



DEGREE PROJECT IN APPLIED MATHEMATICS AND  
INDUSTRIAL ECONOMICS , FIRST LEVEL  
*STOCKHOLM, SWEDEN 2014*

# Drivande faktorer bakom Tv-tittande på den svenska marknaden

EN REGRESSIONSANALYS UTIFRÅN  
DEFINIERADE MÅLGRUPPER

SARA HOLMSÄTER, ANNA LEIJON

KTH ROYAL INSTITUTE OF TECHNOLOGY

SCI SCHOOL OF ENGINEERING SCIENCES



# Drivande faktorer bakom Tv-tittande på den svenska marknaden

En regressionsanalys utifrån definierade målgrupper

SARA HOLMSÄTER  
ANNA LEIJON

Examensarbete inom teknik:  
Tillämpad matematik och industriell ekonomi (15 credits)  
Civilingenjörsutbildning i industriell ekonomi (300 credits)  
Kungliga Tekniska Högskolan 2014  
Handledare på KTH Anna Jerbrant och Tatjana Pavlenko  
Examinator Tatjana Pavlenko

TRITA-MAT-K 2014:07  
ISRN-KTH/MAT/K--14/07--SE

Kungliga Tekniska Högskolan  
*Skolan för Teknikvetenskap*

**KTH** SCI  
SE-100 44 Stockholm, Sweden

URL: [www.kth.se/sci](http://www.kth.se/sci)



## Sammanfattning

Det här kandidatexamensarbetet utreder huruvida multipel linjär regressionsanalys går att tillämpa på TV-tittarhistorik för att avgöra vilka programdynamiska faktorer som bidrar till höga tittarsiffror. Datan som ligger till underlag för den här analysen kommer från Mediamätning i Skandinavien, MMS. Slutsatser som har dragits i det här kandidatexamensarbetet är att det, inte utan svårigheter, går att tillämpa regressionsanalys på MMS-data. Det kräver bland annat omfattande bearbetning av datan, men också formulering av de programkategorierna som ska undersökas. Utöver det är TV-tittarhistorik som  $Y$ -variabel begränsad såväl nedifrån som ovanifrån, vilket kräver viss omtransformering för att kunna användas i regressionsanalysen. I det här arbetet används logaritmen av  $(Y + 1)$ . Slutsatser kring programdynamiska faktorer som kunnat dras är att TV6 når ut bra i dagsläget till sin definierade målgrupp, medan TV3, TV8 och TV10 kan förbättra sitt TV-tablå-utbud för att bättre nå ut till sina respektive målgrupper. Programkategorier som Talkshows, Hjälp-TV-serier och Utslagstävlingar driver höga tittarsiffror, medan väderinslaget i Nyheterna borde uteslutas ur nyhetssändningen då de flesta byter kanal när det tar vid. Fortsatt analys av MMS-data med regressionsanalys rekommenderas eftersom det är ett rikt område där nya regressionsmodeller kan ställas upp för att utreda andra och/eller mer specifika variabler som påverkar TV-tittande.



# Driving factors of TV habits in the Swedish market

## A regression analysis on defined target groups

### Abstract

This bachelor's thesis investigates whether multiple linear regression analysis can be applied to analyse TV audience measurements in order to examine and draw conclusions about which program dynamics factors contribute to high ratings. The data that forms the basis for this analysis comes from Media Measurement in Scandinavia, MMS. The conclusion that has been drawn in this bachelor's thesis is that it is possible, but not without difficulty, to apply regression analysis on MMS data. It requires, among other things, extensive processing of the data, but also a formulation of the program categories which are to be examined. In addition, the TV audience measurement as the  $Y$  variable in the model is limited from above as well as from below, which requires transformation of the  $Y$  variable for it to be used in regression modelling. In this bachelor's thesis, the logarithm of  $(Y + 1)$  is used. Conclusions that could be drawn about program dynamics factors is that TV6 reaches its target audience well, while TV3, TV8 and TV10 can improve their TV guide, in order to better reach out to their target audiences. Program categories such as Talk Shows, Help TV series and Knockout Competitions drive high viewer ratings, while the weather forecast in the Newscast should be excluded, since most people change the channel when it begins. Continued use of regressions analysis on TV audience measurement is recommended, as it is a rich area where new regression models can be set up to investigate further and /or more specific variables which affect television viewing ratings.





## Tacksägelser

Författarna till det här kandidatexamensarbetet skulle vilja passa på att tacka sina handledare på Kungliga Tekniska Högskolan: Anna Jerbrant på Institutionen för Industriell Ekonomi för värdefull feedback och Tatjana Pavlenko på Institutionen för Matematisk Statistik för nya matematiska infallsvinklar. Vi vill även tacka Andreas Nordfors i egenskap av analytiker på Nepa för ett givande samarbete. Utöver dessa vill vi även tacka Misha Wolynski i egenskap av examinerad civilingenjör och civilekonom på Kungliga Tekniska Högskolan respektive Handeshögskolan i Stockholm för eminenta diskussioner.



# Innehållsförteckning

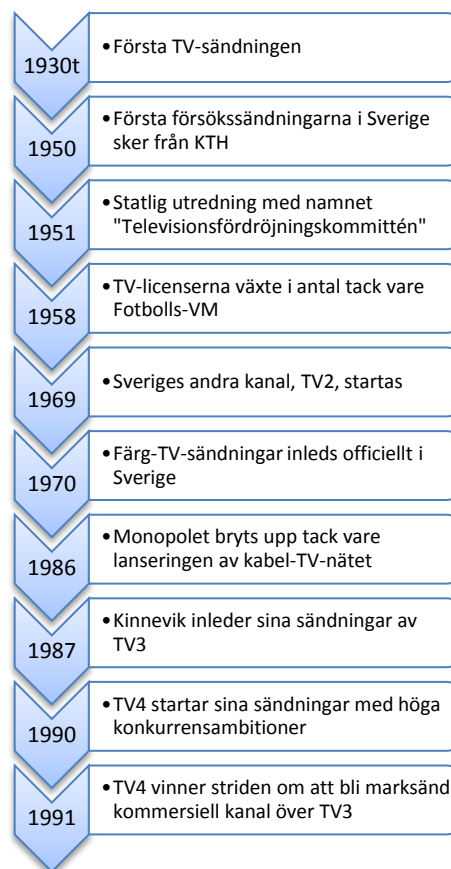
|  |    |
|--|----|
| 1. Inledning.....  | 1  |
| 1.1. Bakgrund.....   | 1  |
| 1.2. Syfte.....  | 3  |
| 1.3. Problemformulering.....   | 3  |
| 1.4. Avgränsningar.....  | 3  |
| 1.4.1. Variabelavgränsning.....  | 3  |
| 1.4.2. Kanal- och målgruppsavgränsning.....                                | 4  |
| 1.4.3. Dataavgränsning.....  | 4  |
| 2. Teori.....  | 5  |
| 2.1. De fem grundantagandena.....  | 6  |
| 2.2. Dummyvariabler.....   | 6  |
| 2.3. Multikollinearitet.....   | 7  |
| 2.3.1. Upptäckt.....   | 8  |
| 2.3.2. Åtgärd.....   | 8  |
| 2.4. Endogeneitet.....   | 8  |
| 2.4.1. Upptäckt.....   | 9  |
| 2.4.2. Åtgärd.....   | 9  |
| 2.5. Heteroskedasticitet.....  | 10 |
| 2.5.1. Upptäckt.....   | 10 |
| 2.5.2. Åtgärd 1 - Transformerings av modellen.....                         | 11 |
| 2.5.3. Åtgärd 2 – White’s Consistent Variance Estimator.....               | 12 |
| 2.5.4. Hypotesprövning då heteroskedacitet identifierats.....              | 12 |
| 2.6. Mått av regressionsmodellens passform.....                            | 13 |
| 2.6.1. $R^2$ -värde och justerat- $R^2$ -värde.....                        | 14 |
| 2.6.2. RMSE – Kvadratroten ur medelvärde av den kvadrerade residualen..... | 14 |
| 3. Mediamätning i Skandinavien.....  | 15 |
| 3.1. Insamling av TV-tittarhistorik.....                                   | 15 |
| 3.1.1. Metod för hanteringen av MMS-datan.....                             | 15 |
| 4. Metod.....  | 17 |
| 4.1. Grundekvation.....  | 17 |
| 4.2. Splittest.....  | 18 |
| 4.3. Heteroskedasticitet och transformering av Y-variabeln.....            | 18 |
| 4.4. Multikollinearitet.....   | 20 |

|      |  |    |
|------|--|----|
| 4.5. | Endogeneitet.....  | 20 |
| 4.6. | Omskrivning av förklaringsmatrisen.....                                      | 21 |
| 4.7. | Matlab.....  | 21 |
| 5.   | Resultat.....  | 22 |
| 5.1. | Alla, 3 – 99 år, kl. 20:00-22:59 .....                                       | 24 |
| 5.2. | Alla, 15 – 49 år, kl. 20:00-22:59 .....                                      | 26 |
| 5.3. | Kvinnor, 25 – 44 år, kl. 20:00-22:59 .....                                   | 28 |
| 5.4. | Kvinnor, 25 – 49 år, kl. 20:00-22:59 .....                                   | 30 |
| 5.5. | Män, 15 – 34 år, kl. 20:00-22:59 .....                                       | 32 |
| 5.6. | Män, 25 – 49 år, kl. 20:00-22:59 .....                                       | 34 |
| 6.   | Matematisk diskussion .....  | 36 |
| 6.1. | Kanalgenerell diskussion.....  | 37 |
| 6.2. | Alla, 3 – 99 år, kl. 20:00-22:59 .....                                       | 39 |
| 6.3. | Alla, 15 – 49 år, kl. 20:00-22:59, kanalgruppens generella målgrupp .....    | 40 |
| 6.4. | Kvinnor, 25 – 44 år, kl. 20:00-22:59, TV3:s målgrupp.....                    | 42 |
| 6.5. | Kvinnor, 25 – 49 år, kl. 20:00-22:59, TV8:s målgrupp.....                    | 44 |
| 6.6. | Män, 15 – 34 år, kl. 20:00-22:59, TV6:s målgrupp.....                        | 46 |
| 6.7. | Män, 25 – 49 år, kl. 20:00-22:59, TV10:s målgrupp.....                       | 48 |
| 6.8. | Rekommendationer .....   | 50 |
| 7.   | Kritik riktad mot regressionsmodellen med stöd ur marknadsföringsteori ..... | 52 |
| 7.1. | Inledning.....   | 52 |
| 7.2. | Teori .....  | 52 |
| 7.3. | Regressionsmodellen – ur ett marknadsföringsperspektiv .....                 | 53 |
|      | Källhänvisning.....  | 57 |
|      | Bilaga 1. Ursprungsplostrar .....  | 59 |
|      | Bilaga 2. Plostrar med transformeringsmetod $\log(Y+1)$ .....                | 60 |
|      | Bilaga 3. Plostrar med transformeringsmetod $(Y+1)^{0,25}$ .....             | 61 |
|      | Bilaga 5. Alla kategorier med tillhörande program kl. 20:00-22:59 .....      | 63 |
|      | Bilaga 6. Alla programkategorier med förklaring .....                        | 65 |
|      | Bilaga 7. Matlab-kod.....  | 66 |
|      | Bilaga 8. Multikollinearitetstest .....                                      | 68 |

## 1. Inledning

TV som massmedium har haft stor genomslagskraft sedan den första TV-sändningen på 1930-talet. I Sverige har TV-tittandet varit utbredd sedan 60-talet med TV-legender som Lennart Hyland i spetsen<sup>1</sup>.

De första försökssändningarna i Sverige skedde emellertid från KTH: ”TV från Teknis” på onsdagar under hösten 1950 av dåvarande Televisionsforskningsnämnden<sup>2</sup>. Även om KTH var i framkant var inte alla positivt inställda till TV-fenomenet och 1951 tillsatte regeringen en statlig utredning med öknamnet ”Televisionsfördröjningskommittén”. Trots i huvudsak socialdemokratiskt motstånd riktat mot TV:n, startades en statlig, licensfinansierad kanal. Långsamt, men stadigt, blev TV allt mer populärt, men det skulle dröja till Fotbolls-VM 1958 innan antalet TV-licenser sköt i höjden. Den första kanalen fick sedan namnet TV1 när TV2 så småningom lanserades 1969, även denna statlig och licensfinansierad. TV:n blev om möjligt än mer populär i och med de första färg-TV-sändningarna 1970. En viss konkurrens efterfrågades och förslaget med en marksänd, kommersiell, kanal lades fram för att bidra till TV-programmens kvalitet. En kombination av tekniska framsteg och politisk vilja ledde till att det tidigare, statliga, TV-monopolet bröts upp 1986 tack vare lanseringen av kabel-TV-nätet, men det skulle dröja innan den första marksända, kommersiella, TV-kanalen lanserades. Kinneviks TV3 kom däremot till Sverige via satellit-sändningar 1987 och det skulle dröja till 1990 innan nästa kommersiella kanal, TV4, påbörjade sina satellitsändningar. 1990 fanns ett 30-tal olika TV-kanaler att tillgå via kabel- och satellit-TV, däribland TV4 som utmärkte sig med sina höga konkurrensambitioner. Konkurrensen bidrog till att debatten återigen blossade upp angående marksänd kommersiell TV, som slutligen landade i beslutet att TV4 blev den första marksända reklamkanalen i Sverige<sup>3</sup>. En tidslinje över svensk TV-kanal-historia återfinns i Figur 1.



Figur 1. Tidslinje över svensk TV-kanal-historia.

### 1.1. Bakgrund

TV:n har följaktligen varit starkt närvarande i de svenska hemmen i drygt 50 år. TV:ns föregångare utgörs av radio och printmedier och dess samtida, tekniska utmanare är bland andra internet och applikationer, som utgör såväl substitut som komplement till traditionellt TV-tittande. Däremot anammar många inte den nya tekniken och även om så kallade ”early adoptors” ligger i framkant, krävs ofta ett generationsskifte för att på riktigt förändra ett grundläggande konsumentbeteende i stor skala.

<sup>1</sup> Joakim Thelander, 'En titt på TV', *Populär Historia*, 09 December 2002, pp. 10-14.

<sup>2</sup> Björn Nilsson, *Boken om TV* (Malmö: Bengt Forsbergs Förlag, 1961), p. 41.

<sup>3</sup> Jan Milld, 'BGF historia', in *Television* <<http://www.bgf.nu/historia/6/tv.html>> [accessed 04 May 2014]

Totalt sett ägnar svenskarna idag mer tid åt TV-tittande än någonsin förut enligt en intervju från 2013 med Magnus Anshelm, VD på MMS-data. MMS står för Mediamätning i Skandinavien, och som namnet antyder, samlar de in och analyserar TV-tittarhistorik, och har så gjort sedan 1993. Det totala TV-tittandet utgörs av traditionellt liksom internetbaserat tittande på TV-program. Med konkurrensen från internet vore det inte förvånansvärt om det traditionella TV-tittandet byts ut mot internettittande i ett längre perspektiv, på grund av den större programtillgängligheten internet medför. Mycket tyder dock på att det traditionella tablåttittandet på TV inte minskar, tack vare att traditionellt TV-tittande och webbaserat TV-tittande går hand i hand. Det finns ett tydligt samband – de format som är populärast på nätet är också de format som är populärast på traditionell TV<sup>4</sup>. Magnus Anshelm tillägger att ”om det inte vore för linjär TV skulle tittarna inte veta vilka format de vill se on demand”.<sup>5</sup>

Oavsett vad statistiken säger antas ofta gamla tekniska lösningar, som tablå-TV:n, behöva lämna plats för och ersättas av andra, nyare lösningar, som datorer, surfplattor och mobiltelefoner och tekniska system som TV-kablar och fjärrkontroller av internet och interaktiva applikationer. Enligt Magnus Anshelm är detta antagande felaktigt: han menar att dessa snarare kommer att leva i symbios. Ett argument som stöder Anshelms uttalande är att TV-branschen omsätter mer och mer pengar. 21 miljarder år 2013, vilket är en miljard mer än 2012<sup>6</sup>. Ytterligare aktörer på TV-marknaden som backar upp Anshelms påstående är Konkurrensverket och Myndigheten för Radio och TV, som rapporterar att 95 procent av konsumtionen av rörlig bild utgörs av traditionell TV<sup>7</sup>. Så även om det är en populär åsikt att den gamla tidens traditionella TV-tittande är på nedgång och den nya tidens webbaserade TV-tittande på uppgång, präglas nätverkstittandet till så pass stor del av det traditionella TV-tittandet att det är aktuellt att analysera det senare.

