgelopen twee weken en het aantal keer dat men alcoholisch gedrag vertoond heeft in het voorgaande jaar. Het probleem van omgekeerde causaliteit lost ze op door toevoeging van instrumenten: de staatswerkloosheid, werkervaring, opleiding ouders en echtelijke kenmerken. Er kan echter worden tegengegaan dat deze instrumenten exogeen zijn.³ De uitkomst van het onderzoek: een positieve verandering van het inkomen verbetert zowel de mentale als psychische gezondheid, maar verhoogt de prevalentie van alcoholgebruik. Ze veronderstelt dat dit erop wijst dat alcohol een normaal goed is; niet dat personen met een hoger inkomen meer kans hebben op alcoholmisbruik.

Apouey & Clark (2009) gebruiken Britse panel data (British Household Panel Survey, BHPS) (1997-2005) om zo het exogene inkomenseffect op een aantal individuele gezondheidsuitkomsten te onderzoeken. Een van de gezondheidsuitkomsten is gezondheidsgedrag, in dit artikel test men de variabele roken en drinken. Als exogene inkomensverandering wordt gebruik gemaakt van loterijprijzen. De uitkomst is dat exogeen

² Ruhm (2003) heeft onderzoek verricht naar het effect van een tijdelijke economische recessie op gezondheidsgedrag. Roken en gewicht verlagen gedurende tijdelijke economische recessies, terwijl lichaamsbeweging in de vrije tijd stijgt. Een reductie in werktijden kan een reden zijn voor gezonder gedrag tijdens een verzwakte economie. Echter is er weinig bewijs gevonden dat een verlaging van het inkomen een belangrijke rol speelt voor de gezondere levensstijl.

³ Meer et al. (2003) merkt op dat de staatswerkloosheid alleen een valide instrument is als regionale variaties in gezondheid enkel de variaties in inkomen weerspiegelen.

inkomen geassocieerd is met groter risicovol gezondheidsgedrag: loterijwinnaars roken meer en zijn meer geneigd tot sociaal drinken. Echter heeft men hier geen positief gezondheidsgedrag meegenomen, zoals sporten.

1.3 Aanpak

Dit onderzoek zal zich richten op loterijwinnaars, om zo een causale relatie tussen exogene veranderingen in het inkomen en veranderingen in gezondheidsgedrag te kunnen onderzoeken.

Het onderzoek van Apouey & Clark (2009) zal als leidraad dienen. De nadruk zal liggen op de geteste invloed van inkomen op de gezondheiduitkomst gezondheidsgedrag. Apouey & Clark (2009) testen voor gezondheidsgedrag de variabelen roken en drinken. Dit onderzoek zal zich voornamelijk richten op het gezondheidsgedrag roken. Daarnaast zullen er een aantal variabelen, die wordt geassocieerd met gezondheidsgedrag, extra worden meegenomen.

Met betrekking tot gezondheid kan gezondheidsgedrag worden onderverdeeld in positief gedrag, bijvoorbeeld sporten, ontspanning, consumptie van groente en fruit, en negatief gedrag, bijvoorbeeld consumptie van sigaretten, alcohol en drugs. Er zal in dit onderzoek een aantal positieve, sporten en ontspanning, en negatieve, roken, gedragingen belicht worden.

In het onderzoek zal van dezelfde dataset als Apouey & Clark (2009) gebruik worden gemaakt; British Household Panel Survey (BHPS). Anders, dan in Apouey & Clark, is een uitbreiding van het onderzoek met beschikbare data tot en met 2008.

De toevoeging van dit onderzoek is zodoende tweeledig:

1. Door gebruik te maken van een langere dataset zal worden getest of de uitkomsten van het rookgedrag van het onderzoek (Apouey & Clark) nog steeds gelden.
2. Positief gezondheidsgedrag wordt, naast negatief gezondheidsgedrag, getoetst op het effect van inkomen. Zo is er een bredere omvatting van gezondheidsgedrag en is er zodoende een betere mogelijkheid om een conclusie te kunnen trekken wat betreft het inkomenseffect op gezondheidsgedrag.

Zal men ten gevolge van een positieve inkomensverandering positief gezondheidsgedrag vertonen of juist negatief gezondheidsgedrag? Of zal men beide, positief en negatief gezondheidsgedrag, vertonen? Oftewel, wat is de invloed van een inkomensverandering op het gezondheidsgedrag?

Dit onderzoek is opgebouwd als volgt. Het theoretisch kader zal het verband tussen nut, gezondheid en gezondheidsgedrag uitlichten. Het kopje data presenteert de aangewende variabelen van het BHPS. Methode geeft weer hoe de analyse is uitgevoerd en hoe de variabelen hierin worden toegepast, waarna het kopje resultaten volgt. De interpretatie geeft een conclusie van de gevonden resultaten en als laatste wordt er afgesloten met de discussie.

2. Theoretisch kader

Grossman (1972a) heeft een model ontwikkeld om het basisproduct gezondheid te analyseren; de vraag naar gezondheid. Het centrale uitgangspunt van het model is dat gezondheid een duurzaam product is. Men erft een gezondheidsniveau, dit gezondheidsniveau deprecieert over tijd en kan worden vergroot door investeringen.⁴

De nutsfunctie van een individu is afgeleid uit de consumptie van goederen en het gezondheidsniveau.

Vergelijking 1

$$U = U(C_t, H_t)$$

C_t is de consumptie van een samengesteld consumptiegoed, goederen anders dan gezondheid.

H_t is het gezondheidsniveau⁵, H_0 is het geërfde gezondheidsniveau. t staat voor tijd.

De netto investering in het gezondheidsniveau is gelijk aan de bruto investering verminderd met de depreciatie.

Vergelijking 2

$$H_{t+1} - H_t = I_t - \delta_t H_t$$

I_t is de bruto investering, δ_t is de exogene depreciatiegraad.

Uit bovenstaande kan worden afgeleid dat gezondheid geen normaal goed is. Men kan gezondheid niet kopen, men kan investeren in gezondheid, tegelijkertijd gaat de gezondheid erop achteruit.

Galama & Van Kippersluis (2010) haken op Grossmans model in. Zij hebben een levenscyclusmodel ontwikkeld waarin meerdere mechanismen zijn opgenomen die een groot deel van de waargenomen verschillen in gezondheid in sociale economische status verklaren.

⁴De vergelijkingen die hier worden gebruikt verschillen van de originele vergelijkingen. De originele vergelijkingen van Grossman luiden; $U = U(\phi_0 H_0, \dots, \phi_n H_n, Z_0, \dots, Z_n)$; $H_{i+1} - H_i = I_i - \delta_i H_i$. Waarbij H_n het gezondheidsniveau is, Z_n de consumptie van een samengesteld consumptiegoed anders dan gezondheid, I_i de bruto investering in gezondheid en δ_i de exogene depreciatiegraad.

⁵Grossman stelt dat het nut dat men uit gezondheid verleent niet direct voortkomt uit gezondheid, maar voortkomt uit de activiteiten die men kan doen doordat men gezond is. Het nut hangt dus niet af van H_t maar van $\phi_t H_t$, Grossman noemt dit "health services". Om het model zo helder mogelijk te maken wordt ϕ_t hier weggelaten.

In bovenstaand model hebben zij rekening gehouden met de invloed van verschillende goederen op gezondheid. Er wordt onderscheid gemaakt tussen gezonde consumptie en ongezonde consumptie.⁶

De nutsfunctie van een individu is afgeleid uit de gezondheid en de consumptie van goederen; gezonde consumptie en ongezonde consumptie.

Vergelijking 3

$$U_t = U(C_t^h, C_t^u, H_t)$$

C_t^h is gezonde consumptie, C_t^u is ongezonde consumptie en H_t is gezondheid, tijd t geeft de tijd weer vanaf het moment dat een individu zijn studie heeft afgerond en de arbeidsmarkt op is gegaan. Het nut is voor zowel gezonde als ongezonde consumptie positief afhankelijk. Gezondheid verbetert door investeringen in medische zorg⁷ en verslechtert met de biologisch verouderingsratio.

Vergelijking 4

$$H_{t+1} - H_t = I_t^{m\alpha} - d_t H_t$$

Met I_t^m medische zorg⁸ en d_t als de endogene biologisch verouderingsratio.⁹ De biologische verouderingsratio heeft men uitgesplitst.

