

# Ekonomiska indikatorer samt systemfaktorer och hur de förutsäger och påverkar dödligheten i prostatacancer

Författare: Oscar Brink Bolin, 19940724-2891

Institution: Nationalekonomiska institutionen

Handledare: Ulf Gerdtham

Uppsatsnivå: C-uppsats

Datum för uppsattsseminarium: oktober 2020

Nyckelord: Prostate cancer death rate; CDR; PSA-test; variab\*; system; faktor; OECD

## **Sammanfattning**

Hur olika ekonomiska indikatorer och faktorer förutsäger och påverkar dödligheten i prostatacancer har i detta arbete undersökts. Tidigare forskning tyder på att där är fler faktorer som påverkar dödligheten i prostatacancer utöver sjukvårdsfinansiering och produktionsfaktorer inom sjukvården. En linjär regressionsmodellen har satts upp för sambandet mellan dödligheten i prostatacancer och total sjukvårdsfinansiering, ginikoefficienten, antal läkare per 1000 personer, åldersfördelningen bland män samt bruttonationalprodukten per capita. För de variabler som sedan inte har ingått i denna modell har det estimerats hur de påverkar den beroende variabeln när modellens oberoende variabler ändras. Resultaten visar att dödligheten i prostatacancer sjunker då bruttonationalprodukten per capita ökar, och att där finns systemskillnader mellan länder. Vidare visar resultaten att en del av denna förändring kan förklaras med att övriga variabler påverkas av ändringar i bruttonationalprodukten per capita och i sin tur påverkar dödligheten i prostatacancer. En del oförklarad effekt kvarstår däremot och har en samlad ökande effekt på dödligheten i prostatacancer då bruttonationalprodukten per capita ökar. Det finns en del forskning på området som förklarar dessa skeenden men det är uppenbart att mer forskning är nödvändig. Både för att kartlägga vilka faktorer som, samt på vilket sätt och varför dessa faktorer, påverkar dödligheten i prostatacancer.

## Innehållsförteckning

1.	Inledning	3
2.	Tidigare forskning	5
3.	Metod	6
3.1.	Val av beroende och oberoende variabler	7
3.2.	Linjär regressionsmodell, arbetsgång	8
4.	Bygge av den linjära regressionsmodellen	9
5.	Ekvationssystemet	28
6.	Numeriska resultat	30
7.	Slutsats och diskussion	32
8.	Reservationer och problem	34
9.	Källhänvisningar	36
10.	Bilaga A - Tabeller	38
11.	Bilaga B - Teori och definitioner	64

## 1. Inledning

“Dödligheten i prostatacancer varierar stort över landet” (Nyhetsbyrån Siren, 2019-11-04) är en rubrik i en artikel publicerad 4:e november 2019 i tidningen Vårdfokus. Man konstaterar i artikeln att prostatacancer är den cancerform som flest män dör av, och att även om dödsfallen sjunker från år till år så skiljer sig dödligheten mellan regionerna i Sverige. Vidare konstaterar man att ålder inte förklarar hela skillnaden mellan regionerna. Man framhåller också bland annat regionernas arbete med PSA-testning (Bratt Ola, 2020-01-20) och befolkningens kunskaper om hur viktigt det är att gå att testa sig. Vidare så tyder forskning på att socioekonomiska faktorer spelar roll för dödligheten i prostatacancer. Jens Klein och Olaf von dem Knesebeck skriver i en artikelstudie *Socioeconomic inequalities in prostate cancer survival: A review of the evidence and explanatory factors* följande: “33 studies indicate a significant association between SES and survival among prostate cancer patients...” (Klein Jens, von dem Knesebeck Olaf, oktober 2015). Subas Neupane, Freddie Bray och Anssi Auvinen kommer i sin studie, “*National economic and development indicators and international variation in prostate cancer incidence and mortality: an ecological analysis*”, fram till att sjukvårdsutgifter har en positiv effekt både på frekvensen av fall i prostatacancer och dödligheten i densamma. De framhåller också att andra makroekonomiska effekter reflekterar ett allt mer västerländskt levnadssätt kopplat till till exempel diet, en stillasittande livsstil och ökande fetma, och att en av indikatorerna för detta är bruttonationalprodukten. Vidare framhåller de också att frekvensen av fall i prostata cancer varierar mycket mellan länder, även för de med liknande nivåer av sjukvårdsutgifter som andel av bruttonationalprodukten, vilket påvisar att frekvensen av fall i prostatacancer inte bara kan förklaras av dessa indikatorer. (Neupane Subas, Bray Freddie, Auvinen Anssi, juni 2017)

I kontrast till detta så landar man ofta i den politiska debatten att lösningen på de problem i sjukvården man adresserar är att tillföra mer pengar till sjukvården. I december 2019 släpper regeringen ett pressmeddelande att man tänker satsa 4,7 miljarder kronor för att stärka svensk hälso- och sjukvård (4,7 miljarder till hälso- och sjukvården, 2019-12-18). Även då det är svårt att tänka sig att mer pengar till sjukvården inte skulle ha någon effekt, eller ha en negativ effekt, så kan man ställa sig frågan om det finns andra faktorer som spelar roll än finansieringen av sjukvården och produktionsfaktorer inom sjukvården. Man skulle kunna tänka sig ett bredare scenario där man inkluderar andra faktorer som möjligtvis bör adresseras utöver finansieringen av

sjukvården och produktionsfaktorer inom sjukvården. Kollar man på dödlighet i prostatacancer så finns det som tidigare diskuterat indikatorer på att andra faktorer utöver dessa som också påverkar dödligheten i prostatacancer.

Jens Klein och Olaf Knesebeck påpekar att även om man kan förklara en del av differenser av dödlighet i prostatacancer med faktorer som livsstil, medvetenhet, tillgång till tidig upptäckt, med flera, så kvarstår en oförklarad variation i dödlighet i prostatacancer. De framhåller vidare att ett viktigt framtida forskningsämne är kring förklaringsfaktorerna av dödligheten i prostatacancer och olika faktors bidrag till dödligheten (Klein Jens, von dem Knesebeck Olaf, oktober 2015). Det skall också framhållas att prostatacancer är den näst mest diagnostiserade cancerformen bland män efter lungcancer, och är den femte vanligaste dödsorsaken världen över (Prashanth Rawla, april 2019).

Syftet med denna studie är att undersöka hur olika faktorer och interaktion mellan dessa påverkar dödligheten i prostatacancer; undersöka huruvida det finns några systemeffekter, eller icke observerbara effektskillnader, som påverkar dödligheten i prostatacancer och som visar sig genom skillnad i dödlighet mellan likvärdiga länder och/eller år; samt resonera kring vad som orsakar de resultat man observerar och vilka implikationer detta i sådana fall leder till.

Detta kommer att genomföras genom en registerstudie där insamlad data används för att bygga en linjär regressionsmodell av sambandet mellan Prostate cancer death rate och Total health spendings per person, GDP/capita, Gini coefficient, Percentage of male population aged 40 +, samt Medical doctors per 1000 people. Därefter sätts ett ekvationssystem upp, innehållande de variabler som ej ingår i regressionsmodellen. Detta ekvationssystem ämnar att beskriva hur förändringar i de oberoende variablerna påverkar de ej i regressionsmodellen ingående variablerna och vidare den beroende variabeln. Utifrån detta förs slutligen en diskussion och slutsatser dras om faktorer och deras påverkan på dödligheten i prostatacancer.

## 2. Tidigare forskning

Martin C.S. Wong , William B. Goggings, Harry H.X. Wang, Franklin D.H. Fung, Colette Leung, Samuel Y.S. Wong, Chi Fai Ng och Joseph J.Y. Sung gjorde 2016 en studie där korrelationen mellan frekvensen och dödligheten i prostatacancer samt de socioekonomiska faktorerna Gross domestic product (GDP) och human development index (HDI). Estimat av frekvensen och dödligheten i prostatacancer hämtade de från GLOBOCAN databas för 184 länder under 2012. Data för GDP och HDI för varje land under 2012 hämtade de från United Nations Human Development Report. HDI är här ett sammansatt index som innefattar förväntad livslängd, utbildningsperiod och olika indikatorer för inkomst per capita. De använde sig i studien av joinpoint regressionsanalys för att analysera trender i frekvensen och dödligheten i prostatacancer. Studien visade att länder med högre DHI och GDP hade högre frekvens av prostatacancer men inte högre dödlighet. De observerade också stor geografisk skillnad i frekvens och dödlighet i prostatacancer. Olika möjliga förklaringar till dessa skillnader som identifierades var bland annat skillnader i prostatacancer-screening, diagnostiskt konstaterande och populationers riskfaktorer (Wong Martin C.S., Goggings William B., Wang Harry H.X., Fung Franklin D.H., Leung Colette, Wong Samuel Y.S., Ng Chi Fai, Sung Joseph J.Y., november 2016).

Czorny Nunes, César Rildo, Maria Helena Pinto, Daniele Alcalá Pompeo, Denise Bereta, Luciana Ventura Cardoso och Dorival Marques da Silva publicerade 2017 en studie där de har undersökt och försökt identifiera riskfaktorer för prostatacancer. Studien de genomförde var en beskrivande, tvärsnitts- och kvantitativ studie, med ett urval av 150 män. De riskfaktorer som visade sig vara mest relevanta var ålder, utbildningsnivå, stillasittande livsstil, alkoholkonsumtion, konsumtion av rött kött, mjölk och fett samt högt BMI. Man kunde också se att andelen män som visade åtminstone en av riskfaktorerna var hög (Nunes Czorny, Rildo César, Pinto Maria Helena, Alcalá Pompeo Daniele, Bereta Denise, Ventura Cardoso Luciana, Marques da Silva Dorival, 2017).

Jens Klein och Olaf von dem Knesebeck publicerade 2015 en studie där de undersökt samband mellan överlevnadsgraden i prostatacancer och socioekonomisk status. Studien var en litteraturstudie där systematisk genomgång av PubMed-databasen där sökningen baserades på en kombination av ett antal sökord (socio\*, inequalit\*, income, education\*, occupation\*, survival, excess mortality, case fatality, prognosis, prostate, cancer, common cancer\*, major cancer\*, cancer

site\*). Studier som ansågs lämpliga kontrollerades för potentiella förklarande faktorer som bidrar till skillnad i överlevnad mellan socioekonomiska grupper. En stor grupp av studierna visade en signifikant skillnad i överlevnad mellan olika socioekonomiska grupper (Klein Jens, von dem Knesebeck Olaf, oktober 2015).

Nancy Fullman, et. al, publicerade 2016 en studie där personlig sjukvårdstillgång och kvalitet utvärderades. De använde sig i denna studie av studien Global Burden of Disease, Injuries, and Risk Factors Study 2016 (GBD 2016) för att studera ovanstående med Healthcare Access and Quality (HAQ) Index för 195 länder och områden. I studien använde de sig av principalkomponentanalys för att skapa sitt HAQ index. Relationer mellan HAQ index-värden och potentiella faktorer för prestation, såsom total health spendings per capita, undersöktes. Studien visade bland annat att nationell prestation på HAQ indexet var positivt associerat med högre nivå av total health spendings per capita och produktionsfaktorer inom hälso- och sjukvårdssystemet (Fullman, Nancy, et.al, juni 2018).

### **3. Metod**

Följande arbete kommer baseras på en registerstudie, där nödvändig data hämtas från Gapminder. All data är aggregerad internationell paneldata på OECD-nivå (Organisation for Economic Cooperation and Development - nivå) . Datasetet innehåller data från OECD-länderna för att erhålla ett dataset som innehåller data från länder med liknande ekonomisk situation. Inledningsvis sorteras datan i excel så att varje datarad representerar en nation vid ett visst år, vilket gör att varje datapunkt är en viss variabel för en viss nation vid en specifik tidpunkt. Vidare så laddas den sorterade datan upp i RStudio. Med programspråket R skrivs i RStudio programkod så att en linjär regressionsmodell för sambandet mellan den beroende och de oberoende variablerna byggs. Därefter sätts ett ekvationssystem upp som beskriver hur förändringar i de oberoende variablerna påverkar de i regressionsmodellen ej ingående variablerna och vidare den beroende variabeln. Detta används sedan för att kontrollera hur stor del av effekten i en viss förändring enligt regressionsmodellen som beror på direkt effekt på den beroende variabeln jämte indirekt effekt på grund av att de ej ingående variablerna påverkas vid en förändring och de i sin tur påverkar den beroende variabeln. Slutligen analyseras och diskuteras de resultat som erhållits.

### 3.1. Val av beroende och oberoende variabler

I datan som samlas in så är dataraderna länder och kolumnerna är: Prostate cancer death rate [st] (Gapminder, Prostate cancer, deaths per 100 000 men, n.d.); Total health spendings per person (US\$) [US\$] (Gapminder, Total health spending per person (US\$)), n.d.); GDP/capita (US\$, inflation-adjusted) [US\$] (Gapminder, GDP/Capita (US\$, inflation-adjusted), n.d.); Gini coefficient [%] (Gapminder, Gini coefficient, n.d.), Percentage of male population aged 40 + [%] (Gapminder, Population aged 40-59, M (%), n.d.), (Gapminder, Population aged 60+, M (%), n.d.) och Medical doctors per 1000 people [st] (Gapminder, Medical Doctors (per 1000 people), n.d.). Samtliga variabler är kontinuerliga och datan är aggregerad data på OECD-nivå.

Den beroende variabeln väljs till Prostate cancer death rate [st] och de oberoende variablerna väljs till: Total health spendings per person [US\$]; GDP/capita (US\$, inflation-adjusted) [US\$], Gini coefficient [%], Percentage of male population aged 40 + [%] och Medical doctors per 1000 people [st].

Total health spendings per person väljs som en av de beroende variablerna då det anses trivialt att finansiering påverkar utkomsten inom hälso- och sjukvård och således påverkar den beroende variabeln. Frågan är snarare i vilken utsträckning finansiering påverkar den beroende variabeln. GDP/capita väljs som en oberoende variabel då bruttonationalprodukten förväntas kunna vara en indikator på externaliteter från landets produktion och anses kunna vara en indikator på andra makroekonomiska effekter som påverkar den beroende variabeln. Gini coefficient väljs som en oberoende variabeln då denne anses kunna påvisa ekonomisk ojämlikhet som påverkar den beroende variabeln. Antal läkare väljs som en oberoende variabel då den anses indikera tillgängligheten på läkare som skulle kunna påverka den beroende variabeln. Slutligen väljs Percentage of male population aged 40 + som en oberoende variabel då en högre andel av befolkningen tillhörande en riskgrupp på grund av ålder anses påverka den beroende variabeln.

Vi har sex variabler med 31 olika kombinationer när man räknar utan samspel mellan variablerna. Detta gör att vi minimalt behöver 217 observationer. Tar vi minimala antalet observationer och ungefär tredubblar det resulterar det i att vi vill ha 651 observationer. Då datan innehåller 576 observationer så är det mindre än önskat men det är mer än det minimala antalet observationer vi

behöver. Det skall påpekas att South Korea, som ingår i OECD-länderna, inte ingår i vår data då denna data var så pass inkomplett att den valdes att inte användas.

### 3.2. Linjär regressionsmodell, arbetsgång

Vid framtagandet av den linjära regressionsmodellen används R som programspråk i RStudio. Under arbetets gång när en regressionsmodell passas in så används least-square-metoden för att beräkna modellens koefficienter. Den slutgiltiga modellen tas fram genom att följande steg genomförs:

- A. En första kontroll av datapunkter genomförs och problematiska datapunkter plockas bort. Detta kan vara punkter som är av fel datatyp eller anses vara orimliga.
- B. En backwards elimination med Bayes Information Criterion (BIC) som kriterium genomförs för att skapa oss en uppfattning om hur modellen skulle kunna se ut och för att få en modell att utgå från i det fortsatta arbetet.
- C. En residualanalys genomförs för den initiala modell erhållen i steg B. Här plottar vi: ett histogram innehållandes modellens residualer; en kvantil-kvantil-plot (Q-Q-plot) med modellens residualer mot de teoretiska residualerna för en normalfördelning; modellens residualer mot modellens predikterade värden; modellens studentiserade residualer mot modellens predikterade värden samt modellens residualer mot modellens ingående oberoende variabler. Detta genomförs också för olika transformationer av modellen för att kontrollera om någon transformation av modellen förbättrar den samma. Transformationerna om kontrolleras är: Log-Log- ; Log-Lin-; Lin-Log-; samt box-cox-transformation. Efter genomförd residualanalys väljs en modell att gå vidare med i det fortsatta arbetet.
- D. Modeller med ytterligare och/eller andra variabler skapas och det genomförs två-vägs ANOVA tester på dessa mot föregående modell. F- och P-värdena kontrolleras och avgör huruvida det har blivit en signifikant förbättring med fler oberoende variabler i modellen. En modell väljs för att gå vidare med i det fortsatta arbetet.
- E. Leverage, studentiserade residualer samt Cook's D beräknas och plottas mot samtliga oberoende variabler ingående i modellen. Datapunkter med leverage  $> 0.026$  samt datapunkter med  $|\text{studentiserade residualer}| > 1.96$  markeras. Problematiska datapunkter är de datapunkter som uppfyller:  $\text{Cook's D} > 4/\text{antal datarader} \ \& \ |\text{studentiserad residual}| > 1.96$



eller Cook's  $D > 4/\text{antal datarader}$  & leverage  $> 0.026$ . Problematiska datapunkter tas bort och modellen omberäknas med det nya datasetet.

- F. Steg C. Genomförs igen med den senaste modellen.
- G.  $R^2$ ,  $R^2$ -adjusted; Akaike Information Criterion (AIC) samt Bayes Information Criterion (BIC) beräknas för den nuvarande modellen samt ytterligare modeller med fler och/eller färre oberoende variabler.  $R^2$ ,  $R^2$ -adjusted, AIC samt BIC plottas därefter och jämförs för samtliga modeller och den mest passande modellen väljs för att gå vidare med i det fortsatta arbetet.
- H. En forward selection med BIC som kriterium genomförs för att kontrollera om där är några variabler som har en signifikant betydelse för modellen som bör tas med.
- I. Steg C. Samt steg D. genomförs återigen, med tillägget att de båda stegen nu ligger till grund för valet av den slutgiltiga modellen.

Vid varje givet steg hanteras en eller flera modeller för vilka man genomför varje steg för. Mellan varje steg väljs vilken eller vilka modeller som man fortsatt arbetar med i kommande steg. Slutligen väljs den mest passande modell till slutgiltig linjär modell. Över arbetets gång följs principen om Ockhams rakkniv, vilken säger att man inte skall anta fler företeelser eller ting än som behövs för att förklara de observationer man gör. Praktiskt tillämpas det under arbetets gång på sådant sätt att den enklaste modellen väljs vid val mellan likvärdiga modeller. Vidare så framhålls det att avsteg från metoden kan uppstå vid behov.

#### **4. Bygge av den linjära regressionsmodellen**

- A. Vid kontroll av datapunkterna upptäcktes en problematisk datapunkt. Denna punkt var en datapunkt för vilken Total health spendings per person (US\$) var 0 vilket är högst otroligt att det stämmer för OECD-länderna och är mest troligt ett fel i källan datan kommer ifrån eller att datan inte var tillgänglig för denna tidpunkt. Denna datapunkt plockades således bort och vi har då 575 kvarvarande datapunkter
- B. Backward elimination med BIC som kriterium genomfördes med

*Prostate cancer death rate ~ Total health spendings per person + GDP/capita + Gini coefficient  
+ Percentage of male population aged 40 plus + Medical doctors per 1000 people*

som övre tillåtna modell vilket gav en modell enligt:

$$\begin{aligned} \text{Prostate cancer death rate} &\sim 30.16 - 0.1523 \cdot \text{Gini coefficient} + 0.00014 \cdot \text{GDP/capita} \\ &- 0.00133 \cdot \text{Total health spendings per person}, \end{aligned}$$

och benämns modell 1

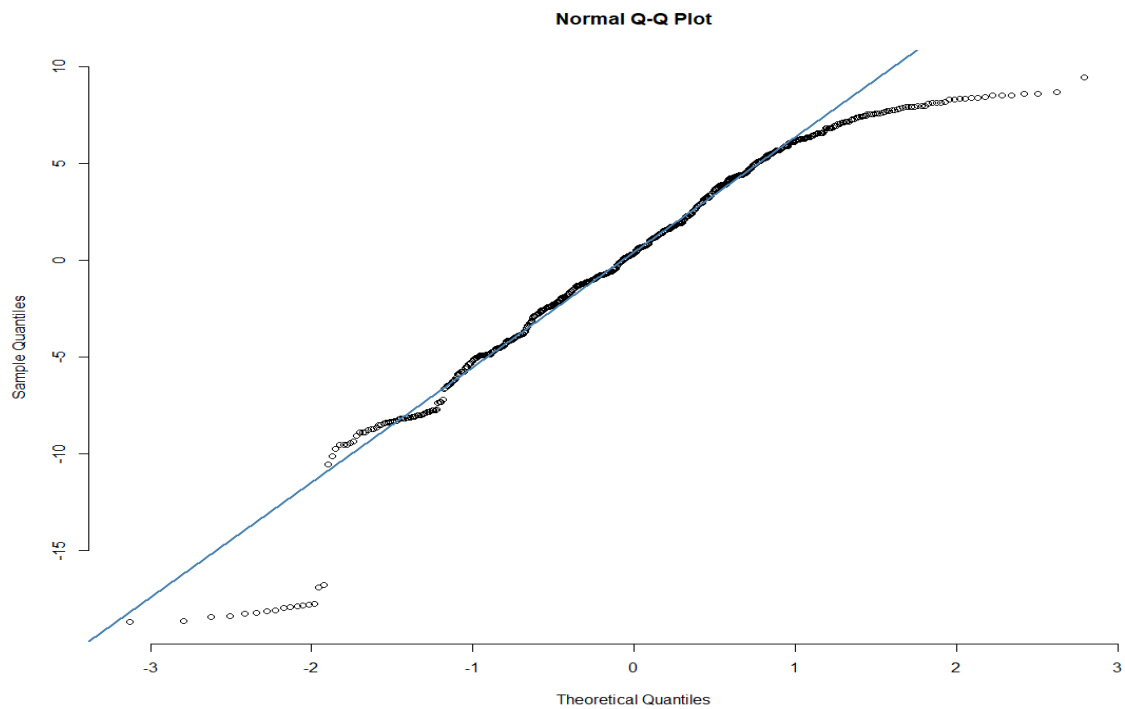
Då de oberoende variabler som för närvarande inte är med i modellen läggs till modellen är dessa inte signifikanta och lämnas för närvarande utanför modellen.