TV:n är följaktligen här för att stanna och det här kandidatexamensarbetet utreder vilka faktorer, som TV-bolagen själva kan påverka, som leder till höga tittarsiffror. Intresset ligger således i att analysera vad som driver TV-tittande, ett intresse som författarna av denna uppsats och researchföretaget Nepa delar. Denna undersökning utformas som ett kandidatexamensarbete i samarbete med Nepa. Nepa grundades 2006 av två studenter vid Handelshögskolan i Stockholm och är idag ett snabbt växande researchbolag<sup>8</sup>. Det här kandidatexamensarbetet skrivs på uppdrag av och i samförstånd med Nepa.

---

<sup>4</sup> Åsa Jamal, 'Spotlight TV4-gruppen' <<https://spotlight.tv4.se/Nyheter/Arkiv/2013/Maj/Framtidens-tv/>> [accessed 06 April 2014]

<sup>5</sup> I det här sammanhanget kallas linjär för traditionell TV och TV on demand för internet- eller webbaserad TV på grund av att vi anser det vara ett tydligare språkbruk.

<sup>6</sup> Madeleine Östlund, 'Dagens Media' <<http://www.dagensmedia.se/nyheter/tv/article3626272.ece>> [accessed 06 April 2014]

<sup>7</sup> Konkurrensverket, 'Rörlig Bild - En Analys av Marknaden för Rörlig Bild' (Report, Myndigheten för Radio och TV, Konkurrensverket, 2012), p. 15.

<sup>8</sup> Nepa, 'Nepa' <<http://www.nepa.se/sidor/om-oss/>> [accessed 06 April 2014]

## 1.2. Syfte

Fokus ligger på att ställa upp och analysera de faktorer som rör TV-tablån och som TV-bolagen kan påverka. Detta utreds med hjälp av matematisk statistik för att på så sätt avgöra vilka faktorer som bidrar mest till höga tittarsiffror och således vad TV-bolagen bör satsa på för att nå ut till tittarna i så stor utsträckning som möjligt. Att analysera MMS-data med hjälp av regressionsanalys har inte gjorts tidigare, varför det här kandidatexamensarbetet ämnar besvara frågan om det ens är möjligt.

När det gäller TV-tablåns konstruktion, som hädanefter benämns som programdynamik, syftar det till de olika sätt på vilka man kan kombinera faktorerna som tablån utgörs av för att leverera ett så attraktivt paket till tittarna som möjligt. Faktorerna går närmare in på under avsnittet om avgränsningar. Slutligen återges en samlad rekommendation för respektive kanal med hänsyn till dess initiala målgrupp både vad gäller dess slutprodukt till kund – dess kanalpaket – det vill säga, vad kanalen har att erbjuda, men också riktlinjer vid såväl inköp som nyproduktion av program. Lämpligen används slutsatsen och diskussionen av denna kandidaatexamensuppsats av personer som arbetar med verksamhetsledning inom respektive kanal.

## 1.3. Problemformulering

Problemformuleringen i det här kandidatexamensarbetet grundar sig i frågeställningen: hur kan regressionsanalys tillämpas för att analysera TV-tittarhistorik och mer specifikt, vilka programdynamiska faktorer leder till höga tittarsiffror? Hypoteser kommer att ställas upp rörande programdynamiken och regressionsanalysen kommer sedan att hjälpa till med att avgöra om dessa ska förkastas eller inte. De hypoteser som sedan ej förkastas kommer att kunna användas som grund för hur TV-kanalerna lägger upp sina TV-tablåer och för att på så vis kunna driva höga tittarsiffror och stärka kanalens position i jämförelse med dess konkurrenter.

## 1.4. Avgränsningar

Generella avgränsningar i det här kandidatexamensarbetet är bland annat att 1% signifikansnivå har valts i regressionsanalysen. Det är främst eftersom mängden data möjliggör en så pass låg nivå, men också då författarna på så sätt garderar sig genom att endast basera rekommendationerna på slutsatser med stort stöd ur både ett matematiskt och ett observationsmässigt perspektiv. På grund av att regressionsanalys av MMS-data inte tidigare använts för att analysera TV-tittarhistorik väljs gärna en mer återhållsam attityd i största allmänhet.

### 1.4.1. Variabelavgränsning

Faktorerna som det kommer att tas hänsyn till i regressionsanalysen kommer att begränsas till: tid på dygnet, dag i veckan, programkategori, programlängd och program i följd. Program i följd innebär att exempelvis en TV-serie sänder två på varandra efterföljande episoder och med den variabeln ligger fokus på att analysera huruvida det har en positiv eller negativ inverkan på det fortsatta TV-tittandet.

### 1.4.2. Kanal- och målgruppsavgränsning

Nepa har formulerat följande målgrupper för de respektive TV-kanalerna. Av den anledningen analyseras TV-kanalerna utifrån hur pass väl de når ut till sin målgrupp under diskussionsdelen. Att målgrupperna överlappar i viss utsträckning är väntat, eftersom kanalerna konkurrerar om TV-tittarna. Regressionsanalysen begränsas således till de sex målgrupperna som framgår av Tabell 1.

| Alla kanaler  | Kanalgruppen   | TV3        | TV6        | TV8        | TV10       |
|---------------|----------------|------------|------------|------------|------------|
| Alla, 3-99 år | Alla, 15-49 år | ♀ 25-44 år | ♂ 15-34 år | ♀ 25-49 år | ♂ 25-49 år |

**Tabell 1.** Målgruppsindelning efter kanal, kanalgrupp och för alla.

Kanalgruppen består av TV3, TV6, TV8 och TV10 och är även den formulerad av Nepa. Dessa är intressanta att analysera i en samlad grupp för att se vad de tillsammans skulle kunna fokusera på för att förbättra sina tittarsiffror. Utöver dessa tas också SVT1, SVT2, TV4 och Kanal 5 med i regressionerna. Detta görs främst för att ha någonting att jämföra kanalerna med samt för att få en större spridning i observationerna.

För övrigt har 99,8% av Sveriges hushåll tillgång till marksänd digital-TV, vilken utgörs av SVT1, SVT2, TV4 och TV6, av de som tas upp här. Det finns även en mängd abonnemang tillgängliga på marknaden som möjliggör tittande på betalkanalerna TV3, Kanal 5, TV8 och TV10. Olika abonnemang är populära i olika utsträckning och följaktligen leder det till att fri-TV-kanalerna finns tillgängliga för de flesta, men övriga kanaler i varierande utsträckning.<sup>9</sup>

### 1.4.3. Dataavgränsning

Datainsamlingen har behövt begränsas då MMS-data insamlats sedan 1993 och det finns en uppsjö av möjlig statistik att hämta på området<sup>10</sup>. För att kunna analysera i ett brett spektrum av olika tablåer och programtider, har månaden november valts som representativ för TV-tittande i största allmänhet och fem år tillbaka i tiden som den period för vilken observationerna fortfarande är aktuella ur TV-kanalernas synvinkel. Detta har flera anledningar. November är relativt opåverkad av säsongsinfluenser som så kallat semestertittande, som främst gör sig gällande under jul och sommarmånaderna. Melodifestivalen under våren och Kalle Anka under julafton är program då SVT1 traditionellt sett, men även kulturellt betingat, inte går att jämföra med några andra program. De är typiska störningsmoment eftersom kandidatexamensarbetet avser analysera normal-TV-tittande.

Månaden november gör det möjligt att undvika dessa program i så stor utsträckning som möjligt, även om det inte helt går att undkomma exempelvis större sportevenemang. Månaden november valdes även då det procentuella tittandet på de olika programmen var snarlikt, baserat på MMS-data.

<sup>9</sup> Konkurrensverket, 'Rörlig Bild - En Analys av Marknaden för Rörlig Bild' (Report, Myndigheten för Radio och TV, Konkurrensverket, 2012), p. 15.

<sup>10</sup> MMS, 'Om MMS - Historik' <<http://www.mms.se/start/om-mms/historik/>> [accessed 06 April 2014]



## 2. Teori

Ekonometri är tillämpning av matematiska och statistiska metoder på empiriskt framtagen data. Det finns en uppsjö av olika typer av statistiska modeller inom ekonometri.<sup>11</sup> Linjär regression är en modell som används då målet är att skapa en linjär funktion som bäst beskriver observerad data och kan användas för prediktion och analys.

Linjär regression strävar efter att förklara hur en beroende variabel varierar med avseende på en utvald mängd av oberoende variabler. De oberoende variablerna representeras av  $X$ -matrisen och den beroende av  $Y$ -vektorn.  $\beta$  i Formel (1) skattas med hjälp av  $X$  och  $Y$ . Den linjära regressionsmodellen ställs upp som Formel (1) och dess ingående delar enligt Formel (2). Som synes beskrivs den beroende variabeln i modellen av en konstantterm samt  $m$  oberoende variabler, där alla variabler har  $n$  observationer vardera.

$$Y = X\beta + e \quad (1)$$

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, X = \begin{bmatrix} 1 & x_{1,1} & \cdots & x_{1,m} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n,1} & \cdots & x_{n,m} \end{bmatrix} \quad (2)$$

$\beta_0$  utgör intercept och övriga  $\beta$ -värden utgör känsligheter för var och en av de  $m$  beroende variablerna i den uppställda modellen, som återges i Formel (3). De estimerade  $\beta$ -värdena betecknas  $\hat{\beta}$ . Vektorn  $e$  är en residualvektor och utgörs av residualerna mellan modellens värden och de faktiska observationerna för den beroende variabeln.<sup>12</sup>

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}, e = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} \quad (3)$$

Ur ett matematiskt perspektiv beräknas  $\hat{\beta}$  genom att tillämpa minsta kvadrat-metoden (även kallad Ordinary Least Square, OLS) på en mängd observationer för att minimera summan av kvadraterna av residualerna  $\hat{e}^t \hat{e} = |\hat{e}|^2$ . Detta görs med matrisoperationer enligt Formel (4).

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (4)$$

Kovariansmatrisen beräknas som Formel (5).<sup>13</sup>

$$\begin{aligned} \text{Cov}(\hat{\beta}|X) &= (X^T X)^{-1} X^T (\sigma^2 I) X (X^T X)^{-1} \\ &= (X^T X)^{-1} \sigma^2 \end{aligned} \quad (5)$$

---

<sup>11</sup> Harald Lang, *Topics on Applied Mathematical Statistics* (Stockholm: Kungliga Tekniska Högskolan, 2013), pp. 17-18.

<sup>12</sup> Ibid. pp. 18-19.

<sup>13</sup> Ibid. p 21.

## 2.1. De fem grundantagandena

Den klassiska linjära regressionsmodellens fem grundantaganden är:

1. Den beroende variabeln är en linjär funktion av de oberoende variablerna plus residual.
2. Väntevärdet av residualen är lika med noll.

$$E[e] = 0 \quad (6)$$

3. Residualen har enhetlig varians och residualer korrelerar inte med varandra.

$$E[e^T e] = \sigma^2 I \quad (7)$$

4. Observationer av de oberoende variablerna kan anses vara fixerade i upprepade stickprov.
5. Det finns inga exakta linjära samband mellan oberoende variabler och det finns fler observationer än oberoende variabler.<sup>14</sup>

$$\sum_{t=1}^m (x_t - \bar{x})^2 \neq 0, \quad (8)$$

$$\text{Rank}(X) = m \leq n, t \in 1, 2, \dots, m$$

Utöver dessa grundantaganden förutsätter regressionsmodellen att residualerna är normalfördelade.<sup>15</sup> Regressionsdiagnostik ämnar besvara frågan hur väl grundantagandena i linjär regression tillfredsställs av den aktuella modellen. Om modellen inte uppfyller antagandena bör såväl slutsatser som resultat som härleds därur ifrågasättas. Det bästa enskilda sättet att testa om antagandena håller är att göra såkallade residualplottar.<sup>16</sup>

## 2.2. Dummyvariabler

Om olika kategorier förväntas ge olika resultat, är det lämpligt att dela upp dem och använda dummyvariabler i regressionsmodellen. Dummyvariabler är konstruerade så att de antar värdet 1 när en observation tillhör kategorin och värdet 0 annars. För att inte få perfekt multikollinearitet, mer om det fenomenet under Kapitel 2.3., utelämnas alltid en kategori och utgör ett slags benchmark. Om modellen behandlar  $k$  kategorier, har den följaktligen  $k - 1$  dummyvariabler. Den utelämnade kategorin, kodad som noll för alla dummyvariablerna i uppsättningen, fungerar som ett slags benchmark och utgörs av intercept, med vilken de andra kategorierna jämförs, se Formel (9). Som även kan utläsas ur Formel (9) är det oväsenligt vilken kategori som utelämnas och används som benchmark då de slutgiltiga intercepten för respektive kategori alltid blir de samma. Ett typexempel är dummyvariabeln kön som kan representeras av 0 om utfallet är ”man” och 1 om utfallet är ”kvinna” i

---

<sup>14</sup> Peter Kennedy, *A Guide to Econometrics*, 6th edn (United States of America: Blackwell Publishing Ltd, 2008), pp. 40-43.

<sup>15</sup> Harald Lang, *Topics on Applied Mathematical Statistics* (Stockholm: Kungliga Tekniska Högskolan, 2013), pp. 18-19.

<sup>16</sup> Oxford Journals, 'Regression Diagnostics', *Our Journals*, 5 (2014), 49-51.

en ekvation som ämnar besvara hypotesen att män tjänar mer än kvinnor. När dummyvariabeln representeras av fler kategorier än två, kan en uppsättning av dummyvariabler representera de olika utfallen i regressionen.

Modell: 
$$Y_i = \alpha + X_{i1}\beta_1 + X_{i2}\beta_2 + D_{i1}\gamma_1 + D_{i2}\gamma_2 + \dots + D_{in}\gamma_n + e_i$$

Benchmark: 
$$Y_i = \alpha + X_{i1}\beta_1 + X_{i2}\beta_2 + e_i$$

Kategori 1: 
$$Y_i = (\alpha + \gamma_1) + X_{i1}\beta_1 + X_{i2}\beta_2 + e_i \tag{9}$$

Kategori 2: 
$$Y_i = (\alpha + \gamma_2) + X_{i1}\beta_1 + X_{i2}\beta_2 + e_i$$

⋮

Kategori n: 
$$Y_i = (\alpha + \gamma_n) + X_{i1}\beta_1 + X_{i2}\beta_2 + \varepsilon_i$$

I Tabell 2. nedan är kategori ”m” benchmark. Modellen representerar parallella regressionsytor, en för varje kategori, som återges i Formel (9).<sup>17</sup>

| Kategori | $D_1$ | $D_2$ | ... | $D_{m-1}$ |
|----------|-------|-------|-----|-----------|
| 1        | 1     | 0     | ... | 0         |
| 2        | 0     | 1     | ... | 0         |
| ⋮        | ⋮     | ⋮     |     | ⋮         |
| $m - 1$  | 0     | 0     | ... | 1         |
| $m$      | 0     | 0     | ... | 0         |

**Tabell 2.** *Multipla dummyvariablers uppställning i en regressionsmodell.*

### 2.3. Multikollinearitet

Om det femte grundantagandet, som återges i Formel (8), inte håller i regressionsmodellen beror det antingen på att man har färre observationer än oberoende variabler eller att det finns ett eller flera linjära samband mellan de oberoende variablerna. Vid färre observationer än oberoende variabler korrigeras det lätt genom kompletterande inhämtning av fler observationer. Exakta linjära samband gör det däremot omöjligt att utföra regressionen med OLS (se Formel (4)). Approximativa linjära samband utgör också ett problem i modellen, men detta går att åtgärda. Förekomsten av approximativa linjära samband mellan de oberoende variablerna i modellen benämns multikollinearitet och är ett vanligt förekommande problem inom regressionsanalys. Det går att förebygga på olika sätt (se stycket 2.3.2 om Åtgärd på det här området).