Vergelijking 5

$$d_t = d(t, C_t^h, C_t^u, z_t, I_t^p; \varepsilon_t)$$

De biologische verouderingsratio is endogeen afhankelijk van gezonde consumptie C_t^h , ongezonde consumptie C_t^u , werk gerelateerde gezondheidsstress z_t , investeringen in preventieve zorg I_t^p en een vector van exogene functies ε_t . De verouderingsratio is negatief afhankelijk van gezonde consumptie en preventieve zorg en is positief afhankelijk van ongezonde consumptie en werk gerelateerde gezondheidsstress.

⁶De originele vergelijkingen van Galama & Kippersluis luiden; $U(t) = U[C_h(t), C_u(t), H(t)]; \dot{H}(t) = I_m(t)^\alpha - d(t)H(t); d(t) = d[t, C_h(t), C_u(t), z(t), I_p(t); \varepsilon(t)]$. Waarbij $C_h(t)$ gezonde consumptie is, $C_u(t)$ ongezonde consumptie, $I_m(t)$ medische zorg, $d(t)$ het endogene biologisch verouderingsratio, $z(t)$ werkgerelateerde gezondheidsstress, $I_p(t)$ investeringen in de preventieve zorg en $\varepsilon(t)$ een vector van exogene functies.

⁷Naast onderscheid in gezonde en ongezonde consumptie, maken Galama & Van Kippersluis ook onderscheid in investeringen in de gezondheid: investeringen in medische zorg en investeringen in preventieve zorg.

⁸Er wordt aangenomen dat de gezondheidsproductie functie $I_t^{m\alpha}$ afnemende meeropbrengsten vertoont ($0 < \alpha < 1$).

⁹Dit is een verschil met het model van Grossman. Grossman gaat uit van een exogene depreciatiegraad δ_t .

Beide modellen houden rekening met de budgetrestrictie. De hoeveelheid die men consumeert aan gezonde en ongezonde consumptie en investeert in gezondheid hangt af van en is beperkt door het inkomen. Bovendien hangt de hoeveelheid af van de prijzen van de verschillende goederen.

Uit het model van Galama & Van Kippersluis kan het volgende worden opgemerkt: Gezonde consumptie heeft een directe en een indirecte positieve invloed op het nut.

Gezonde consumptie verlaagt de verouderingsratio waardoor dit een positief effect heeft op de gezondheid en dus indirect een positief effect op het nut.

Ongezonde consumptie heeft een directe positieve invloed op het nut, maar tegelijkertijd indirect een negatieve invloed op het nut. Ongezonde consumptie verhoogt de verouderingsratio en heeft daardoor een negatieve invloed op de gezondheid, dus een indirecte negatieve invloed op het nut.

Vanuit dit perspectief rijzen de volgende vragen op: Zal ten gevolge van een positieve inkomensverandering de gezonde consumptie worden verhoogd of zal juist de ongezonde consumptie worden verhoogd? Of zullen beide, gezonde en ongezonde consumptie, verhogen?

Een verhoging van gezonde consumptie zou het meest vanzelfsprekend zijn. Gezonde consumptie heeft namelijk een direct positieve en tegelijkertijd een indirect positieve invloed op het nut.

Theoretisch is de invloed van een inkomensverandering op ongezonde consumptie onduidelijk. Ongezonde consumptie heeft namelijk een direct positieve en tegelijkertijd een indirect negatieve invloed op het nut. Dit effect zal empirisch worden getoetst.

3. Data

In dit onderzoek wordt gebruik gemaakt van het British Household Panel Survey (BHPS). Deze paneldata begon in 1991. De eerste golf bestaat uit 5.500 huishoudens en 10.300 individuen uit 250 gebieden. In 1999 en 2001 werden huishoudens uit Schotland, Wales en Noord-Ierland toegevoegd. De panel is zo geschikt geworden voor onderzoek van het gehele Verenigd Koninkrijk. Nu bestaat het panel uit ongeveer 9.000 huishoudens en 16.000 individuen. De golven 7 tot en met 18 (1997-2008) worden in dit onderzoek benut, in eerdere golven was er nog geen informatie over gestandaardiseerde loterijen¹⁰. De BHPS omvat een brede omvang van informatie over individuele en huishoudelijke demografie, geestelijke en lichamelijke gezondheid, de status van arbeidsparticipatie en banen. Alle volwassenen in de dezelfde huishoudens worden apart geïnterviewd. Nadere details van het onderzoek kan men vinden op de volgende website: <http://www.iser.essex.ac.uk/ulsc/bhps/>.

De variabelen die in het onderzoek worden gebruikt, worden weergegeven in Tabel 1 in de appendix weergegeven. De voornaamste variabelen worden hieronder in het kort toegelicht. De letterlijke vragen uit de BHPS worden in tabel 2 in de appendix weergegeven.

3.1 Gezondheidsgedrag

Er zal onderscheid worden gemaakt in positief gezondheidsgedrag en negatief gezondheidsgedrag.

3.1.1 Positief gezondheidsgedrag

De eerste variabele voor positief gezondheidsgedrag test het aantal keer dat men een sport speelt, loopt of zwemt, genaamd “Sport”. De respondent is gevraagd aan te geven hoe frequent men een van deze activiteiten doet. De antwoorden zijn gecodeerd van waardes 1 tot en met 5. Waarde 1 geeft nooit/bijna nooit aan, waarde 2 geeft één keer per jaar aan, waarde 3 een aantal keer per jaar, waarde 4 geeft één keer per maand aan; hoe hoger de waarde hoe meer men aan deze activiteiten doet, met waarde 5 als ten minste één keer per week.¹¹ Deze variabele komt voor in golven met even cijfers, om het jaar dus, beginnend in golf 6 en eindigend in golf 18.

De tweede variabele voor positief gezondheidsgedrag test de hoeveelheid geld dat men in een gemiddelde maand aan vrije tijd besteed, genaamd “Bedragvrijetijd”. De respondent is gevraagd aan te geven hoeveel men persoonlijk spendeert in een gemiddelde maand aan

¹⁰ In 1994 werd de nationale loterij in het Verenigd Koninkrijk ingewijd.

¹¹ Deze variabele is vanwege een betere interpretatie in de analyse hergecodeerd. In de BHPS staat waarde 1 voor ten minste één keer per week en geeft waarde 5 nooit/bijna nooit aan.

vrijtijdsactiviteiten, amusement en hobby's, anders dan uit eten (exclusief vakanties, inclusief alle activiteiten zowel binnen- als buitenshuis). De antwoorden in de BHPS zijn gecodeerd van 0 tot en met 12. Hoe meer men spendeert hoe hoger de waarde, waarde 0 staat voor £0,- en waarde 12 voor £160,- of meer. Voor dit onderzoek is de codering van 0 tot en met 12 hergecodeerd naar de gemiddelde hoeveelheid geld in deze categorieën. In de BHPS stelt waarde 1 een bedrag tussen de 1 en 10 voor, waarde 1 is hergecodeerd naar een gemiddeld bedrag van £5,-. Waarde 2 stelt in de BHPS een bedrag tussen de 10 en 20 voor, waarde 2 is hergecodeerd naar een gemiddeld bedrag van £15,-, etc. De laatste waarde, waarde 12, staat voor een bedrag van £160,- of meer, deze is hergecodeerd naar een bedrag van £170,-. De variabele komt vanaf golf 7 in elke golf voor.

3.1.2 Negatief gezondheidsgedrag

Als negatief gezondheidsgedrag wordt het rookgedrag getest. Er wordt gebruik gemaakt van twee verschillende variabelen voor roken. De eerste is een binaire variabele die aangeeft of de respondent een "huidige roker" is of niet. Deze variabele wordt "Roker" genoemd, het gaat hier alleen om sigarettengebruik, het gebruik van sigaren en pijpen is zodoende uitgesloten. De vraag die de respondent is gesteld, is: "Rook je sigaretten?". Waarde 0 geeft nee aan, waarde 1 ja.¹²

De tweede variabele, genaamd "Sig", geeft het aantal sigaretten dat een individu per dag rookt weer. De waarde geeft aan hoeveel sigaretten men per dag rookt. Beide variabelen komen in alle golven¹³ voor.

3.2 Loterijprijzen

Dit onderzoek richt zich op de relatie tussen inkomen en gezondheidsgedrag. Als exogene verandering in het inkomen wordt er gebruik gemaakt van twee BHPS vragen over loterijprijzen. Deze vragen verschijnen vanaf 1997 elk jaar. De eerste vraag luidt:

"Heeft u sinds 1 september vorig jaar enige betalingen ontvangen, of betalingen in natura ontvangen, door het winnen van voetbalpools, het winnen van de nationale loterij of door een andere vorm van gokken?"