- C. När vi plottade ett histogram innehållandes modellens residualer, en Q-Q-plot, modellens residualer mot modellens predikterade värden, modellens studentiserade residualer mot modellens predikterade värden samt modellens residualer mot modellens ingående oberoende variabler ansågs en transformation kunna förbättra modellen (se figur 1-7). När ovanstående återigen plottades för modell 1 transformerad enligt Log-Log- ; Log-Lin-; Lin-Log-; samt box-cox-transformation ansågs modell 1 box-cox-transformerad med  $\lambda = 1.84$  vara den bästa transformationen samt bättre än modell 1 otransformerad (se figur 8-15).

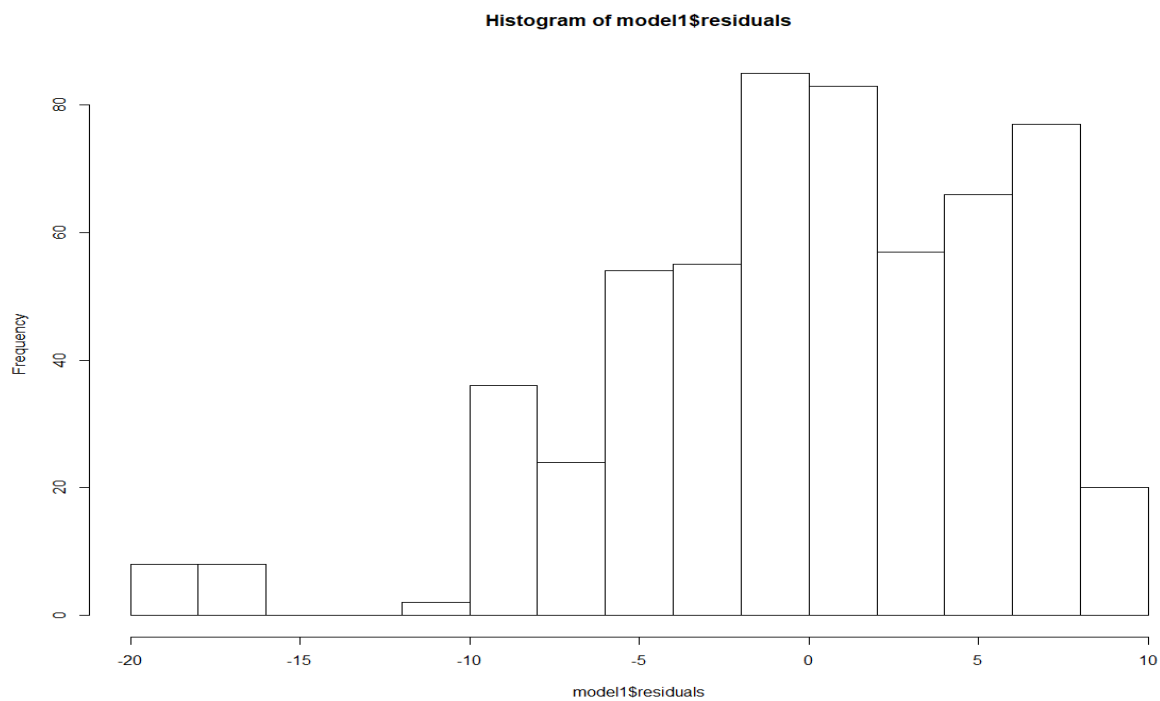
Den nya modellen som vi går vidare med är:

$$\begin{aligned} (\text{Prostate cancer death rate})^{1.84} &\sim 30.16 - 0.1523 \cdot \text{Gini coefficient} + 0.00014 \cdot \text{GDP/capita} - 0.00133 \cdot \\ &\text{Total health spendings per person}, \end{aligned}$$

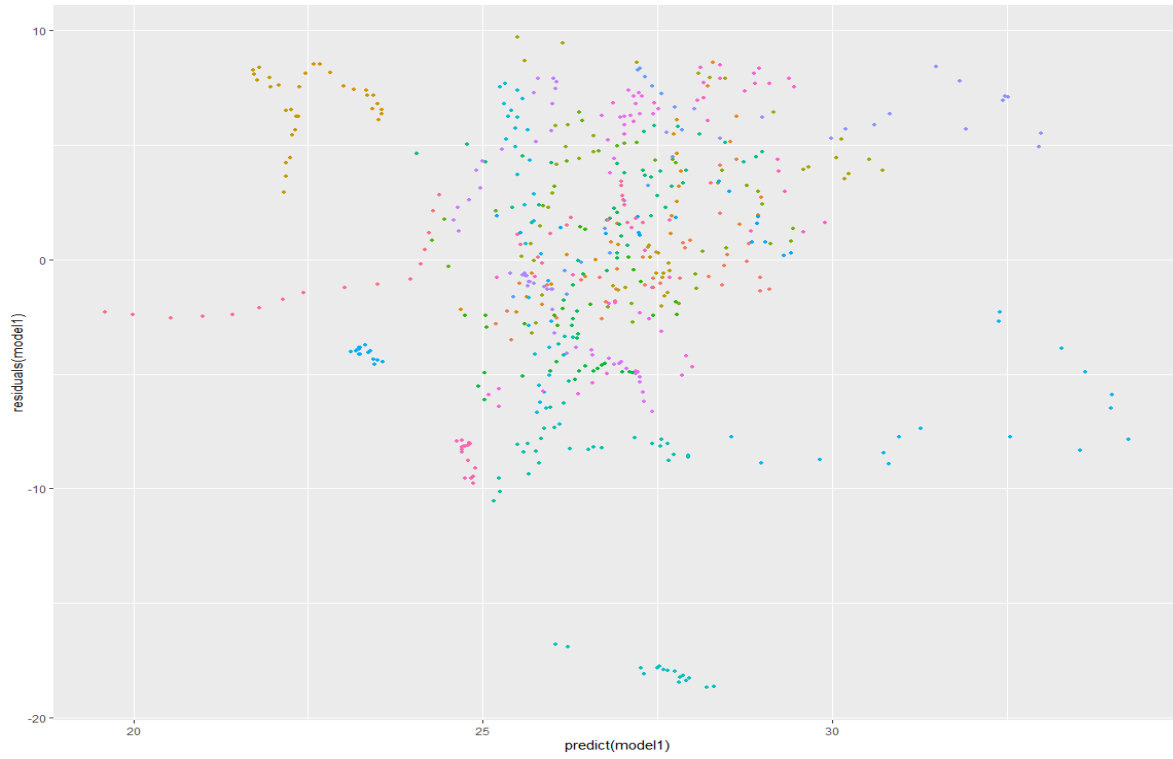
och benämns modell 2.



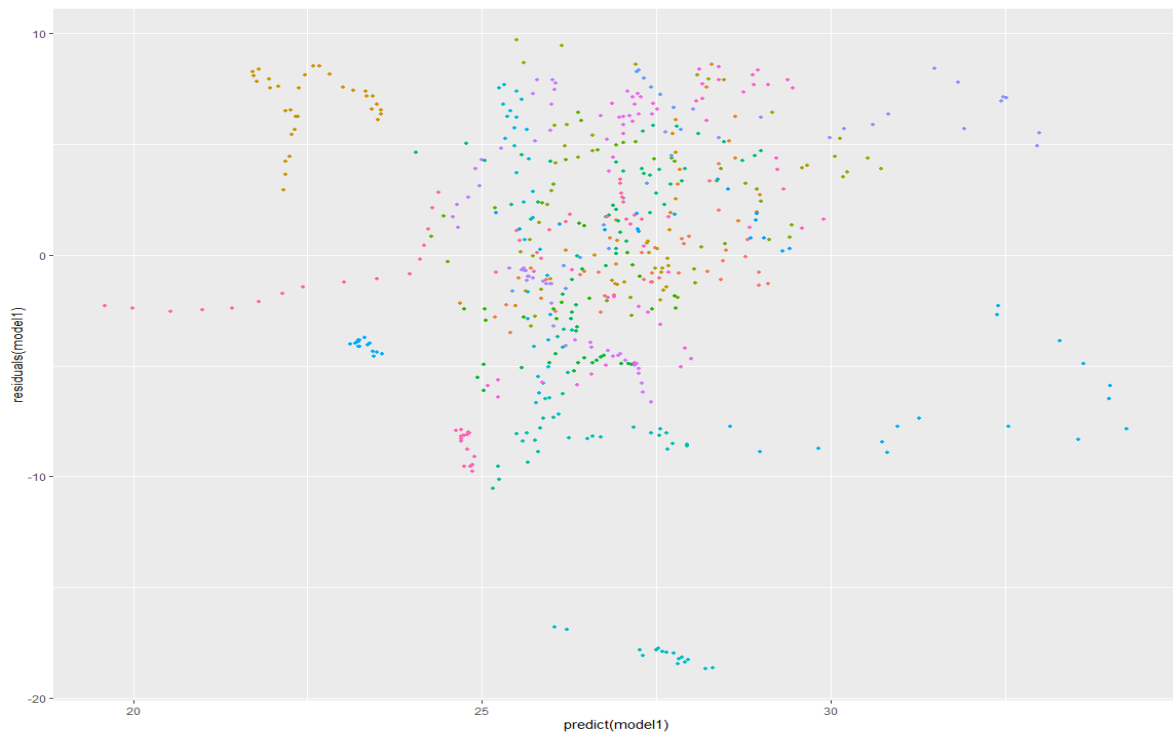
Figur 1: Q-Q-plot av teoretiska samt beräknade residualer för modell 1



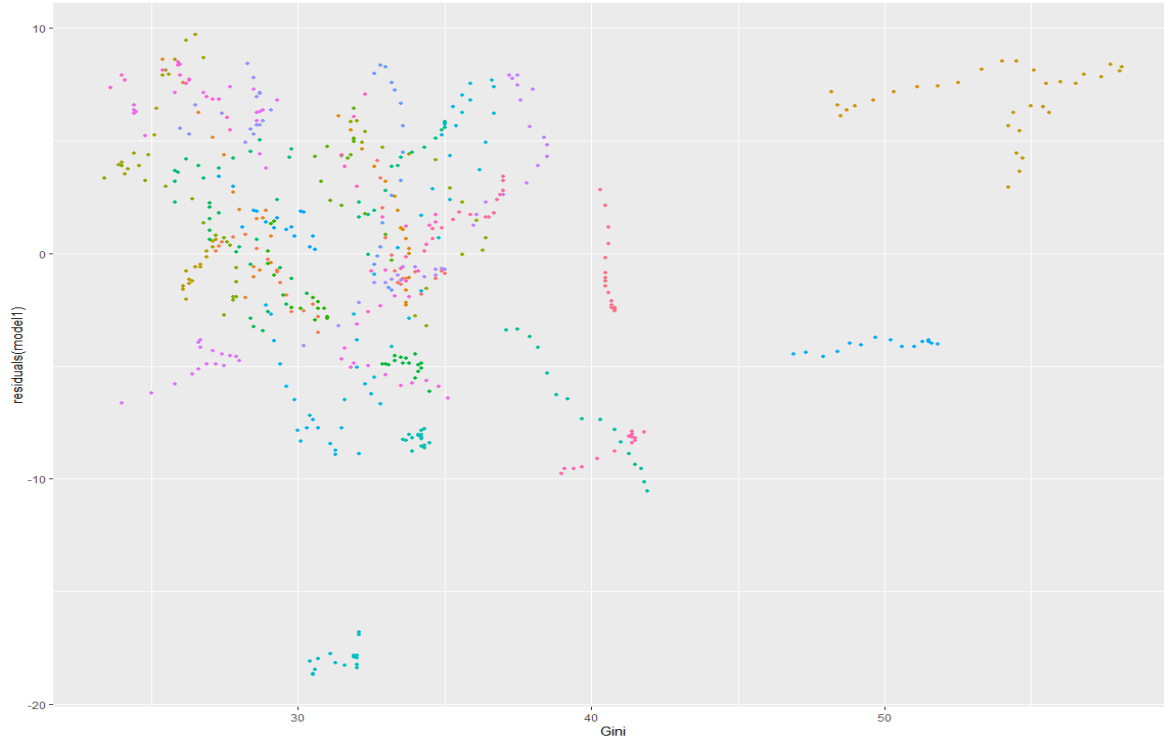
Figur 2: Histogram över residualer för modell 1



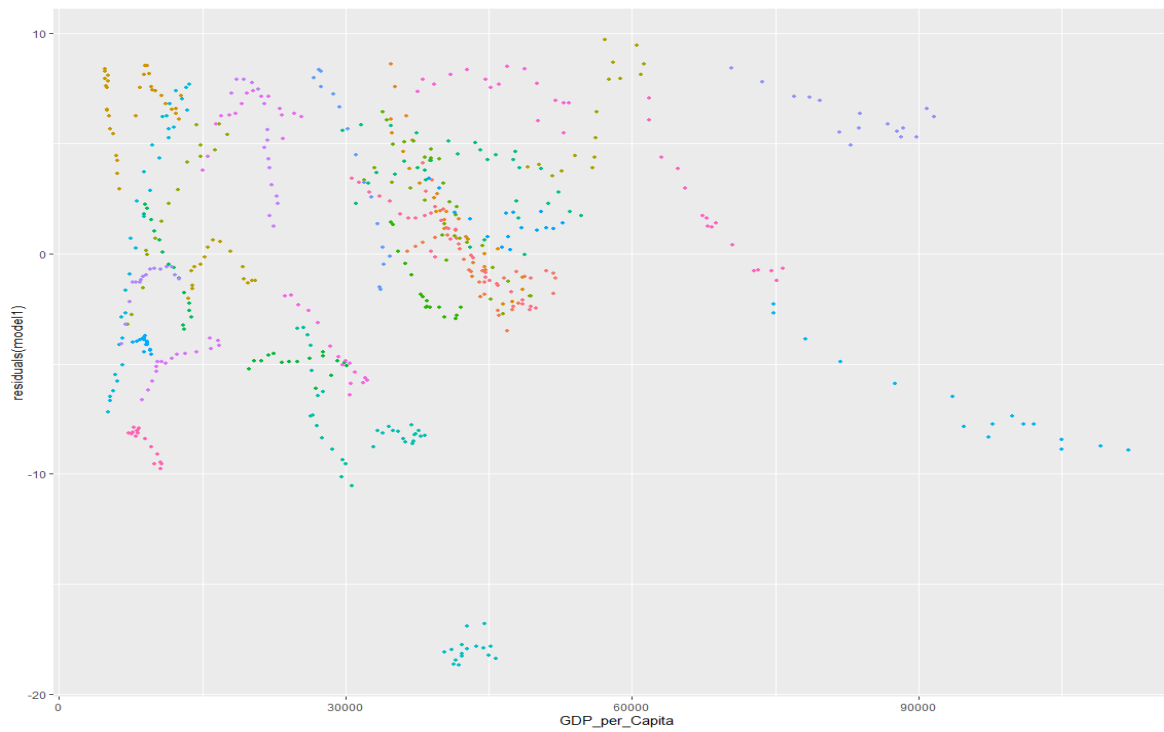
Figur 3: Plott med residualer mot predikterade värden för modell 1



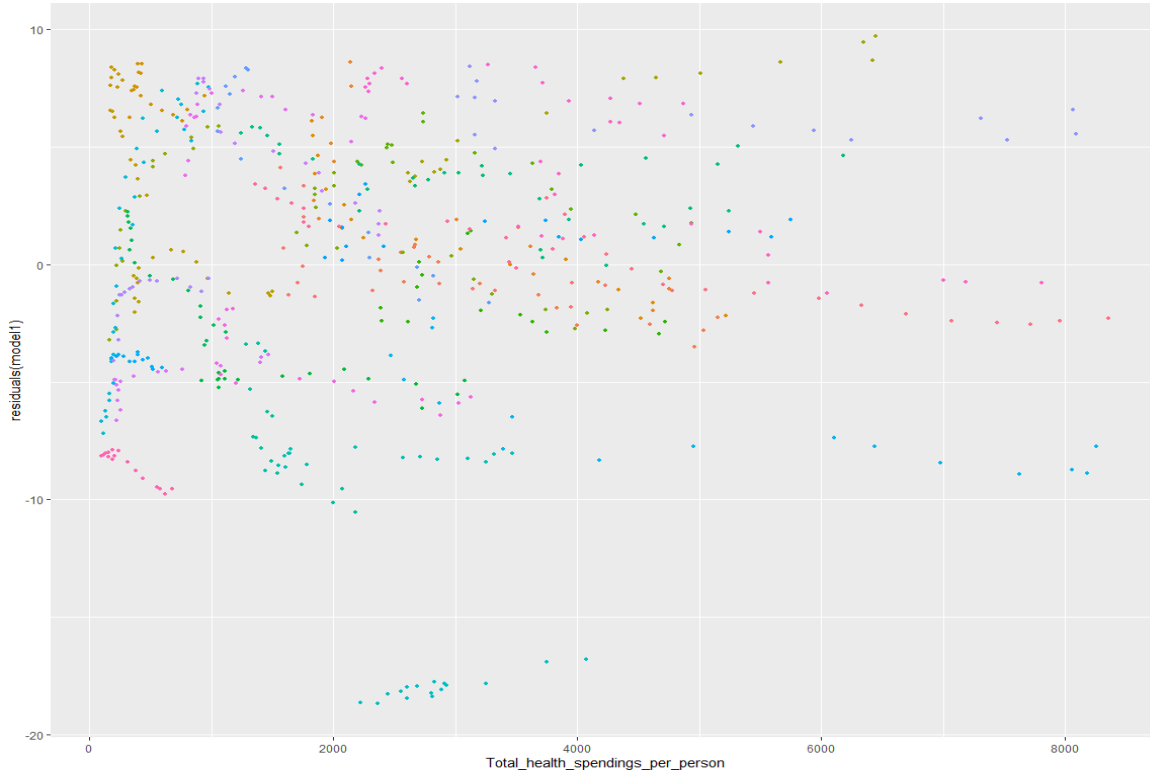
Figur 4: Plott med studentiserade residualer mot predikterade värden för modell 1



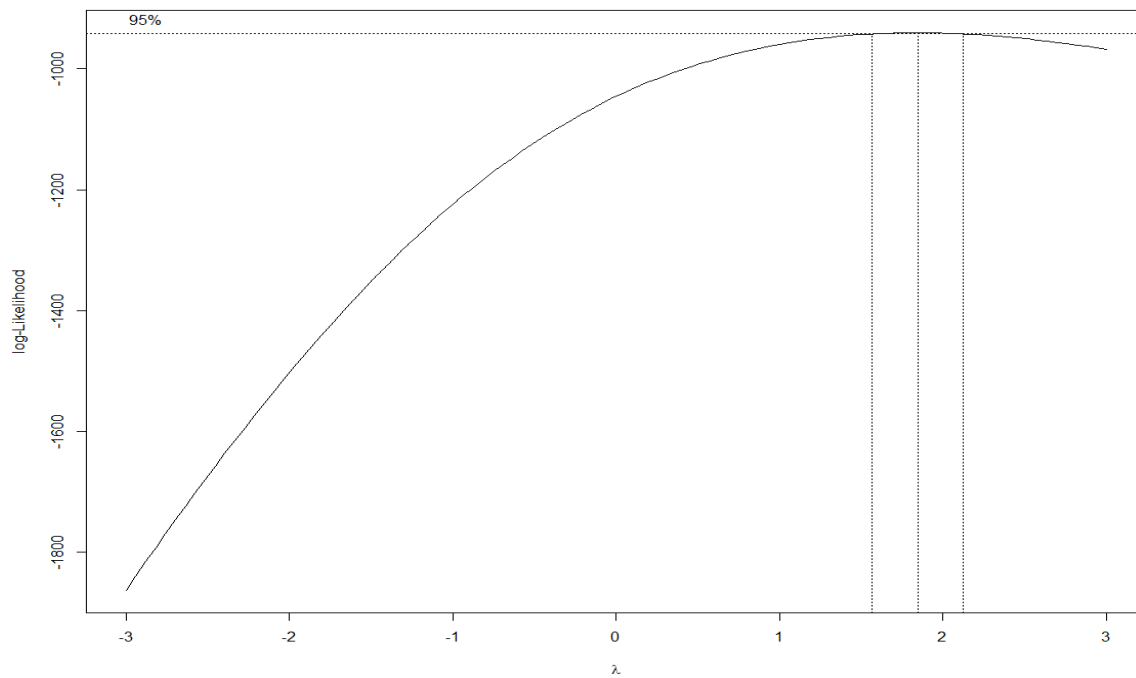
Figur 5: Plott med residualer mot Gini coefficient för modell 1