Multikollinearitet är således ett problem som följer av att två eller flera förklaringsvariabler i  $X$ -matrisen är korrelerade och att variationen dem emellan följaktligen är liten (eller icke-existerande, vilket leder till ett exakt linjärt samband och perfekt korrelation). Då det inte går att helt särskilja förklaringsvariablerna åt går det heller inte att särskilja deras respektive påverkan på den beroende variabeln  $Y$  åt, och även om modellen som helhet fortfarande kan ha hög tillförlitlighet, går det inte att göra någon utsago rörande de individuella förklaringsvariablernas påverkan på  $Y$ -variabeln. Multikollinearitet får således som följd att den linjära regressionen fortfarande går väl att använda vid

<sup>17</sup> John Fox, *Applied Regression Analysis and Generalized Linear Models*, Second Edition edn (Washington?: SAGE Publications, Inc, 2008), pp. 126-27.

prediktion, under förutsättningen att de linjära sambanden kommer att bestå, men desto sämre vid tolkning av förklaringsvariablerna var för sig.<sup>18</sup>

### 2.3.1. Upptäckt

Ett sätt att upptäcka multikollinearitet på är att titta på de estimerade standardavvikelserna tillhörande de olika  $\beta$ -estimaterna i den linjära regressionsmodellen. Korrelerade förklaringsvariabler har nämligen mycket höga estimerade standardavvikelser<sup>19</sup>. Ett annat sätt att undersöka om multikollinearitet existerar i modellen, som inte är lika vanligt förekommande, men mer tillfredsställande, är att använda ett så kallat "condition index", eller konditionsindex. Konditionsindex är roten ur matrisen  $X^T X$ :s egenvärden i ordningen det största till det minsta, där  $X$  är tagen ur Formel (2). Tekniskt sett är detta ett mått på hur pass nära matrisen  $X^T X$  är singularitet<sup>20</sup>. En tumregel med konditionsindex är att ett index över 30 indikerar stark kollinearitet.<sup>21</sup>

### 2.3.2. Åtgärd

Om regressionsmodellen ska användas för analys behöver multikollineariteten inte åtgärdas under förutsättning att de korrelerade variablerna ska analyseras tillsammans eller att de koefficienterna som ska analyseras inte är berörda av multikollineariteten. Multikollineariteten bör emellertid *inte* förbises om de förklaringsvariabler som ska analyseras var för sig är korrelerade.

Rekommenderade tillvägagångssätt för att åtgärda multikollinearitet är dels att omformulera modellen, och dels att tillföra ytterligare observationer. Att inhämta fler observationer är rekommenderat i de fall multikollinearitet tros ha uppstått på grund av för lite variation i den befintliga observationsmängden. När det gäller omformulering av modellen kan den skrivas om så att relationer formuleras mellan de oberoende variablerna. Modellen kan också förändras genom att utesluta en av de korrelerade variablerna, vilket möjliggör analys av de respektive återstående förklaringsvariablerna var för sig.<sup>22</sup>

## 2.4. Endogeneitet

Om observationsmängden är endogen leder till att det fjärde grundantagandet i Kapitel 2.1. inte håller. Det kan bero på ett systematiskt mätfel hos observationerna, att ett eftersläpande värde av den beroende variabeln används som oberoende variabel eller att den beroende variabeln i vissa situationer bestäms av den simultana interaktionen mellan flera relationer.<sup>23</sup>

---

<sup>18</sup> Peter Kennedy, *A Guide to Econometrics*, 6th edn (United States of America: Blackwell Publishing Ltd, 2008), pp. 192-94.

<sup>19</sup> Harald Lang, *Topics on Applied Mathematical Statistics* (Stockholm: Kungliga Tekniska Högskolan, 2013), pp. 29-30.

<sup>20</sup> Singularitet: Innebär att en matris inte är inverterbar.

<sup>21</sup> Peter Kennedy, *A Guide to Econometrics*, 6th edn (United States of America: Blackwell Publishing Ltd, 2008), p. 199.

<sup>22</sup> Ibid. pp. 196-98.

<sup>23</sup> Peter Kennedy, *A Guide to Econometrics*, 6th edn (United States of America: Blackwell Publishing Ltd, 2008), pp. 40-43.

Endogenitet innebär att minst en av förklaringsvariablerna i regressionsmodellen är korrelerad med residualen och leder till att regressionen kommer att generera icke-konsekventa estimat. Detta är ett problem då regressionskvationen ska användas för tolkning, men inte i de fall den ska användas för förutsägelser.<sup>24</sup>

### 2.4.1. Upptäckt

Några vanliga fall som leder till endogenitet är då den insamlade datan är vinklad, simultan, om det saknas relevanta förklaringsvariabler eller i de fall då det uppstår mätfel. Vinklad data uppstår om urvalet av data beror av några andra faktorer än slumpen.

Ett exempel på vinklad data är ”Self selection bias”. Detta uppstår när de mätpunkter som tas med i regressionen tillkommit av en faktor som inte tas hänsyn till i ekvationen. Ett exempel på ”Self selection bias” skulle kunna vara att genomföra en opinionsundersökning i enbart höginkomstområden eftersom dess invånare kan tänkas rösta mer ekonomiskt liberalt. Av den anledningen återspeglar inte opinionsundersökningen den totala befolkningen och observationerna blir vinklade.

Simultan data uppstår när den beroende variabeln är korrelerad med minst en oberoende variabel. Detta innebär att orsak och verkan går i mer än en riktning. Ett exempel på detta är utbud- och efterfrågekurvan. Priset är endogent i ekvationen eftersom tillverkarna ändrar sitt pris som en motreaktion på en förändring i efterfråga, medan konsumenterna i sin tur ändrar sin efterfråga till följd av pris.

Endogeneitet kan också vara en följd av att det saknas relevanta kovariater i regressionsmodellen. Ett exempel är att utelämnat motorstorlek i en ekvation som ska analysera en bils slutpris. En bil som gör av med lite bensin borde vara dyr, eftersom man kan tjäna in pengar på att den drar mindre. Om inte motorstorlek tas med i ekvationen kommer högre bensinkonsumtion leda till högre pris istället, eftersom motorstorlek och bensinkonsumtion är korrelerade och motorstorlek också är en bidragande faktor till en dyrare bil. Således kommer  $\beta$ -värdet för bensinkonsumtion att vara missvisande.<sup>25</sup>

### 2.4.2. Åtgärd

För att motverka endogeneitet vad det gäller ”Self selection bias” gäller det att hitta observationer som är representativa för det som ska undersökas samt att stickprovet verkligen är slumpmässigt draget inom den population som avses undersökas. I det fall då det saknas relevanta förklaringsvariabler i regressionen är åtgärden intuitiv och går ut på att införa den förklarande variabeln som ej inkluderats i modellen hittills. Utöver detta kan endogeneitet motverkas genom införandet av så kallade instrumentvariabler.<sup>26</sup>

---

<sup>24</sup> Harald Lang, *Topics on Applied Mathematical Statistics* (Stockholm: Kungliga Tekniska Högskolan, 2013), pp. 29-30.

<sup>25</sup> Ibid. pp. 30-32.

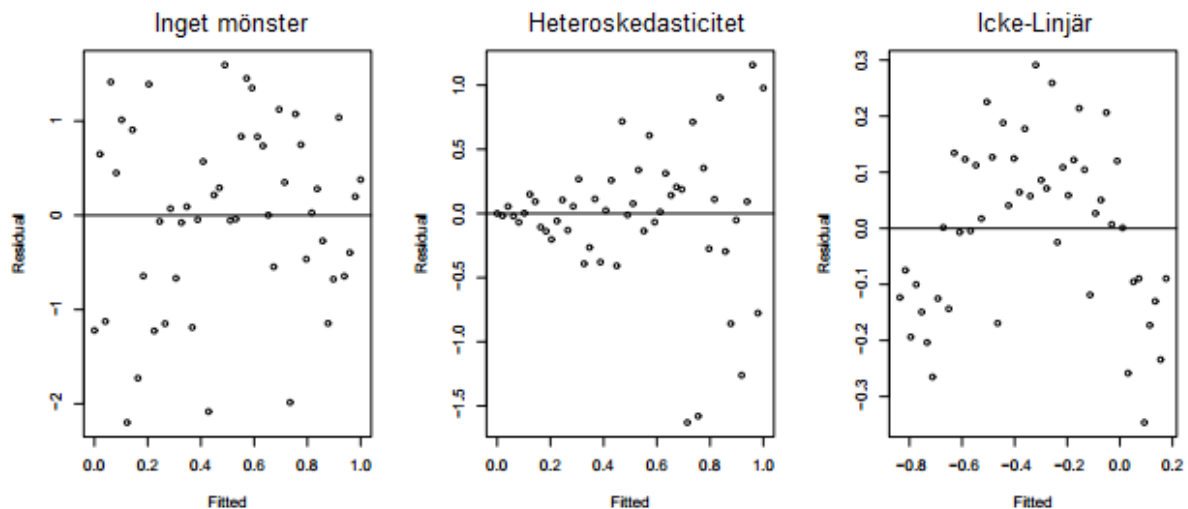
<sup>26</sup> Harald Lang, *Topics on Applied Mathematical Statistics* (Stockholm: Kungliga Tekniska Högskolan, 2013), p. 38.

## 2.5. Heteroskedasticitet

Om det tredje grundantagandet i regressionsmodellen inte håller beror detta på heteroskedasticitet eller autokorrelation. Som termen antyder är heteroskedasticitet motsatsen till homoskedasticitet. Homoskedasticitet håller om en observationsmängd har samma varians i regressionsmodellen, det vill säga att residualens varians är likformig. Om det inte gäller är observationsmängden heteroskedastisk. Problem uppstår i det heteroskedastiska fallet, eftersom hypotesprövningen invalideras då F-test inte går att använda och standardavvikelseerna för de estimerade parametrarna är inkonsekventa.<sup>27</sup> Regressionsanalys med heteroskedastisk observationsmängd medför *inte* snedvriden estimering av relationen mellan förklaringsvariablerna och utfallet. Dock blir standardavvikelse och slutsatserna som kan dras från regressionen inte nödvändigtvis tillförlitliga.<sup>28</sup>

### 2.5.1. Upptäckt

Vanligaste sättet att undersöka datan för att upptäcka heteroskedasticitet är att utföra olika typer av residualplottar. Den viktigaste residualplotten, som inte bara går att använda för detektion av heteroskedasticitet, utan även för upptäckt av ”outliers” och icke-linjäritet, är ”Residuals vs. Fitted Values”. En gestaltning av hur denna plott kan se ut och vad som är relevant att titta efter återfinns i Figur 2.<sup>29</sup>



**Figur 2.** Bilden representerar ”Residual vs. Fitted Values”-plottingar av olika typer av data och hämtades ur Julian J Faraways *Practical Regression and Anova using R* från 2002.<sup>30</sup>

<sup>27</sup> Harald Lang, *Topics on Applied Mathematical Statistics* (Stockholm: Kungliga Tekniska Högskolan, 2013), pp. 32-33.

<sup>28</sup> Peter Kennedy, *A Guide to Econometrics*, 6th edn (United States of America: Blackwell Publishing Ltd, 2008), pp. 112-16.

<sup>29</sup> Julian J Faraway, *Practical Regression and Anova using R* ([n.p]: [n.pub.], 2002), p. 81.

<sup>30</sup> Ibid.

## 2.5.2. Åtgärd 1 - Transformeringsmetoden

Det tredje av de fem grundantagandena för regressionsmodellen under Kapitel 2.1. förutsätter att residualerna är normalfördelade och deras varians likformig. För att undersöka om så är fallet måste först en vanlig regressionsanalys göras på den råa observationsmängden. Residualerna plottas för att avgöra om mönstret är slumpmässigt eller inte. Visar det sig att mönstret är slumpmässigt, det vill säga jämnt utspritt över intervallet, går det att använda modellen (se exempelplottar i Figur 2). Om det däremot visar sig att mönstret inte är slumpmässigt i residualplottarna (se Figur 2) tyder det på en avvikelse från grundantagandena. Residualernas varians är således inte likformig och någon form av transformeringsmetod bör väljas för att skriva om observationsmängd så att den blir bättre lämpad för linjär regressionsanalys.

Transformeringsmetoden av den råa observationsmängden kan tillämpas på den oberoende variabeln, den beroende variabeln, eller båda. När transformeringen har genomförts görs en ny regressionsanalys med de transformerade variablerna och om den här regressionsmodellens  $R^2$  är högre än den ursprungliga modellen, med den otransformerade observationsmängden, är transformeringen lyckad. I sedvanlig ordning reflekterar  $R^2$ -värdet den uppställda modellens förklaringsgrad, som rimligen bör vara högre vid tillämpningen av en bättre lämpad modell, för mer om  $R^2$ -värden, läs stycke 2.6.1. om  $R^2$ -värde och justerat- $R^2$ -värde. Valet av transformeringsmetod beror på den råa observationsmängdens natur och det transformeringen bidrar med är att underlätta användandet av linjär regression med en observationsmängd som inte ursprungligen uppfyller de fem grundantagandena.<sup>31</sup>

I specialfallet när residualerna inte har likformig varians finns det olika sätt att transformera datan för att få bättre residualplottar.

1. Log(Y): Logaritmering av Y-variabeln är en metod som stabiliserar residualernas varians. Denna transformering tillämpas med fördel när spridningen av Y-värdena är mycket större bland stora Y-värden än bland små.<sup>32</sup>
2. BoxCox-transformering av Y: används för att transformera en icke-normalfördelad observationsmängd till en normalfördelad. Originalversionen av BoxCox ser ut som Formel (10).<sup>33</sup>

$$y(\lambda) = \begin{cases} \frac{y^\lambda - 1}{\lambda}, & \text{if } \lambda \neq 0; \\ \log y, & \text{if } \lambda = 0. \end{cases} \quad (10)$$

3.  $Y^{0,25}$ : En fjärdedels roten ur är rekommenderat att använda när ursprungsfördelningen är gravt positivt skev.<sup>34</sup>

Alla tre transformeringsmetoder ovan har gemensamt att tal mellan 0 och 1 hanteras annorlunda jämfört med över 1, varför en fördelning som ska transformeras borde ankras vid 1 eller högre.<sup>35</sup> Box-

<sup>31</sup> stattrek.com, 'Transformations to Achieve Linearity', in *Stat Trek* <<http://stattrek.com/regression/linear-transformation.aspx>> [accessed 19 May 2014]

<sup>32</sup> Oxford Journals, 'Regression Diagnostics', *Our Journals*, 5 (2014), 49-51.

<sup>33</sup> Jason W. Osborne, 'Improving your data transformations: Applying the Box-Cox transformation', *Practical Assessment, Research & Evaluation*, 15 (2010), 2-9.

<sup>34</sup> Phil Ender, 'Linear Statistical Models: Regression', in *Philender* <<http://www.philender.com/courses/linearmodels/notes1/trans1.html>> [accessed 24 May 2014]

Cox och logaritmfunktionen är bara definierad på intervallet strikt större än noll, vilket är ytterligare är ett argument för att addera 1 vid sådana observationsmängder.  $Y$ -vektorn adderas i enlighet med Formel (11).

$$y_i = y + 1, \text{ för } i = 1, 2, \dots, n. \quad (11)$$

### 2.5.3. Åtgärd 2 – White's Consistent Variance Estimator

I de fall där heteroskedasticitet återfinns och transformering av modellen inte fullt ut lyckats råda bot på detta fenomen bör man använda sig av "White's Consistent Variance Estimator". Whites konsekventa varians estimator är, som namnet antyder, en konsekvent estimator för kovariansmatrisen i det heteroskedastiska fallet. Estimatorn återfinns i Formel (12).

$$\begin{aligned} \text{Cov}(\hat{\beta}) &= (X^t X)^{-1} X^t D(\hat{e}^2) X (X^t X)^{-1} \\ &= (X^t X)^{-1} \left( \sum_{i=1}^n \hat{e}_i^2 x_i^t x_i \right) (X^t X)^{-1} \end{aligned} \quad (12)$$

Där  $D(\hat{e}^2)$  är en  $n \times n$ -diagonalmatrix vars  $i$ :te element är  $\hat{e}_i^2$ . De flesta regressionsprogram är försedda med en inbyggd funktion för att kunna lägga till valet "robust" i regressionen, vilket innebär att regressionen beräknas med Whites konsekventa varians estimator. Det är alltid rekommenderat att använda Whites konsekventa varians estimator när möjlighet finns.<sup>36</sup>

### 2.5.4. Hypotesprövning då heteroskedasticitet identifierats

Som nämnts i Kapitel 2.5. om Heteroskedasticitet invalideras F-test när residualerna är heteroskedastiska och man vill testa en linjär restriktion som innehåller mer än ett  $\beta$ -värde.<sup>37</sup> För att kunna utföra hypotesprövning kan Wald-test användas. Wald-test använder sig av Whites konsekventa varians estimator och beräknas på följande vis i de två skilda fallen vid testning av dels en grupp koefficienter, dels en enskild koefficient:

Fall 1: Vi vill testa att en grupp av koefficienter i den linjära regressionen är lika med noll och av den anledningen benämns denna grupp för  $\beta_2$  och övriga koefficienter  $\beta_1$ , enligt Formel (13).

$$Y = x_1 \beta_1 + x_2 \beta_2 + e \quad (13)$$

Hypotesen  $H_0$  ställs upp och testas om  $\beta_2 = 0$ . I Formel (12) är  $x_1$  och  $x_2$  radmatriser och  $\beta_1$  och  $\beta_2$  kolumnmatriser.  $\beta_2$  består av "r" parametrar motsvarande antalet  $\beta$ :an som tillsammans testas mot att de är lika med noll (om antalet  $\beta$ :an är lika med 3 blir således  $r = 3$ ).