¹² In de BHPS ontbreekt de variabele Roker in golf 9. De variabele is in dit onderzoek voor golf 9 opgevuld met een andere variabele. Deze variabele komt enkel in golf 9 voor. De vraag die de respondent wordt gesteld is: rook je überhaupt sigaretten tegenwoordig? De precieze vraag in het Engels luidt: Do you smoke cigarettes at all nowadays? Waarde 0 is nee, waarde 1 is ja.

¹³ Apouey & Clark (2009) hebben ervoor gekozen om deze variabele in schalen her te coderen. Hier is opzettelijk niet voor hercodering gekozen. De uitkomst van de analyse kan op deze wijze namelijk in aantal sigaretten worden uitgedrukt. Wanneer het aantal sigaretten wordt hergecodeerd naar een schaalverdeling, kan dit niet.

Als het antwoord positief werd beantwoord, werd de respondent gevraagd:

“Hoeveel heeft u in totaal ontvangen?”

De variabele is uitgedrukt in ponden en wordt “Lotprijs” genoemd.

Bekend is dus welke individuen er wonnen en hoeveel zij aan prijzen ontvingen. Er zijn echter twee zwaktes van de loterijdata in de BHPS.¹⁴ Ten eerste bevat het geen directe informatie over het aantal keer dat een individu met een loterij meespeelt.¹⁵ Zodanig kan er geen onderscheid worden gemaakt tussen niet-spelers en onsuccesvolle spelers. Daarnaast is niet bekend hoeveel geld er is vergokt door zowel de loterijwinnaars als de onsuccesvolle spelers. Er is dus wel sprake van kennis van de bruto-opbrengst maar geen sprake van kennis van de netto-opbrengst.

Aan de andere kant zijn er ook voordelen in het gebruik van loterijprijzen. De ontvangst van een loterijprijs is exogeen; willekeurig worden de winnaars uit de groep spelers gekozen. Bovendien spelen in Groot-Brittannië veel mensen mee in loterijen.¹⁶

De nominale loterijprijzen zijn gecorrigeerd voor inflatie via het consumenten prijs indexcijfer (CPI) met als basisjaar 2005. In de empirische analyse zal het logaritme van de loterijprijzen worden gebruikt. In empirische analyses van gezondheid en welzijn wordt inkomen veelvuldig in logaritmische vorm gebruikt.¹⁷ Met gebruik van het logaritme brengt men de extreme waarden terug, zodat de extreme waarden de uitkomst niet vertekenen.

¹⁴ Deze zwaktes zijn opgemerkt in Apouey & Clark (2009).

¹⁵ Deze zwakte verschijnt ook in de Zweedse loterij data die Lindahl (2005) gebruikt.

¹⁶ Wardle et al. (2007) schat dat twee derde van de Britten deelneemt aan een vorm van gokken, waarvan 57% van de populatie meespeelt in de nationale loterij.

¹⁷ Apouey & Clark (2009) en Lindahl (2005) nemen het logaritme van loterijprijzen. Ettner (1996) neemt het logaritme van het inkomen.

4. Methode

Om het inkomenseffect op gezondheidsgedrag te testen is er voor gekozen om gebruik te maken van een regressieanalyse, met loterijprijzen als exogene inkomensverandering. Naast de onafhankelijke variabele loterijprijzen wordt er gebruik gemaakt van een aantal controlevariabelen: burgerlijke staat¹⁸, etniciteit, geboortejaar, geslacht, golf, grootte huishouden, leeftijd, opleiding, regio en werkstatus. Zodanig wordt er gecorrigeerd voor variabelen die beide inkomen en gezondheidsgedrag beïnvloeden. De afhankelijke variabelen voor gezondheidsgedrag zijn sporten, bedrag dat aan ontspanning wordt uitgegeven en roken. De methode die wordt gebruikt komt grotendeels overeen met het onderzoek van Apouey & Clark (2009).

4.1 Gebruik van de loterijdata

Aangezien de positieve relatie tussen inkomen en gezondheid voor volwassenen open staat voor een aantal interpretaties, vereist het testen van een causale relatie van inkomen op gezondheid exogene bewegingen in inkomen. Om deze reden is ervoor gekozen om loterijprijzen als exogene inkomensverandering op te nemen. Loterijprijzen worden namelijk willekeurig verdeeld.

Om met behulp van loterijdata het effect op gezondheidsgedrag te testen, beperkt de wijze waarop de loterijdata wordt gebruikt zich tot de winnaars. Winnaars van hoge prijzen worden vergeleken met winnaars van lage prijzen.

Deze methode is niet de enige methode om loterijdata te gebruiken. Met de beschikbare data maakt deze methode het meest een exogeen onderscheid in de verklarende variabele.

4.1.1. Winnaars tegenover niet-winnaars

De methode had zich onder andere kunnen richten op winnaars en niet-winnaars; winnaars vergelijken met niet-winnaars. Deze methode heeft echter niet de voorkeur.

Het meespelen aan de loterij is endogeen afhankelijk: niet-spelers en spelers verschillen in zowel observeerbare als in niet observeerbare karakteristieken van elkaar. Zo kan het meespelen aan de loterij afhangen van risicoaversie, inkomen en leeftijd. De groep niet-winnaars bestaat enerzijds uit onsuccesvolle spelers en anderzijds uit niet-spelers, binnen de groep bestaat er dus onderscheid tussen spelers en niet-spelers. De groep winnaars bestaat uit alleen spelers. Het vergelijken van winnaars met niet-winnaars zou een vertekend beeld geven

¹⁸ De BHPS heeft in de golven 16 tot en met 18 de opties voor het beantwoorden van de controlevariabele burgerlijk staat uitgebreid met een optie geregistreerd partnerschap. Deze optie is in dit onderzoek gelijkgesteld aan de optie getrouwd. In golf 18 heeft een tweede uitbreiding plaats gevonden met de optie ontbonden geregistreerd partnerschap. Deze optie is in dit onderzoek gelijk gesteld aan de optie gescheiden.

van het effect aangezien de groep winnaars (spelers) verschilt van de groep niet-winnaars (spelers en niet-spelers). Apouey & Clark (2009) tonen dit aan door middel van een regressievergelijking; het winnen van de loterij is significant gecorreleerd met het inkomen, etniciteit, opleiding, de status van arbeidsparticipatie, aantal kinderen en leeftijd.

4.1.2 Winnaars tegenover onsuccesvolle spelers

De methode die beter geschikt is voor gebruik van loterijprijzen, is het vergelijken van de winnaars met de onsuccesvolle spelers.

Echter heeft de loterijdata van het BHPS een nadeel; de loterijdata bevat geen directe informatie over het aantal keer dat een individu meespeelt. Dit heeft als gevolg dat er in de groep niet-winnaars geen onderscheid kan worden gemaakt tussen niet-spelers en onsuccesvolle spelers. Het onderscheid tussen winnaars en onsuccesvolle spelers kan daarmee evenzo niet worden gemaakt.

4.1.3 Kleine winnaars tegenover grote winnaars

Om het probleem van endogeniteit zo klein mogelijk te houden, is er zodoende voor gekozen de winnaars met elkaar te vergelijken. Winnaars van hoge prijzen worden vergeleken met winnaars van lage prijzen.

Apouey & Clark (2009) illustreren dat deze methode het meest exogene onderscheid maakt in de variabele door middel van een regressievergelijking; minder variabelen zijn gecorreleerd met de gewonnen bedragen, grote en kleine winnaars blijken gelijk te zijn wat betreft de status van arbeidsparticipatie en leeftijd.

4.2 Tijdsaspect

De regressie richt zich op het effect van loterijprijzen in jaar t op het gezondheidsgedrag in jaar t . Dit tijdsaspect verschilt met opzet van Apouey & Clark (2009). Apouey & Clark (2009) testen het effect van loterijprijzen op de gezondheidstatus. Een inkomensverandering zal geen direct gevolg hebben op de gezondheid; effecten van investeringen in gezondheid zullen pas na een aantal jaar zichtbaar zijn. Dit is de reden waarom zij het effect van loterijprijzen in jaar t op de gezondheidstatus in jaar $t + 2$ testen.

In dit onderzoek wordt gezondheidsgedrag los gezien van de gezondheidstatus. Dit onderzoek richt zich in zijn geheel op gezondheidsgedrag. Een inkomensverandering kan direct een effect hebben op het gedrag van personen, dus ook op gezondheidsgedrag. Vandaar dat de variabelen in de regressie zich richten op dezelfde tijdseenheid.