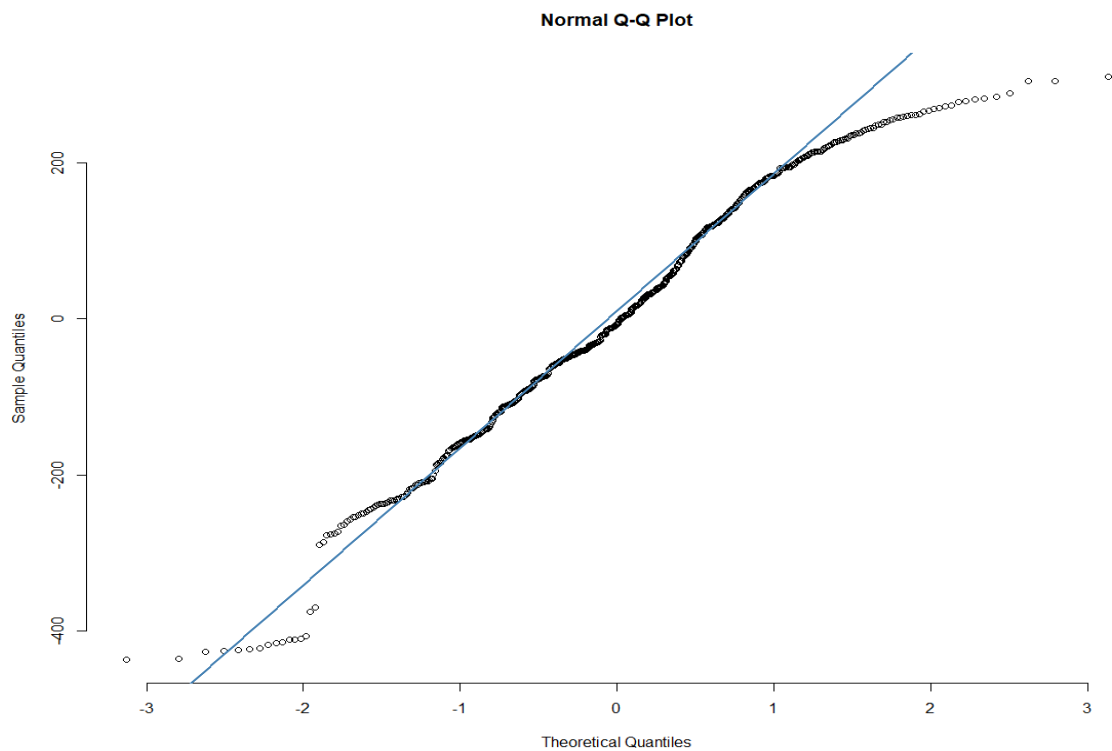
Figur 6: Plott med residualer mot GDP/Capita för modell 1



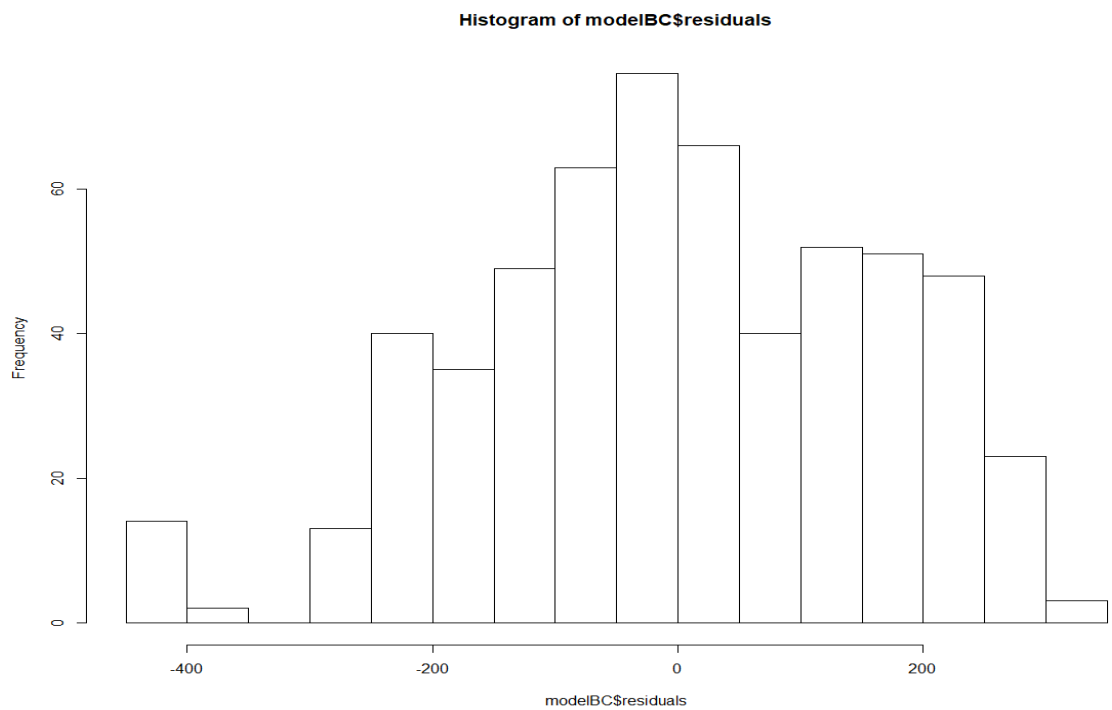
Figur 7: Plott med residualer mot Total health spendings per person för modell 1



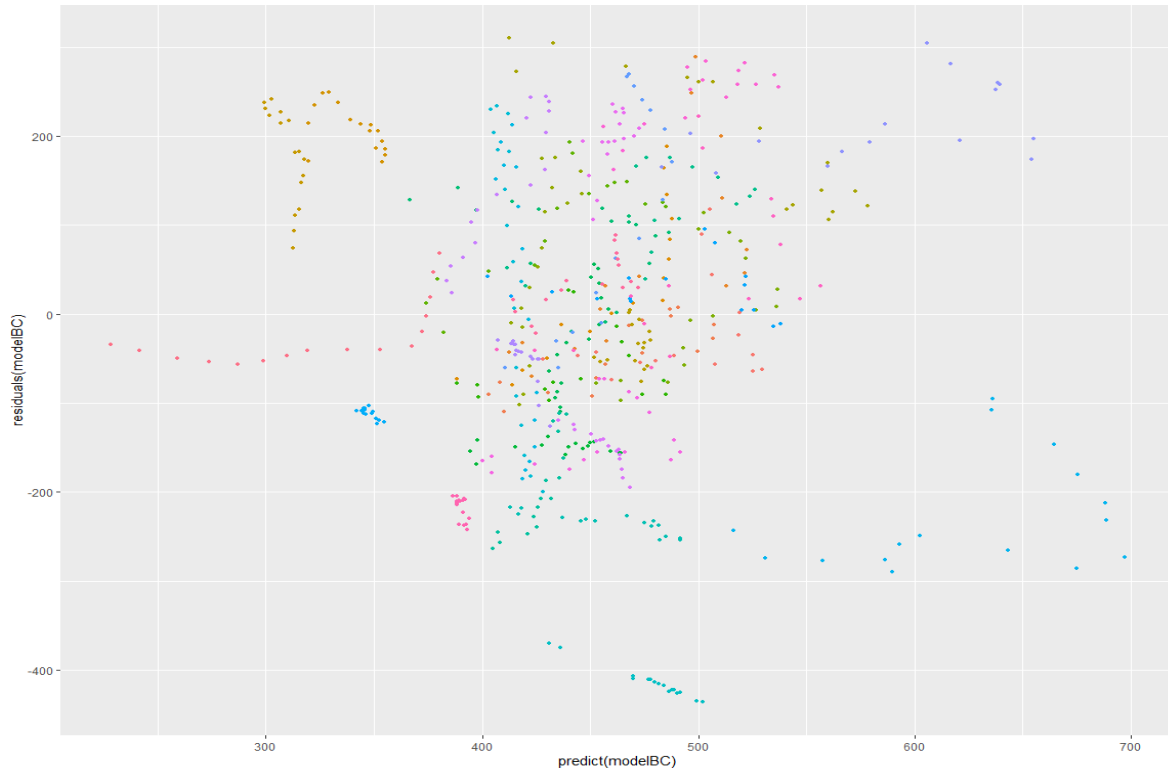
Figur 8: Plott av log-likelihood-funktionen för box-cox-transformationen av modell 1



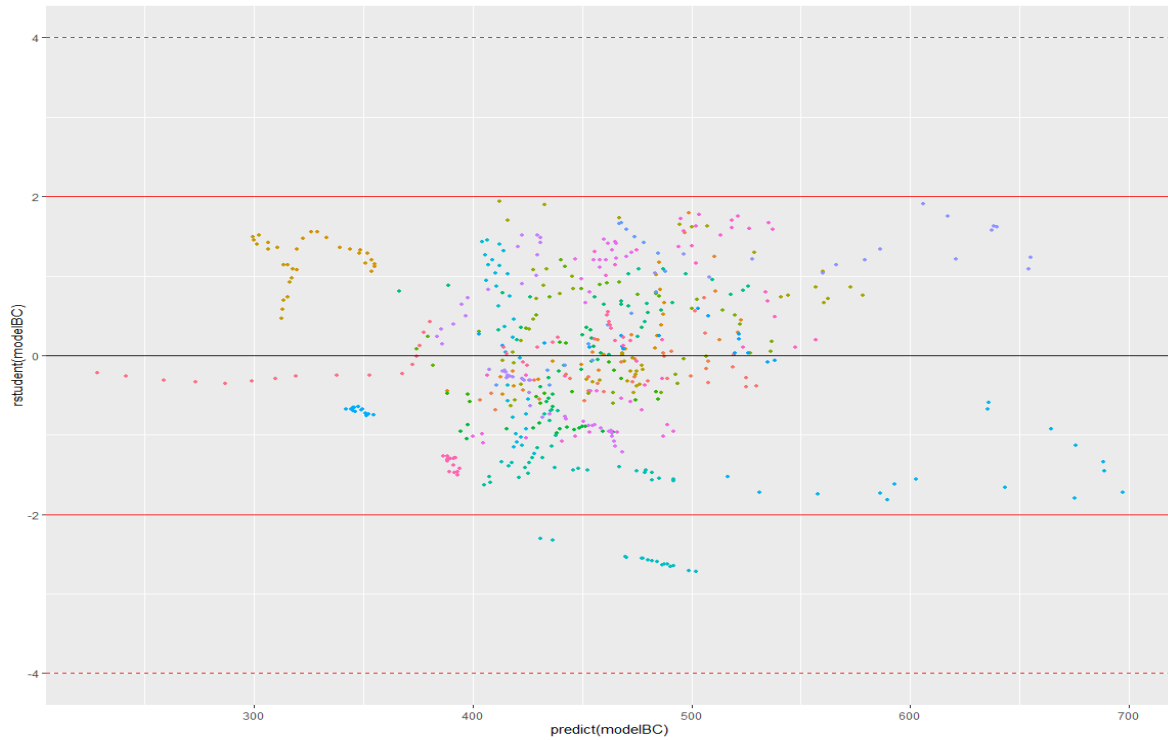
Figur 9: Q-Q-plot av teoretiska samt beräknade residualer för modell 1 box-cox-transformerad



Figur 10: Histogram över residualer för modell 1 box-cox-transformerad

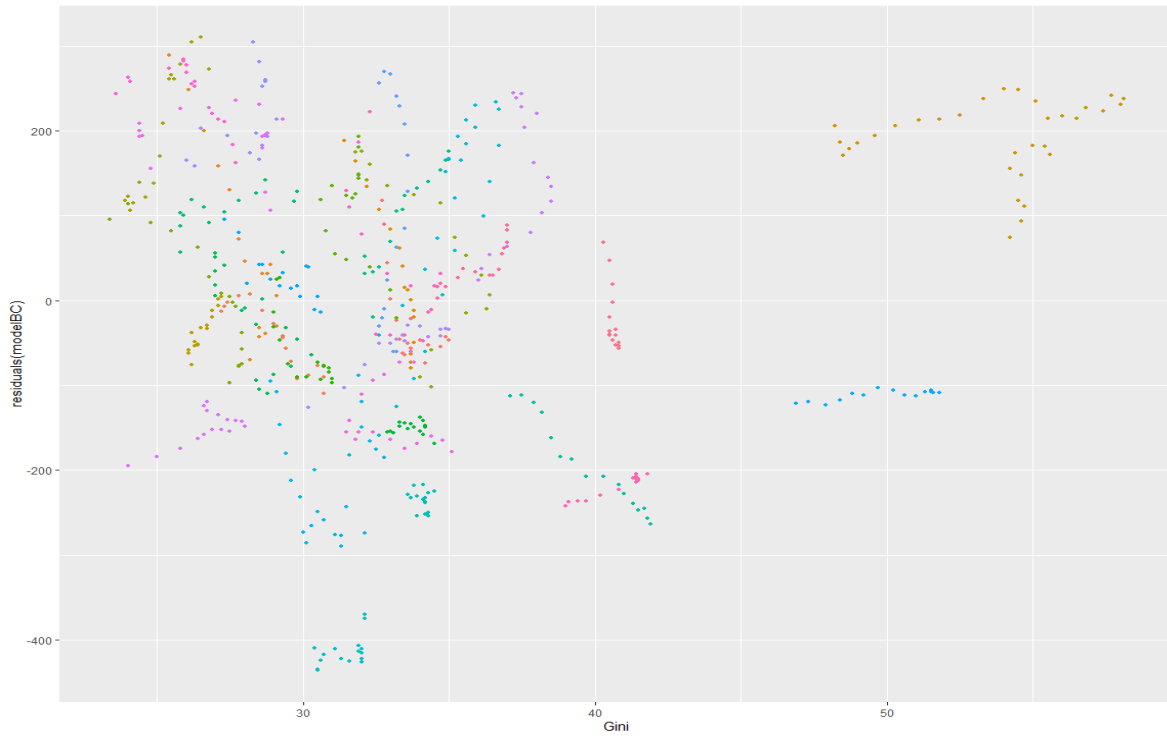


Figur 11: Plott med residualer mot predikterade värden för modell 1 box-cox-transformerad

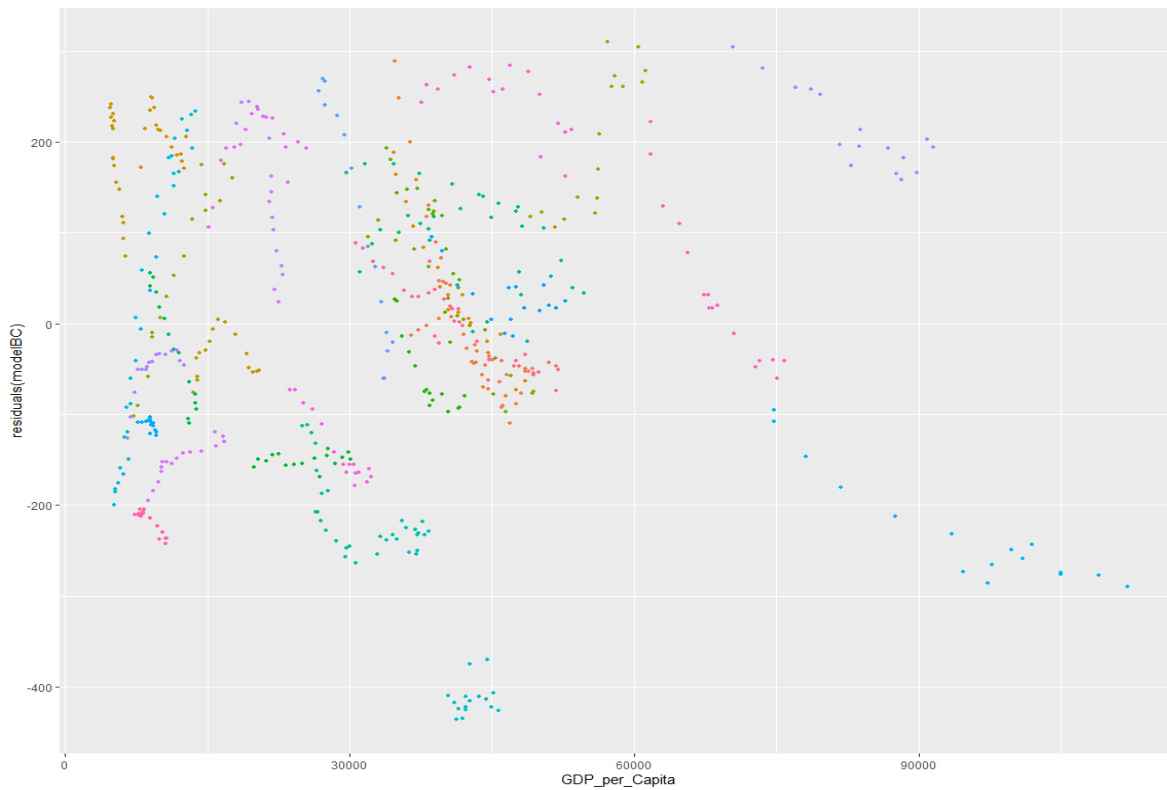


Figur 12: Plott med studentiserade residualer mot predikterade värden för modell 1 box-cox-transformerad

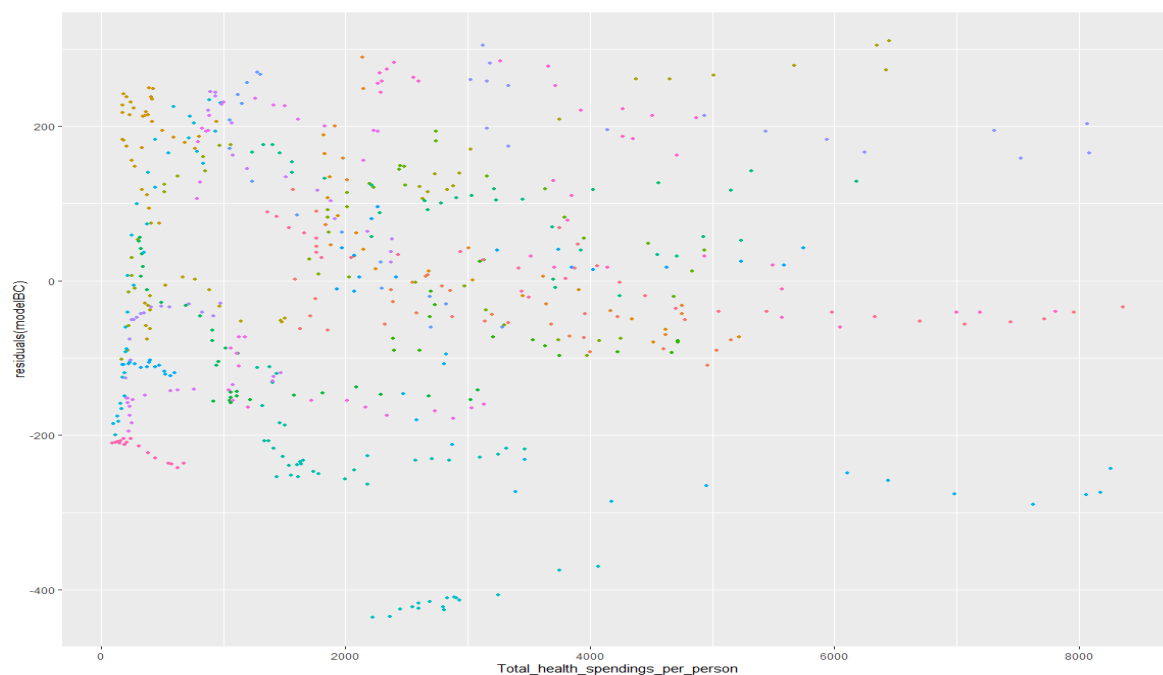




Figur 13: Plott med residualer mot Gini coefficient för modell 1 box-cox-transformerad



Figur 14: Plott med residualer mot GDP/Capita för modell 1 box-cox-transformerad



Figur 15: Plott med residualer mot Total health spendings per person för modell 1 box-cox-transformerad

D. När de oberoende variablerna Medical doctors per 1000 people och Percentage of male population aged 40+ testades att läggas till modell 2 och ett tvåvägs-ANOVA kördes mellan modell 2 och dessa nya modeller gav det resultat enligt tabell 1. Dessa resultat visade att ingen av de variabler som testades gav upphov till någon förbättrad modell och därför fortsatte modell 2 användas.

Tabell 1: P-värde och F-värde från tvåvägs-ANOVA mellan modell 2 och modell 2 med ytterligare en oberoende variabel tillagd

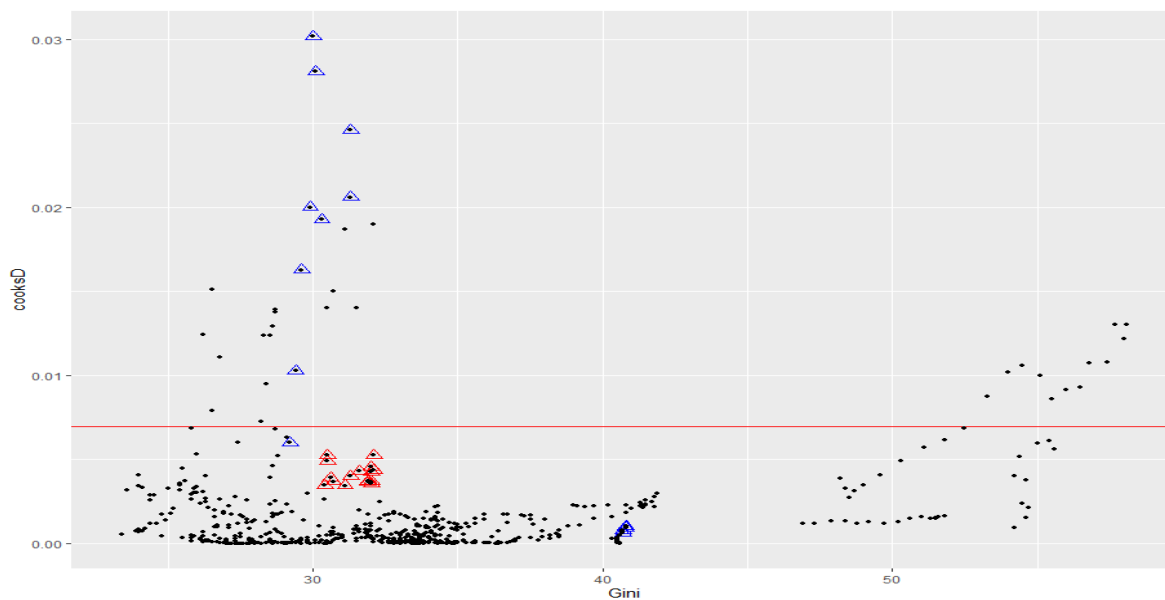
Tillagd oberoende variabel	P-värde	F-värde
Percentage of male population aged 40 +	0.1108	2.5513
Medical doctors per 1000 people	0.1256	2.3535

När dummyvariabler testades att läggas till modell 2 och ett tvåvägs-ANOVA kördes mellan modell 2 och dess nya modeller gav det ett resultat enligt tabell 2. Det visar på att vi bör lägga till dummyvariabler för länderna, vilket görs i ett senare skede.