<sup>35</sup> Jason W. Osborne, 'Improving your data transformations: Applying the Box-Cox transformation', *Practical Assessment, Research & Evaluation*, 15 (2010), 2-9.

<sup>36</sup> Harald Lang, *Topics on Applied Mathematical Statistics* (Stockholm: Kungliga Tekniska Högskolan, 2013), pp. 34-35.

<sup>37</sup> Ibid. p. 34.



Vidare är  $\hat{V}_2$  den estimerade kovariansmatrisen för  $\hat{\beta}_2$ , enligt kovariansmatrisen i Formel (12), vilket innebär att de ”r” sista kolumnerna och de ”r” sista raderna i den estimerade kovariansmatrisen behålls och de andra tas bort. Wald-testet utförs slutligen enligt Formel (14).

$$W = \hat{\beta}_2^t \hat{V}_2^{-1} \hat{\beta}_2 \quad (14)$$

Formel (14) har en approximativ (asymptotisk) fördelning enligt  $\chi^2(r)$  under nollhypotesen. I detta fall kan även F-fördelningen ( $F(r, n - k - 1)$ , där  $n$  är antalet observationer och  $k$  är antalet koefficienter i den fulla modellen) användas på  $r^{-1}W$ .

Fall 2: Vi vill testa en enskild koefficient,  $\hat{\beta}_n$ . Detta beräknas på följande sätt:

Först beräknas Whites standardavvikelser. Detta görs genom att köra den fulla modellen och spara dess residualer  $\hat{e}_i$ . Därefter körs regressionen av  $x_n$  (som är en enskild variabel) på de andra koefficienterna enligt Formel (15).

$$x_n = x_1 \gamma + u \quad (15)$$

Där  $x_1$  är en radmatris och  $\gamma$  en kolumnmatris. Residualerna sparas i  $\hat{u}_i$ . Whites estimat av standardavvikelsen för  $\hat{\beta}_n$  beräknas slutligen som Formel (16).

$$SE(\hat{\beta}_n) = \frac{\sqrt{\sum(\hat{u}_i \hat{e}_i)^2}}{\sum \hat{u}_i^2} \quad (16)$$

Denna standardavvikelse kan sedan användas i F-testet, i enlighet med Formel (18).<sup>38</sup>

### 2.5.5. F-test

För hypotestest där  $\hat{\beta}_j$  ska testas vara lika med ett värde, kan F-test användas. Testet ställs upp enligt Formel (17) och  $F$  beräknas enligt Formel (18).

$$H_0: \hat{\beta}_j = \beta_j^0 \quad (17)$$

$$F = \left( \frac{\hat{\beta}_j - \beta_j^0}{SE(\hat{\beta}_j)} \right)^2 \quad (18)$$

Där P-värdet för hypotesen är  $\Pr(X > F)$  och där  $X$  följer F-fördelningen;  $F(1, n - k - 1)$ .  $n$  står för antalet observationer och  $k$  för antalet förklarande variabler.<sup>39</sup>

## 2.6. Mått av regressionsmodellens passform

Det går att mäta regressionsmodellens ”goodness of fit” på olika sätt, vilka är en del av regressionsdiagnostiken. Vilket tillvägagångssätt som fungerar bäst beror på syftet med modellen. Två vanligt förekommande mått är  $R^2$ , och dess tillhörande justerade- $R^2$ , samt RMSE, vilket står för

<sup>38</sup> Harald Lang, *Topics on Applied Mathematical Statistics* (Stockholm: Kungliga Tekniska Högskolan, 2013), pp. 34-35.

<sup>39</sup> Ibid. p. 23.

kvadratroten ur medelvärde av den kvadrerade residualen, fritt översatt till svenska. Ursprungsnamn på måttet är ”Root Mean Square Error”, i arbetet hädanefter kallat RMSE.

### 2.6.1. $R^2$ -värde och justerat- $R^2$ -värde

$R^2$  är ett relativt mått på hur pass bra regressionsmodellen är konstruerad, ett mått på så kallad ”goodness of fit”. Det är ett sätt att avgöra modellens validitet.  $R^2$  beräknas enligt Formel (19).<sup>40</sup>

$$R^2 = \frac{Var(X\hat{\beta})}{Var(Y)} = 1 - \frac{Var(\hat{\beta})}{Var(Y)} \quad (19)$$

Utöver  $R^2$ -värde finns det även ett justerat- $R^2$ -värde som ofta betecknas som  $\bar{R}^2$ . Då  $R^2$  inte kan sjunka i och med tillägget av fler oberoende variabler har det justerade- $R^2$ -värdet inkluderats för att minimera risken att onödiga oberoende variabler förs in i modellen. Det justerade- $R^2$ -värdet tar således hänsyn till frihetsgraden i regressionsmodellen.<sup>41</sup>

### 2.6.2. RMSE – Kvadratroten ur medelvärde av den kvadrerade residualen

RMSE är ett absolut mått av regressionsmodellens passform mot en observationsmängd, vilket innebär att det indikerar hur nära de observerade värdena är modellens förutspådda värden. I matematiska termer är RMSE roten av en varians<sup>42</sup> och beräknas som Formel (20).<sup>43</sup>

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (20)$$

RMSE har samma enhet som den förklarande variabeln och är bättre, ju lägre värde måttet har. Denna typ av analys av passformen av regressionsmodellen används speciellt för att avgöra hur pass bra modellen är lämpad för prediktion.<sup>44</sup> Om RMSE i stora drag är samma för stickprov som är tagna både ur den observationsmängd som modellen baseras på, ett så kallat in sample, och ett representativt stickprov utom den mängden, ett så kallat out sample, är modellen konsekvent. Om det är mycket lägre inom än utom observationsmängden antyder det att modellen är bra anpassad efter just den observationsmängden, och kan verka bättre än kvaliteten på de prediktioner som den faktiskt kan producera.

---

<sup>40</sup> Harald Lang, *Topics on Applied Mathematical Statistics* (Stockholm: Kungliga Tekniska Högskolan, 2013), p. 23.

<sup>41</sup> Peter Kennedy, *A Guide to Econometrics*, 6th edn (United States of America: Blackwell Publishing Ltd, 2008), p. 79.

<sup>42</sup> Karen, 'Assessing the Fit of Regression Models', in *The Analysis Factor* <<http://www.theanalysisfactor.com/assessing-the-fit-of-regression-models/>> [accessed 20 May 2014]

<sup>43</sup> Kaggle Inc, 'Root Mean Squared Error (RMSE)', in *Kaggle* <<https://www.kaggle.com/wiki/RootMeanSquaredError>> [accessed 23 May 2014]

<sup>44</sup> Karen, 'Assessing the Fit of Regression Models', in *The Analysis Factor* <<http://www.theanalysisfactor.com/assessing-the-fit-of-regression-models/>> [accessed 20 May 2014]

### 3. Mediamätning i Skandinavien

Analysen i detta kandidatexamensarbete bygger på MMS-data. MMS står för Mediamätning Skandinavien och MMS etablerade sig på den svenska marknaden 1993<sup>45</sup> och har genomfört statistiskt säkerställda mätningar av Sveriges TV-tittande sedan dess.<sup>46</sup> Datinhämtningen går till så att en tittarpanel, representativ för den svenska TV-befolkningen, har elektroniska mätare hemma hos sig som registrerar all TV-trafik per automatik, såkallade People Meters. MMS är ett vedertaget verktyg inom branschen, enligt Andreas Nordfors, analytiker på Nepa. MMS används i många sammanhang. Exempelvis använder sig Konkurrensverket av MMS vid sina undersökningar.<sup>47</sup>

#### 3.1. Insamling av TV-tittarhistorik

MMS tittarpanel består av 1200 rapporterande hushåll motsvarande cirka 2600 personer. Hur panelens sammansättning ser ut avgörs utifrån resultaten från en basundersökning som genomförs av MMS två gånger om året. Dessa undersökningar fungerar som rekryteringskälla för panelhushållen och omfattas av totalt 9000 intervjuer per år. När en basundersökning skett uppdateras panelens sammansättning och universumtalen som ligger till grund för beräkningen av tittarsiffrorna justeras för att stämma överrens med den nya tittarpanelen. För att säkerställa att datan som samlas in från de olika panelhushållen stämmer görs kontinuerligt telefonuppföljningar för att studera hur väl tittardata överensstämmer med panelmedlemmarnas inloggningar. Tittande registreras inte i de fallen då tittande sker på en icke identifierad kanal, panelindividerna har glömt logga in sig som TV-tittare eller om demografiska data gällande tillfälliga gäster saknas. Utöver detta sker dagligen kontroller som säkerställer att paneldeltagarna registrerar sitt TV-tittande på rätt sätt.

Själva People Metern fungerar på så vis att varje panelhushåll har en People Meter per TV-apparat. Till varje People Meter hör en fjärkontroll via vilken hushållsmedlemmarna var och en loggar in med en personlig knapp. Om hushållet har några tillfälliga gäster registreras även deras TV-tittande samt kön och ålder. För hushållets ordinarie medlemmar finns, förutom kön och ålder, ytterligare en stor mängd bakgrundsdata inkodade. Det registrerade TV-tittandet överförs sedan varje natt till Nielsen Audience Measurement,<sup>48</sup> som har i uppdrag av MMS att sköta hanteringen av panelen och insamlingen av data från People Meter-mätarna,<sup>49</sup> för att sedan leverera det till MMS.

##### 3.1.1. Metod för hanteringen av MMS-datan

De observationer som ligger till grund för regressionsmodellen i detta kandidatexamensarbete utgörs av cirka 6000-7000 observationer per månad. November har valts som en representativ månad för normal-tittar-beteende (för motivering se Kapitel 1.4.3. om Dataavgränsningar) och data som spänner

---

<sup>45</sup> MMS, 'Om MMS - Historik' <<http://www.mms.se/start/om-mms/historik/>> [accessed 06 April 2014]

<sup>46</sup> MMS, 'Om MMS' <<http://mms.se/start/om-mms/>> [accessed 06 May 2014]

<sup>47</sup> Konkurrensverket, 'Rörlig Bild - En Analys av Marknaden för Rörlig Bild' (Report, Myndigheten för Radio och TV, Konkurrensverket, 2012), p. 15.

<sup>48</sup> MMS, 'Våra tjänster - Mediamätning - Så mäts TV-tittandet' <<http://mms.se/start/vara-tjanster/mediamatning/sa-mats-tv-tittandet/>> [accessed 06 May 2014]

<sup>49</sup> MMS, 'Glossary - Nielsen Audience Measurement' <<http://mms.se/glossary/nielsen-audience-measurement/>> [accessed 06 May 2014]

över fem år har samlats in. Det har resulterat i att analysen bygger på upp emot 21 000 observationer för november månad från 2009, 2012 och 2013. 2011 visade sig vara ett avbrott i den annars uppåtstigande trenden i antalet TV-tittare, varför det året uteslöts ur regressionsmodellen för säkerhets skull. Detta år kom istället att användas som ett stickprov mot vilket modellens prediktion testas.

Själva datan hämtades ur MMS egna program HotProg. Problem uppstod när datan skulle anpassas för regressionen, eftersom dess inputs är siffror, medan HotProgs output främst är text. Outputen från HotProg var en Excel-fil med kategorierna: Kanal, Datum, Starttid, Sluttid, Längd, Beskrivning, Veckodag, Genre, A3-99 Rat [L] %, A15-49 Rat [L] %, M15-34 Rat [L] %, M25-49 Rat [L] %, K25-49 Rat [L] %, K25-44 Rat [L] %, för förklaring av benämningarna se Tabell 3. Då uppdelningen såg ut på detta vis har det krävts omfattande bearbetning av datan för att kunna anpassa den till regressionsmodellen. Genre har behövt översättas, och i många fall omdefinierats, till 34 programkategorier. Dessa 34 kategorier återfinns i Bilaga 6. och har definierats helt självständigt av författarna. Varje enskild sändning har kategoriseras för sig och denna kategorisering måste givetvis stämma över de tre åren. Detta har genomförts genom att titta på kategorin Beskrivning samt att söka på programmet på [www.tv.nu](http://www.tv.nu) och [www.imdb.com](http://www.imdb.com) för att få underlag för programmets egentliga genre. HotProgs genrebeskrivning utgjordes av många fler kategorier och stämde i många fall inte överens med varken [www.tv.nu](http://www.tv.nu) eller [www.imdb.com](http://www.imdb.com). Därav har kategoriseringen behövts göras utan tidigare underlag. Detta har således varit mycket tidskrävande.

På samma vis har kategorin för kanalerna behövt översättas till dummyvariabler. Detta kunde dock ske på ett mycket mer systematiskt tillvägagångssätt än för programkategorierna då Excels filterfunktion gjorde mestadelen av arbetet. På samma sätt hanterades dummyomskrivningen av dagar i veckan. Program i följd är inte en output ur HotProg, utan har definierats fristående utifrån programbeskrivning, kanal, datum och tid på dygnet. I Tabell 3. nedan följer en förklaring av HotProgs output.

| <b>Namn</b>      | <b>Beskrivning</b>   |
|------------------|--|
| Kanal            | Vilken kanal som sändningen gick på, det vill säga SVT1, SVT2, TV3, TV4, Kanal 5, TV6, TV8 eller TV10.   |
| Datum            | Det datum som sändningen gick.   |
| Starttid         | Tidpunkt när sändningen gick igång.  |
| Sluttid          | Tidpunkt då sändningen slutade.  |
| Längd            | Sändningens längd i minuter.   |
| Beskrivning      | Namnet på sändningen. Till exempel "Vem vet mest?" eller "Gladiatorerna".  |
| Veckodag         | Dag i veckan som sändningen ägde rum.  |
| Genre 1          | Vilken typ av genre som sändningen utgjorde.<br>Exempel är: Långfilm eller Nyheter .   |
| A3-99 Rat [L] %  | Andelen som tittar på en viss sändning i spannet Alla, 3 till 99 år, under minst tre minuter. Anges i procent i intervallet 0-100 med tre decimalers noggrannhet.<br>Exempel: 23,01 motsvarar då att 23,01 % av alla i åldern 3 - 99 år tittade på programmet. |
| A15-49 Rat [L] % | Gestaltas på samma sätt som ovan kategori, dock gäller denna alla mellan 15 till 49 år.  |
| M15-34 Rat [L] % | -"- Gäller män i ålder 15 till 34 år.  |
| M25-49 Rat [L] % | -"- Gäller män i ålder 25 till 39 år.  |
| K25-49 Rat [L] % | -"- Gäller kvinnor i ålder 25 till 49 år.  |
| K25-44 Rat [L] % | -"- Gäller kvinnor i ålder 25 till 44 år.  |

**Tabell 3.** Output från MMS:s HotProg-program.

## 4. Metod

Regressionsanalys av MMS-data har inte gjorts tidigare, enligt Andreas Nordfors, analytiker på Nepa. Av den anledningen finns inget tidigare underlag för hur en sådan modell borde se ut. Med andra ord har modellen i det är kandidatexamensarbetet formulerats helt fristående.

$Y$ -variabeln i ekvationen har ställts upp utifrån andelen tittare i procent i en viss målgrupp. Anledningen till att det har valts som  $Y$ -variabel är för att kunna urskilja vilka typer av program, vilka dagar i veckan och vilka kanaler som attraherar en viss målgrupp, som således ställs upp som förklaringsvariabler. Utöver dessa förklaringsvariabler läggs programföljd och –längd till för att se hur de påverkar tittarna. De här faktorerna bygger upp kanalernas TV-tablåer och är följaktligen faktorer som TV-bolagen kan påverka själva, varför dessa är intressanta.