4.3 Regressie

Voor alle afhankelijke variabelen voor gezondheidsgedrag wordt het volgende regressiemodel gebruikt:

$$G_t = \alpha + \beta \text{Log}(\text{Lotprijs})_t + \gamma x_t + \varepsilon_t$$

G_t geeft de afhankelijke variabelen voor gezondheidsgedrag weer; sporten, bedrag dat aan vrije tijd wordt uitgegeven en roken. x_t geeft de vector van alle controlevariabelen weer; burgerlijke staat, etniciteit, geboortjaar, geslacht, golf, grootte huishouden, leeftijd, opleiding, regio en werkstatus. Er wordt gebruik gemaakt van een significantieniveau van 10%

In sommige controlevariabelen ontbreekt een groot aantal gegevens.¹⁹ De regressies worden daarom ook zonder controlevariabelen uitgevoerd. Er wordt echter verwacht dat er variabelen aanwezig zijn die beide gezondheidsgedrag en loterijprijzen beïnvloeden. De originele regressie, met toevoeging van controlevariabelen, heeft zodoende de voorkeur aangezien deze corrigeert voor deze variabelen en corrigeert voor omgekeerde causaliteit.

Naast bovenstaande regressiemodel zal een tweede regressiemodel worden getest:

$$G_t = \alpha + \beta \text{Log}(\text{Hhinkomen})_t + \gamma x_t + \varepsilon_t$$

In dit model is de variabele Lotprijs vervangen door de variabele Hhinkomen. De variabele Hhinkomen geeft het jaarlijkse inkomen in een huishouden aan en is, net als loterijprijzen, gecorrigeerd voor inflatie.²⁰ De vraag is of de resultaten van dit model verschillen tonen met betrekking tot de resultaten van bovenstaand model. Zo kan er worden vastgesteld of er werkelijk sprake is van gemeenschappelijke factoren en omgekeerde causaliteit in de regressie Hhinkomen op gezondheidsgedrag.

¹⁹ Het percentage wat niet beschikbaar is voor de variabelen: burgerlijke staat 45% niet beschikbaar, voor etniciteit geldt 10%, grootte huishouden 40%, leeftijd 40%, opleiding 40%, regio 45% en werkstatus 45%. De percentages zijn afgerond.

²⁰ De variabele Hhinkomen is gecorrigeerd voor inflatie via het consumenten prijs indexcijfer (CPI) met als basisjaar 2005.

5. Resultaten

De resultaten van de geteste invloed van een inkomensverandering op gezondheidsgedrag worden hieronder per afhankelijke variabele besproken. De waarde van de coëfficiënten en de R-squared worden toegelicht. De uitkomsten zijn weergegeven in de appendix. Tabel 3 geeft de resultaten voor het effect van loterijprijzen op gezondheidsgedrag zonder opname van controlevariabelen weer. Tabel 4 geeft de resultaten voor het effect van loterijprijzen op gezondheidsgedrag met opname van controlevariabelen weer. Tabel 5 geeft de resultaten voor het effect van inkomen van huishoudens op gezondheidsgedrag met opname van controlevariabelen weer.

5.1 Inkomenseffect op positief gezondheidsgedrag

Om de invloed van een positieve inkomensverandering op positief gezondheidsgedrag te testen, is er gebruik gemaakt van het positieve gezondheidsgedrag sporten en vrije tijd.

Voor sporten is de variabele Sport gebruikt.

Met opname van controlevariabelen is er een significante ($p=0,0115$) positieve invloed ($\beta=0,029803$) gevonden van loterijprijzen op Sport. Als een loterijprijs met 1% stijgt, betekent dit dat men met 0,00029803 stijgt naar een hogere schaal²¹. Het positieve effect is van kleine omvang. De controlevariabelen zijn niet allen significant. De R-squared laat een waarde zien van 0,288874. Dit betekent dat de onafhankelijke variabelen, logLotprijs plus de controlevariabele, 28,9% van de variatie in de afhankelijke variabele Sport verklaren. Een percentage van 100 zou betekenen dat het model perfect door de onafhankelijke variabelen kan worden verklaard. Hoe dichterbij de 100%, hoe groter de verklaringskracht. In dit geval zijn er waarschijnlijk nog andere variabelen die een grotere verklaringskracht tot gevolg hebben.

Het significante ($p=0,0000$) effect van Hhinkomen op Sport, met toevoeging van controlevariabelen, is positief ($\beta=0,076992$). Dit effect is van grotere omvang dan het bovenstaande effect, gevonden bij de exogene inkomensvariabele Loterijprijzen. Dit duidt erop dat er in dit regressiemodel verborgen gemeenschappelijke variabelen voorkomen die beide inkomen en sporten beïnvloeden en/of dat er sprake is van omgekeerde causaliteit, waardoor een hogere uitkomst van het effect van inkomen op sporten wordt veroorzaakt.

Voor ontspanning wordt de variabele Bedragvrijetijd gebruikt. Er is een significante ($p=0,0000$) positieve invloed ($\beta=4,506510$) gevonden van loterijprijzen op Bedragvrijetijd.

²¹ De schaalverdeling van Sport loopt van waarde 1 tot en met 5, zie data.

Als een loterijprijs met 1% stijgt, betekent dit dat men circa €0,05 meer aan vrije tijd besteed. Het positieve effect is van redelijke omvang. De controlevariabelen zijn niet allen significant. De R-squared geeft een waarde van 0,190265. De onafhankelijke variabelen, logLotprijs plus de controlevariabele, verklaren 19% van de variatie in de afhankelijke variabele Bedragvrijetijd.

Het effect van Hhinkomen op Bedragvrijetijd, met toevoeging van controlevariabelen, is positief ($\beta=8,994961$) significant ($p=0,0000$). Ook bij deze variabele is het effect van grotere omvang dan het bovenstaande effect, gevonden bij de exogene inkomensvariabele Loterijprijzen. Dit duidt erop dat er in dit regressiemodel verborgen gemeenschappelijke variabelen bestaan die beide inkomen en Bedragvrijetijd beïnvloeden en/of dat er sprake is van omgekeerde causaliteit, waardoor een hogere uitkomst van inkomen op Bedragvrijetijd wordt veroorzaakt.

5.2 Inkomenseffect op negatief gezondheidsgedrag

Om de invloed van een positieve inkomensverandering op negatief gezondheidsgedrag te testen, is er gebruik gemaakt van het negatieve gezondheidsgedrag roken. De analyse is gemaakt met behulp van de twee afhankelijke variabelen Roker en Sig.

Voor de variabele Roker wordt een significante ($p=0,0000$) positieve invloed ($\beta=0,010972$) gevonden van loterijprijzen op Roker. Als een loterijprijs met 1% stijgt, betekent dit dat de kans dat men rookt met 0,01% stijgt. Hier moet worden opgemerkt dat het positieve effect van een kleine omvang is. De controlevariabelen zijn niet allen significant. De R-squared is gelijk aan 0,081494. De onafhankelijke variabelen, logLotprijs plus de controlevariabele, verklaren 8,1% van de variatie in de afhankelijke variabele Roker.

Voor de variabele Sig wordt een significante ($p=0,0021$) positieve invloed ($\beta=0,265449$) gevonden van loterijprijzen op Sig. een loterijprijs met 1% stijgt, rookt men 0,00265449 sigaretten meer. De controlevariabelen zijn niet allen significant. De R-squared is gelijk aan 0,122030. De onafhankelijke variabelen, logLotprijs plus de controlevariabele, verklaren 12,2% van de variatie in de afhankelijke variabele Sig.

Het effect van Hhinkomen op Roker en van Hhinkomen op Sig geeft, ten opzichte van bovenstaand resultaat, een tegenovergestelde uitkomst. Het effect van Hhinkomen op Roker, met toevoeging van controlevariabelen, is negatief ($\beta=-0,037930$) significant ($p=0,0000$). Het effect van Hhinkomen op Sig is tevens negatief ($\beta=-0,198228$) significant ($p=0,0018$).

Dit beduidt dat er in dit regressiemodel verborgen gemeenschappelijke variabelen bestaan die beide inkomen en roken beïnvloeden en/of dat er sprake is van omgekeerde causaliteit, waardoor een andere uitkomst van inkomen op roken wordt veroorzaakt.