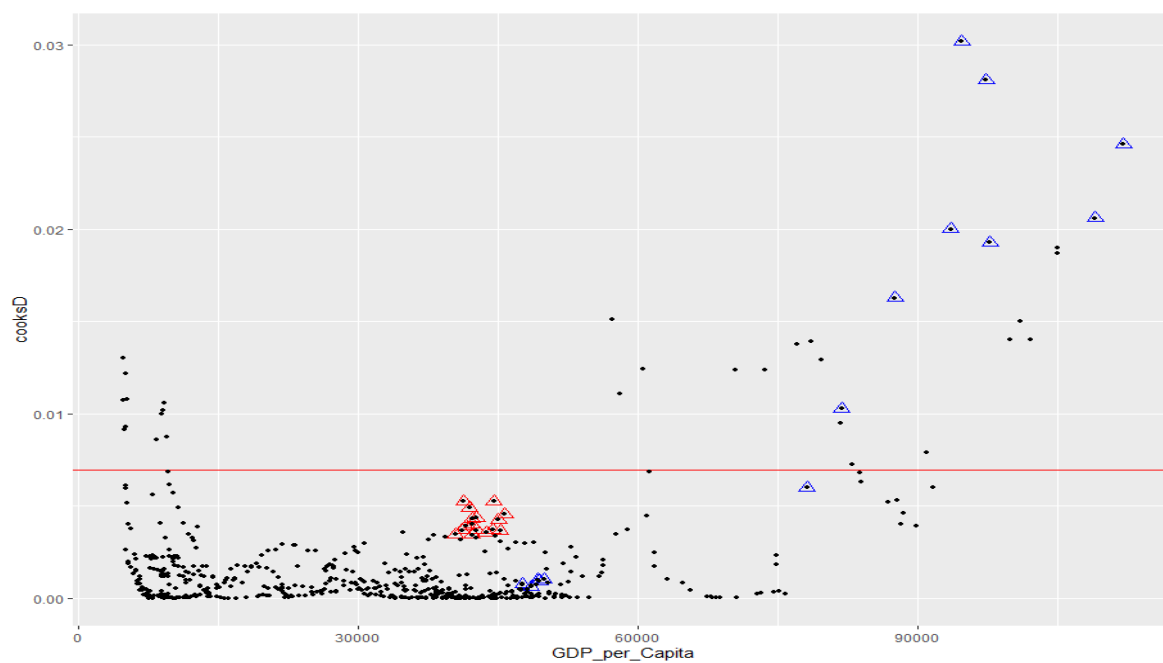
Tabell 2: P-värde och F-värde från tvåvägs-ANOVA mellan modell 2 och modell 2 med ytterligare en eller två dummyvariabler tillagda

Tillagd dummyvariabel	P-värde	F-värde
Land	$< 2.2 \cdot 10^{-16}$	94.86
År	0.8366	0.6465
Land och År	$< 2.2 \cdot 10^{-16}$	66.065

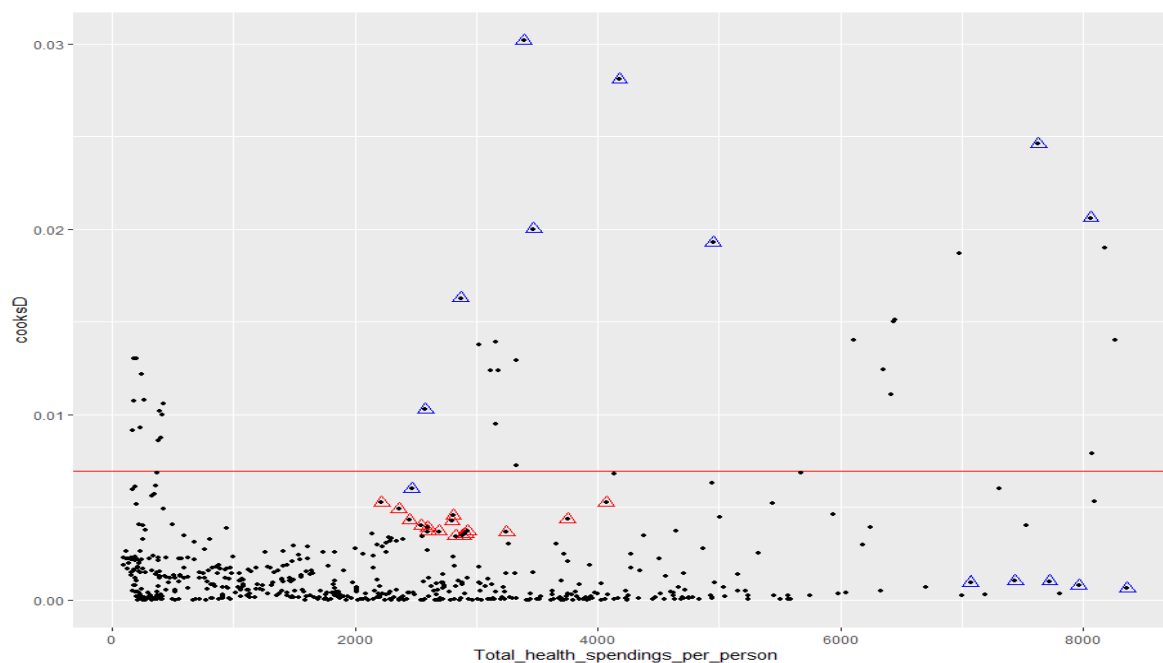
E. När leverage, studentiserade residualer samt Cook's distance beräknades för samtliga datapunkter och Cook's distance plottades mot samtliga av modell 2 ingående oberoende variabler (se figur 16-18) ser man att där är ett antal datapunkter som anses problematiska och därför plockas bort ur datasetet för det fortsatta arbetet. Modell 2 omberäknades därefter med det nya datasetet för det fortsatta arbetet, denna nya modell kallas modell 3.



Figur 16: Plott med Cook's distance mot Gini coefficient, med datapunkter där leverage  $> 0.026$  markerade med blå triangel och datapunkter med  $|\text{studentiserad residual}| > 1.96$  markerade med röd triangel



Figur 17: Plott med Cook's distance mot GDP/Capita, med datapunkter där leverage > 0.026 markerade med blå triangel och datapunkter med  $|\text{studentiserad residual}| > 1.96$  markerade med röd triangel



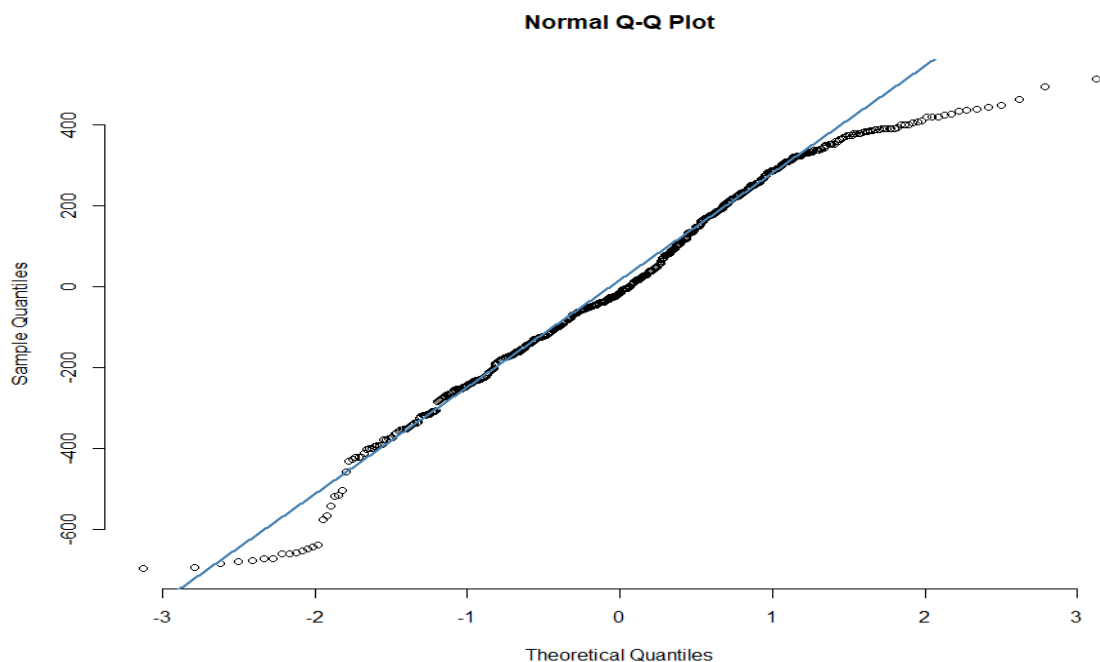
Figur 18: Plott med Cook's distance mot Total health spendings per person, med datapunkter där leverage > 0.026 markerade med blå triangel och datapunkter med  $|\text{studentiserad residual}| > 1.96$  markerade med röd triangel

F. När vi för modell 3 plottade ett histogram innehållandes modellens residualer, en Q-Q-plot, modellens residualer mot modellens predikterade värden, modellens studentiserade residualer mot modellens predikterade värden samt modellens residualer mot modellens ingående oberoende variabler ansågs fortfarande box-cox-transformationen vara den bästa (se figur 19-25). Beräkningar av box-cox-transformationens lambda-värde gav  $\lambda = 1.96$  (se figur 26).

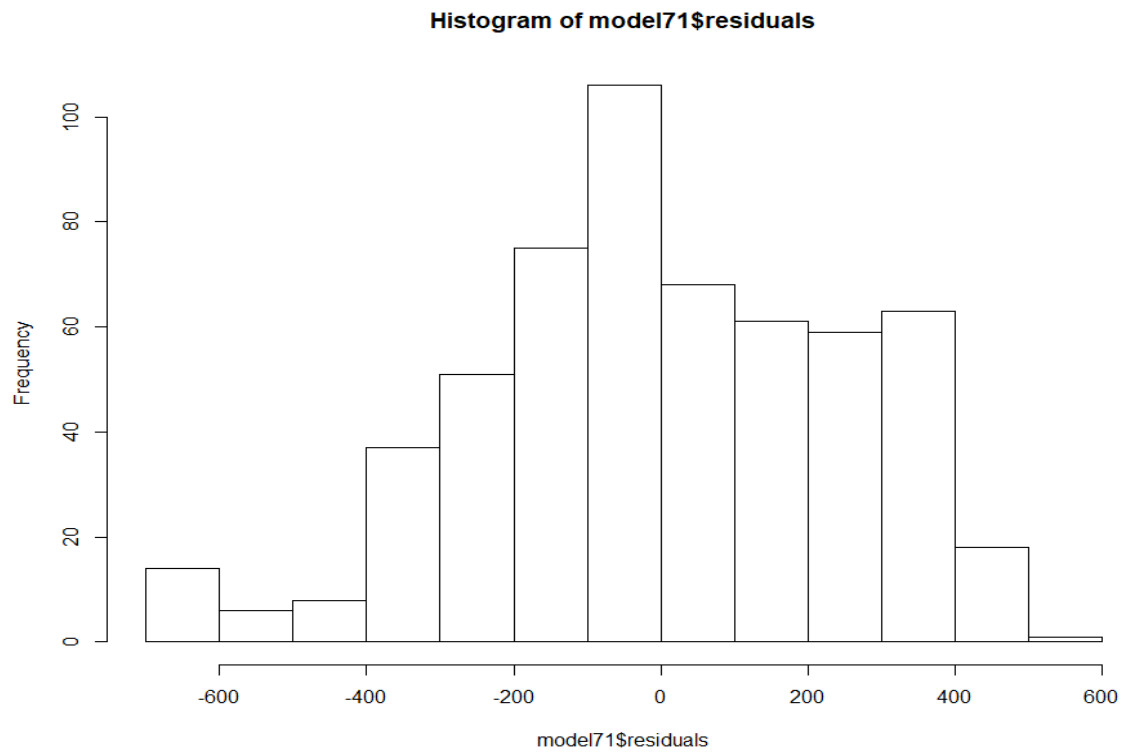
Den nya modellen som vi går vidare med är:

$$(\text{Prostate cancer death rate})^{1.96} \sim 750.3 - 5.894 \cdot \text{Gini coefficient} + 0.010 \cdot \frac{\text{GDP}}{\text{capita}} - 0.086 \cdot \text{Total health spendings per person},$$

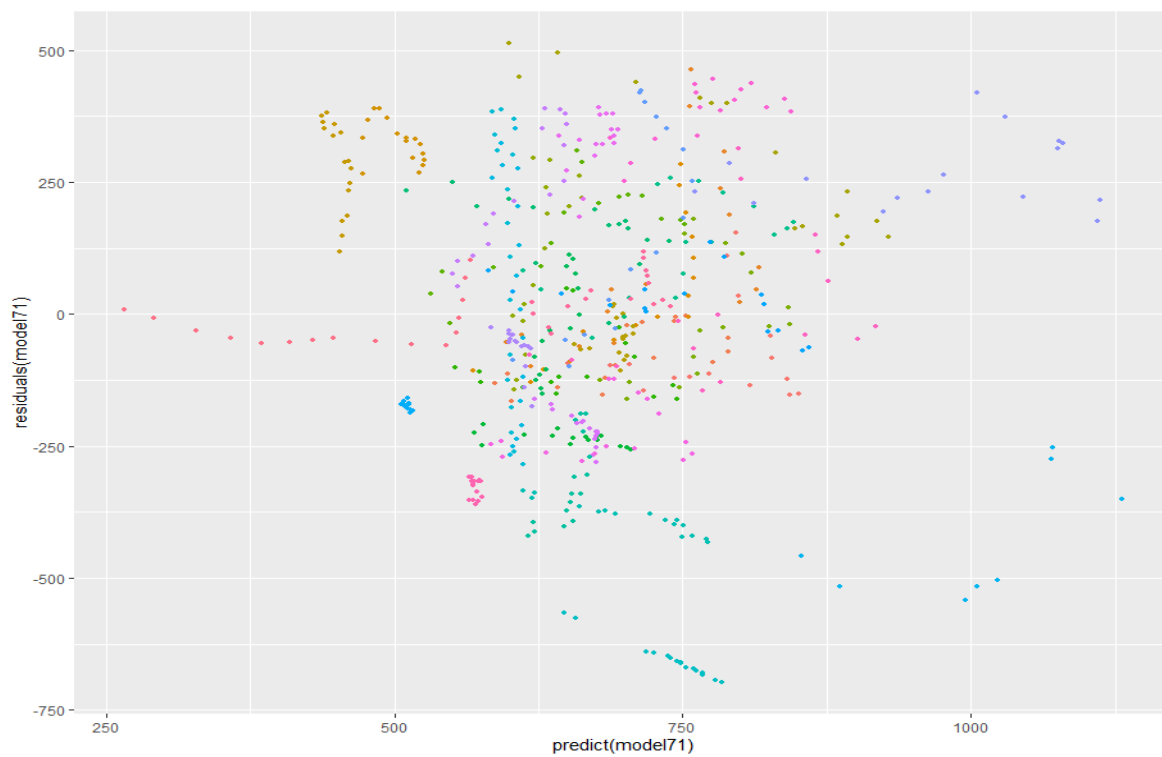
och benämns modell 4.



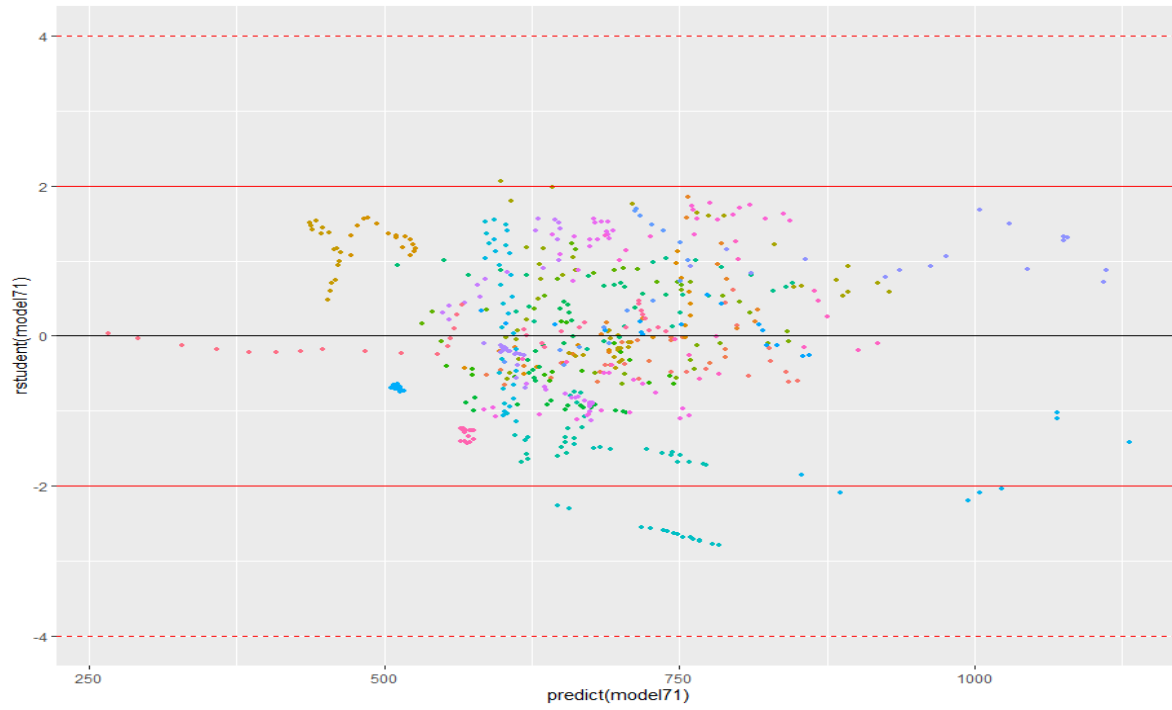
Figur 19: Q-Q-plot av teoretiska samt beräknade residualer för modell 3



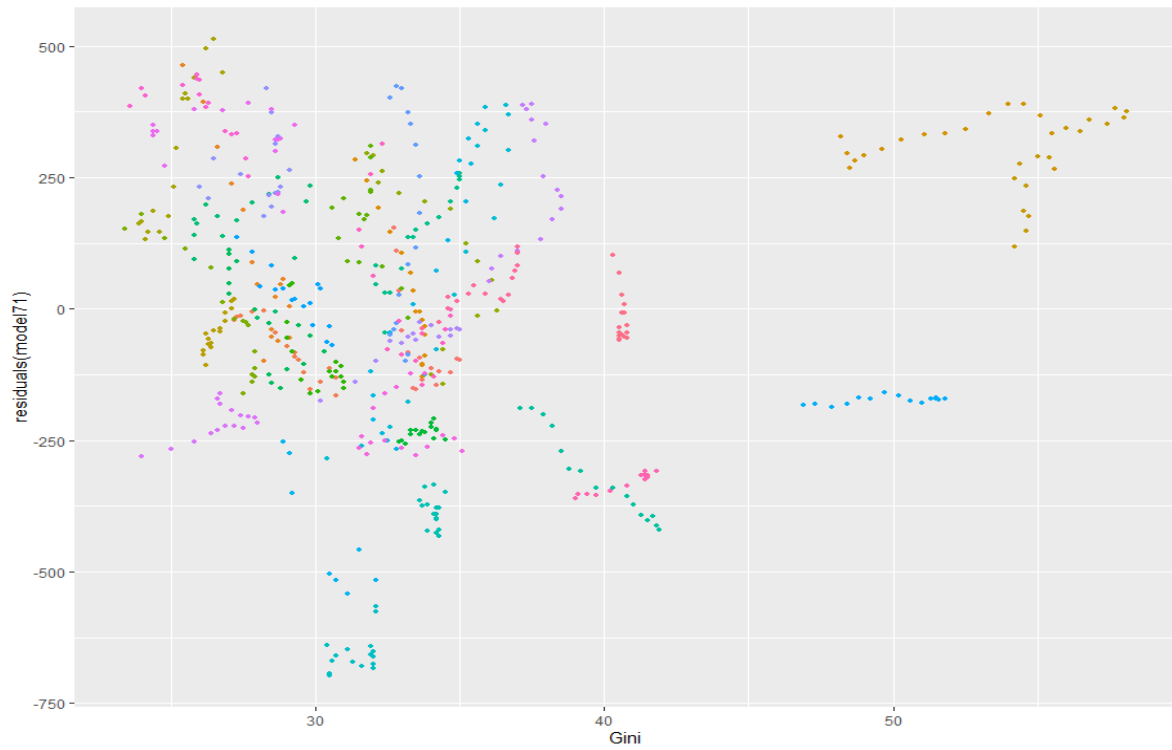
Figur 20: Histogram över residualer för modell 3