### 4.1. Grundekvation

Utifrån hur MMS-data är upplagd och tack vare att Nepa formulerade målgrupper för kanalerna av intresse, föranledde det en formulering av regressionsmodellen enligt Formel (21).

$$\begin{aligned} Y = & \beta_0 + \text{Längd} * \beta_1 + \text{Programkategorier} * (\beta_2, \beta_3, \dots, \beta_{44})^T + \\ & + \text{Veckodagar} * (\beta_{45}, \beta_{46}, \dots, \beta_{50})^T + \\ & + \text{Kanaler} * (\beta_{51}, \beta_{52}, \dots, \beta_{57})^T + \text{Programföljd} * \beta_{58} + e \end{aligned} \quad (21)$$

Nedan följer en förklaring av variablerna i den linjära regressionsmodellen:

- $Y$  – motsvarar den oberoende variabeln och är andelen tittare i procent i en viss målgrupp som ser på ett TV-program och är därför en kontinuerlig variabel mellan 0 och 100 och mäts med tre decimalers noggrannhet.
- $Längd$  – står för TV-programmets längd i minuter och är en kontinuerlig variabel.
- $Programkategorier$  – utgörs av en radvektor med 42 stycken av totalt 43 stycken olika indelningar av programmen i TV-tablåerna och är således dummyvariabler i regressionen. För alla programkategorier se Bilaga 6.
- $Veckodagar$  – en radvektor som består av måndag, onsdag, torsdag, fredag, lördag och söndag och även dessa utgörs av dummyvariabler.
- $Kanaler$  – en radvektor som avgränsas till kanalerna SVT1, TV3, TV4, TV5, TV6, TV8 och TV10, som representeras av dummyvariabler i regressionen.
- $Programföljd$  – representerar programmet efter det aktuella programmet i regressionen och består av en dummyvariabel. Är det samma program innan ett program, exempelvis en efterföljande episod i en TV-serie, motsvarar det dummyvariabeln ett, medan om det inte är samma program innan motsvaras det av dummyvariabeln noll.

Benchmark i regressionsmodellen väljs som Nyheter, tisdagar, SVT2 och programföljden noll, det vill säga att det inte är två på varandra efterföljande avsnitt av samma program.

## 4.2. Splittest

I den första regressionen med ovan grundekvation togs det inte hänsyn till tid på dygnet över huvudtaget och som väntat resulterade det i ett  $R^2$  på 0,25, vilket är lågt. Det är väntat att tiden på dygnet påverkar tittarna. Exempelvis borde det vara många färre under natten. Tidsaspekten kan fångas upp i modellen med olika tillvägagångssätt.

Tid kan tas hänsyn till genom att antingen införa ytterligare en dummykategori, som representerar olika tidsspann, eller genom att dela upp observationerna, ett så kallat splittest. Då införandet av en ny dummykategori, i form av tid, skulle kräva att korsvariabler mellan de olika tidsspannen och programkategorierna införs, leder det till en onödigt komplex och lång modell. Utöver detta har införandet av många förklaringsvariabler en tendens att blåsa upp  $R^2$ . Således var det bättre att utföra sex regressioner (vilket motsvarar en regression för varje målgrupp) för vart och ett av de sex tidsspannen, och på så vis fånga upp de olika drivande faktorerna bakom tittarbeteendet, än att endast införa dummyvariabler i den ursprungliga regressionsmodellen. På så sätt undviks också multikollinearitet, eftersom tid på dygnet kan antas vara korrelerat med vissa typer av program.

Tidsspannen inom vilka det antogs att det var likartat TV-tittande är som i Tabell 4. nedan.

|                        |             |
|------------------------|-------------|
| Morgon-TV-passet       | 06:00-08:59 |
| Dags-TV-passet         | 09:00-12:59 |
| Eftermiddags-TV-passet | 13:00-15:59 |
| Förkvälls-TV-passet    | 16:00-19:59 |
| Prime Time-TV-passet   | 20:00-22:59 |
| Natt-TV-passet         | 23:00-05:59 |

**Tabell 4.** *Tidsuppdelningar av olikartat TV-tittande sinsemellan, men likartat TV-tittande inom, ett spann.*

## 4.3. Heteroskedasticitet och transformering av Y-variabeln

Det observerades på ett tidigt stadium att residualerna var positivt skeva, det vill säga ”positively skewed”, se de fyra näst sista bilderna uppfifrån i Bilaga 1., dels för att alla  $Y$ -värden är positiva, och dels för att majoriteten av  $Y$ -värdena är noll eller nära noll. Att  $Y$ -värdena är noll eller nära noll gällde alla målgrupper i allmänhet och Män, 15-34 år i synnerhet. Det finns olika sätt att gå tillväga för att korrigera problemet. Tre sätt identifierades genom vilka det är möjligt att tillrättalägga problemet och transformera  $Y$ -variabeln:

$$\log(Y)$$

$$Y^{0,25}$$

BoxCox-transformering av  $Y$

Om observationsmängden antar värden mellan 0 och 1, ska dessa ankras om i enlighet med teoridelen. Eftersom  $Y$ -variabeln är definierad i procent som 0 till 100 och inte i decimalform skulle lägre värden än 1 hanteras annorlunda än de över och av den anledningen utfördes alla transformeringarna på  $(Y + 1)$  i enlighet med Tabell 5 och 6. Som nämnts under teoridel 2.5.2. väljs den transformering som bidrar till högst  $R^2$ -värde. Nedan följer modellernas  $R^2$ -värde och justerat  $R^2$ -värde för respektive modell.

| R <sup>2</sup> -värde för regressionerna: | Y     | Log(Y+1) | (Y+1) <sup>0,25</sup> | Boxcox(Y+1) |
|---|-------|----------|-----------------------|-------------|
| A399                                      | 0,555 | 0,676    | 0,652                 | 0,691       |
| A1549                                     | 0,516 | 0,634    | 0,609                 | 0,647       |
| M1534                                     | 0,416 | 0,531    | 0,509                 | 0,546       |
| M2549                                     | 0,465 | 0,585    | 0,561                 | 0,593       |
| K2544                                     | 0,521 | 0,625    | 0,606                 | 0,632       |
| K2549                                     | 0,552 | 0,660    | 0,640                 | 0,667       |

**Tabell 5.** Tabell över R<sup>2</sup>-värden för olika transformeringar för tidsspannet 20:00-22:59.

| Justerat-R <sup>2</sup> -värde för regressionerna: | Y     | Log(Y+1) | (Y+1) <sup>0,25</sup> | Boxcox(Y+1) |
|--|-------|----------|-----------------------|-------------|
| A399   | 0,548 | 0,671    | 0,647                 | 0,686       |
| A1549  | 0,509 | 0,628    | 0,603                 | 0,642       |
| M1534  | 0,407 | 0,524    | 0,509                 | 0,54        |
| M2549  | 0,457 | 0,579    | 0,555                 | 0,586       |
| K2544  | 0,514 | 0,620    | 0,600                 | 0,626       |
| K2549  | 0,546 | 0,655    | 0,635                 | 0,662       |

**Tabell 6.** Tabell över justerade-R<sup>2</sup>-värden för olika transformeringar för tidsspannet 20:00-22:59.

$Boxcox(Y + 1)$  gav någorlunda bättre R<sup>2</sup>-värden än  $Log(Y + 1)$ , medan  $(Y + 1)^{0,25}$  genomgående resulterade i sämre R<sup>2</sup>-värden. Skillnaden i R<sup>2</sup>-värdena mellan  $Boxcox(Y + 1)$  och  $Log(Y + 1)$  är marginell. Utöver det går kursen, som ligger till grund för det här kandidatexamensarbetet, Tillämpad Matematisk Statistik med kurskod SF2950 på KTH, igenom logaritmttransformering, men inte de andra två transformeringarna som var lämpliga i det här fallet. Av dessa anledningar, tillsammans med att logaritmeringstransformering är mer intuitivt tolkningsbar, valdes logaritmering som transformeringsmetod i regressionsmodellen. För att sedan transformera tillbaka till rätt enhet används  $e^{X\hat{\beta}} - 1$ .

Som kan ses i de sista bilderna uppifrån i Bilaga 2., vilket motsvarar ”Residual vs Fitted Values”-plotten för  $Log(Y + 1)$ -modellen, finns fortfarande en ”konformig” trend hos residualernas varians, även om de är mycket bättre efter logaritmeringen. Eftersom problemet kvarstod mer eller mindre, bestämdes det att regressionen skulle köras i robust form i enlighet med Whites konsekventa varians estimator. Detta gjordes genom att första använda *LinearModel.fit* i Matlab och därefter beräkna standaravvikelsena för  $\beta$ -värdena genom att använda funktionen *hac* (för exakt kod, se Bilaga 7.), som använder sig av just Whites konsekventa varians estimator, se Formel (12) under teoridelen.

Tack vare att residualernas fördelning blev mer lik en normalfördelning blev det en tydlig förbättring av plottarna efter logaritmeringen av  $Y$ -variabeln (se bilderna två och tre uppifrån i Bilaga 2. i jämförelse mot bilderna två och tre uppifrån i Bilaga 1.). Dock gäller det endast i splittestet som utgörs av tidsspannet 20:00-22:59 och således beslutades det att övriga tidsspann inte är lämpliga att analysera med den uppställda linjära regressionsmodellen i det här kandidatexamensarbetet. Det beror sannolikt på att det under övriga tider på dygnet inte är särskilt många TV-tittare.  $Y$ -variabeln blir således extra tydligt begränsad underifrån (det vill säga, ”positively skewed”) och svår att transformera och ännu svårare att analysera. Resultatdelen kommer därför endast att behandla regressionerna som utfördes i tidsspannet 20:00-22:59.

|                        |             |
|------------------------|-------------|
| Morgon-TV-passet       | 06:00-08:59 |
| Dags-TV-passet         | 09:00-12:59 |
| Eftermiddags-TV-passet | 13:00-15:59 |
| Förkvälls-TV-passet    | 16:00-19:59 |
| Prime Time-TV-passet   | 20:00-22:59 |
| Natt-TV-passet         | 23:00-05:59 |

**Tabell 7.** *Det enda tidsspannet som lämpar sig för regressionsmodellen, enligt grundekvationen i Formel (21).*

#### 4.4. Multikollinearitet

När beslutet fattats att resultatdelen enbart kommer att behandla splittestet 20:00-22:59 (se anledning under 4.2 Transformering av Y-variabel) påbörjades undersökningen efter multikollinearitet bland förklaringsvariablerna. Till att börja med uteslöts benchmark för respektive dummykategori bland förklaringsvariablerna. Därefter beräknades konditionsindex och då tumregeln är att det högsta konditionsindexet ska vara under 30 och det högsta indexet för förklaringsmatrisen låg på 9,83 så är nivån av multikollinearitet förbisebar (se teorin i Kapitel 2.3.1. om Upptäckt). Fullständig redogörelse av konditionsindexet för förklaringsmatrisen återfinns i Bilaga 8.

För att vidare undvika multikollinearitet hämtades tittarhistorik från tre år under november månad. De valda åren var 2012 och 2013, då dessa ligger närmast i tiden, samt 2009, då det ger en större spridning i observationerna att också gå något längre bak i tiden. Det bör understrykas att detta inte har gjorts för att analysera över tid, utan för att få ett bredare spann av olika programkategorier på olika dagar och tidpunkter. Under en månad återkommer många program på samma tidpunkt, samma dag och samma kanal. Genom att hämta tittarhistorik från olika år fås en bättre spridning eftersom TV-tablåerna inte ser likadana ut. I den utsträckning det går undviks på så sätt multikollinearitet med fler, men också mer varierade observationer.

#### 4.5. Endogeneitet

Det antas att endogeneitet inte finns i de observationer som detta kandidatexamensarbete grundar sin regressionsanalys på. Detta antagande görs eftersom de paneler som observationerna hämtats ifrån ska vara representativa för populationen av TV-tittare i Sverige. Systematiska mätfel anses inte heller finnas i dessa observationer, då MMS, som sköter denna observationsinsamling, har stora kontrollprocesser kring vad för typ av data som tas med och vilka som är med i panelerna samt att de som är med i panelerna registrerar sitt TV-tittande på rätt sätt. För mer djupgående genomgång om observationsinsamlingsprocessen se Kapitel 3.1. om Insamling av TV-tittarhistorik.

I den mån det går har även de förklaringsvariabler som kan antas vara relevanta för att förklara andelen tittarsiffror i procent i en viss målgrupp tagits med i regressionsmodellen. Att observationerna skulle vara simultana kan också uteslutas, då en viss typ av tittarsiffror inte driver en viss programkategori, dag i veckan, en viss programföljd eller kanal, utan tvärtom. Att höga tittarsiffror skulle driva längre program kan även det antas vara inkorrekt.



## 4.6. Omskrivning av förklaringsmatrisen

Förklaringsmatrisen utgjordes från början av 42 programkategorier, 6 veckodagar, 7 kanaler, en variabel för programlängd och en variabel för de fall då programmet i observationen sändes direkt följande efter ett likadant program. Då det saknades observationer i vissa av programkategorierna i splittestet 20:00-22:59, eftersom alla program inte går alla tider på dygnet, togs dessa kategorier bort. Även kategorier med färre än 20 observationer plockades bort, eftersom det inte skulle finnas tillräckligt med underlag för estimeringen av  $\beta$ -värdena. 20 är en godtycklig siffra i sammanhanget, men det viktigaste är att de programkategorier som endast har några få observationer utesluts ur modellen på grund av för lite variation i den befintliga observationsmängden, enligt Kapitel 2.3.2. om Åtgärd vad gäller multikollinearitet. Det blev till slut 27 programkategorier kvar. Den slutgiltiga förklaringsmatrisen bestod således av 42 förklaringsvariabler och den slutgiltiga regressionsmodellen avviker således från grundekvationen enligt Formel (21). Det som skiljer sig från grundekvationen är dels att  $\beta$ -värdena är färre till antalet för programkategorierna och att transformering av  $Y$  till  $\text{Log}(Y + 1)$  har genomförts.

## 4.7. Matlab

Regressionsmodellen ställdes upp och beräknades i Matlab. De funktioner som användes för beräkning av den linjära regressionsmodellen är *LinearModel.fit* och funktionen *hac* användes eftersom heteroskedasticitet upptäckts i datan. Det *hac* gör är att beräkna Whites konsekventa variansestimater i enlighet med Formel (12). Detta användes för att ta fram nya standardavvikelser för  $\beta$ -värdena och för att kunna utföra Wald-test i enlighet med Formel (14). För de olika plottarna används följande funktioner i Matlab:

- *qqplot(residual)*: Plottar residualerna mot en teoretisk normalfördelad kvantitet. Om residualplottarna ligger väl längs linjen kan residualerna antas vara normalfördelade.<sup>50</sup>
- *histfit(residual)*: Visar residualernas fördelning relativt en normalfördelningskurva.<sup>51</sup>
- *plotResiduals(mdl,'fitted')*: Plottar residualernas variation utifrån ”Fitted Values”. På så vis kan icke-linjäritet, att residualerna inte har samma varians, och ”outliers” upptäckas.<sup>52</sup>

För beräkning av konditionsindex används funktionen *collintest* och för logaritmering av  $Y$ -variablerna användes *log*. För fullständiga Matlab-kod se Bilaga 7.

---

<sup>50</sup> Inc. The MathWorks, 'qqplot', in *MathWorks* <<http://www.mathworks.se/help/stats/qqplot.html>> [accessed 22 May 2014]

<sup>51</sup> Inc. The MathWorks, 'histfit', in *MathWorks* <<http://www.mathworks.se/help/stats/histfit.html>> [accessed 22 May 2014]

<sup>52</sup> The Pennsylvania State University, 'A residuals vs. fits plot', in *Online Courses* <<https://onlinecourses.science.psu.edu/stat501/node/36>> [accessed 22 May 2014]

## 5. Resultat

Observera att genomgående i hela resultatdelen redovisas endast regressionsresultat och tillhörande test för tidsspännat 20:00-22:59, se metoddel Kapitel 4.3. för motivering.