5.3 Regressiemodel exclusief controlevariabelen

Zonder opname van controlevariabelen zijn de volgende resultaten gevonden: een positieve invloed ($\beta=0,018699$) van loterijprijzen op Sport, deze relatie is niet significant ($p=0,1531$); een positieve significante ($p=0,0000$) invloed ($\beta=5,641744$) van loterijprijzen op BedragVrijetijd; een positieve significante ($p=0,0000$) invloed ($\beta=0,011383$) van loterijprijzen op Roker; en een positieve significante ($p=0,0005$) invloed ($\beta=5,641744$) gevonden van loterijprijzen op Sig. Deze effecten komen weliswaar in de richting van de resultaten met toevoeging van controlevariabelen, maar de effecten zijn niet identiek aan elkaar. Dit wil zeggen dat de verdeling van lage en hoge loterijprijzen niet geheel willekeurig is.

In het regressiemodel met opname van controlevariabelen is dit te zien aan het effect van de controlevariabelen op loterijprijzen. Controlevariabelen die een aanzienlijk effect hebben op loterijprijzen zijn Geslacht, Opleiding, Dumregio14, Dumwerkstat03 (werkloos) en Dumwerkstat08 (gedurende lange tijd ziek). Deze controlevariabelen tonen voor elke regressie een significante invloed op loterijprijzen.

6. Interpretatie

Dit onderzoek richt zich op het effect van een inkomensverandering op gezondheidsgedrag. De deelvragen waar in dit onderzoek een antwoord op wordt gezocht, zijn: Zal men ten gevolge van een positieve inkomensverandering positief gezondheidsgedrag vertonen of juist negatief gezondheidsgedrag? Of zal men beide, positief en negatief gezondheidsgedrag, vertonen?

Voor elke afhankelijke variabele voor gezondheidsgedrag is getoetst of een inkomensverandering hier effect op heeft. De toetsing vond plaats in twee regressiemodellen; een met de onafhankelijke variabele Hhinkomen en een met de onafhankelijke variabele Lotprijs. Na het vergelijken van de uitkomsten van de twee regressiemodellen kan worden geconcludeerd dat er in het regressiemodel met Hhinkomen verborgen gemeenschappelijke variabelen bestaan die beide inkomen en gezondheidsgedrag beïnvloeden en/of dat er sprake is van omgekeerde causaliteit, waardoor een andere uitkomst van inkomen op gezondheidsgedrag wordt veroorzaakt dan getest met loterijprijzen. Dit komt overeen met de verwachting benoemd in de inleiding. Om deze reden richtte dit onderzoek zich op loterijprijzen als exogene inkomensverandering.

Als positief gezondheidsgedrag zijn sporten en vrije tijd geanalyseerd. Er is een positieve invloed van het logaritme van loterijprijzen op de variabele Sport gevonden. De positieve invloed is van kleine omvang maar wel degelijk significant. Dit betekent dat personen met hoge loterijprijzen vaker aan sport doen dan personen met lage loterijprijzen. Een positieve exogene inkomensverandering leidt tot meer sporten, dus positief gezondheidsgedrag. Bovendien is er een significante positieve invloed geconstateerd van het logaritme van loterijprijzen op de variabele Bedragvrijetijd. Als een loterijprijs met 1% stijgt, betekent dit dat men circa £0,05 meer aan vrije tijd besteed; personen met hoge loterijprijzen spenderen meer aan vrije tijd dan personen met lage loterijprijzen. Een positieve exogene inkomensverandering zal leiden tot grotere uitgaven aan vrije tijd. Bij deze variabele kan dus ook worden geconstateerd dat een positieve inkomensverandering zal leiden tot positief gezondheidsgedrag. De verklaringskracht van beide regressies is redelijk sterk, 28,8% voor de regressie met Sport en 19% voor de regressie met Bedragvrijetijd. De regressies geven zodoende een adequate benadering van de werkelijkheid.

De bevinding dat een positieve inkomensverandering zal leiden tot positief gezondheidsgedrag komt overeen met de verwachting uitgesproken in het theoretisch kader: positief gezondheidsgedrag heeft een direct positieve en een indirect positieve invloed op het nut.

Roken is geanalyseerd als negatief gezondheidsgedrag. Er is een significante positieve invloed waargenomen van het logaritme van loterijprijzen op de variabele Roker. Personen met hoge loterijprijzen hebben een hogere kans om roker te zijn dan personen met lage loterijprijzen. Daarnaast is er een significante positieve invloed gesignaleerd van het logaritme van loterijprijzen op de variabele Sig. Als het logaritme van de loterijprijs met één pond stijgt, zal men 0,00265449 meer sigaretten gaan roken; binnen het aantal rokers, roken personen met hoge loterijprijzen meer sigaretten dan personen met lage loterijprijzen. Een positieve exogene inkomensverandering zal resulteren in negatief gezondheidsgedrag. De verklaringskracht van beide regressies is redelijk tot zwak, 8,1% voor de regressie met Roker en 12,2% voor de regressie met Sig, maar de verbanden die zijn aangetroffen zijn wel significant.

Het positieve verband dat is waargenomen voor inkomen en negatief gezondheidsgedrag komt overeen met Ettner (1996) en Apouey & Clark (2009). Ettner (1996) concludeert dat een positieve verandering van het inkomen de prevalentie van alcoholgebruik verhoogt. Apouey & Clark (2009) concluderen dat exogeen inkomen geassocieerd is met groter risicovol gezondheidsgedrag: loterijwinnaars roken meer.

Als van nutsmaximalisatie wordt uitgegaan betekent dit, zoals in het theoretisch kader naar voren komt, dat de directe positieve invloed van ongezonde consumptie op het nut groter is dan de indirecte negatieve invloed die ongezonde consumptie op gezondheid en dus op het nut heeft.

Kort samengevat kan er dus vanuit dit onderzoek geconcludeerd worden dat een positieve inkomensverandering enerzijds zal leiden tot positief gezondheidsgedrag, meer sporten en meer uitgaven aan vrije tijd. Anderzijds zal een positieve inkomensverandering leiden tot negatief gezondheidsgedrag anderzijds, meer roken.

7. Discussie

7.1 Toevoeging variabelen

Er moet worden opgemerkt dat een statistisch onderzoek een benadering van de werkelijkheid is. De conclusie uit dit onderzoek geldt alleen voor die variabelen die in dit onderzoek getest zijn. Voor een betere benadering van de werkelijkheid kunnen er meer variabelen worden toegevoegd.

Het gezondheidsgedrag wordt gemeten door middel van roken, sporten en vrijetijdsbesteding. Het onderzoek kan verbeterd worden door het opnemen van meer variabelen voor gezondheidsgedrag in de analyse. Op deze manier kan een beter beeld worden geschetst van het inkomenseffect op gezondheidsgedrag in bredere zin. Mogelijkheden voor negatief gezondheidsgedrag zijn alcoholgebruik en drugsgebruik. Mogelijkheden voor positief gezondheidsgedrag zijn slapen en de consumptie van groente en fruit.

In het hoofdstuk resultaten komt naar voren dat de verdeling van lage en hoge loterijprijzen niet geheel willekeurig is. Het gebruik van loterijprijzen is dus niet volledig exogeen. In het onderzoek zijn elf controlevariabelen toegevoegd om zo te corrigeren voor omgekeerde causaliteit en voor variabelen die beide loterijprijzen en gezondheidsgedrag beïnvloeden. Er zullen meer variabelen zijn die beide loterijprijzen en gezondheidsgedrag beïnvloeden. Door toevoeging van meer controlevariabelen kan het onderzoek verbeterd worden. Onmogelijk is het om voor al deze, bekende en onboekende, variabelen te corrigeren. Een methode die dit automatisch doet is fixed effects. In deze methode corrigeert het individuele specifieke effect voor alle effecten die tegelijkertijd loterijprijzen en gezondheidsgedrag beïnvloeden die constant zijn over de tijd. Een betere analyse zal kunnen worden gemaakt als er gebruik wordt gemaakt van het fixed effects model. Met gebruik hiervan kan er echter nog steeds sprake zijn van factoren die beide loterijprijzen en gezondheidsgedrag beïnvloeden die variëren over de tijd.

7.2 Paneldata

Het British Household Panel Survey (BHPS) is een uitgebreide dataset met paneldata. Helaas zitten er een aantal nadelen aan paneldata vast. In paneldata bestaat een grote kans van uitval van personen over de tijd. Bovendien is er sprake van toevoeging van personen in de BHPS. Toevoeging en uitval in het panel leiden tot onevenwichtige gegevens. Een ander nadeel is dat mensen zich mogelijk anders gaan gedragen, omdat zij in een panel zitten. Hierdoor zijn ze niet meer representatief. Ook kunnen persoonlijke veranderingen ervoor zorgen dat de

representativiteit van het panel vermindert. Bijvoorbeeld de verandering van burgerlijke staat of het krijgen van kinderen.