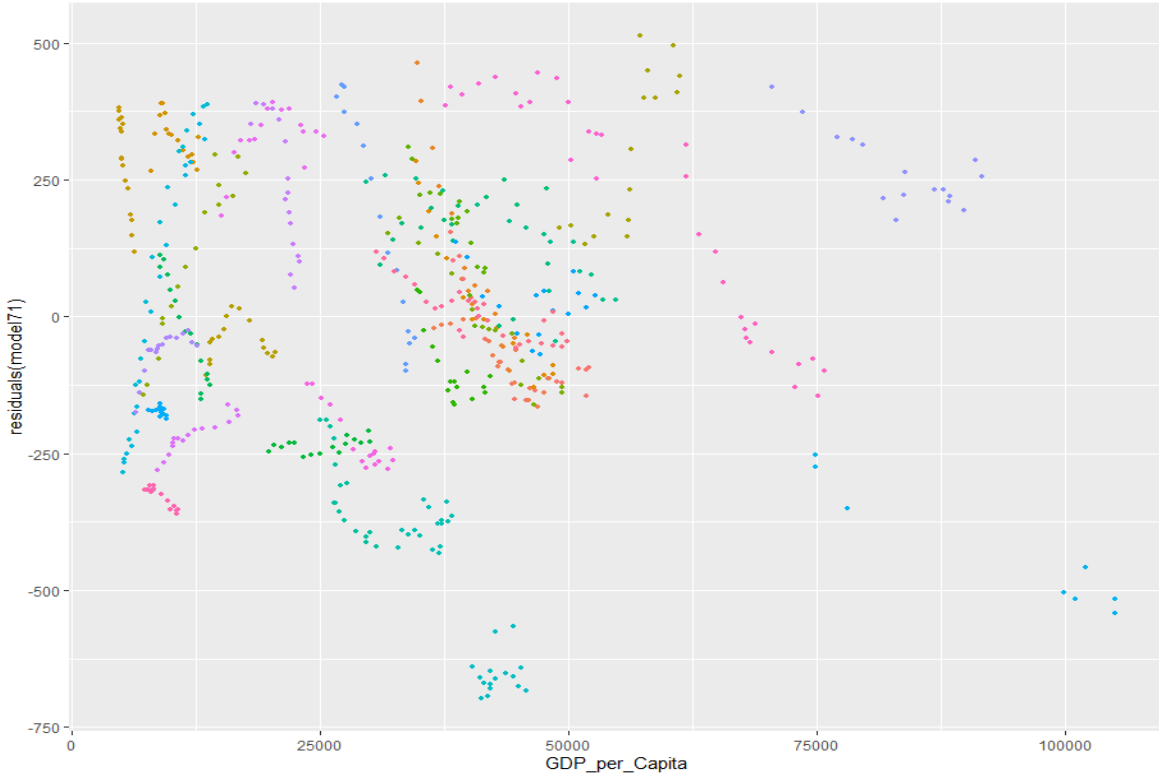
Figur 21: Plott med residualer mot predikterade värden för modell 3



Figur 22: Plott med studentiserade residualer mot predikterade värden för modell 3



Figur 23: Plott med residualer mot Gini coefficient för modell 3

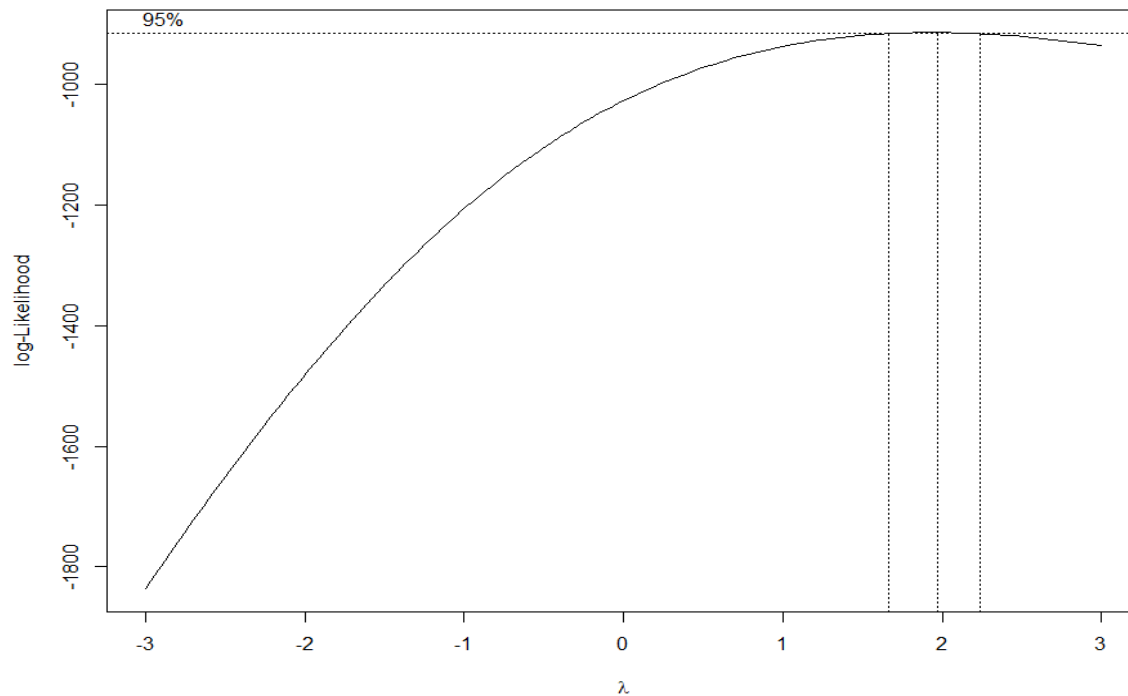


Figur 24: Plott med residualer mot GDP/Capita för modell 3



Figur 25: Plott med residualer mot Total health spendings per person för modell 1





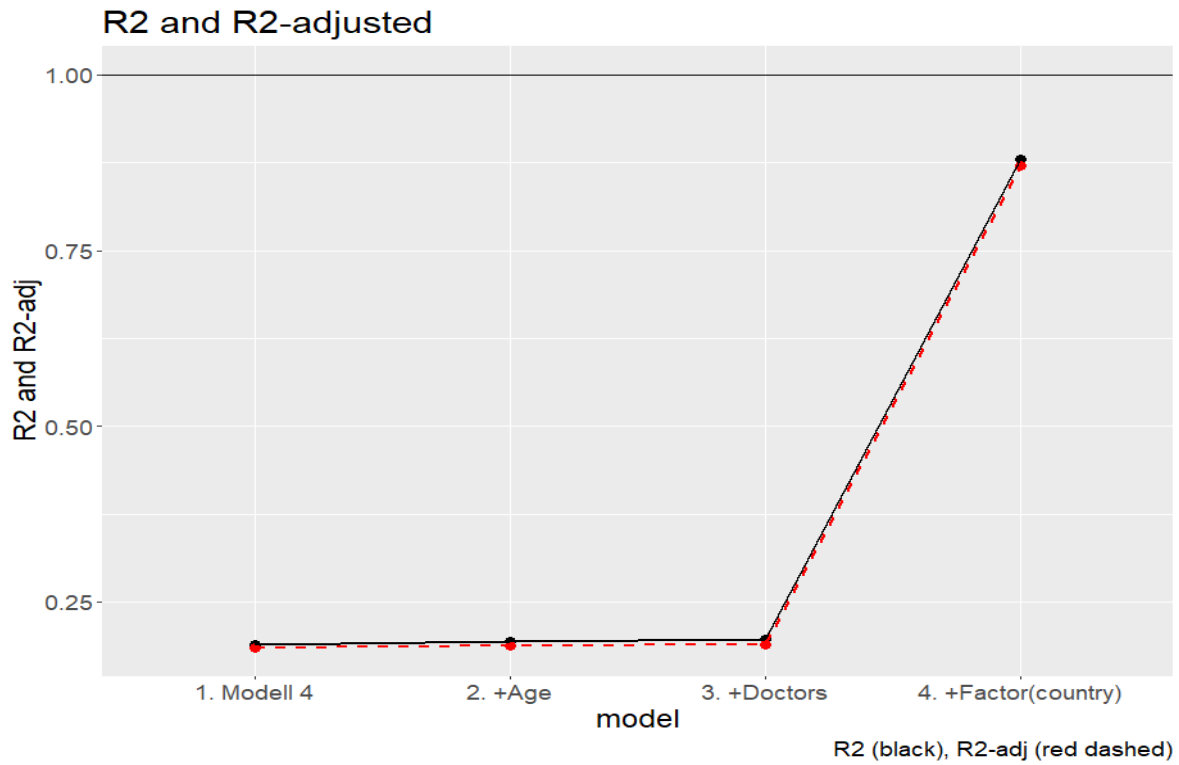
Figur 26: Plott av log-likelihood-funktionen för box-cox-transformationen av modell 3

G. Beräkningar av R<sup>2</sup>, R-adjusted, AIC och BIC för modell 4 samt modell 4 med Medical doctors per 1000 people och Percentage of male population aged 40+ samt en dummyvariabel för länderna indikerar det att man får en förbättring i modellen om man lägger till dummyvariabler för länderna (se figur 27-28).

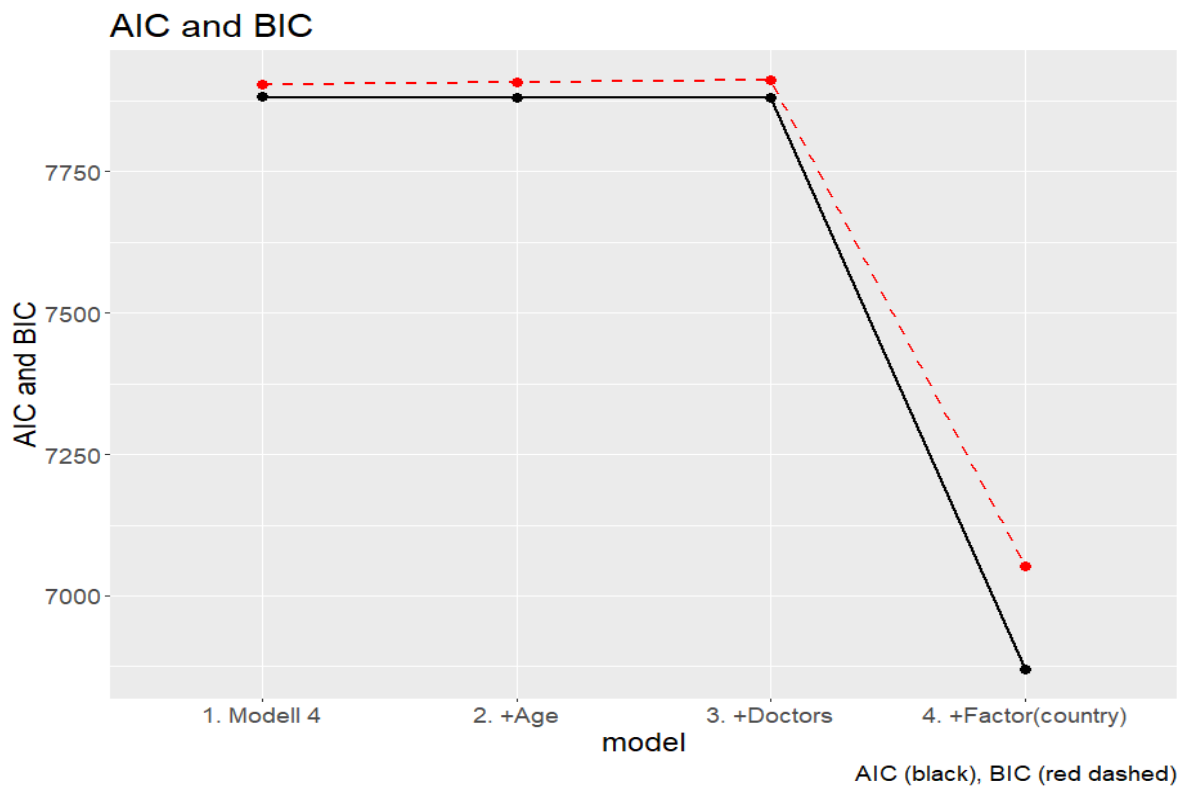
Den nya modellen som vi går vidare med är:

$$(Prostate\ cancer\ death\ rate)^{1.96} \sim 750.3 - 5.894 \cdot Gini\ coefficient + 0.010 \cdot GDP/capita - 0.086 \cdot Total\ health\ spendings\ per\ person + factor(countries),$$

och benämns modell 5.



Figur 27: R2 och R2-adjusted plottad för modell 4 samt varianter av modell 4



Figur 28: AIC och BIC plottad för modell 4 samt varianter av modell 4

H. När en forward selection körs med BIC som kriterium och

$$\begin{aligned} \text{Prostate cancer death rate} &\sim \text{Total health spendings per person} + \text{GDP/capita} \\ &+ \text{Gini coefficient} + \text{Percentage of male population aged 40} \\ &+ \text{Medical doctors per 1000 people} + \text{factor(country)} + \text{factor(year)} \end{aligned}$$

som övre tillåtna modell vilket gav en modell enligt:

$$(\text{Prostate cancer death rate})^{1.96} \sim 32.02 - 0.002 \cdot \text{Total health spendings per person} + \text{factor(countries)}$$

vilket blir den modell vi går vidare med, och benämns modell 5.

I. En residualanalys på modell 5 och transformationer på denna visade inget nytt och samma transformation som tidigare fortsätter användas. Om man lägger till en oberoende variabel till modell 5 och genomför ett tvåvägs-ANOVA-test syns inga förbättringar på modell 5 (se tabell 3). Genomför man ett envägs-ANOVA-test på modell 5 och på modell 5 där man bytt ut den oberoende variabeln mot var en av de andra, ej ingående, oberoende variablerna tyder dessa tester på att den modell som innehåller GDP/capita istället för Total health spendings per person är den bästa med högst F-värde lägst P-värde (se tabell 4). Ett korrelationstest visar att GDP/capita och Total health spendings är högt korrelerade med en korrelation på 0.86. Envägs-ANOVA-testerna visar högst F-värde för modellen med GDP/capita som oberoende variabel (se tabell 4). Detta gör att vi vill se Total health spendings som en mekanism av GDP/capita, och leder till att den slutliga modellen som väljs ser ut enligt följande:

$$(\text{Prostate cancer death rate})^{1.96} \sim 1.401 \cdot 10^3 - 1.518 \cdot 10^{-2} \cdot \text{GDP/capita} + \alpha^T \mathbf{D},$$

som benämns modell 6. Här är  $\mathbf{D}$  en vektor med dummyvariabler för länderna med Australien som referensland, och  $\alpha$  är en vektor med dummyvariablernas koefficienter (se tabell 41, bilaga A).

Tabell 3: P-värde och F-värde från tvåvägs-ANOVA mellan modell 5 och modell 5 med ytterligare en oberoende variabel

Tillagd oberoende variabler	P-värde	F-värde
GDP/capita	0.0924	2.8414
Medical doctors per 1000 people	0.9194	0.0102
Gini coefficient	0.4243	0.6393
Percentage of male population aged 40 +	0.0925	2.8414

Tabell 4: P-värde och F-värde från envägs-ANOVA för modell 5 och modell 5 med den oberoende variabeln utbytt mot var en av de andra oberoende variablerna

Oberoende variabel utöver dummyvariabler för länder	P-värde	F-värde
GDP/Capita	$<2.2 \cdot 10^{-16}$	292.373
Medical doctors per 1000 people	0.6487	0.2078
Gini coefficient	$<2.2 \cdot 10^{-16}$	204.415
Percentage of male population aged 40 +	$2.904 \cdot 10^{-14}$	61.124
Total health spendings per person	$1.044 \cdot 10^{-9}$	38.619

## 5. Ekvationssystemet

Ekvationssystemets uppbyggnad utgår från hur den erhållna regressionsmodellen (modell 6) ser ut. Samtliga förändringskoefficienter är estimerade genom att i RStudio uppskatta en linjär modell för hur varje variabel beror varje annan variabel, samt en ingående vektor med dummyvariabler för att kontrollera för systemeffekter.

Vi antar att vi tittar på ett så pass kort tidsintervall att variabler endast hinner svara på förändringar i GDP/capita (US\$, inflation-adjusted) samt hur andra variabler initialt ändras med förändringar i GDP/capita (US\$, inflation-adjusted). Vi antar också att förändringar i kommande tidsintervall på grund av någon annan förändring än nyss nämnda är så pass små att de är försumbara. Detta gör att vi undgår att få ett dynamiskt system med variabler över tid.

Ekvationssystemet sätts upp enligt:

$\beta_{i,j}$ : Förändringskoefficient där  $X_j$  beror på  $X_i$

$\delta_k$ : Förändringskoefficient där Prostate cancer death rate beror på variabel  $X_k$

$D$ : Vektor med dummyvariabler som beror på länder, där Australien är referensland

$\alpha$ : Vektor med dummyvariablernas koefficienter

$Y_i$ : Prostate cancer death rate

$X_1$ : GDP/capita (US\$, inflation-adjusted)

$X_2$ : Total health spendings per person (US\$)

$X_3$ : Medical doctors per 1000 people

$X_4$ : Gini coefficient

$X_5$ : Percentage of male population aged 40 +

Hur de oberoende variablerna ändras när GDP/capita ändras kan beskrivas enligt:

$$dX_j = \sum_{i=1, i \neq j}^5 \beta_{i,j} \frac{dX_i}{X_1}, j = 2,3,4,5$$

Vilket ger:

$$dX_2 = \beta_{1,2} + \beta_{3,2} \frac{dX_3}{dX_1} + \beta_{4,2} \frac{dX_4}{dX_1} + \beta_{5,2} \frac{dX_5}{dX_1} = \beta_{1,2} + \beta_{3,2}\beta_{1,3} + \beta_{4,2}\beta_{1,4} + \beta_{5,2}\beta_{1,5}$$

$$dX_3 = \beta_{1,3} + \beta_{2,3} \frac{dX_2}{dX_1} + \beta_{4,3} \frac{dX_4}{dX_1} + \beta_{5,3} \frac{dX_5}{dX_1} = \beta_{1,3} + \beta_{2,3}\beta_{1,2} + \beta_{4,3}\beta_{1,4} + \beta_{5,3}\beta_{1,5}$$

$$dX_4 = \beta_{1,4} + \beta_{2,4} \frac{dX_2}{dX_1} + \beta_{3,4} \frac{dX_3}{dX_1} + \beta_{5,4} \frac{dX_5}{dX_1} = \beta_{1,4} + \beta_{2,4}\beta_{1,2} + \beta_{3,4}\beta_{1,3} + \beta_{5,4}\beta_{1,5}$$

$$dX_5 = \beta_{1,5} + \beta_{2,5} \frac{dX_2}{dX_1} + \beta_{3,5} \frac{dX_3}{dX_1} + \beta_{4,5} \frac{dX_4}{dX_1} = \beta_{1,5} + \beta_{2,5}\beta_{1,2} + \beta_{3,5}\beta_{1,3} + \beta_{4,5}\beta_{1,4}$$

Hur Prostate cancer death rate ändras beroende på  $X_2, X_3, X_4$  och  $X_5$ :

$$dY_2 = \delta_2 dX_2 + \delta_3 dX_3 + \delta_4 dX_4 + \delta_5 dX_5$$

Hur Prostate cancer death rate beror på GDP/capita (US\$, inflation-adjusted):

$$\text{Modell 6 : } Y_1^{1.96} = \text{Intercept} + \delta_1 X_1 + \alpha^T \mathbf{D} \Leftrightarrow Y_1 = (\text{Intercept} + \delta_1 X_1 + \alpha^T \mathbf{D})^{1/1.96}$$

Hur Prostate cancer death rate ändras beroende på de oberoende variablerna när GDP/capita ändras kan skrivas som:

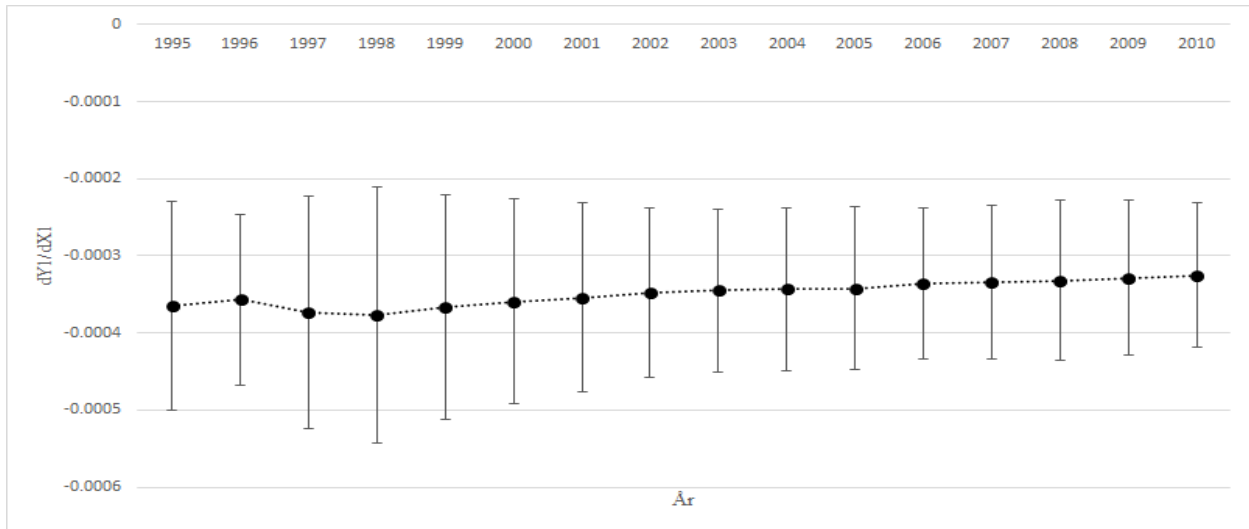
$$\frac{dY_1}{dX_1} = \frac{\delta_1}{1.96} (\text{Intercept} + \delta_1 X_1 + \alpha^T \mathbf{D})^{(1/1.96)-1}, \frac{dY_1}{dX_1} = dY_2 + dW, \text{ eller } \frac{dY_1}{dX_1} - dY_2 = dW$$

där  $dW$  är skillnaden mellan förändring i Prostate cancer death rate för  $\frac{dY_1}{dX_1}$  och  $dY_2$ .

Beräkningar för 1995-2010 års GDP/capita med en förändring på 1 GDP/capita (US\$), samt beräkningar av förändringskoefficienterna ger att:  $dY_2 = -0.001020188594$  och  $\frac{dY_1}{dX_1}$ ,  $dW$  enligt tabell 5- 36 (se bilaga A) och förändringskoefficienter enligt tabell 37 och 38 (se bilaga B).

## 6. Numeriska resultat

Beräkningar ger att  $dY_2 = -0.001020188594$ , vilket betyder att då GDP/capita ökar med 1 så bör Prostate cancer death rate gå ner med cirka 0.001 person per 100,000 på grund av att GDP/capita ändras vilket påverkar de i regressionsmodellen ej ingående variablerna som i sin tur påverkar Prostate cancer death rate. Med beräkningar av  $\frac{dY_1}{dX_1}$ , för varje år 1995-2010 och land (se tabell 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35; bilaga A), beräknas  $\frac{dY_1}{dX_1}$  för varje enskilt år med 95%-konfidensintervall (se tabell 39, bilaga A; figur 29), samt ett 95%-konfidensintervall för  $\frac{dY_1}{dX_1}$  över samtliga år ( $I_{\frac{dY_1}{dX_1}} = (-0.00036; -0.00034)$ ).