RMSE har beräknats för regressionensmodellen med såväl out sample som flera in samples. Som gicks igenom i teoridelen i Kapitel 2.6.2. om RMSE och som följer i Tabell 8. innebär in sample stickprov från den ursprungliga observationsmängden, vilka  $\beta$ -värdena genererats ifrån, och out sample förslagsvis ett godtyckligt år, som inte tagits med i regressionsmodellen tidigare. Stickprovet som utgörs av det godtyckliga året är 2011, där "Typ av sample" är benämnt som "Out" i Tabell 8. För in sample görs RMSE-beräkningen på dragningar, som varje gång slumpmässigt tar ut 5% av in sample-observationerna och testar de skattade  $\beta$ -värdena på dem. Alla år representeras av månaden november.

| Målgrupp:         |           |               | Alla,<br>3-99 år | Alla,<br>15-49 år | Kvinnor,<br>25-44 år | Kvinnor,<br>25-49 år | Män,<br>15-34 år | Män,<br>25-49 år |
|-------------------|-----------|---------------|------------------|-------------------|----------------------|----------------------|------------------|------------------|
| Data från år      | Omfång    | Typ av sample | RMSE             | RMSE              | RMSE                 | RMSE                 | RMSE             | RMSE             |
| 2011              | Hela året | Out           | 1,9726           | 1,6264            | 2,4189               | 2,4421               | 1,1942           | 1,7979           |
| 2009, 2012 & 2013 | 5%        | In            | 2,0205           | 1,9344            | 2,6632               | 2,4943               | 1,4608           | 2,0179           |
| 2009, 2012 & 2013 | 5%        | In            | 2,1092           | 2,3743            | 3,1141               | 2,7246               | 1,5947           | 2,4726           |
| 2009, 2012 & 2013 | 5%        | In            | 2,2141           | 2,6408            | 2,7993               | 2,5507               | 1,3608           | 2,594            |
| 2009, 2012 & 2013 | 5%        | In            | 2,686            | 1,5949            | 2,8811               | 2,4743               | 2,1026           | 1,9383           |
| 2009, 2012 & 2013 | 5%        | In            | 2,97             | 1,8411            | 2,9043               | 2,6709               | 1,6158           | 2,8455           |

**Tabell 8.** Tabellen representerar RMSE-beräkningar för modellen på in och out samples.

Den viktigaste slutsatsen av RMSE-beräkningarna ovan är att modellen är relativt konsekvent mellan olika stickprov och inom olika målgrupper. Den förklarande variabeln varierar i genomsnitt som mest mellan noll och 30 procent TV-tittare inom en specificerad målgrupp. Då RMSE ligger på runt 2 procentenheter, lite beroende av målgrupp, kan det konstateras att modellen är bra för prediktion, och framförallt konsekvent mellan både in och out samples. 2011 är, som tidigare nämnts, ett år med avbrott i den stigande TV-tittartrenden, eftersom tidigare år konsekvent ökat i den totala tiden som tittas på TV. Av den anledningen är det extra anmärkningsvärt att RMSE:n är så pass jämn mellan out och in samples.

Utöver RMSE-testen ovan följer de slutgiltiga resultaten för de sex målgrupperna under prime time i kommande kapitel. Plottarna bygger på data som genomgått transformeringen  $\log(Y + 1)$  i regressionsmodellen. Ursprungsplotten, det vill säga plottning av regressionsmodellen utan omtransformering, finns att tillgå i Bilaga 1. Benchmark utgörs av interceptet och för alla regressioner är det SVT2, Nyheter, tisdagar och inte program i följd. Utöver det testades huruvida veckodagarna torsdag, fredag och lördag kunde slås ihop med tisdag i benchmark för alla målgrupper, eftersom dessa dagar saknade signifikans i samtliga regressioner (se respektive Wald-test-tabell nummer 10, 12, 14, 16, 18 och 20 för varje målgrupp). Som synes går det inte att förkasta den hypotesen i något av fallen.

Det är även motiverat att testa dummykategorierna var för sig för att se om hela kategorin har en påverkan på TV-tittandet, eller om den kan uteslutas ur modellen. I regressionstabellerna, 9, 11, 13, 15, 17 och 19 nedan, går det att utläsa på P-värdena att de olika dummykategorierna troligtvis har en statistisk signifikans, men för att säkerställa det tillämpades även Wald-test på de respektive kategorierna enligt tabellerna för varje målgrupp. Som synes i de här fallen går det att förkasta

hypoteserna att de olika dummykategorierna inte skulle ha en inverkan på TV-tittandet. För regressioner, plottar och Wald-test, se efterföljande sex delkapitel.

I regressionstabellerna 9, 11, 13, 15, 17 och 19 är variablerna med statistiskt signifikans fetmarkerade. Utöver dessa återfinns även  $R^2$ -värden och andra mått av relevans för regressionen i tabellerna och är på godtagbara nivåer. I översta plotten i Figur 3, 4, 5, 6, 7 och 8 nedan, ses fortfarande en konformig trend, trots transformeringen av modellen, men som nämnts tidigare har standardavvikelseerna för modellen beräknats med Whites konsekventa varians estimator för att ta hänsyn till denna heteroskedasticitet och för att kunna utföra tillförlitliga hypotestest. När det gäller QQ-plotten, den mellersta bilden i figurerna, och histogramplotten, den nedersta bilden i figurerna, av residualerna mot normalfördelningen kan residualerna antas vara normalfördelade. Vidare har regressionsanalyserna resulterat i tämligen liknande resultat för alla målgrupper, men i den mån de avviker har det kommenterats för varje målgrupp och annars hänvisats till den första målgruppen med lika resultat. Resultatet rörande programkategorierna och kanalerna diskuteras ingående under diskussionsdelen.

## 5.1. Alla, 3 – 99 år, kl. 20:00-22:59

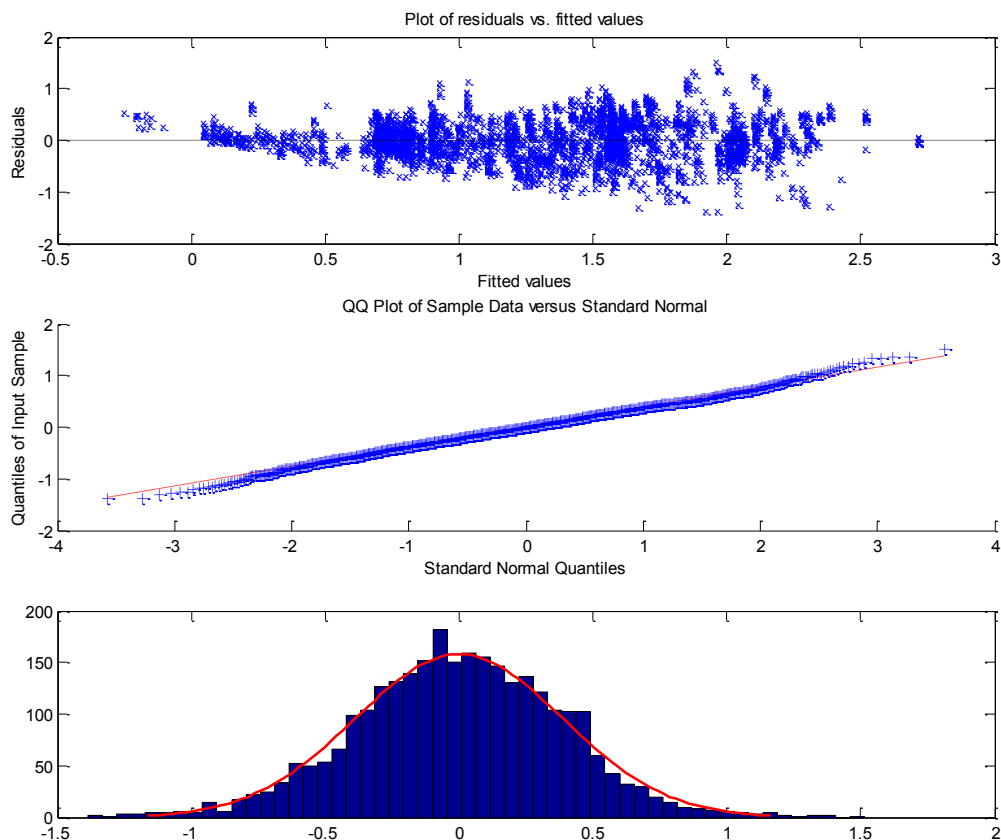
Regressionsanalysen för målgruppen Alla, 3-99 år, följer nedan och de viktigaste resultaten sammanfattas och diskuteras vad gäller de respektive  $\beta$ -värdena under diskussionsdelen i Kapitel 6.-6.7. Just programkategorierna och kanalernas påverkan på tittarsiffrorna diskuteras inte under detta kapitel, utan under diskussionskapitlet.

| Regression för målgruppen Alla, 3-99 år* |                 |                |                |                | Fortsättning av Regressionen* |                 |                |                |                |
|--|-----------------|----------------|----------------|----------------|-------------------------------|-----------------|----------------|----------------|----------------|
| Alla, 3-99 år                            | Estimat         | White SE       | F-värde        | P-värde        | Alla, 3-99 år                 | Estimat         | White SE       | F-värde        | P-värde        |
| I (Intercept)                            | 1,55            | 2,81E-02       | 3,07E+03       | 0              | K Svt1                        | <b>0,302633</b> | <b>0,04164</b> | <b>52,8195</b> | <b>4,7E-13</b> |
| P Debatt                                 | <b>-0,23773</b> | <b>0,04295</b> | <b>30,6409</b> | <b>3,4E-08</b> | a TV3                         | <b>-0,462</b>   | <b>0,04271</b> | <b>117,022</b> | <b>0</b>       |
| r Dokumentär                             | <b>-0,71709</b> | <b>0,06643</b> | <b>116,526</b> | <b>0</b>       | n TV4                         | <b>0,445806</b> | <b>0,029</b>   | <b>236,299</b> | <b>0</b>       |
| o Film                                   | <b>-0,33132</b> | <b>0,05724</b> | <b>33,507</b>  | <b>7,9E-09</b> | a Kanal 5                     | <b>-0,40232</b> | <b>0,04291</b> | <b>87,9261</b> | <b>0</b>       |
| g Kriminalmagasin                        | <b>0,192638</b> | <b>0,0692</b>  | <b>7,7491</b>  | <b>0,00541</b> | l TV6                         | <b>-0,61077</b> | <b>0,04147</b> | <b>216,951</b> | <b>0</b>       |
| r Magasin, humor                         | -0,09772        | 0,05514        | 3,14048        | 0,07648        | TV8                           | <b>-1,02106</b> | <b>0,0419</b>  | <b>593,949</b> | <b>0</b>       |
| a Matlagingsprogram                      | -0,09811        | 0,06271        | 2,44819        | 0,11777        | TV10                          | <b>-1,14056</b> | <b>0,04144</b> | <b>757,649</b> | <b>0</b>       |
| m Nyheter, kultur                        | -0,13645        | 0,07505        | 3,30573        | 0,06914        | D Måndag                      | 0,044188        | 0,02415        | 3,34714        | 0,06743        |
| k Nyheter, sport                         | -0,04374        | 0,03747        | 1,36302        | 0,24311        | a Onsdag                      | 0,032131        | 0,02578        | 1,55327        | 0,21276        |
| a Pausprogram                            | <b>-0,91187</b> | <b>0,06799</b> | <b>179,88</b>  | <b>0</b>       | g Torsdag                     | 0,001592        | 0,02617        | 0,0037         | 0,95151        |
| t Sitcom                                 | <b>-0,29941</b> | <b>0,04312</b> | <b>48,2164</b> | <b>4,7E-12</b> | Fredag                        | -0,00393        | 0,02963        | 0,01763        | 0,89439        |
| e Sport/betting                          | -0,12668        | 0,06588        | 3,69728        | 0,0546         | Lördag                        | 0,013637        | 0,03084        | 0,19548        | 0,65843        |
| g Talkshow                               | <b>0,564719</b> | <b>0,08728</b> | <b>41,8629</b> | <b>1,2E-10</b> | Söndag                        | <b>0,101079</b> | <b>0,02963</b> | <b>11,6399</b> | <b>0,00065</b> |
| o Talkshow, kultur                       | <b>-0,56981</b> | <b>0,11189</b> | <b>25,9353</b> | <b>3,8E-07</b> | F Program i följd             | <b>0,064097</b> | <b>0,02324</b> | <b>7,60381</b> | <b>0,00586</b> |
| r Talkshow, diskussion                   | <b>0,214615</b> | <b>0,06344</b> | <b>11,4428</b> | <b>0,00073</b> | L Programlängd                | <b>0,001712</b> | <b>0,00046</b> | <b>13,9739</b> | <b>0,00019</b> |
| i TV-serie, criminal                     | 0,07061         | 0,0471         | 2,24788        | 0,13391        |                               |                 |                |                |                |
| TV-serie, dokumentär                     | <b>-0,42676</b> | <b>0,04381</b> | <b>94,8998</b> | <b>0</b>       |                               |                 |                |                |                |
| TV-serie, drama                          | <b>-0,25331</b> | <b>0,05883</b> | <b>18,5376</b> | <b>1,7E-05</b> |                               |                 |                |                |                |
| TV-serie, historia                       | <b>-0,29857</b> | <b>0,06384</b> | <b>21,8743</b> | <b>3E-06</b>   |                               |                 |                |                |                |
| TV-serie, hjälp                          | <b>0,285926</b> | <b>0,06789</b> | <b>17,7399</b> | <b>2,6E-05</b> |                               |                 |                |                |                |
| TV-serie, inredning                      | 0,007728        | 0,05006        | 0,02383        | 0,87733        |                               |                 |                |                |                |
| TV-serie, kultur                         | <b>-0,40949</b> | <b>0,06002</b> | <b>46,5544</b> | <b>1,1E-11</b> |                               |                 |                |                |                |
| TV-serie, reality                        | 0,015089        | 0,09987        | 0,02283        | 0,87992        |                               |                 |                |                |                |
| TV-serie, tävling                        | <b>0,193108</b> | <b>0,07483</b> | <b>6,65965</b> | <b>0,00991</b> |                               |                 |                |                |                |
| TV-serie, utslagstävling                 | <b>0,149743</b> | <b>0,0496</b>  | <b>9,11396</b> | <b>0,00256</b> |                               |                 |                |                |                |
| TV-serie, yrke                           | <b>-0,32855</b> | <b>0,05901</b> | <b>31,0032</b> | <b>2,8E-08</b> |                               |                 |                |                |                |
| TV-serie, övernaturlig                   | <b>-0,57909</b> | <b>0,0714</b>  | <b>65,7797</b> | <b>0</b>       |                               |                 |                |                |                |
| Värdet                                   | <b>-0,45027</b> | <b>0,04116</b> | <b>119,661</b> | <b>0</b>       |                               |                 |                |                |                |

Antalet observationer: 2857, Frihetsgrader: 2814  
 RMSE: 0,39  
 R<sup>2</sup>: 0,676, Justerat-R<sup>2</sup>-värde: 0,671

\* De som med signifikans skiljer sig från interceptet är fetmarkerade.

Tabell 9. Regressionsanalys med beräknade  $\beta$ -estimat, Whites standardavvikelser, F-värde och P-värde för alla variabler i X-matrisen. Detta för målgruppen Alla, 3-99 år.



**Figur 3.** Översta plotten utgörs av "Residuals Vs. Fitted", mellersta plotten är en QQ-plot och den nedersta plotten visar ett histogram av residualerna mot det röda strecket, vilken utgörs av normalfördelningen.

Programlängd och -följd har båda en signifikant påverkan enligt Tabell 9. i målgruppen Alla, 3-99 år, med 1%-signifikansnivån. Alla dagar utom söndag har inte en signifikant påverkan enligt Tabell 9., och i tabellen nedan följer hypotestestet för de olika dummykategorierna och ett test för att slå ihop torsdag, fredag och lördag med interceptet. Hypotesen att de ovan nämnda dagarna kan slås ihop med tisdagen som benchmark, kan inte förkastas och även om det inte gjorts i den här regressionsanalysen är det ett rekommenderat tillvägagångssätt vid fortsatta studier av TV-tittarvanor.

| Wald-tester för Alla, 3-99 år                            |         |       |                           |
|--|---------|-------|---------------------------|
| Hypoteser  | Utfall  |       | Slutsats                  |
| Programkategori har ingen inverkan                       | W       | 965   | kan förkasta hypotesen    |
|  | P-värde | 0     |                           |
| Dagar i veckan har ingen inverkan                        | W       | 18,6  | kan förkasta hypotesen    |
|  | P-värde | 0     |                           |
| TV-kanalen har ingen inverkan                            | W       | 2720  | kan förkasta hypotesen    |
|  | P-värde | 0     |                           |
| Torsdag, fredag och lördag kan slås ihop med interceptet | W       | 0,330 | kan ej förkasta hypotesen |
|  | P-värde | 0,194 |                           |

**Tabell 10.** Hypotestest vid heteroskedasticitet för målgruppen Alla, 3-99 år.

## 5.2. Alla, 15 – 49 år, kl. 20:00-22:59

Tabellerna för regressionsanalysen för målgruppen Alla, 15-49 år, följer nedan och de viktigaste resultaten sammanfattas och diskuteras vad gäller de respektive  $\beta$ -värdena under diskussionsdelen 6.-6.7. tillsammans med analysen. Utöver detta återfinns även  $R^2$  och andra mått av relevans för regressionen i tabellen nedan.