Een ander nadeel is dat er een groot aantal gegevens ontbreekt in de BHPS. De grote omvang van de dataset kan hier de oorzaak van zijn; niet elke vraag kan (uitgebreid) aan elke respondent gesteld worden plus correct beantwoord worden door de respondent..

7.3 Variabele Bedragvrijetijd gezondheidsbevorderend?

Om het inkomenseffect op positief gezondheidsgedrag te testen, is in dit onderzoek vrije tijd getest. De variabele die hiervoor is aangewend is Bedragvrijetijd, deze variabele geeft weer hoeveel geld men aan vrije tijd besteed. Voor normale goederen geldt: hoe meer uitgaven men besteedt aan een goed, hoe meer men van dit goed consumeert. Geldt dit echter ook voor vrijetijdsbesteding? Het besteden van vrije tijd kan geld kosten, maar dit hoeft niet. Men kan een dierentuin bezoeken maar een kosteloos alternatief kan het bezoeken van een kinderboerderij zijn. Beide vrijetijdsbesteding, maar niet voor hetzelfde bedrag. Het hoeft dus niet zo te zijn dat personen die veel uitgaven doen aan vrijetijdsbesteding per definitie meer vrije tijd besteden. Er is een significante positieve invloed geconstateerd van het logaritme van loterijprijzen op de variabele Bedragvrijetijd. Dit initieert dat hoe hoger de loterijprijs, hoe meer geld men besteedt aan vrije tijd. Dit positieve verband kan niet aangeven hoeveel tijd men daadwerkelijk aan vrije tijd besteedt. Is de variabele Bedragvrijetijd dan wel een betrouwbare variabele voor het testen van ontspanning; is deze variabele werkelijk gezondheidsbevorderend? Een betere variabele zou een variabele zijn die de hoeveelheid tijd die men aan vrijetijdsbesteding besteedt weergeeft.

Overigens moet worden opgemerkt dat de variabele Bedragvrijetijd niet accuraat het bedrag weergeeft wat men aan vrije tijd besteed. De antwoorden voor deze variabelen zijn in categorieën verdeeld. Wanneer men aangeeft tot de laatste categorie te behoren, geeft men aan £160,- of meer uit te geven. Hoeveel “meer” is, is niet bekend. Het effect van het logaritme van de loterijprijzen op Bedragvrijetijd zou dus hoger kunnen zijn dan hier getest is.

7.4 Loterijdata

Zoals in Data naar voren komt zijn er twee zwaktes van de loterijdata in de BHPS: het bevat geen directe informatie over het aantal keer dat een individu met een loterij meespeelt, daarnaast is niet bekend hoeveel geld er is vergokt door zowel de loterijwinnaars als de onsuccesvolle spelers.

In dit onderzoek is er wel sprake van kennis van de bruto-opbrengst maar geen sprake van kennis van de netto-opbrengst. Het is mogelijk dat personen meer geld hebben vergokt dan zij

bruto aan prijzen ontvangen hebben. Over het algemeen zijn de netto-opbrengsten kleiner dan de bruto-opbrengsten.²² De variabele Lotprijs, bruto-opbrengst, overschat de waarde van de werkelijke inkomensstijging, netto-opbrengst. De coëfficiënten van het logaritme van loterijprijzen zullen dus naar beneden moeten worden bijgesteld.²³

Bovendien is het zo dat personen die vaker meespelen of meer geld spenderen aan loterijen een grotere kans hebben om te winnen. Er kan dus worden tegengegaan dat het winnen van de loterij geheel exogeen verdeeld is.

Vanwege de ontbrekende informatie is ervoor gekozen om de winnaars van loterijen met elkaar te vergelijken; winnaars van lage prijzen met winnaars van hoge prijzen. In Methode wordt verklaard dat de groep winnaars verschilt van de groep niet winnaars, aangezien het meespelen aan de loterij endogeen afhankelijk is. De resultaten van dit onderzoek gelden voor loterijwinnaars. Maar gelden deze resultaten ook voor de groep niet winnaars? Is de groep winnaars representatief voor de gehele bevolking?

Wanneer het aantal keer dat een individu meespeelt en de hoeveelheid geld dat men vergoet wel bekend zou zijn, zou een meer betrouwbare analyse kunnen worden gemaakt.

²² Netto-opbrengsten zijn niet per definitie kleiner dan de bruto-opbrengsten. Het is mogelijk een bedrag te winnen zonder enkele inleg van geld. Bijvoorbeeld door het cadeau krijgen van een lot of door invulling van een enquête.

²³ Deze constatering maken Apouey en Clark (2009) ook; zij maken gebruik van dezelfde loterijdata.

8. Referenties

Apouey, B., Clark, A.E., 'Winning big but feeling no better? The effect of lottery prizes on physical and mental health', Working paper, (2009).

Arendt, J.N., 'Does education cause better health? A panel data analysis using school reforms for identification', *Economics of Education Review*, (2005), vol. 24 (2), 149-160

Card, D.E., 'The causal effect of education on earnings', In: O. Ashenfelter and D. E. Card (red) *Handbook of Labor Economics*, (1999), Volume 3A Amsterdam: Elsevier.

Ettner, S., 'New Evidence on the Relationship Between Income and Health', *Journal of Health Economics*, (1996), januari, 15, 67-86.

Galama, T., Van Kippersluis, H., 'A Theory of Socioeconomic Disparities in Health over the Life Cycle', RAND working paper, (2010), WR-773.

Grossman, M., 'The Demand for Health-A theoretical and Empirical Investigation', *New York: National Bureau of Economic Research*, (1972).

Lindahl, M., 'Estimating the Effect of Income on Health Using Lottery Prizes as Exogenous Source of Variation in Income', *Journal of Human Resources*, (2005), januari, 40, 144-168.

Meer, J., Miller, D.L., Rosen H.S., 'Exploring the Health-Wealth Nexus', *Journal of Health Economics*, (2003), mei, 22, 713-730.

Ruhm, C., 'Healthy Living in Hard Times', *Journal of Health Economics*, (2003), februari, 24, 341-363.

Wardle, H., Sproston, K., Orford, J., Erens, B., Griffiths, M. Constantine, R., Pigott, S., 'British Gambling Prevalence Survey 2007', *National Centre for Social Research*, (2007).

9. Appendix

Tabel 1: Definitie variabelen

Gezondheidsgedrag	
Sport	= 1 als het individu nooit/bijna nooit sport tot = 5 als het individu ten minste één keer per week
Bedragvrijetijd	bedrag wat men aan vrije tijd besteed per maand
Roker	= 1 als het individu rookt
Sig	aantal sigaretten wat het individu per dag rookt
Loterij	
Log(Lotprijs)	logaritme van loterijprijzen, gecorrigeerd voor inflatie.
Inkomen	
Log(Hhinkomen)	logaritme van het jaarlijkse inkomen van een huishouden, gecorrigeerd voor inflatie
Controlevariabelen	
Dumburgstat..	dummyvariabelen voor elke vorm van burgerlijkstaat: 01 getrouwd, 02 uit elkaar(separated), 03 gescheiden (divorced), 04, weduwe/weduwenaar, 05 nooit getrouwd
Wit	= 1 als het individu blank is
Geboortejr	geboortejaar van het individu
Geslacht	= 0 als het individu een vrouw is = 1 als het individu een man is
Dumgolf	dummyvariabelen voor elk jaar
Hhgrootte	aantal personen in het huishouden van het individu
Leeftijd	leeftijd van het individu
Opleiding	opleiding van het individu, = 1 laagste level tot = 13 hoogste level
Dumregio..	dummyvariabelen voor elke regio
Dumwerkstat..	dummyvariabelen voor elke vorm van werkstatus 01 zelfstandig, 02 in dienst 03 werkloos 04, gepensioneerd, 05 zwangerschapsverlof, 06 mantelzorg, 07 studie, 08 lange termijn ziek, 09 bijscholing (government training scheme), 10 anders