Figur 29:  $\frac{dY_1}{dX_1}$  för varje enskilt år med konfidensintervall

Beräkningarna gjorda av  $\frac{dY_1}{dX_1}$  ger vidare att  $dW \neq 0$  och positiv för samtliga länder, för alla år 1995-2010, förutom vid 5 tillfällen då  $dW$  istället var negativ (se tabell 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36; bilaga A). Med dessa beräkningar beräknas  $dW$  för varje enskilt år med 95%- konfidensintervall (se tabell 40, bilaga A; figur 30), samt ett 95%-konfidensintervall för  $dW$  över samtliga år ( $I_{dW} = (0.00066; 0.00068)$ ).

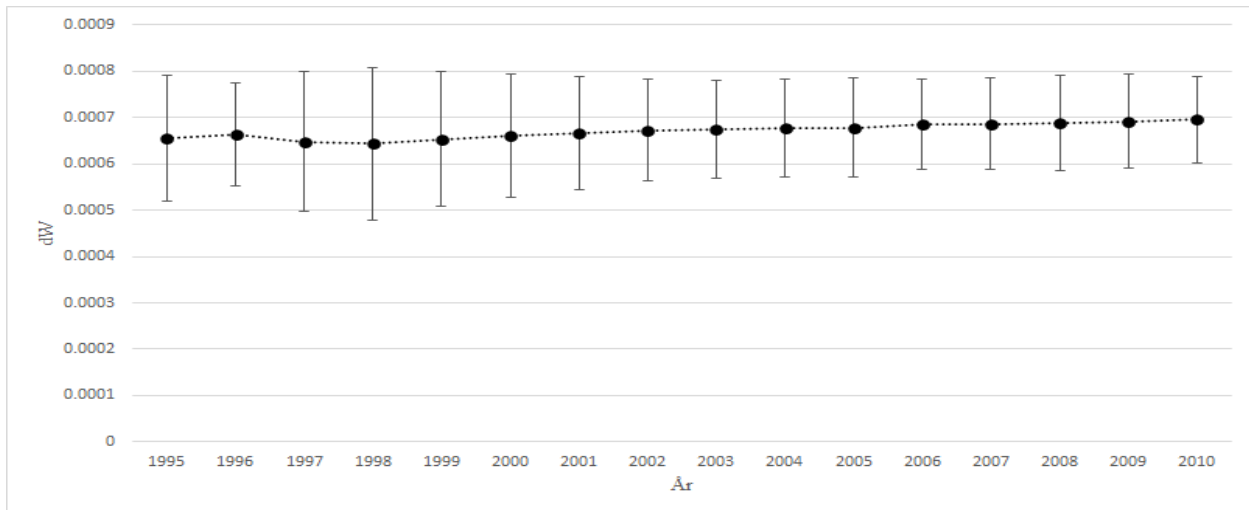


Figure 30:  $dW$  för varje enskilt år med konfidensintervall

Dessa resultat tyder på att då GDP ökar så minskar Prostate cancer death rate bland annat till följd av att de i regressionsmodellen ej ingående variablerna förändras och i sin tur påverkar prostate

cancer death rate, men att förändringar i GDP också har en motsatt effekt på Prostate cancer death rate och att det leder till att då GDP/capita ökar med 1 så går Prostate cancer death rate ner med mindre än 0.001 person per 100,000. Dummyvariablernas koefficienter  $\alpha$  (tabell 41, bilaga A) tyder på existerande systemeffekter hos länderna.

## 7. Slutsats och diskussion

Resultaten tyder på att där finns ett negativt samband mellan Prostate cancer death rate (CDR) och variablerna Total health spendings per person (THS), Gini coefficient (GC), Medical doctors per 1000 people (MD) och Percentage of male population aged 40 + (PMP). Vidare så visar regressionsmodellen att där finns ett negativt samband mellan CDR och GDP/capita (GDP). Dummyvariabeln i regressionsmodellen tyder också på att där finns oförklarade systemeffekter som gör att CDR skiljer sig mellan länder med samma nivå av GDP. Att  $\frac{dY_1}{dX_1}$  och  $dY_2$  inte är ekvivalenta ( $dW \neq 0$ ) indikerar vidare på att där finns en inneboende effekt i GDP, utöver de effekter som får CDR att sjunka då GDP ökar, som snarare verkar i motsatt riktning. Vi konstaterar också att när GDP ändras så beror storleken på  $dW$  inte enbart på hur mycket GDP ändras utan också vilken nivå GDP ligger på från början och dummyvariabeln för landet i fråga. Med andra ord så gör dummyvariabeln att förändringar i CDR då GDP ändras antingen förstärks eller försvagas. Detta betyder att länder med olika GDP kan svara likadant, och länder med samma GDP kan svara helt olika, på förändringar i GDP.

Att CDR har ett negativt samband med THS och MD (se tabell 38, bilaga A) är inte ett förvånande resultat. Här blir en intressant följdfråga vad det är inom sjukvården som påverkar CDR, vad som är mest kostnadseffektivt när det kommer till hur man ska använda pengar inom sjukvården och hur man skall använda sig av det humankapital man besitter i MD. Det är mer överraskande att CDR också uppvisar ett negativt samband med GC och PMP (se tabell 38, bilaga A). Vad gäller GC så är den intuitiva tanken att CDR bör ha ett positivt samband med denna. Jens Klein och Olaf von dem Knesebeck skriver bland annat att "... ,*survival analyses regarding the socioeconomic status (SES) suggest inequalities indicating worse prognosis for lower SES groups*" (Klein Jens, von dem Knesebeck Olaf, oktober 2015). Detta motsägelsefulla resultat kontra vad Jens Klein och Olaf von dem Knesebeck konstaterar tyder på att där är mer som påverkar när det kommer till socioekonomisk tillhörighet än inkomstfördelning. Vidare forskning på ekonomisk ojämlikhet och



hur det påverkar hälsa är ett ämne för sig för vidare forskning. Detta för att klargöra vad som sker vid ekonomisk ojämlikhet och vad som skulle kunna vara förklaringsfaktorerna till att sambandet skulle kunna se ut enligt resultaten. En förklaring skulle kunna vara att vid stor ekonomisk ojämlikhet så dominerar andra dödsorsaker över dödligheten i prostatacancer. När det kommer till att CDR skulle sjunka med PMP så skulle man kunna föra ett liknande resonemang. Man kan bland annat tänka sig att en äldre befolkning beror på en bättre befolkningshälsa som har en större effekt än att antal dödsfall ökar då fler blir äldre och hamnar i riskgrupper. Rawla Prashanth framhåller bland annat att antalet fall ökar kraftigt i världen, och kommer fortsätta att göra det, samtidigt som antalet bekräftade fall också är kopplat till längre förväntad livslängd för länder med högre åldersfördelning (Prashanth Rawla, april 2019). Man kan resonera kring hur orsakssambandet ser ut här: är det så att CDR sjunker med hög ålder på grund av någon okänd faktor som beror på PMP; eller är det så att högre PMP beror på bättre folkhälsa och sjukvård till exempel och med det följer lägre CDR. Vad som är klart är att det är ett ämne av intresse att forska vidare på. När det kommer till den skillnad  $dW$  ( $I_{dW} = (0.00066; 0.00068)$ ) påvisar med hur CDR ändras med GDP, och hur de i regressionsmodellen ej ingående variablerna (THS, MD, GC och PMP) påverkar CDR på grund av att GDP ändras tyder  $dW$  på existerande inneboende effekter vid högre GDP som får CDR att öka. Så som påpekat tidigare så kan ett mer västerländskt levnadssätt vara en bidragande faktor till detta ( Neupane Subas, Bray Freddie, Auvinen Anssi, juni 2017). Vidare framför Rawla Prashanth framför att bland annat att: en diet innehållande mycket rött kött och processat fett; hög alkoholkonsumtion; övervikt och fetma samt dåligt med fysisk aktivitet är riskfaktorer. Detta är också saker som ofta är förknippat med ett västerländskt levande eller levande för högutvecklade länder (Prashanth Rawla, april 2019). Dessa är däremot förmodligen inte de enda signifikanta faktorerna och leder till en intressant fortsatt frågeställning om vad som mer sker i högutvecklade länder som har en påverkan på CDR. Kollar man på de dummyvariabler som ingår i regressionsmodellen så indikerar de på systemskillnader mellan länder som påverkar CDR. Rawla Prashanth framhåller bland annat att etnicitet kan spela en viktig roll här och att man bland annat genom amerikanska studier visar skillnader i antal fall bland de amerikanska medborgarnas etniska grupper. Det framhålls här också att det både kan, och förmodligen gör, bero på både genetiska och socioekonomiska faktorer. Det framgår också i samma studie att PSA-tester kan vara en bidragande faktor (Prashanth Rawla, april 2019). Detta går i linje med det som framhålls i *Vårdfokus* att man ser regionala skillnader mellan CDR där ålder inte kan förklara hela

skillnaden, och att PSA-testning kan vara en del av förklaringen till detta (Nyhetsbyrån Siren, 2019-11-04). Utöver detta så framhåller Rawla Prashanth att exponering för miljöfaktorer kan spela en viktig roll. Detta innefattar till exempel insektsmedel, ogräsmedel och andra organiska föreningar (Prashanth Rawla, april 2019). Detta stödjer det beteende som regressionsmodellen förutsäger när det kommer till dummyvariablens påverkan som nämnts tidigare. Man kan konstatera att regionala skillnader spelar roll för CDR och att man har god koll på en del av förklaringsfaktorerna, men att det behövs mer forskning på området.

Sammanfattningsvis så konstaterar vi att CDR sjunker med högre GDP, och att detta bland annat beror på hur THS, MD, GC och PMP påverkar CDR när GDP ändras. Vidare så konstaterar vi också att där är en inneboende, delvis oförklarad, effekt i GDP som inte kan förklaras av hur THS, MD, GC och PMP påverkar CDR och att det existerar systemeffekter som skapar regionala skillnader mellan länder med liknande GDP-nivå, eller liknande effekter för länder med olika GDP-nivå. Slutligen konstaterar vi också att det behövs mycket mer forskning på området både för att kartlägga vilka ekonomiska faktorer som påverkar CDR och vad som förklarar deras påverkan på CDR.

## **8. Reservationer och problem**

Under denna genomförda studie så har en linjär regressionsmodell satts upp och i en del av denna processen finns det inga enskilda svar på hur man väljer sin modell. De val man här gör påverkar självfallet de slutgiltiga resultaten. Utifrån vad studien syftar till och till vilken nivå av noggrannhet som har krävts av modellen så anses den dock fullt tillräcklig. Ekvationssystemet som har satts upp i studien skall ses som ett försök att estimerade förklaringsfaktorer utifrån rådande förutsättningar. Det skall tilläggas här att detta hade gått att göra noggrannare till exempel genom att inte göra antaganden som leder till att systemet är stationärt och att varje enskild estimation av förhållandena mellan de olika variablerna hade kunnat genomföras med samma metod som den linjära regressionsmodellen satts upp med. Detta har varit ett medvetet val att inte göra på detta vis på grund av arbetets storlek och att ekvationssystemet oavsett detta har uppfyllt sitt syfte. Antal observationer som datan innehåller hade önskats vara fler. Däremot så är antalet observationer över det estimerade minsta antal observationer som behövs. Fler observationer hade förmodligen påverkat resultaten, men inte på ett sådant vis att slutsatserna hade ändrats. Fortsatt så kan valet

av variabler diskuteras. Framför allt framhålls valet av variablerna Gini coefficient och Total health spendings per person samt Medical doctors per 1000 people kunna diskuteras. Istället för Gini coefficient skulle man kunna ha valt en variabel som bättre mäter socioekonomisk ojämlikhet i stort i samhället, och inte enbart inkomstfördelning. Total health spendings per person samt Medical doctors per 1000 people anses i sig inte vara särskilt problematiska men man skulle kunna ha använt sig av fler variabler av liknande slag som också är kopplade till sjukvården för att få en bättre överblick av hur olika produktionsfaktorer påverkar sjukvården. Man hade också kunnat tänka sig att använda sig av ett antal variabler som inte bara visar hur mycket Total health spendings per person är utan också vilken form av finansiering det handlar om för att skapa sig en uppfattning om det har någon betydelse. Slutligen framhålls att man hade velat replikera denna studie för andra cancerformer för att kontrollera för liknande eller samma resultat. Detta framhålls som viktigt då det kan påvisa om resultaten är specifika för prostatacancer eller om det gäller mer generellt samt för att påvisa de skillnader som möjligtvis existerar mellan olika cancerformer. Detta har dock inte kunnat göras inom ramen för denna studie på grund av studiens storlek.

## 9. Källhänvisningar

1. Bratt Ola, PSA-prov för prostatacancer, *Cancerfonden*, 2020-01-20.  
[https://www.cancerfonden.se/om-cancer/undersokningar/psa-test?gclid=EAIaIQobChMI1uTTtoqSs6wIVhvuyCh3ncQMrEAAYASAAEgKoU\\_D\\_BwE](https://www.cancerfonden.se/om-cancer/undersokningar/psa-test?gclid=EAIaIQobChMI1uTTtoqSs6wIVhvuyCh3ncQMrEAAYASAAEgKoU_D_BwE)
2. Nunes Czorny, Rildo César, Pinto Maria Helena, Alcalá Pompeo Daniele, Bereta Denise, Ventura Cardoso Luciana, Marques da Silva Dorival, FATORES DE RISCO PARA O CÂNCER DE PRÓSTATA: POPULAÇÃO DE UMA UNIDADE BÁSICA DE SAÚDE, *Cogitare Enfermagem*, Vol. 22, 2017. <http://dx.doi.org/10.5380/ce.v22i4.51823>
3. Fullman, Nancy, et.al, Measuring performance on the Healthcare Access and Quality Index for 195 countries and territories and selected subnational locations: A systematic analysis from the Global Burden of Disease Study 2016, *The Lancet*, Vol. 391, juni 2018.  
[https://doi.org/10.1016/S0140-6736\(18\)30994-2](https://doi.org/10.1016/S0140-6736(18)30994-2)
4. Gapminder, GDP/Capita (US\$, inflation-adjusted), n.d., <https://www.gapminder.org/data/>
5. Gapminder, Gini coefficient, n.d., <https://www.gapminder.org/data/>
6. Gapminder, Medical Doctors (per 1000 people), n.d., <https://www.gapminder.org/data/>
7. Gapminder, Population aged 40-59, M (%), n.d., <https://www.gapminder.org/data/>
8. Gapminder, Population aged 60+, M (%), n.d., <https://www.gapminder.org/data/>
9. Gapminder, Prostate cancer, deaths per 100 000 men, n.d.,  
<https://www.gapminder.org/data/>
10. Gapminder, Total health spending per person (US\$), n.d.,  
<https://www.gapminder.org/data/>

11. Klein Jens, von dem Knesebeck Olaf, Socioeconomic inequalities in prostate cancer survival: A review of the evidence and explanatory factors, *Social Science & Medicine*, Vol. 142, oktober 2015. <https://doi.org/10.1016/j.socscimed.2015.07.006>
12. Neupane Subas, Bray Freddie, Auvinen Anssi, National economic and development indicators and international variation in prostate cancer incidence and mortality: an ecological analysis, *World Journal of Urology*, Vol. 35, juni 2017. <https://doi-org.ludwig.lub.lu.se/10.1007/s00345-016-1953-9>
13. Nyhetsbyrån Siren, Dödligheten i prostatacancer varierar stort över landet, *Vårdfokus*, 2019-11-04. <https://www.vardfokus.se/webbnyheter/2019/november/dodligheten-i-prostatacancer-varierar-stort-over-landet/>
14. Prashanth Rawla, Epidemiology of prostate cancer, *World Journal of Oncology*, Vol. 10, Apr 2019. <https://doi.org/10.14740/wjon1191>
15. Rawlings John O., Pantula Sastry G., Dickey David A., Applied Regression Analysis: A research tool, second edition, *Springer-Verlag New York, Inc.*, 1998, ISBN 0-387-98454-2
16. Socialdepartementet, 4,7 miljarder till hälso- och sjukvården, *Regeringskansliet*, 2019-12-18. <https://www.regeringen.se/pressmeddelanden/2019/12/47-miljarder-till-halso--och-sjukvarden/>
17. Wong Martin C.S., Goggings William B., Wang Harry H.X., Fung Franklin D.H., Leung Colette, Wong Samuel Y.S., Ng Chi Fai, Sung Joseph J.Y., Global Incidence and Mortality for Prostate Cancer: Analysis of Temporal Patterns and Trends in 36 Countries, *European Urology*, Vol. 70, November 2016. <https://doi.org/10.1016/j.eururo.2016.05.043>

## 10. Bilaga A - Tabeller

Tabell 5:  $\frac{dY_1}{dX_1}$  för varje land år 2010

Land	$\frac{dY_1}{dX_1}$	Land	$\frac{dY_1}{dX_1}$
Australia	-3.34E-04	Japan	-1.03E-03
Austria	-3.49E-04	Latvia	-3.23E-04
Belgium	-3.00E-04	Lithuania	-3.39E-04
Canada	-3.25E-04	Luxembourg	-4.17E-04
Chile	-2.94E-04	Mexico	-4.50E-04
Colombia	-3.09E-04	Netherlands	-3.11E-04
Czech Republic	-3.38E-04	New Zealand	-2.88E-04
Denmark	-2.54E-04	Norway	-2.43E-04
Estonia	-3.15E-04	Poland	-3.72E-04
Finland	-3.16E-04	Portugal	-2.90E-04
France	-2.96E-04	Slovak Republic	-4.23E-04
Germany	-3.64E-04	Slovenia	-2.70E-04
Greece	-4.17E-04	Spain	-8.31E-04
Hungary	-3.39E-04	Sweden	-3.08E-04
Iceland	-3.01E-04	Switzerland	-2.04E-04
Ireland	-2.88E-04	Turkey	-5.60E-04
Israel	-4.87E-04	United Kingdom	-3.53E-04
Italy	-4.58E-04	United States	-4.18E-04

Tabell 6: dW för varje land år 2010

Land	dW	Land	dW
Australia	6.86E-04	Japan	-1.07E-05
Austria	6.71E-04	Latvia	6.97E-04
Belgium	7.20E-04	Lithuania	6.82E-04
Canada	6.95E-04	Luxembourg	6.03E-04
Chile	7.26E-04	Mexico	5.70E-04

Colombia	7.11E-04	Netherlands	7.09E-04
Czech Republic	6.82E-04	New Zealand	7.32E-04
Denmark	7.66E-04	Norway	7.77E-04
Estonia	7.05E-04	Poland	6.48E-04
Finland	7.05E-04	Portugal	7.31E-04
France	7.24E-04	Slovak Republic	5.98E-04
Germany	6.57E-04	Slovenia	7.50E-04
Greece	6.03E-04	Spain	1.89E-04
Hungary	6.81E-04	Sweden	7.12E-04
Iceland	7.19E-04	Switzerland	8.17E-04
Ireland	7.32E-04	Turkey	4.60E-04
Israel	5.33E-04	United Kingdom	6.67E-04
Italy	5.63E-04	United States	6.02E-04

Tabell 7:  $\frac{dY_1}{dX_1}$  för varje land år 2009

Land	$\frac{dY_1}{dX_1}$	Land	$\frac{dY_1}{dX_1}$
Australia	-3.34E-04	Japan	-8.66E-04
Austria	-3.46E-04	Latvia	-3.25E-04
Belgium	-2.98E-04	Lithuania	-3.37E-04
Canada	-3.22E-04	Luxembourg	-3.97E-04
Chile	-2.93E-04	Mexico	-4.48E-04
Colombia	-3.09E-04	Netherlands	-3.10E-04
Czech Republic	-3.37E-04	New Zealand	-2.89E-04
Denmark	-2.53E-04	Norway	-2.45E-04
Estonia	-3.15E-04	Poland	-3.70E-04
Finland	-3.12E-04	Portugal	-2.89E-04
France	-2.95E-04	Slovak Republic	-4.16E-04
Germany	-3.56E-04	Slovenia	-2.70E-04
Greece	-4.32E-04	Spain	-3.97E-04

Hungary	-3.39E-04	Sweden	-2.52E-04
Iceland	-3.01E-04	Switzerland	-3.06E-04
Ireland	-2.87E-04	Turkey	-5.53E-04
Israel	-4.75E-04	United Kingdom	-3.15E-04
Italy	-4.53E-04	United States	-4.12E-04

Tabell 8: dW för varje land år 2009

Land	dW	Land	dW
Australia	6.86E-04	Japan	1.55E-04
Austria	6.74E-04	Latvia	6.95E-04
Belgium	7.22E-04	Lithuania	6.83E-04
Canada	6.98E-04	Luxembourg	6.24E-04
Chile	7.27E-04	Mexico	5.73E-04
Colombia	7.11E-04	Netherlands	7.10E-04
Czech Republic	6.83E-04	New Zealand	7.32E-04
Denmark	7.67E-04	Norway	7.76E-04
Estonia	7.06E-04	Poland	6.50E-04
Finland	7.08E-04	Portugal	7.31E-04
France	7.25E-04	Slovak Republic	6.04E-04
Germany	6.64E-04	Slovenia	7.50E-04
Greece	5.88E-04	Spain	6.23E-04
Hungary	6.81E-04	Sweden	7.68E-04
Iceland	7.19E-04	Switzerland	7.14E-04
Ireland	7.33E-04	Turkey	4.67E-04
Israel	5.45E-04	United Kingdom	7.05E-04
Italy	5.67E-04	United States	6.08E-04