Regression för målgruppen Alla, 15-49 år\*

| Alla, 15-49 år           | Estimat         | White SE       | F-värde        | P-värde        |
|--------------------------|-----------------|----------------|----------------|----------------|
| l (Intercept)            | 1,11            | 2,74E-02       | 1,65E+03       | 0              |
| P Debatt                 | <b>-0,24382</b> | <b>0,03921</b> | <b>38,6645</b> | <b>5,8E-10</b> |
| r Dokumentär             | <b>-0,47468</b> | <b>0,05906</b> | <b>64,6037</b> | <b>0</b>       |
| o Film                   | -0,05207        | 0,05065        | 1,05677        | 0,30404        |
| g Kriminalmagasin        | <b>0,142663</b> | <b>0,05333</b> | <b>7,15675</b> | <b>0,00751</b> |
| r Magasin, humor         | 0,009307        | 0,05496        | 0,02868        | 0,86553        |
| a Matlagningsprogram     | -0,04922        | 0,05865        | 0,70423        | 0,40144        |
| m Nyheter, kultur        | -0,08145        | 0,05557        | 2,14824        | 0,14284        |
| k Nyheter, sport         | 0,02846         | 0,04153        | 0,46951        | 0,49327        |
| a Pausprogram            | <b>-0,72723</b> | <b>0,05743</b> | <b>160,35</b>  | <b>0</b>       |
| t Sitcom                 | -0,05786        | 0,04091        | 2,00031        | 0,15738        |
| e Sport/betting          | 0,019836        | 0,06444        | 0,09477        | 0,75822        |
| g Talkshow               | <b>0,626637</b> | <b>0,08433</b> | <b>55,21</b>   | <b>1,4E-13</b> |
| o Talkshow, kultur       | <b>-0,48223</b> | <b>0,08565</b> | <b>31,6993</b> | <b>2E-08</b>   |
| r Talkshow, diskussion   | <b>0,212878</b> | <b>0,06091</b> | <b>12,2153</b> | <b>0,00048</b> |
| i TV-serie, kriminal     | 0,093801        | 0,04394        | 4,55708        | 0,03287        |
| TV-serie, dokumentär     | <b>-0,26469</b> | <b>0,03697</b> | <b>51,2498</b> | <b>1E-12</b>   |
| TV-serie, drama          | -0,10095        | 0,05399        | 3,49524        | 0,06165        |
| TV-serie, historia       | -0,09373        | 0,06599        | 2,0176         | 0,1556         |
| TV-serie, hjälp          | <b>0,274581</b> | <b>0,06394</b> | <b>18,4387</b> | <b>1,8E-05</b> |
| TV-serie, inredning      | 0,114906        | 0,0467         | 6,05353        | 0,01394        |
| TV-serie, kultur         | <b>-0,25124</b> | <b>0,04555</b> | <b>30,4179</b> | <b>3,8E-08</b> |
| TV-serie, reality        | 0,009359        | 0,08412        | 0,01238        | 0,91142        |
| TV-serie, tävling        | <b>0,248828</b> | <b>0,06737</b> | <b>13,6408</b> | <b>0,00023</b> |
| TV-serie, utslagstävling | <b>0,287175</b> | <b>0,04639</b> | <b>38,3252</b> | <b>6,9E-10</b> |
| TV-serie, yrke           | -0,17151        | 0,051          | 11,3083        | 0,00078        |
| TV-serie, övernaturlig   | <b>-0,25922</b> | <b>0,06988</b> | <b>13,7606</b> | <b>0,00021</b> |
| Värdet                   | <b>-0,38904</b> | <b>0,04281</b> | <b>82,595</b>  | <b>0</b>       |

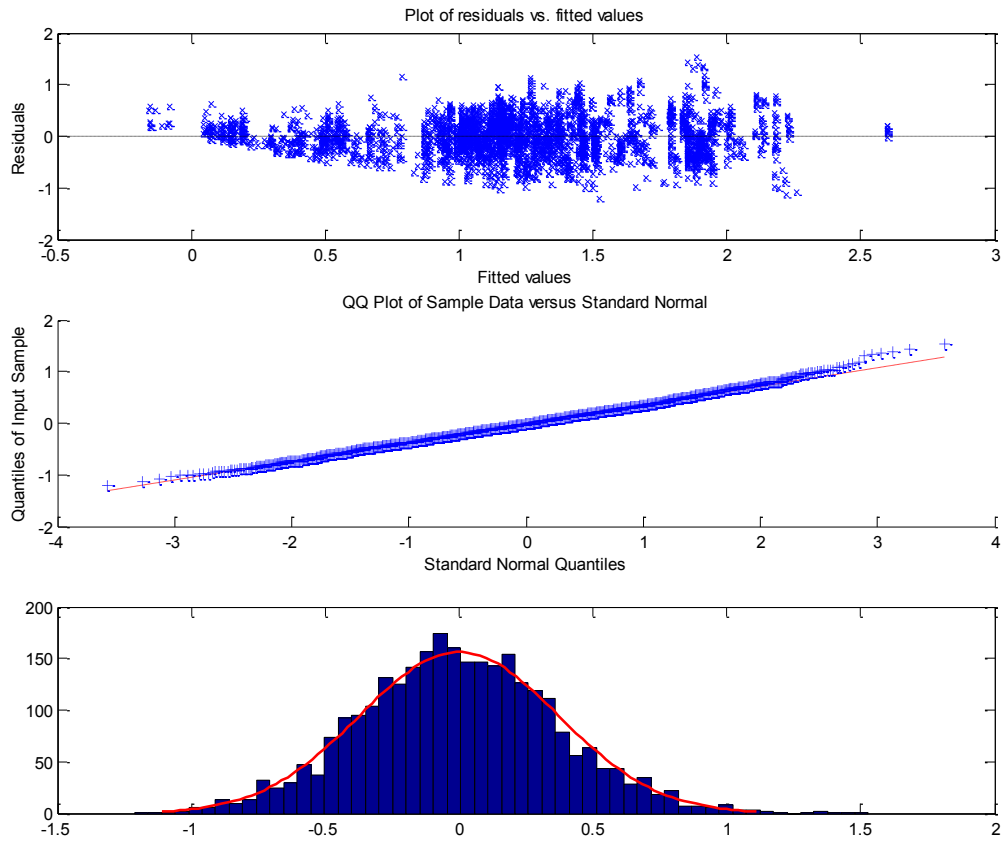
Fortsättning av Regressionen\*

| Alla, 15-49 år    | Estimat         | White SE       | F-värde        | P-värde        |
|-------------------|-----------------|----------------|----------------|----------------|
| K Svt1            | <b>0,249789</b> | <b>0,03462</b> | <b>52,0722</b> | <b>6,8E-13</b> |
| a TV3             | 0,017997        | 0,03986        | 0,20383        | 0,65168        |
| n TV4             | <b>0,735216</b> | <b>0,02787</b> | <b>695,784</b> | <b>0</b>       |
| a Kanal 5         | <b>0,159527</b> | <b>0,04091</b> | <b>15,2033</b> | <b>9,9E-05</b> |
| l TV6             | <b>-0,10845</b> | <b>0,03852</b> | <b>7,92676</b> | <b>0,0049</b>  |
| TV8               | <b>-0,74208</b> | <b>0,03675</b> | <b>407,687</b> | <b>0</b>       |
| TV10              | <b>-0,81595</b> | <b>0,03664</b> | <b>495,795</b> | <b>0</b>       |
| D Måndag          | 0,042237        | 0,02399        | 3,09917        | 0,07844        |
| a Onsdag          | 0,049936        | 0,02527        | 3,90574        | 0,04822        |
| g Torsdag         | 9,28E-05        | 0,0254         | 1,3E-05        | 0,99709        |
| Fredag            | -0,01963        | 0,0301         | 0,42525        | 0,51438        |
| Lördag            | -0,01927        | 0,03023        | 0,40634        | 0,52388        |
| Söndag            | <b>0,103421</b> | <b>0,02904</b> | <b>12,6852</b> | <b>0,00037</b> |
| F Program i följd | 0,042094        | 0,0279         | 2,27695        | 0,13142        |
| L Programlängd    | 0,000762        | 0,00038        | 3,99939        | 0,04561        |

Antalet observationer: 2857, Frihetsgrader: 2814  
 RMSE: 0,373  
 $R^2$ : 0,634, Justerat- $R^2$ -värde: 0,628

\* De som med signifikans skiljer sig från interceptet är fetmarkade.

Tabell 11. Regressionsanalys med beräknade  $\beta$ -estimat, Whites standardavvikelser, F-värde och P-värde för alla variabler i X-matrisen. Detta för målgruppen Alla, 15-49 år.



**Figur 4.** Översta plotten utgörs av "Residuals Vs. Fitted", mellersta plotten är en QQ-plott och den nedersta plotten visar ett histogram av residualerna mot det röda strecket, vilken utgörs av normalfördelningen.

Programlängd och –följd har inte en signifikant påverkan enligt Tabell 11. i målgruppen Alla, 15-49 år, med 1%-signifikansnivån. För Wald-testet, se målgruppen Alla, 3-99 år, då samma slutsatser kan dras från det testet.

| Wald-tester för Alla, 15-49 år                           |         |       |                           |
|--|---------|-------|---------------------------|
| Hypoteser  | Utfall  |       | Slutsats                  |
| Programkategori har ingen inverkan                       | W       | 721   | kan förkasta hypotesen    |
|  | P-värde | 0     |                           |
| Dagar i veckan har ingen inverkan                        | W       | 29,0  | kan förkasta hypotesen    |
|  | P-värde | 0     |                           |
| TV-kanalen har ingen inverkan                            | W       | 2764  | kan förkasta hypotesen    |
|  | P-värde | 0     |                           |
| Torsdag, fredag och lördag kan slås ihop med interceptet | W       | 0,803 | kan ej förkasta hypotesen |
|  | P-värde | 0,239 |                           |

**Tabell 12.** Hypotestest vid heteroskedasticitet för målgruppen Alla, 15-49 år. Programlängd och –följd har båda en signifikant påverkan.

### 5.3. Kvinnor, 25 – 44 år, kl. 20:00-22:59

Tabellerna för regressionsanalysen för målgruppen Kvinnor, 25-44 år, följer nedan och de viktigaste resultaten sammanfattas och diskuteras vad gäller de respektive  $\beta$ -värdena under diskussionsdelen 6.-6.7. tillsammans med analysen. Regressionsanalysen presenteras nedan, medan slutsatser som följd av de här resultaten tas upp under diskussionsdelen. Utöver detta återfinns även  $R^2$  och andra mått av relevans för regressionen i tabellen nedan.

Regression för målgruppen Kvinnor, 25-44 år\*

| Kvinnor, 25-44 år        | Estimat         | White SE       | F-värde        | P-värde        |
|--------------------------|-----------------|----------------|----------------|----------------|
| I (Intercept)            | 1,18            | 3,32E-02       | 1,27E+03       | 0              |
| P Debatt                 | <b>-0,34599</b> | <b>0,04453</b> | <b>60,3778</b> | <b>1,1E-14</b> |
| r Dokumentär             | <b>-0,58997</b> | <b>0,06462</b> | <b>83,3661</b> | <b>0</b>       |
| o Film                   | -0,13216        | 0,05539        | 5,6933         | 0,0171         |
| g Kriminalmagasin        | 0,026979        | 0,0632         | 0,18223        | 0,6695         |
| r Magasin, humor         | -0,10518        | 0,06157        | 2,91899        | 0,08765        |
| a Matlagingsprogram      | -0,10217        | 0,06638        | 2,36881        | 0,12389        |
| m Nyheter, kultur        | -0,06278        | 0,07371        | 0,72549        | 0,39442        |
| k Nyheter, sport         | -0,07152        | 0,04383        | 2,66233        | 0,10286        |
| a Pausprogram            | <b>-0,88269</b> | <b>0,07095</b> | <b>154,769</b> | <b>0</b>       |
| t Sitcom                 | <b>-0,12273</b> | <b>0,04622</b> | <b>7,05034</b> | <b>0,00797</b> |
| e Sport/betting          | <b>-0,16764</b> | <b>0,06292</b> | <b>7,09837</b> | <b>0,00776</b> |
| g Talkshow               | <b>0,685388</b> | <b>0,09132</b> | <b>56,3345</b> | <b>8,1E-14</b> |
| o Talkshow, kultur       | <b>-0,51343</b> | <b>0,10174</b> | <b>25,4683</b> | <b>4,8E-07</b> |
| r Talkshow, diskussion   | 0,157178        | 0,07504        | 4,38759        | 0,03629        |
| i TV-serie, criminal     | 0,116386        | 0,05008        | 5,40106        | 0,0202         |
| TV-serie, dokumentär     | <b>-0,34066</b> | <b>0,04162</b> | <b>66,9844</b> | <b>0</b>       |
| TV-serie, drama          | -0,06766        | 0,06576        | 1,05839        | 0,30367        |
| TV-serie, historia       | <b>-0,25031</b> | <b>0,08327</b> | <b>9,03668</b> | <b>0,00267</b> |
| TV-serie, hjälp          | <b>0,258152</b> | <b>0,0727</b>  | <b>12,608</b>  | <b>0,00039</b> |
| TV-serie, inredning      | <b>0,168023</b> | <b>0,05663</b> | <b>8,80475</b> | <b>0,00303</b> |
| TV-serie, kultur         | <b>-0,31404</b> | <b>0,04967</b> | <b>39,9706</b> | <b>3E-10</b>   |
| TV-serie, reality        | -0,05291        | 0,0919         | 0,33144        | 0,56486        |
| TV-serie, tävling        | <b>0,202385</b> | <b>0,07627</b> | <b>7,04071</b> | <b>0,00801</b> |
| TV-serie, utslagstävling | <b>0,310053</b> | <b>0,05121</b> | <b>36,663</b>  | <b>1,6E-09</b> |
| TV-serie, yrke           | <b>-0,24245</b> | <b>0,06001</b> | <b>16,3238</b> | <b>5,5E-05</b> |
| TV-serie, övernaturlig   | -0,25611        | 0,10056        | 6,48593        | 0,01093        |
| Värdet                   | <b>-0,48504</b> | <b>0,04904</b> | <b>97,8083</b> | <b>0</b>       |

Fortsättning av Regressionen\*

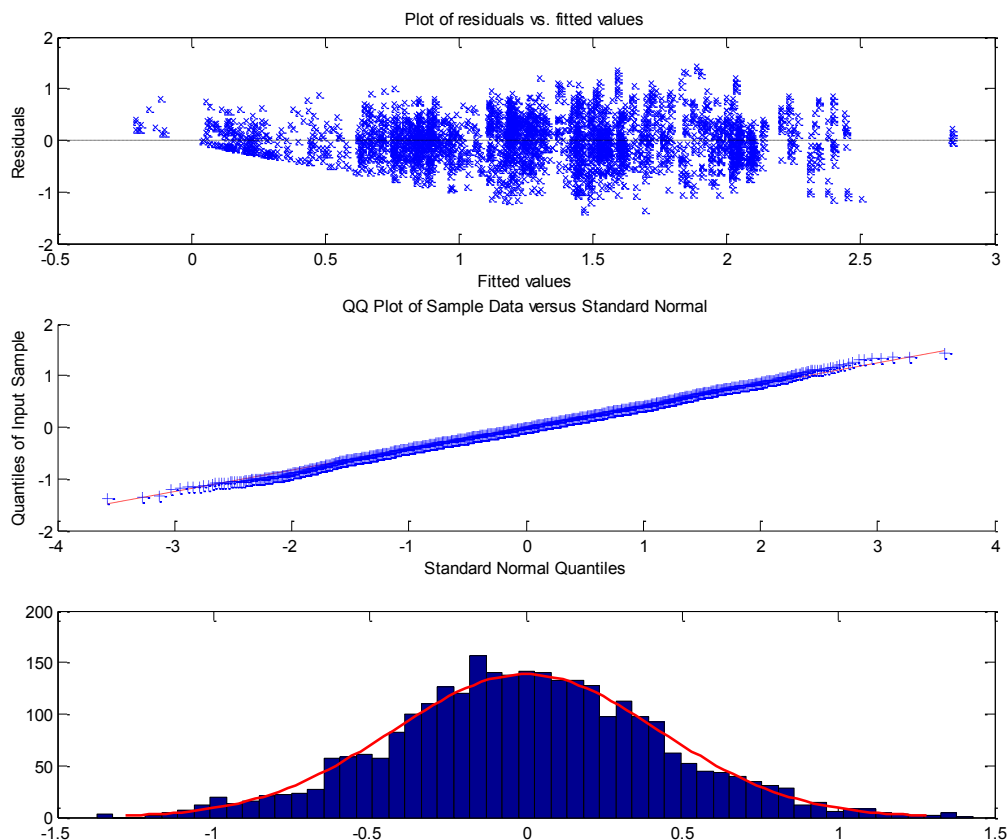
| Kvinnor, 25-44 år | Estimat         | White SE       | F-värde        | P-värde        |
|-------------------|-----------------|----------------|----------------|----------------|
| K Svt1            | <b>0,291623</b> | <b>0,03983</b> | <b>53,6184</b> | <b>3,2E-13</b> |
| a TV3             | <b>0,137344</b> | <b>0,04545</b> | <b>9,13156</b> | <b>0,00253</b> |
| n TV4             | <b>0,832791</b> | <b>0,03091</b> | <b>726,007</b> | <b>0</b>       |
| a Kanal 5         | <b>0,361716</b> | <b>0,04831</b> | <b>56,0558</b> | <b>9,4E-14</b> |
| l TV6             | <b>-0,3093</b>  | <b>0,04375</b> | <b>49,9816</b> | <b>1,9E-12</b> |
| TV8               | <b>-0,70606</b> | <b>0,04337</b> | <b>264,988</b> | <b>0</b>       |
| TV10              | <b>-0,82493</b> | <b>0,04057</b> | <b>413,439</b> | <b>0</b>       |
| D Måndag          | <b>0,076274</b> | <b>0,02873</b> | <b>7,04908</b> | <b>0,00798</b> |
| a Onsdag          | 0,064293        | 0,03041        | 4,47096        | 0,03456        |
| g Torsdag         | 0,023884        | 0,02974        | 0,6449         | 0,42201        |
| Fredag            | -0,01001        | 0,03544        | 0,07977        | 0,77763        |
| Lördag            | -0,02331        | 0,03466        | 0,45233        | 0,50129        |
| Söndag            | <b>0,114244</b> | <b>0,03349</b> | <b>11,6363</b> | <b>0,00066</b> |
| F Program i följd | 0,054966        | 0,03624        | 2,30011        | 0,12948        |
| L Programlängd    | 0,000926        | 0,00042        | 4,97257        | 0,02583        |

Antalet observationer: 2857, Frihetsgrader: 2814  
 RMSE: 0,431  
 $R^2$ : 0,625, Justerat- $R^2$ -värde: 0,62

\* De som med signifikans skiljer sig från interceptet är fetmarkade.