Tabel 2: Precieze vragen BHPS

Positief gezondheidsgedrag
<p>Hoe vaak: lopen/zwemmen/sporten: “We are interested in the things people do in their leisure time, I'm going to read out a list of some leisure activities. Please look at the card and tell me how frequently you do each one... Play sport or go walking or swimming.”</p> <ol style="list-style-type: none"> 1. Never/almost never 2. Once a year or less 3. Several times a year 4. At least once a month 5. At least once a weak
<p>Bedrag gespendeerd aan vrije tijd per maand: “Please look at this card and tell me about how much you personally spend in an average month on leisure activities, and entertainment and hobbies, other than eating out? (Excludes holidays. Includes all activities either at home or outside home)”</p>
Negatief gezondheidsgedrag
<p>Roker: “Do you smoke cigarettes?”</p>
<p>Aantal gerookte sigaretten: “Approximately how many cigarettes a day do you usually smoke, including those you roll yourself?”</p>
Loterijprijzen
<p>Overwinning op voetbalpools/ nationale loterij etc.: “Since September 1st (year before) have you received any payments, or payment in kind, from a win on the football pools, national lottery or other form of gambling?”</p>
<p>Bedrag van voetbalpools/loterijprijzen etc.: “About how much in total did you receive? (win on the football pools, national lottery or other form of gambling)”</p>

Tabel 3: Resultaten $G_t = \alpha + \beta \text{Log}(\text{Lotprijs})_t + \varepsilon_t$

	Afhankelijke variabelen			
	Sport	Bedragvrijetijd	Roker	Sig
α	3,498411 (0,0000)*	33,31606 (0,0000)*	0,234905 (0,0000)*	14,35564 (0,0000)*
Log(lotprijs)	0,018699 (0,1531)	5,641744 (0,0000)*	0,011383 (0,0000)*	0,295572 (0,0005)*
R-squared	0,000250	0,028754	0,001327	0,002641

Opmerking: voor elke uitkomst geeft het bovenste getal de coëfficiënt weer en het onderste getal tussen haakjes geeft de p-waarde, de overschrijdingskans, weer.

* Significant bij het significantieniveau van 10%.

Tabel 4: Resultaten $G_t = \alpha + \beta \text{Log}(\text{Lotprijs})_t + \gamma x_t + \varepsilon_t$

	Afhangelijke variabelen			
	Sport	Bedragvrijetijd	Roker	Sig
α	125,8708 (0,0919)*	-621,2559 (0,7144)	-38,35996 (0,0150)*	586,0478 (0,3225)
Log(lotprijs)	0.029803 (0.0115)*	4,506510 (0,0000)*	0,010972 (0,0000)*	0,265449 (0,0021)*
Geboortejr	-0,060716 (0,1044)	0,323527 (0,7036)	0,019577 (0,0132)*	-0,292820 (0,3237)
Geslacht	0,199056 (0,0000)*	18,59126 (0,0000)*	-0,023144 (0,0020)*	1,886326 (0,0000)*
Hhgrootte	-0,041160 (0,0114)*	-0,273213 (0,4482)	0,000628 (0,8533)	0,254331 (0,0270)*
Leeftijd	-0,075408 (0,0434)*	-0,212455 (0,8025)	0,014888 (0,0592)*	-0,161727 (0,5850)
Opleiding	0,045207 (0,0000)*	0,705606 (0,0000)*	-0,018777 (0,0000)*	-0,084575 (0,0642)*
Wit	-0,204287 (0,1846)	11,51029 (0,0006)*	-0,074049 (0,0176)*	5,528366 (0,0000)*
Dum1998		-2,466986 (0,1873)	-0,016065 (0,3535)	-0,828744 (0,1826)
Dum1999		-0,581695 (0,8055)	-0,024961 (0,2515)	-0,146147 (0,8540)
Dum2000	0,167753 (0,0566)*	0,229170 (0,9387)	-0,037334 (0,1759)	-0,744533 (0,4620)
Dum2001		3,457612 (0,3647)	-0,083715 (0,0181)*	-0,129500 (0,9216)
Dum2002	0,420925 (0,0082)*	4,700299 (0,3048)	-0,097587 (0,0220)*	-0,490718 (0,7567)
Dum2003		7,280333 (0,1794)	-0,123811 (0,0140)*	0,989092 (0,6019)
Dum2004	0,511467	6,828225	-0,139638	0,898281

	(0,0278)*	(0,2740)	(0,0159)*	(0,6784)
Dum2005		6,926927 (0,3267)	-0,177226 (0,0069)*	0,435973 (0,8594)
Dum2006	0,705473 (0,0209)*	5,783306 (0,4633)	-0,190069 (0,0094)*	0,931780 (0,7346)
Dum2007		8,949841 (0,3056)	-0,221860 (0,0062)*	0,489192 (0,8723)
Dum2008	-1,738151 (0,0000)*	4,060137 (0,6708)	-0,241755 (0,0065)*	0,769378 (0,8181)
Dumburgstat01	-0,103857 (0,4471)	0,414967 (0,8925)	-0,083836 (0,0037)*	-3,274934 (0,0004)*
Dumburgstat03	-0,180443 (0,2185)	11,11769 (0,0007)*	0,059470 (0,0541)*	-2,152681 (0,0274)*
Dumbyrgstat04	-0,331420 (0,0341)*	5,51978 (0,1530)	-0,044556 (0,1775)	-3,997388 (0,0005)*
Dumburgstat05	-0,171441 (0,2182)	15,39649 (0,0000)*	-0,025591 (0,3839)	-3,210598 (0,0007)*
Dumregio01	-0,156590 (0,3123)	11,48076 (0,0005)*	-0,003461 (0,9128)	-1,405271 (0,2144)
Dumregio02	-0,055751 (0,5526)	9,988297 (0,0000)*	-0,033198 (0,0824)*	1,255331 (0,0827)*
Dumregio04	0,159286 (0,0369)*	-4,958489 (0,0038)*	-0,071124 (0,0000)*	-0,833765 (0,1881)
Dumregio05	0,263856 (0,0075)*	-4,712706 (0,0386)*	-0,062453 (0,0027)*	0,054979 (0,9466)
Dumregio06	-0,140252 (0,0522)*	-2,233473 (0,1746)	-0,032279 (0,0357)*	1,914397 (0,0006)*
Dumregio07	0,068955 (0,5621)	-6,493250 (0,0141)*	-0,123674 (0,0000)*	1,351210 (0,2238)
Dumregio08	-0,063738 (0,5039)	-1,163041 (0,5851)	-0,047982 (0,0145)*	2,293077 (0,0018)*
Dumregio09	-0,341759 (0,0004)*	-1,997101 (0,3617)	-0,040334 (0,0460)*	0,812199 (0,2782)

Dumregio10	-0,014273 (0,9109)	6,267634 (0,0282)*	-0,076710 (0,0043)*	-0,776175 (0,4785)
Dumregio11	-0,046843 (0,6386)	-0,316946 (0,8888)	-0,060114 (0,0043)*	1,880342 (0,0232)*
Dumregio12	-0,291023 (0,0160)*	-1,203603 (0,6571)	-0,146032 (0,0000)*	0,282025 (0,8169)
Dumregio13	0,040908 (0,7529)	-2,948429 (0,2923)	-0,085940 (0,0012)*	-0,651736 (0,5340)
Dumregio14	0,211312 (0,0555)*	-6,687254 (0,0070)*	-0,054016 (0,0177)*	1,603419 (0,0656)*
Dumregio15	-0,161573 (0,1907)	7,109067 (0,0157)*	-0,024768 (0,3608)	3,343597 (0,0004)*
Dumregio16	0,013252 (0,8926)	-0,110829 (0,9601)	-0,045261 (0,0304)*	0,766437 (0,3315)
Dumregio17	-0,003661 (0,9550)	-6,247255 (0,0000)*	-0,045190 (0,0008)*	1,414276 (0,0055)*
Dumregio18	0,062715 (0,2743)	-3,581381 (0,0055)*	0,023657 (0,0485)*	2,800461 (0,0000)*
Dumregio19	0,022352 (0,8326)	-4,435934 (0,0472)*	0,006550 (0,7546)	1,342185 (0,0622)*
Dumwerkstat01	-0,309190 (0,0003)*	11,00475 (0,0000)*	0,014851 (0,4005)	6,318055 (0,0000)*
Dumwerkstat02	-0,264242 (0,0001)*	1,989169 (0,1967)	0,027225 (0,0507)*	4,447390 (0,0000)*
Dumwerkstat03	-0,305868 (0,0191)*	-20,39927 (0,0000)*	0,168997 (0,0000)*	5,588491 (0,0000)*
Dumwerkstat05	-0,459499 (0,1028)	-8,316008 (0,2389)	-0,116537 (0,0628)*	0,031988 (0,9914)
Dumwerkstat06	-0,104085 (0,2850)	-7,844433 (0,0009)*	0,072475 (0,0003)*	5,155894 (0,0000)*
Dumwerkstat07	0,123954 (0,3002)	-16,02012 (0,0000)*	-0,147288 (0,0000)*	2,709942 (0,0062)*
Dumwerkstat08	-0,845068	-11,08326	0,209589	6,855228

	(0,0000)*	(0,0000)*	(0,0000)*	(0,0000)*
Dumwerkstat09	-0,436515 (0,4261)	-13,24115 (0,2581)	0,243357 (0,0296)*	0,906306 (0,7615)
Dumwerkstat10	0,121835 (0,7390)	-14,42758 (0,0441)*	-0,009090 (0,8881)	1,175629 (0,6438)
R-squared	0,288874	0,190265	0,081494	0,122030

Opmerking: voor elke uitkomst geeft het bovenste getal de coëfficiënt weer en het onderste getal tussen haakjes geeft de p-waarde, de overschrijdingskans, weer.