Tabell 9:  $\frac{dY_1}{dX_1}$  för varje land år 2008

Land	$\frac{dY_1}{dX_1}$	Land	$\frac{dY_1}{dX_1}$



Australia	-3.34E-04	Japan	-1.14E-03
Austria	-3.56E-04	Latvia	-3.31E-04
Belgium	-3.02E-04	Lithuania	-3.45E-04
Canada	-3.29E-04	Luxembourg	-4.56E-04
Chile	-2.94E-04	Mexico	-4.54E-04
Colombia	-3.09E-04	Netherlands	-3.17E-04
Czech Republic	-3.41E-04	New Zealand	-2.90E-04
Denmark	-2.59E-04	Norway	-2.49E-04
Estonia	-3.23E-04	Poland	-3.68E-04
Finland	-3.27E-04	Portugal	-2.91E-04
France	-2.99E-04	Slovak Republic	-4.24E-04
Germany	-3.68E-04	Slovenia	-2.74E-04
Greece	-4.45E-04	Spain	-4.08E-04
Hungary	-3.43E-04	Sweden	-2.57E-04
Iceland	-3.05E-04	Switzerland	-3.15E-04
Ireland	-2.96E-04	Turkey	-5.65E-04
Israel	-4.80E-04	United Kingdom	-3.22E-04
Italy	-4.78E-04	United States	-4.26E-04

Tabell 10: dW för varje land år 2008

Land	dW	Land	dW
Australia	6.86E-04	Japan	-1.22E-04
Austria	6.65E-04	Latvia	6.89E-04
Belgium	7.18E-04	Lithuania	6.75E-04
Canada	6.91E-04	Luxembourg	5.65E-04
Chile	7.26E-04	Mexico	5.66E-04
Colombia	7.11E-04	Netherlands	7.03E-04
Czech Republic	6.79E-04	New Zealand	7.30E-04
Denmark	7.61E-04	Norway	7.71E-04
Estonia	6.97E-04	Poland	6.52E-04

Finland	6.93E-04	Portugal	7.29E-04
France	7.21E-04	Slovak Republic	5.96E-04
Germany	6.52E-04	Slovenia	7.46E-04
Greece	5.75E-04	Spain	6.12E-04
Hungary	6.77E-04	Sweden	7.63E-04
Iceland	7.15E-04	Switzerland	7.06E-04
Ireland	7.25E-04	Turkey	4.55E-04
Israel	5.40E-04	United Kingdom	6.98E-04
Italy	5.42E-04	United States	5.94E-04

Tabell 11:  $\frac{dY_1}{dX_1}$  för varje land år 2007

Land	$\frac{dY_1}{dX_1}$	Land	$\frac{dY_1}{dX_1}$
Australia	-3.31E-04	Japan	-1.24E-03
Austria	-3.53E-04	Latvia	-3.33E-04
Belgium	-3.02E-04	Lithuania	-3.43E-04
Canada	-3.29E-04	Luxembourg	-4.90E-04
Chile	-2.93E-04	Mexico	-4.54E-04
Colombia	-3.09E-04	Netherlands	-3.14E-04
Czech Republic	-3.40E-04	New Zealand	-2.91E-04
Denmark	-2.61E-04	Norway	-2.50E-04
Estonia	-3.26E-04	Poland	-3.66E-04
Finland	-3.26E-04	Portugal	-2.91E-04
France	-2.99E-04	Slovak Republic	-4.17E-04
Germany	-3.65E-04	Slovenia	-2.73E-04
Greece	-4.47E-04	Spain	-4.09E-04
Hungary	-3.42E-04	Sweden	-2.58E-04
Iceland	-3.11E-04	Switzerland	-3.12E-04
Ireland	-3.06E-04	Turkey	-5.65E-04
Israel	-4.76E-04	United Kingdom	-3.24E-04

Italy	-4.85E-04	United States	-4.31E-04
-------	-----------	---------------	-----------

Tabell 12: dW för varje land år 2007

Land	dW	Land	dW
Australia	6.89E-04	Japan	-2.18E-04
Austria	6.68E-04	Latvia	6.87E-04
Belgium	7.18E-04	Lithuania	6.77E-04
Canada	6.91E-04	Luxembourg	5.30E-04
Chile	7.27E-04	Mexico	5.66E-04
Colombia	7.11E-04	Netherlands	7.06E-04
Czech Republic	6.80E-04	New Zealand	7.29E-04
Denmark	7.60E-04	Norway	7.70E-04
Estonia	6.94E-04	Poland	6.55E-04
Finland	6.94E-04	Portugal	7.29E-04
France	7.21E-04	Slovak Republic	6.04E-04
Germany	6.55E-04	Slovenia	7.48E-04
Greece	5.73E-04	Spain	6.11E-04
Hungary	6.78E-04	Sweden	7.62E-04
Iceland	7.09E-04	Switzerland	7.08E-04
Ireland	7.14E-04	Turkey	4.55E-04
Israel	5.44E-04	United Kingdom	6.96E-04
Italy	5.35E-04	United States	5.89E-04

Tabell 13:  $\frac{dY_1}{dX_1}$  för varje land år 2006

Land	$\frac{dY_1}{dX_1}$	Land	$\frac{dY_1}{dX_1}$
Australia	-3.25E-04	Japan	-1.11E-03
Austria	-3.46E-04	Latvia	-3.27E-04
Belgium	-2.98E-04	Lithuania	-3.37E-04
Canada	-3.20E-04	Luxembourg	-4.19E-04

Chile	-2.92E-04	Mexico	-4.54E-04
Colombia	-3.08E-04	Netherlands	-3.09E-04
Czech Republic	-3.36E-04	New Zealand	-2.89E-04
Denmark	-2.60E-04	Norway	-2.47E-04
Estonia	-3.21E-04	Poland	-3.62E-04
Finland	-3.18E-04	Portugal	-2.89E-04
France	-2.97E-04	Slovak Republic	-4.05E-04
Germany	-3.59E-04	Slovenia	-2.69E-04
Greece	-4.38E-04	Spain	-4.05E-04
Hungary	-3.42E-04	Sweden	-2.55E-04
Iceland	-3.07E-04	Switzerland	-3.05E-04
Ireland	-3.02E-04	Turkey	-5.59E-04
Israel	-4.64E-04	United Kingdom	-3.22E-04
Italy	-4.80E-04	United States	-4.27E-04

Tabell 14: dW för varje land år 2006

Land	dW	Land	dW
Australia	6.95E-04	Japan	-8.91E-05
Austria	6.74E-04	Latvia	6.93E-04
Belgium	7.22E-04	Lithuania	6.83E-04
Canada	7.01E-04	Luxembourg	6.01E-04
Chile	7.28E-04	Mexico	5.66E-04
Colombia	7.12E-04	Netherlands	7.12E-04
Czech Republic	6.84E-04	New Zealand	7.31E-04
Denmark	7.60E-04	Norway	7.73E-04
Estonia	6.99E-04	Poland	6.58E-04
Finland	7.02E-04	Portugal	7.31E-04
France	7.23E-04	Slovak Republic	6.15E-04
Germany	6.61E-04	Slovenia	7.51E-04
Greece	5.82E-04	Spain	6.15E-04

Hungary	6.78E-04	Sweden	7.65E-04
Iceland	7.14E-04	Switzerland	7.15E-04
Ireland	7.18E-04	Turkey	4.61E-04
Israel	5.56E-04	United Kingdom	6.99E-04
Italy	5.40E-04	United States	5.93E-04

Tabell 15:  $\frac{dY_1}{dX_1}$  för varje land år 2005

Land	$\frac{dY_1}{dX_1}$	Land	$\frac{dY_1}{dX_1}$
Australia	-3.23E-04	Japan	-1.03E-03
Austria	-3.40E-04	Latvia	-3.22E-04
Belgium	-2.96E-04	Lithuania	-3.33E-04
Canada	-3.15E-04	Luxembourg	-3.90E-04
Chile	-2.90E-04	Mexico	-4.51E-04
Colombia	-3.07E-04	Netherlands	-3.04E-04
Czech Republic	-3.31E-04	New Zealand	-2.88E-04
Denmark	-2.56E-04	Norway	-2.45E-04
Estonia	-3.16E-04	Poland	-3.59E-04
Finland	-3.13E-04	Portugal	-2.89E-04
France	-2.95E-04	Slovak Republic	-3.97E-04
Germany	-3.51E-04	Slovenia	-2.67E-04
Greece	-4.25E-04	Spain	-4.00E-04
Hungary	-3.39E-04	Sweden	-2.52E-04
Iceland	-3.02E-04	Switzerland	-2.98E-04
Ireland	-2.99E-04	Turkey	-5.48E-04
Israel	-4.52E-04	United Kingdom	-3.19E-04
Italy	-4.73E-04	United States	-4.20E-04

Tabell 16: dW för varje land år 2005

Land	dW	Land	dW
------	----	------	----

Australia	6.97E-04	Japan	-6.10E-06
Austria	6.81E-04	Latvia	6.98E-04
Belgium	7.24E-04	Lithuania	6.87E-04
Canada	7.05E-04	Luxembourg	6.30E-04
Chile	7.30E-04	Mexico	5.69E-04
Colombia	7.13E-04	Netherlands	7.17E-04
Czech Republic	6.89E-04	New Zealand	7.32E-04
Denmark	7.64E-04	Norway	7.75E-04
Estonia	7.04E-04	Poland	6.62E-04
Finland	7.08E-04	Portugal	7.31E-04
France	7.25E-04	Slovak Republic	6.23E-04
Germany	6.69E-04	Slovenia	7.53E-04
Greece	5.95E-04	Spain	6.20E-04
Hungary	6.81E-04	Sweden	7.68E-04
Iceland	7.18E-04	Switzerland	7.22E-04
Ireland	7.21E-04	Turkey	4.72E-04
Israel	5.68E-04	United Kingdom	7.01E-04
Italy	5.47E-04	United States	6.01E-04

Tabell 17:  $\frac{dY_1}{dX_1}$  för varje land år 2004

Land	$\frac{dY_1}{dX_1}$	Land	$\frac{dY_1}{dX_1}$
Australia	-3.19E-04	Japan	-9.50E-04
Austria	-3.37E-04	Latvia	-3.18E-04
Belgium	-2.94E-04	Lithuania	-3.29E-04
Canada	-3.09E-04	Luxembourg	-3.82E-04
Chile	-2.89E-04	Mexico	-4.50E-04
Colombia	-3.06E-04	Netherlands	-3.01E-04
Czech Republic	-3.27E-04	New Zealand	-2.86E-04
Denmark	-2.54E-04	Norway	-2.42E-04

Estonia	-3.11E-04	Poland	-3.57E-04
Finland	-3.09E-04	Portugal	-2.88E-04
France	-2.94E-04	Slovak Republic	-3.91E-04
Germany	-3.50E-04	Slovenia	-2.65E-04
Greece	-4.24E-04	Spain	-3.96E-04
Hungary	-3.37E-04	Sweden	-2.50E-04
Iceland	-3.00E-04	Switzerland	-2.93E-04
Ireland	-2.94E-04	Turkey	-5.36E-04
Israel	-4.47E-04	United Kingdom	-3.15E-04
Italy	-4.72E-04	United States	-4.10E-04

Tabell 18: dW för varje land år 2004

Land	dW	Land	dW
Australia	7.01E-04	Japan	7.04E-05
Austria	6.83E-04	Latvia	7.02E-04
Belgium	7.26E-04	Lithuania	6.91E-04
Canada	7.11E-04	Luxembourg	6.38E-04
Chile	7.31E-04	Mexico	5.70E-04
Colombia	7.14E-04	Netherlands	7.19E-04
Czech Republic	6.93E-04	New Zealand	7.34E-04
Denmark	7.66E-04	Norway	7.78E-04
Estonia	7.09E-04	Poland	6.63E-04
Finland	7.11E-04	Portugal	7.32E-04
France	7.26E-04	Slovak Republic	6.29E-04
Germany	6.71E-04	Slovenia	7.55E-04
Greece	5.96E-04	Spain	6.25E-04
Hungary	6.83E-04	Sweden	7.70E-04
Iceland	7.20E-04	Switzerland	7.27E-04
Ireland	7.27E-04	Turkey	4.84E-04
Israel	5.73E-04	United Kingdom	7.05E-04

Italy	5.48E-04	United States	6.10E-04
-------	----------	---------------	----------

Tabell 19:  $\frac{dY_1}{dX_1}$  för varje land år 2003

Land	$\frac{dY_1}{dX_1}$	Land	$\frac{dY_1}{dX_1}$
Australia	-3.15E-04	Japan	-8.66E-04
Austria	-3.33E-04	Latvia	-3.15E-04
Belgium	-2.91E-04	Lithuania	-3.27E-04
Canada	-3.05E-04	Luxembourg	-3.70E-04
Chile	-2.88E-04	Mexico	-4.48E-04
Colombia	-3.06E-04	Netherlands	-2.99E-04
Czech Republic	-3.24E-04	New Zealand	-2.84E-04
Denmark	-2.52E-04	Norway	-2.38E-04
Estonia	-3.08E-04	Poland	-3.55E-04
Finland	-3.04E-04	Portugal	-2.87E-04
France	-2.92E-04	Slovak Republic	-3.88E-04
Germany	-3.48E-04	Slovenia	-2.64E-04
Greece	-4.14E-04	Spain	-3.93E-04
Hungary	-3.35E-04	Sweden	-2.47E-04
Iceland	-2.95E-04	Switzerland	-2.90E-04
Ireland	-2.88E-04	Turkey	-5.25E-04
Israel	-4.41E-04	United Kingdom	-3.13E-04
Italy	-4.68E-04	United States	-4.01E-04

Tabell 20: dW för varje land år 2003

Land	dW	Land	dW
Australia	7.05E-04	Japan	1.55E-04
Austria	6.87E-04	Latvia	7.05E-04
Belgium	7.30E-04	Lithuania	6.93E-04
Canada	7.15E-04	Luxembourg	6.50E-04



Chile	7.33E-04	Mexico	5.72E-04
Colombia	7.15E-04	Netherlands	7.21E-04
Czech Republic	6.96E-04	New Zealand	7.36E-04
Denmark	7.68E-04	Norway	7.82E-04
Estonia	7.12E-04	Poland	6.66E-04
Finland	7.16E-04	Portugal	7.33E-04
France	7.28E-04	Slovak Republic	6.33E-04
Germany	6.72E-04	Slovenia	7.57E-04
Greece	6.07E-04	Spain	6.27E-04
Hungary	6.85E-04	Sweden	7.74E-04
Iceland	7.25E-04	Switzerland	7.31E-04
Ireland	7.33E-04	Turkey	4.95E-04
Israel	5.79E-04	United Kingdom	7.07E-04
Italy	5.52E-04	United States	6.19E-04

Tabell 21:  $\frac{dY_1}{dX_1}$  för varje land år 2002

Land	$\frac{dY_1}{dX_1}$	Land	$\frac{dY_1}{dX_1}$
Australia	-3.12E-04	Japan	-8.31E-04
Austria	-3.33E-04	Latvia	-3.13E-04
Belgium	-2.90E-04	Lithuania	-3.23E-04
Canada	-3.02E-04	Luxembourg	-3.68E-04
Chile	-2.87E-04	Mexico	-4.48E-04
Colombia	-3.05E-04	Netherlands	-2.99E-04
Czech Republic	-3.22E-04	New Zealand	-2.82E-04
Denmark	-2.52E-04	Norway	-2.38E-04
Estonia	-3.05E-04	Poland	-3.53E-04
Finland	-3.02E-04	Portugal	-2.88E-04
France	-2.92E-04	Slovak Republic	-3.84E-04
Germany	-3.49E-04	Slovenia	-2.62E-04

Greece	-4.04E-04	Spain	-3.91E-04
Hungary	-3.33E-04	Sweden	-2.45E-04
Iceland	-2.92E-04	Switzerland	-2.91E-04
Ireland	-2.86E-04	Turkey	-5.20E-04
Israel	-4.43E-04	United Kingdom	-3.10E-04
Italy	-4.71E-04	United States	-3.95E-04

Tabell 22: dW för varje land år 2002

Land	dW	Land	dW
Australia	7.08E-04	Japan	1.89E-04
Austria	6.88E-04	Latvia	7.08E-04
Belgium	7.30E-04	Lithuania	6.97E-04
Canada	7.18E-04	Luxembourg	6.52E-04
Chile	7.33E-04	Mexico	5.72E-04
Colombia	7.15E-04	Netherlands	7.21E-04
Czech Republic	6.98E-04	New Zealand	7.38E-04
Denmark	7.69E-04	Norway	7.82E-04
Estonia	7.15E-04	Poland	6.67E-04
Finland	7.18E-04	Portugal	7.32E-04
France	7.29E-04	Slovak Republic	6.36E-04
Germany	6.71E-04	Slovenia	7.58E-04
Greece	6.16E-04	Spain	6.29E-04
Hungary	6.87E-04	Sweden	7.75E-04
Iceland	7.29E-04	Switzerland	7.29E-04
Ireland	7.34E-04	Turkey	5.01E-04
Israel	5.77E-04	United Kingdom	7.10E-04
Italy	5.50E-04	United States	6.26E-04

Tabell 23:  $\frac{dY_1}{dX_1}$  för varje land år 2001

Land	$\frac{dY_1}{dX_1}$	Land	$\frac{dY_1}{dX_1}$
Australia	-3.08E-04	Japan	-8.31E-04
Austria	-3.31E-04	Latvia	-3.11E-04
Belgium	-2.89E-04	Lithuania	-3.21E-04
Canada	-2.99E-04	Luxembourg	-3.55E-04
Chile	-2.86E-04	Mexico	-4.49E-04
Colombia	-3.05E-04	Netherlands	-3.00E-04
Czech Republic	-3.21E-04	New Zealand	-2.80E-04
Denmark	-2.51E-04	Norway	-2.37E-04
Estonia	-3.03E-04	Poland	-3.52E-04
Finland	-3.00E-04	Portugal	-2.88E-04
France	-2.91E-04	Slovak Republic	-3.81E-04
Germany	-3.50E-04	Slovenia	-2.61E-04
Greece	-3.98E-04	Spain	-3.88E-04
Hungary	-3.30E-04	Sweden	-2.44E-04
Iceland	-2.88E-04	Switzerland	-2.92E-04
Ireland	-2.81E-04	Turkey	-5.14E-04
Israel	-4.48E-04	United Kingdom	-3.07E-04
Italy	-4.70E-04	United States	-3.92E-04

Tabell 24: dW för varje land år 2001

Land	dW	Land	dW
Australia	7.12E-04	Japan	1.89E-04
Austria	6.90E-04	Latvia	7.10E-04
Belgium	7.31E-04	Lithuania	6.99E-04
Canada	7.21E-04	Luxembourg	6.65E-04
Chile	7.34E-04	Mexico	5.71E-04
Colombia	7.15E-04	Netherlands	7.20E-04
Czech Republic	6.99E-04	New Zealand	7.40E-04

Denmark	7.69E-04	Norway	7.83E-04
Estonia	7.17E-04	Poland	6.68E-04
Finland	7.20E-04	Portugal	7.32E-04
France	7.29E-04	Slovak Republic	6.39E-04
Germany	6.71E-04	Slovenia	7.59E-04
Greece	6.23E-04	Spain	6.32E-04
Hungary	6.90E-04	Sweden	7.76E-04
Iceland	7.32E-04	Switzerland	7.28E-04
Ireland	7.39E-04	Turkey	5.06E-04
Israel	5.72E-04	United Kingdom	7.13E-04
Italy	5.51E-04	United States	6.28E-04

Tabell 25:  $\frac{dY_1}{dX_1}$  för varje land år 2000

Land	$\frac{dY_1}{dX_1}$	Land	$\frac{dY_1}{dX_1}$
Australia	-3.07E-04	Japan	-8.31E-04
Austria	-3.29E-04	Latvia	-3.09E-04
Belgium	-2.88E-04	Lithuania	-3.20E-04
Canada	-2.99E-04	Luxembourg	-3.49E-04
Chile	-2.86E-04	Mexico	-4.51E-04
Colombia	-3.05E-04	Netherlands	-2.98E-04
Czech Republic	-3.19E-04	New Zealand	-2.78E-04
Denmark	-2.51E-04	Norway	-2.35E-04
Estonia	-3.01E-04	Poland	-3.52E-04
Finland	-2.97E-04	Portugal	-2.87E-04
France	-2.90E-04	Slovak Republic	-3.78E-04
Germany	-3.47E-04	Slovenia	-2.60E-04
Greece	-3.92E-04	Spain	-3.83E-04
Hungary	-3.29E-04	Sweden	-2.43E-04
Iceland	-2.86E-04	Switzerland	-2.91E-04

Ireland	-2.78E-04	Turkey	-5.23E-04
Israel	-4.54E-04	United Kingdom	-3.05E-04
Italy	-4.62E-04	United States	-3.92E-04

Tabell 26: dW för varje land år 2000

Land	dW	Land	dW
Australia	1.89E-04	Japan	7.13E-04
Austria	7.11E-04	Latvia	6.91E-04
Belgium	7.01E-04	Lithuania	7.32E-04
Canada	6.71E-04	Luxembourg	7.22E-04
Chile	5.70E-04	Mexico	7.34E-04
Colombia	7.23E-04	Netherlands	7.15E-04
Czech Republic	7.42E-04	New Zealand	7.01E-04
Denmark	7.85E-04	Norway	7.69E-04
Estonia	6.69E-04	Poland	7.19E-04
Finland	7.33E-04	Portugal	7.23E-04
France	6.42E-04	Slovak Republic	7.30E-04
Germany	7.60E-04	Slovenia	6.73E-04
Greece	6.37E-04	Spain	6.28E-04
Hungary	7.77E-04	Sweden	6.91E-04
Iceland	7.30E-04	Switzerland	7.34E-04
Ireland	4.97E-04	Turkey	7.43E-04
Israel	7.15E-04	United Kingdom	5.66E-04
Italy	6.28E-04	United States	5.58E-04

Tabell 27:  $\frac{dY_1}{dX_1}$  för varje land år 1999

Land	$\frac{dY_1}{dX_1}$	Land	$\frac{dY_1}{dX_1}$
Australia	-3.04E-04	Japan	-7.68E-04
Austria	-3.24E-04	Latvia	-3.07E-04

Belgium	-2.85E-04	Lithuania	-3.18E-04
Canada	-2.94E-04	Luxembourg	-3.24E-04
Chile	-2.85E-04	Mexico	-4.48E-04
Colombia	-3.05E-04	Netherlands	-2.93E-04
Czech Republic	-3.17E-04	New Zealand	-2.77E-04
Denmark	-2.48E-04	Norway	-2.32E-04
Estonia	-2.98E-04	Poland	-3.49E-04
Finland	-2.91E-04	Portugal	-2.86E-04
France	-2.87E-04	Slovak Republic	-3.78E-04
Germany	-3.42E-04	Slovenia	-2.59E-04
Greece	-3.87E-04	Spain	-3.75E-04
Hungary	-3.27E-04	Sweden	-2.40E-04
Iceland	-2.84E-04	Switzerland	-2.85E-04
Ireland	-2.70E-04	Turkey	-5.17E-04
Israel	-4.42E-04	United Kingdom	-3.01E-04
Italy	-4.48E-04	United States	-3.84E-04

Tabell 28: dW för varje land år 1999

Land	dW	Land	dW
Australia	7.16E-04	Japan	2.52E-04
Austria	6.96E-04	Latvia	7.13E-04
Belgium	7.36E-04	Lithuania	7.02E-04
Canada	7.26E-04	Luxembourg	6.96E-04
Chile	7.35E-04	Mexico	5.73E-04
Colombia	7.15E-04	Netherlands	7.27E-04
Czech Republic	7.03E-04	New Zealand	7.43E-04
Denmark	7.72E-04	Norway	7.88E-04
Estonia	7.22E-04	Poland	6.71E-04
Finland	7.29E-04	Portugal	7.34E-04
France	7.33E-04	Slovak Republic	6.42E-04

Germany	6.78E-04	Slovenia	7.61E-04
Greece	6.33E-04	Spain	6.45E-04
Hungary	6.93E-04	Sweden	7.80E-04
Iceland	7.36E-04	Switzerland	7.35E-04
Ireland	7.50E-04	Turkey	5.03E-04
Israel	5.78E-04	United Kingdom	7.19E-04
Italy	5.72E-04	United States	6.37E-04

Tabell 29:  $\frac{dY_1}{dX_1}$  för varje land år 1998

Land	$\frac{dY_1}{dX_1}$	Land	$\frac{dY_1}{dX_1}$
Australia	-2.99E-04	Japan	-7.78E-04
Austria	-3.19E-04	Latvia	-3.07E-04
Belgium	-2.82E-04	Lithuania	-3.19E-04
Canada	-2.91E-04	Luxembourg	-3.05E-04
Chile	-2.85E-04	Mexico	-4.47E-04
Colombia	-3.06E-04	Netherlands	-2.88E-04
Czech Republic	-3.16E-04	New Zealand	-2.74E-04
Denmark	-2.46E-04	Norway	-2.31E-04
Estonia	-2.99E-04	Poland	-3.48E-04
Finland	-2.88E-04	Portugal	-2.84E-04
France	-2.84E-04	Slovak Republic	-3.78E-04
Germany	-3.40E-04	Slovenia	-2.57E-04
Greece	-3.83E-04	Spain	-3.69E-04
Hungary	-3.26E-04	Sweden	-2.38E-04
Iceland	-3.03E-04	Switzerland	-2.83E-04
Ireland	-2.63E-04	Turkey	-5.23E-04
Israel	-4.41E-04	United Kingdom	-2.98E-04
Italy	-4.43E-04	United States	-3.75E-04

Tabell 30: dW för varje land år 1998

Land	dW	Land	dW
Australia	7.21E-04	Japan	2.42E-04
Austria	7.01E-04	Latvia	7.13E-04
Belgium	7.39E-04	Lithuania	7.02E-04
Canada	7.29E-04	Luxembourg	7.15E-04
Chile	7.35E-04	Mexico	5.74E-04
Colombia	7.14E-04	Netherlands	7.32E-04
Czech Republic	7.04E-04	New Zealand	7.46E-04
Denmark	7.75E-04	Norway	7.89E-04
Estonia	7.22E-04	Poland	6.72E-04
Finland	7.33E-04	Portugal	7.36E-04
France	7.36E-04	Slovak Republic	6.42E-04
Germany	6.80E-04	Slovenia	7.63E-04
Greece	6.37E-04	Spain	6.51E-04
Hungary	6.94E-04	Sweden	7.82E-04
Iceland	7.17E-04	Switzerland	7.37E-04
Ireland	7.57E-04	Turkey	4.97E-04
Israel	5.79E-04	United Kingdom	7.22E-04
Italy	5.77E-04	United States	6.46E-04

Tabell 31:  $\frac{dY_1}{dX_1}$  för varje land år 1997

Land	$\frac{dY_1}{dX_1}$	Land	$\frac{dY_1}{dX_1}$
Australia	-2.95E-04	Japan	-8.12E-04
Austria	-3.15E-04	Latvia	-3.05E-04
Belgium	-2.80E-04	Lithuania	-3.17E-04
Canada	-2.89E-04	Luxembourg	-2.94E-04
Chile	-2.85E-04	Mexico	-4.44E-04
Colombia	-3.06E-04	Netherlands	-2.84E-04



Czech Republic	-3.16E-04	New Zealand	-2.74E-04
Denmark	-2.44E-04	Norway	-2.29E-04
Estonia	-2.97E-04	Poland	-3.46E-04
Finland	-2.83E-04	Portugal	-2.82E-04
France	-2.82E-04	Slovak Republic	-3.76E-04
Germany	-3.36E-04	Slovenia	-2.56E-04
Greece	-3.79E-04	Spain	-3.64E-04
Hungary	-3.24E-04	Sweden	-2.35E-04
Iceland	-3.01E-04	Switzerland	-2.79E-04
Ireland	-2.58E-04	Turkey	-5.22E-04
Israel	-4.38E-04	United Kingdom	-2.95E-04
Italy	-4.38E-04	United States	-3.67E-04

Tabell 32: dW för varje land år 1997

Land	dW	Land	dW
Australia	7.25E-04	Japan	2.08E-04
Austria	7.05E-04	Latvia	7.15E-04
Belgium	7.40E-04	Lithuania	7.03E-04
Canada	7.31E-04	Luxembourg	7.26E-04
Chile	7.36E-04	Mexico	5.77E-04
Colombia	7.14E-04	Netherlands	7.36E-04
Czech Republic	7.04E-04	New Zealand	7.46E-04
Denmark	7.76E-04	Norway	7.91E-04
Estonia	7.23E-04	Poland	6.74E-04
Finland	7.37E-04	Portugal	7.38E-04
France	7.38E-04	Slovak Republic	6.45E-04
Germany	6.84E-04	Slovenia	7.64E-04
Greece	6.41E-04	Spain	6.56E-04
Hungary	6.96E-04	Sweden	7.85E-04
Iceland	7.20E-04	Switzerland	7.41E-04

Ireland	7.62E-04	Turkey	4.98E-04
Israel	5.82E-04	United Kingdom	7.25E-04
Italy	5.82E-04	United States	6.53E-04

Tabell 33:  $\frac{dY_1}{dX_1}$  för varje land år 1996

Land	$\frac{dY_1}{dX_1}$	Land	$\frac{dY_1}{dX_1}$
Australia	-2.92E-04	Japan	-7.89E-04
Austria	-3.12E-04	Latvia	-3.04E-04
Belgium	-2.77E-04	Lithuania	-3.15E-04
Canada	-2.87E-04	Luxembourg	-2.86E-04
Chile	-2.83E-04	Mexico	-4.40E-04
Colombia	-3.06E-04	Netherlands	-2.81E-04
Czech Republic	-3.16E-04	New Zealand	-2.73E-04
Denmark	-2.42E-04	Norway	-2.25E-04
Estonia	-2.94E-04	Poland	-3.44E-04
Finland	-2.78E-04	Portugal	-2.80E-04
France	-2.80E-04	Slovak Republic	-3.73E-04
Germany	-3.34E-04	Slovenia	-2.55E-04
Greece	-3.74E-04	Spain	-3.60E-04
Hungary	-3.23E-04	Sweden	-2.34E-04
Iceland	-2.95E-04	Switzerland	-2.76E-04
Ireland	-2.53E-04	Turkey	-5.15E-04
Israel	-4.34E-04	United Kingdom	-2.92E-04
Italy	-4.32E-04	United States	-3.61E-04

Tabell 34: dW för varje land år 1996

Land	dW	Land	dW
Australia	7.28E-04	Japan	2.31E-04
Austria	7.08E-04	Latvia	7.16E-04

Belgium	7.43E-04	Lithuania	7.05E-04
Canada	7.34E-04	Luxembourg	7.35E-04
Chile	7.37E-04	Mexico	5.81E-04
Colombia	7.14E-04	Netherlands	7.40E-04
Czech Republic	7.04E-04	New Zealand	7.47E-04
Denmark	7.78E-04	Norway	7.96E-04
Estonia	7.26E-04	Poland	6.76E-04
Finland	7.42E-04	Portugal	7.40E-04
France	7.40E-04	Slovak Republic	6.48E-04
Germany	6.86E-04	Slovenia	7.65E-04
Greece	6.46E-04	Spain	6.60E-04
Hungary	6.97E-04	Sweden	7.87E-04
Iceland	7.25E-04	Switzerland	7.44E-04
Ireland	7.68E-04	Turkey	5.05E-04
Israel	5.86E-04	United Kingdom	7.28E-04
Italy	5.88E-04	United States	6.59E-04

Tabell 35:  $\frac{dY_1}{dX_1}$  för varje land år 1995

Land	$\frac{dY_1}{dX_1}$	Land	$\frac{dY_1}{dX_1}$
Australia	-2.90E-04	Japan	-7.35E-04
Austria	-3.10E-04	Latvia	-3.03E-04
Belgium	-2.76E-04	Lithuania	-3.14E-04
Canada	-2.86E-04	Luxembourg	-2.86E-04
Chile	-2.82E-04	Mexico	-4.36E-04
Colombia	-3.06E-04	Netherlands	-2.78E-04
Czech Republic	-3.15E-04	New Zealand	-2.72E-04
Denmark	-2.40E-04	Norway	-2.21E-04
Estonia	-2.93E-04	Poland	-3.43E-04

Finland	-2.76E-04	Portugal	-2.79E-04
France	-2.79E-04	Slovak Republic	-3.69E-04
Germany	-3.34E-04	Slovenia	-2.54E-04
Greece	-3.72E-04	Spain	-3.57E-04
Hungary	-3.23E-04	Sweden	-2.33E-04
Iceland	-2.99E-04	Switzerland	-2.76E-04
Ireland	-2.49E-04	Turkey	-5.09E-04
Israel	-4.29E-04	United Kingdom	-2.90E-04
Italy	-4.29E-04	United States	-3.56E-04

Tabell 36: dW för varje land år 1995

Land	dW	Land	dW
Australia	7.31E-04	Japan	2.86E-04
Austria	7.11E-04	Latvia	7.17E-04
Belgium	7.44E-04	Lithuania	7.06E-04
Canada	7.34E-04	Luxembourg	7.35E-04
Chile	7.38E-04	Mexico	5.84E-04
Colombia	7.14E-04	Netherlands	7.42E-04
Czech Republic	7.05E-04	New Zealand	7.48E-04
Denmark	7.80E-04	Norway	7.99E-04
Estonia	7.27E-04	Poland	6.78E-04
Finland	7.44E-04	Portugal	7.41E-04
France	7.41E-04	Slovak Republic	6.51E-04
Germany	6.87E-04	Slovenia	7.66E-04
Greece	6.49E-04	Spain	6.63E-04
Hungary	6.97E-04	Sweden	7.87E-04
Iceland	7.22E-04	Switzerland	7.44E-04
Ireland	7.71E-04	Turkey	5.11E-04
Israel	5.91E-04	United Kingdom	7.30E-04
Italy	5.91E-04	United States	6.64E-04

Tabell 37: Förändringskoefficienter för hur  $X_j$  beror på  $X_i$

Förändringskoefficienter	GDP/capita	Health spendings per person	Medica doctors per 1000 people	Gini coefficient	Percentage of male population aged 40 +
$\beta_{i,2}$	0.1984	0	398.72	113.55	306.03
$\beta_{i,3}$	0.00003978	0.0001575	0	-0.002361	0.069886
$\beta_{i,4}$	0.00007081	0.0001883	-0.009914	0	0.07222
$\beta_{i,5}$	0.0004376	0.001724	0.9969	0.24531	0

Tabell 38: Förändringskoefficient för hur  $Y_2$  beror på variabel  $X_k$

$\delta_1$	$\delta_2$	$\delta_3$	$\delta_4$	$\delta_5$
0	-0.000571734	-5.73213E-05	-1.85568E-05	-0.000372577

Tabell 39:  $\frac{dY_1}{dX_1}$  för varje enskilt år med 95%-konfidensintervall

År	$\frac{dY_1}{dX_1}$	$I \frac{dY_1}{dX_1}$
1995	-0.000364819	(-0.00036, -0.00029)
1996	-0.000357055	(-0.00036, -0.00030)
1997	-0.000372825	(-0.00037, -0.00030)
1998	-0.00037672	(-0.00037, -0.00030)
1999	-0.000366772	(-0.00037, -0.00030)
2000	-0.000359128	(-0.00038, -0.00031)
2001	-0.000353792	(-0.00038, -0.00031)
2002	-0.000347608	(-0.00038, -0.00031)
2003	-0.000345001	(-0.00038, -0.00031)

2004	-0.000343162	(-0.00039, -0.00031)
2005	-0.000341797	(-0.00040, -0.00032)
2006	-0.000335641	(-0.00041, -0.00032)
2007	-0.000333808	(-0.00043, -0.00032)
2008	-0.000331951	(-0.000421, -0.00032)
2009	-0.000328237	(-0.00039, -0.00032)
2010	-0.000324936	(-0.00041, -0.00032)

Tabell 40:  $dW$  för varje enskilt år med 95%-konfidensintervall

År	$dW$	$I_{dW}$
1995	0.000655369	(0.00066, 0.00073)
1996	0.000663133	(0.00066, 0.00072)
1997	0.000647364	(0.00065, 0.00072)
1998	0.000643468	(0.00065, 0.00072)
1999	0.000653416	(0.00065, 0.00072)
2000	0.00066106	(0.00064, 0.00071)
2001	0.000666396	(0.00064, 0.00071)
2002	0.000672581	(0.00064, 0.00071)
2003	0.000675187	(0.00064, 0.00071)
2004	0.000677026	(0.00063, 0.00071)
2005	0.000678392	(0.00062, 0.00070)
2006	0.000684548	(0.00061, 0.00070)
2007	0.00068638	(0.00059, 0.00070)
2008	0.000688238	(0.00060, 0.00070)
2009	0.000691951	(0.00063, 0.00070)
2010	0.000695252	(0.00061, 0.00070)

Tabell 41: Dummyvariablernas koefficienter

Land	$\alpha_i$	Land	$\alpha_i$
Australia	-	Japan	-6.641e+02
Austria	-1.287e+02	Latvia	-6.641e+02
Belgium	3.284e+01	Lithuania	-6.226e+02
Canada	-3.066e+01	Luxembourg	5.816e+02
Chile	-4.132e+02	Mexico	-9.266e+02
Colombia	-5.883e+02	Netherlands	8.166e+01
Czech Republic	-5.019e+02	New Zealand	-6.246e+01
Denmark	5.488e+02	Norway	1.104e+03
Estonia	-4.877e+02	Poland	-7.176e+02
Finland	-7.227e+00	Portugal	-2.392e+02
France	9.579e-01	Slovak Republic	-7.684e+02
Germany	-2.556e+02	Slovenia	-9.619e+01
Greece	-6.036e+02	Spain	-5.042e+02
Hungary	-6.065e+02	Sweden	4.531e+02
Iceland	8.881e+00	Switzerland	4.460e+02
Ireland	1.655e+02	Turkey	-1.025e+03
Israel	-6.516e+02	United Kingdom	-1.172e+02
Italy	-5.322e+02	United States	-2.780e+02

## 11. Bilaga B - Teori och definitioner

Multivariabel linjär regressionsmodell:

Den multivariabla linjära regressionsmodell definieras som  $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$  där,  $p' = p + 1$

$\mathbf{Y}$ : en  $n \times 1$  kolumnvektor med observationer av den beroende variabeln

$\mathbf{X}$ : en  $n \times p'$  matris bestående av en kolumn med ettor, följt av  $p$  kolumnvektorer med observationer av de oberoende variablerna

$\boldsymbol{\beta}$ : en  $p' \times 1$  kolumnvektor med parametrar att bli estimerade

$\boldsymbol{\epsilon}$ : en  $n \times 1$  kolumnvektor med slumpmässiga fel

$\hat{\boldsymbol{\beta}}$  bestimeras med normalekvationen enligt:  $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{Y})$

(Rawlings John O., Pantula Sastry G., Dickey David A., Applied Regression Analysis: A research tool, second edition, Springer-Verlag New York, Inc., 1998, ISBN 0-387-98454-2, sid 76-78)

Power-transformationer:

Power-transformationer definieras som transformationer där  $X^* = X^k$ , eller  $Y^* = Y^k$  där  $k$  vanligtvis kan vara  $k = -1, -1/2, 0, 1/2, 1, 2$  där 1 skall tolkas som att det inte genomförs en transformation och 0 skall tolkas som den logaritmiska transformationen (Rawlings John O., Pantula Sastry G., Dickey David A., 1998, sid 399), vilket ger oss tre kombinationer av logaritmiska transformationer:

Lin-Log-transformation:  $\mathbf{Y} = \boldsymbol{\beta} \ln(\mathbf{X})$

Log-Lin-transformation:  $\ln(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$

Log-Log-transformation:  $\ln(\mathbf{Y}) = \boldsymbol{\beta} \ln(\mathbf{X})$

Box-Cox-transformationer:

Box-cox-transformation för en linjär regressionsmodell definieras enligt:

$$y_i^\lambda = \alpha + \beta \mathbf{X}$$

$$y_i^\lambda = \frac{y_i^\lambda - 1}{\hat{y}^\lambda}, \text{ för } \lambda \neq 0$$

$$y_i^\lambda = \hat{Y} \ln(y_i), \text{ för } \lambda = 0$$



Där  $\hat{Y}$  är det geometriska medelvärdet av de observerade värdena:  $\hat{Y} = \exp \frac{\sum(\ln(y_i))}{n}$

Optimalt lär det som bäst estimerar en normaldistribution genom att hitta maximum-sannolikhetslösningen, och estimeras med att göra en analys av minsta kvadratsumman på flera val av  $\lambda$ . Man låter  $SS(Res(\lambda))$  vara residualkvadratsumman från att passa in modellen för  $y_i^\lambda$  för givet  $\lambda$  och låter  $\sigma^2(\lambda) = SS(Res(\lambda))/n$ . Sannolikheten för varje  $\lambda$  ges av:  $L_{max} = -\frac{1}{2} \ln(\sigma^2(\lambda))$ .

Maximera sannolikhetsfunktionen är ekvivalent med att minimera residualkvadratsumman.

(Rawlings John O., Pantula Sastry G., Dickey David A., 1998, sid 409-410)

Akaike Information Criterion:

Akaike Information Criterion (AIC) definieras som  $AIC(p') = n \ln(SS(Res)_p) + 2p' - n \ln(n)$  där  $p'$  är antalet parametrar estimerade i modellen. (Rawlings John O., Pantula Sastry G., Dickey David A., 1998, sid 225)

Bayes Information Criterion:

Bayes Information Criterion (BIC) definieras som  $BIC(p') = n \ln(SS(Res)_p) + \ln(n)p' - n \ln(n)$  där  $p'$  är antalet parametrar estimerade i modellen. (Rawlings John O., Pantula Sastry G., Dickey David A., 1998, sid 225)

Forward och backwards selection:

Forward selection eller forward stepwise selection väljer bland subset av modeller genom att addera en variabel i taget till föregående modell. Metoden börjar med att välja en oberoende variabel för att förbättra modellen enligt ett kriterium, och fortsätter att addera variabler till det inte längre går att förbättra modellen. Backward selection fungerar omvänt och börjar med den största möjliga modellen (Rawlings John O., Pantula Sastry G., Dickey David A., 1998, sid 213)

Determinationskoefficienter:

Determinationskoefficienten (R2) är den del av den totala kvadratsumman av den beroende variabeln som kan förklaras med de oberoende variablerna i modellen och definieras som:  $R2 = \frac{SS(Regr)}{SS(Total)}$ . Omskalad R2 (R2- adjusted) är en omskalning av R2 med tanke på frihetsgraden och definieras som:  $R2 - adjusted = 1 - \frac{(1-R2)(n-1)}{(n-p')}$  där  $p'$  är antalet estimerade variabler (Rawlings John O., Pantula Sastry G., Dickey David A., 1998, sid 220).

Leverage:

Minsta kvadratmetoden ger samtliga punkter lika mycket vikt till varje observation, men varje observation har inte samma betydelse för resultaten för minsta kvadratresultaten. En enskild punkt långt ifrån resten av datapunkterna kan ha nästan lika påverkan på modellen som resten av punkterna tillsammans. Sådana punkter kallas punkter med hög leverage. Dess punkter beräknas genom sin relativa placering i X-planet relativt resterande datapunkter (Rawlings John O., Pantula Sastry G., Dickey David A., 1998, sid 330).

Studentiserade residualer:

Studentiserade residualer är residualer som är standardiserade med ett estimat av dess standardavvikelse som är oberoende av residualen själv och definieras enligt:  $r_i^* = \frac{e_i}{s_{(i)}\sqrt{1-v_{ii}}}$ , där  $e_i$ : är den i:th residualen  
 $s_{(i)}$ : är den estimerade variansen utan den i:th residualen  
 $v_{ii} = x_i^T (X^T X)^{-1} x_i$ : där  $x_i$  är den i:th raden i  $X$   
(Rawlings John O., Pantula Sastry G., Dickey David A., 1998, sid 342, 361).

Cook's Distance:

Cook's distance (Cook's D) mäter påverkan av i:th observationen på de estimerade koefficienterna i modellen och definieras som  $D_i = \frac{(\hat{Y}_{(i)} - \bar{Y})^T (\hat{Y}_{(i)} - \bar{Y})}{p' s^2}$ , där  $\hat{Y}_{(i)} = X \hat{\beta}_{(i)}$  och  $p'$  är antalet parametrar estimerade i modellen (Rawlings John O., Pantula Sastry G., Dickey David A., 1998, sid 363-364).