* Significant bij het significantieniveau van 10%.

Tabel 5: Resultaten $G_t = \alpha + \beta \text{Log}(\text{Hhinkomen})_t + \gamma x_t + \varepsilon_t$

	Afhankelijke variabelen			
	Sport	Bedragvrijetijd	Roker	Sig
α	7.721683 (0,7386)	-1083.738 (0,0269)	-10,71384 (0,0146)*	102,7976 (0,5592)
Log(Hhinkomen)	0.076992 (0,0000)*	8.994961 (0,0000)*	-0,037930 (0,0000)*	-0,198228 (0,0018)*
Geboortejr	-0.002111 (0,8554)	0.524845 (0,0324)	0,005882 (0,0074)*	-0,047756 (0,5580)
Geslacht	0.109979 (0,0000)*	16.55472 (0,0000)*	0,021204 (0,0000)*	1,885854 (0,0000)*
Hhgrootte	-0.063225 (0,0000)*	-2.193997 (0,0000)*	0,005929 (0,0000)*	0,368949 (0,0000)*
Leeftijd	-0.018792 (0,1046)	0.064555 (0,7923)	0,001421 (0,5177)	0,032967 (0,7085)
Opleiding	0.064449 (0,0000)*	0.539848 (0,0000)*	-0,019533 (0,0000)*	-0,232018 (0,0000)*
Wit	0.081456 (0,0202)*	5.922103 (0,0000)*	0,047167 (0,0000)*	4,027093 (0,0000)*
Dum1998		-1.600425 (0,0254)*	-0,003703 (0,5815)	0,097593 (0,6931)
Dum1999		0.087252 (0,9103)	-0,007038 (0,3239)	-0,022548 (0,9331)
Dum2000	0.001378 (0,9631)	-0.230569 (0,8042)	-0,019143 (0,0246)*	0,032295 (0,9216)
Dum2001		2.621800 (0,0213)*	-0,023972 (0,0203)*	0,070132 (0,8621)
Dum2002	0.104286 (0,0384)*	1.135883 (0,4032)	-0,031008 (0,0119)*	-0,255848 (0,5975)
Dum2003		3.870346 (0,0153)*	-0,039369 (0,0063)*	-0,083854 (0,8831)
Dum2004	0.080110	1.383726	-0,047289	-0,251900

	(0,2719)	(0,4491)	(0,0041)	(0,7006)
Dum2005		4.688702 (0,0229)*	-0,062604 (0,0007)*	0,161885 (0,8269)
Dum2006	0.172171 (0,0706)*	1.877966 (0,4130)	-0,065761 (0,0014)*	0,022446 (0,9783)
Dum2007		4.731546 (0,0620)*	-0,081591 (0,0003)*	-0,358318 (0,6943)
Dum2008	-2.231707 (0,0000)*	-0.157771 (0,9546)	-0,088684 (0,0004)*	-0,335630 (0,7363)
Dumburgstat01	0.101942 (0,0078)*	-6.361992 (0,0000)*	-0,170606 (0,0000)*	-2,204767 (0,0000)*
Dumburgstat03	0.002938 (0,9435)	-0.482230 (0,5789)	-0,008794 (0,20802)	-0,531450 (0,0289)*
Dumbyrgstat04	-0.164961 (0,0001)*	-1.236260 (0,1862)	-0,138881 (0,0000)*	-3,320277 (0,0000)*
Dumburgstat05	-0.030509 (0,4369)	5.598196 (0,0000)*	-0,102252 (0,0000)*	-2,876869 (0,0000)*
Dumregio01	-0.063302 (0,1581)	7.939641 (0,0000)*	0,031128 (0,0005)*	0,848340 (0,0112)*
Dumregio02	-0.216059 (0,0000)*	5.457039 (0,0000)*	-0,007230 (0,2617)	0,816688 (0,0018)*
Dumregio04	0.045901 (0,0850)*	-4.569293 (0,0000)*	-0,051498 (0,0000)*	-0,653241 (0,0040)*
Dumregio05	0.047714 (0,1691)	-4.911179 (0,0000)*	-0,026917 (0,0001)*	-1,006610 (0,0003)*
Dumregio06	-0.159547 (0,0000)*	-0.992581 (0,0817)*	0,004435 (0,4187)	1,692489 (0,0000)*
Dumregio07	-0.172741 (0,0000)*	-3.868552 (0,0000)*	-0,017110 (0,0284)*	0,366203 (0,2298)
Dumregio08	-0.123992 (0,0002)*	-3.520525 (0,0000)*	-0,018183 (0,0057)*	0,584998 (0,0241)*
Dumregio09	-0.333827 (0,0000)*	3.748295 (0,0000)*	0,002012 (0,7848)	0,100329 (0,7260)

Dumregio10	0.016895 (0,7196)	5.982268 (0,0000)*	-0,020888 (0,0263)*	-0,536836 (0,1378)
Dumregio11	-0.120312 (0,0005)*	-0.638112 (0,3780)	0,029240 (0,0000)*	1,244344 (0,0000)*
Dumregio12	-0.312620 (0,0000)*	1.781624 (0,0450)*	-0,050131 (0,0000)*	0,713584 (0,0470)
Dumregio13	-0.120372 (0,0024)*	5.427845 (0,0000)*	0,017691 (0,0253)*	0,271148 (0,3504)
Dumregio14	-0.070509 (0,0725)*	-1.611120 (0,0474)*	-0,004352 (0,5771)	1,550382 (0,0000)*
Dumregio15	-0.276475 (0,0000)*	4.040983 (0,0000)*	0,024537 (0,0099)*	2,456192 (0,0000)*
Dumregio16	-0.097951 (0,0081)*	-0.425000 (0,5836)	-0,042316 (0,0000)*	0,419407 (0,1597)
Dumregio17	-0.080079 (0,0001)*	-3.857293 (0,0000)*	-0,009076 (0,0256)*	0,693654 (0,0000)*
Dumregio18	0.056769 (0,0046)*	-1.959818 (0,0000)*	0,017393 (0,0000)*	1,909820 (0,0000)*
Dumregio19	-0.021433 (0,3341)	-1.755272 (0,0001)*	-0,017770 (0,0000)*	1,144806 (0,0000)*
Dumwerkstat01	-0.215935 (0,0000)*	7.243494 (0,0000)*	0,021923 (0,0001)*	4,578622 (0,0000)*
Dumwerkstat02	-0.232313 (0,0000)*	-0.081037 (0,8611)	0,044389 (0,0000)*	3,583313 (0,0000)*
Dumwerkstat03	-0.209170 (0,0000)*	-18.58718 (0,0000)*	0,186104 (0,0000)*	4,582110 (0,0000)*
Dumwerkstat05	-0.375357 (0,0000)*	-16.50899 (0,0000)*	-0,084313 (0,0000)*	0,584526 (0,4985)
Dumwerkstat06	-0.109595 (0,0001)*	-7.656101 (0,0000)*	0,066230 (0,0000)*	4,894289 (0,0000)*
Dumwerkstat07	0.109003 (0,0018)*	-17.49780 (0,0000)*	-0,206566 (0,0000)*	0,789175 (0,0073)
Dumwerkstat08	-0.881754	-9.205555	0,189365	5,793508

	(0,0000)*	(0,0000)*	(0,0000)*	(0,0000)*
Dumwerkstat09	-0.109783 (0,3979)	-18.68070 (0,0000)*	0,043879 (0,0786)*	1,545173 (0,0433)*
Dumwerkstat10	0.017851 (0,8218	-12.49702 (0,0000)*	0,029004 (0,0609)*	3,641609 (0,0000)*
R-squared	0.304418	0.163956	0,101235	0,101222

Opmerking: voor elke uitkomst geeft het bovenste getal de coëfficiënt weer en het onderste getal tussen haakjes geeft de p-waarde, de overschrijdingskans, weer.

* Significant bij het significantieniveau van 10%.