Tabell 13. Regressionsanalys med beräknade  $\beta$ -estimat, Whites standardavvikelser, F-värde och P-värde för alla variabler i X-matrisen. Detta för målgruppen Kvinnor, 25-44 år.





**Figur 5.** Översta plotten utgörs av "Residuals Vs. Fitted", mellersta plotten är en QQ-plot och den nedersta plotten visar ett histogram av residualerna mot det röda strecket, vilken utgörs av normalfördelningen.

Programlängd och -följd har inte en signifikant påverkan enligt Tabell 13., vilket är samma som för målgruppen Alla, 15-49 år, med 1%-signifikansnivån. För Wald-testet, se målgruppen Alla, 3-99 år, då samma slutsatser kan dras från det testet.

| Wald-tester för Kvinnor, 25-44 år                        |         |       |                           |
|--|---------|-------|---------------------------|
| Hypoteser  | Utfall  |       | Slutsats                  |
| Programkategori har ingen inverkan                       | W       | 718   | kan förkasta hypotesen    |
|  | P-värde | 0     |                           |
| Dagar i veckan har ingen inverkan                        | W       | 29,8  | kan förkasta hypotesen    |
|  | P-värde | 0     |                           |
| TV-kanalen har ingen inverkan                            | W       | 2690  | kan förkasta hypotesen    |
|  | P-värde | 0     |                           |
| Torsdag, fredag och lördag kan slås ihop med interceptet | W       | 2,360 | kan ej förkasta hypotesen |
|  | P-värde | 0,188 |                           |

**Tabell 14.** Hypotestest vid heteroskedasticitet för målgruppen Kvinnor, 25-44 år. Programlängd och -följd har båda en signifikant påverkan.

## 5.4. Kvinnor, 25 – 49 år, kl. 20:00-22:59

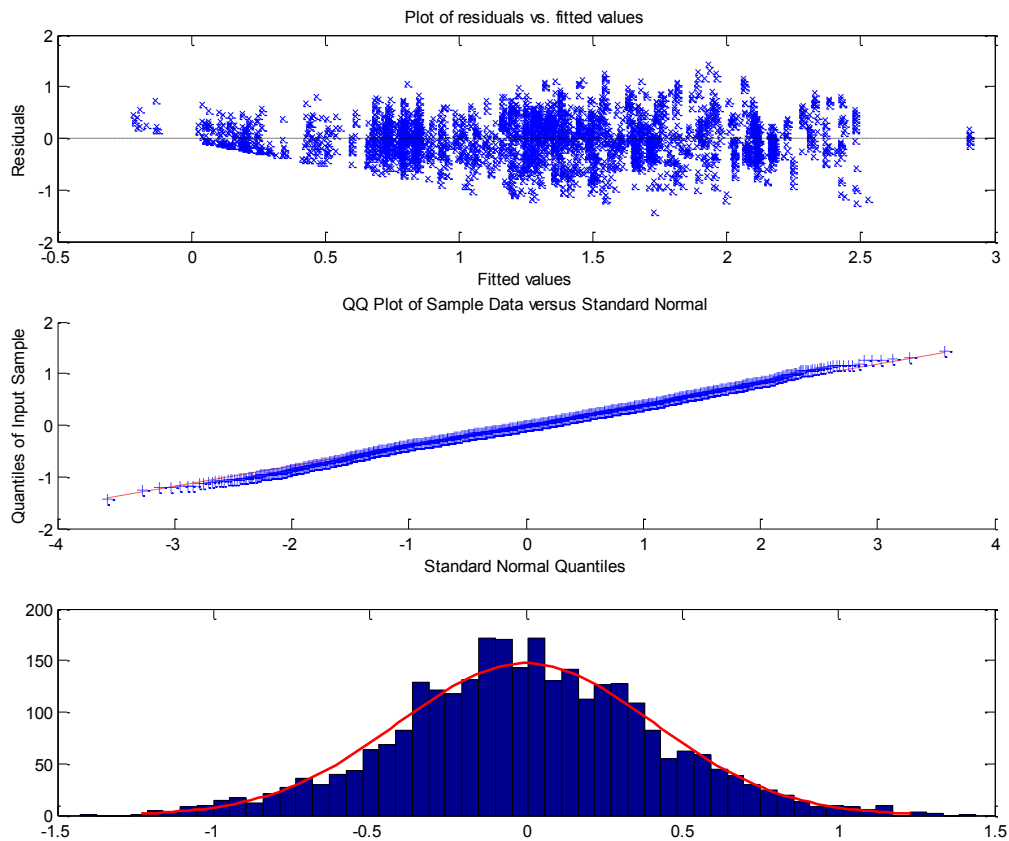
Tabellerna för regressionsanalysen för målgruppen Kvinnor, 25-49 år, följer nedan och de viktigaste resultaten sammanfattas och diskuteras vad gäller de respektive  $\beta$ -värdena under diskussionsdelen 6.-6.7. tillsammans med analysen. Dessa tas således inte upp här. Utöver detta återfinns även  $R^2$  och andra mått av relevans för regressionen i tabellen nedan.

| Regression för målgruppen Kvinnor, 25-49 år* |                 |                |                |                | Fortsättning av Regressionen* |                 |                |                |                |
|--|-----------------|----------------|----------------|----------------|-------------------------------|-----------------|----------------|----------------|----------------|
| Kvinnor, 25-49 år                            | Estimat         | White SE       | F-värde        | P-värde        | Kvinnor, 25-49 år             | Estimat         | White SE       | F-värde        | P-värde        |
| l (Intercept)                                | 1,27            | 3,03E-02       | 1,76E+03       | 0              | K Svt1                        | <b>0,323179</b> | <b>0,03982</b> | <b>65,869</b>  | <b>0</b>       |
| P Debatt                                     | <b>-0,3524</b>  | <b>0,04256</b> | <b>68,5694</b> | <b>0</b>       | a TV3                         | 0,088535        | 0,04523        | 3,83135        | 0,0504         |
| r Dokumentär                                 | <b>-0,6177</b>  | <b>0,06485</b> | <b>90,7336</b> | <b>0</b>       | n TV4                         | <b>0,836726</b> | <b>0,02933</b> | <b>813,603</b> | <b>0</b>       |
| o Film                                       | <b>-0,15358</b> | <b>0,05448</b> | <b>7,94551</b> | <b>0,00485</b> | a Kanal 5                     | <b>0,261155</b> | <b>0,04729</b> | <b>30,4915</b> | <b>3,7E-08</b> |
| g Kriminalmagasin                            | 0,030168        | 0,06235        | 0,23409        | 0,62855        | l TV6                         | <b>-0,38251</b> | <b>0,04268</b> | <b>80,3205</b> | <b>0</b>       |
| r Magasin, humor                             | <b>-0,15772</b> | <b>0,06103</b> | <b>6,67863</b> | <b>0,00981</b> | TV8                           | <b>-0,7501</b>  | <b>0,0425</b>  | <b>311,441</b> | <b>0</b>       |
| a Matlagningsprogram                         | -0,11483        | 0,06529        | 3,0935         | 0,07871        | TV10                          | <b>-0,88872</b> | <b>0,0391</b>  | <b>516,713</b> | <b>0</b>       |
| m Nyheter, kultur                            | -0,11417        | 0,07167        | 2,53719        | 0,1113         | D Måndag                      | 0,057677        | 0,02682        | 4,62398        | 0,03161        |
| k Nyheter, sport                             | -0,08801        | 0,03993        | 4,85896        | 0,02758        | a Onsdag                      | 0,050898        | 0,02833        | 3,22857        | 0,07247        |
| a Pausprogram                                | <b>-0,82687</b> | <b>0,06483</b> | <b>162,688</b> | <b>0</b>       | g Torsdag                     | -0,00148        | 0,02798        | 0,00281        | 0,95773        |
| t Sitcom                                     | <b>-0,17654</b> | <b>0,04478</b> | <b>15,5438</b> | <b>8,3E-05</b> | Fredag                        | -0,02576        | 0,033          | 0,60908        | 0,4352         |
| e Sport/betting                              | <b>-0,19392</b> | <b>0,06004</b> | <b>10,4322</b> | <b>0,00125</b> | Lördag                        | -0,01125        | 0,03336        | 0,11376        | 0,73593        |
| g Talkshow                                   | <b>0,666132</b> | <b>0,0901</b>  | <b>54,6638</b> | <b>1,9E-13</b> | Söndag                        | <b>0,108327</b> | <b>0,03189</b> | <b>11,5361</b> | <b>0,00069</b> |
| o Talkshow, kultur                           | <b>-0,53812</b> | <b>0,09623</b> | <b>31,2674</b> | <b>2,5E-08</b> | F Program i följd             | 0,058189        | 0,03389        | 2,94772        | 0,08611        |
| r Talkshow, diskussion                       | 0,156764        | 0,07419        | 4,46539        | 0,03468        | L Programlängd                | 0,000807        | 0,00041        | 3,85188        | 0,04979        |
| i TV-serie, kriminal                         | 0,119786        | 0,04842        | 6,12079        | 0,01342        |                               |                 |                |                |                |
| TV-serie, dokumentär                         | <b>-0,36125</b> | <b>0,04108</b> | <b>77,3475</b> | <b>0</b>       |                               |                 |                |                |                |
| TV-serie, drama                              | -0,08803        | 0,06399        | 1,89231        | 0,16905        |                               |                 |                |                |                |
| TV-serie, historia                           | <b>-0,28355</b> | <b>0,07882</b> | <b>12,9427</b> | <b>0,00033</b> |                               |                 |                |                |                |
| TV-serie, hjälp                              | <b>0,270382</b> | <b>0,07122</b> | <b>14,4138</b> | <b>0,00015</b> |                               |                 |                |                |                |
| TV-serie, inredning                          | <b>0,15363</b>  | <b>0,05388</b> | <b>8,12902</b> | <b>0,00439</b> |                               |                 |                |                |                |
| TV-serie, kultur                             | <b>-0,3521</b>  | <b>0,04933</b> | <b>50,9416</b> | <b>1,2E-12</b> |                               |                 |                |                |                |
| TV-serie, reality                            | -0,05206        | 0,09425        | 0,30503        | 0,58079        |                               |                 |                |                |                |
| TV-serie, tävling                            | 0,198297        | 0,07696        | 6,63823        | 0,01003        |                               |                 |                |                |                |
| TV-serie, utslagstävling                     | <b>0,273482</b> | <b>0,05028</b> | <b>29,5824</b> | <b>5,8E-08</b> |                               |                 |                |                |                |
| TV-serie, yrke                               | <b>-0,26062</b> | <b>0,0571</b>  | <b>20,8348</b> | <b>5,2E-06</b> |                               |                 |                |                |                |
| TV-serie, övernaturlig                       | <b>-0,29813</b> | <b>0,09295</b> | <b>10,2888</b> | <b>0,00135</b> |                               |                 |                |                |                |
| Värdet                                       | <b>-0,45405</b> | <b>0,04203</b> | <b>116,729</b> | <b>0</b>       |                               |                 |                |                |                |

Antalet observationer: 2857, Frihetsgrader: 2814  
 RMSE: 0,413  
 $R^2$ : 0,66, Justerat- $R^2$ -värde: 0,655

\* De som med signifikans skiljer sig från interceptet är fetmarkerade.

**Tabell 15.** Regressionsanalys med beräknade  $\beta$ -estimat, Whites standardavvikelse, F-värde och P-värde för alla variabler i X-matrisen. Detta för målgruppen Kvinnor, 25-49 år.



**Figur 6.** Översta plotten utgörs av "Residuals Vs. Fitted", mellersta plotten är en QQ-plot och den nedersta plotten visar ett histogram av residualerna mot det röda strecket, vilken utgörs av normalfördelningen.

Se målgruppen Alla, 15-49 år vad gäller programlängd och -följd samt Wald-testet.

| Wald-tester för Kvinnor, 25-49 år                        |         |       |                           |
|--|---------|-------|---------------------------|
| Hypoteser  | Utfall  |       | Slutsats                  |
| Programkategori har ingen inverkan                       | W       | 787   | kan förkasta hypotesen    |
|  | P-värde | 0     |                           |
| Dagar i veckan har ingen inverkan                        | W       | 28,4  | kan förkasta hypotesen    |
|  | P-värde | 0     |                           |
| TV-kanalen har ingen inverkan                            | W       | 34400 | kan förkasta hypotesen    |
|  | P-värde | 0     |                           |
| Torsdag, fredag och lördag kan slås ihop med interceptet | W       | 0,748 | kan ej förkasta hypotesen |
|  | P-värde | 0,237 |                           |

**Tabell 16.** Hypotestest vid heteroskedasticitet för målgruppen Kvinnor, 25-49 år. Programlängd och -följd har båda en signifikant påverkan.

## 5.5. Män, 15 – 34 år, kl. 20:00-22:59

Tabellerna för regressionsanalysen för målgruppen Män, 15-34 år, följer nedan och de viktigaste resultaten sammanfattas under diskussionsdelen 6.-6.7. tillsammans med analysen. Programkategorierna och kanalernas resultat från regressionsanalysen återfinns som sagt under diskussionen. Utöver detta ges även  $R^2$  och andra mått av relevans för regressionen i tabellen nedan.

Regression för målgruppen Män, 15-34 år\*

| Män, 15-34 år                   | Estimat         | White SE       | F-värde        | P-värde        |
|---------------------------------|-----------------|----------------|----------------|----------------|
| I (Intercept)                   | 0,675           | 2,91E-02       | 539            | 0              |
| P Debatt                        | -0,05765        | 0,0519         | 1,23384        | 0,26676        |
| r <b>Dokumentär</b>             | <b>-0,2029</b>  | <b>0,06095</b> | <b>11,0807</b> | <b>0,00088</b> |
| o <b>Film</b>                   | <b>0,124236</b> | <b>0,05222</b> | <b>5,66001</b> | <b>0,01742</b> |
| g <b>Kriminalmagasin</b>        | <b>0,321085</b> | <b>0,05996</b> | <b>28,6759</b> | <b>9,2E-08</b> |
| r <b>Magasin, humor</b>         | <b>0,186522</b> | <b>0,05688</b> | <b>10,7537</b> | <b>0,00105</b> |
| a Matlagningsprogram            | 0,07126         | 0,06389        | 1,24387        | 0,26482        |
| m Nyheter, kultur               | -0,04693        | 0,04837        | 0,94128        | 0,33203        |
| k Nyheter, sport                | 0,082792        | 0,04872        | 2,8875         | 0,08938        |
| a <b>Pausprogram</b>            | <b>-0,41858</b> | <b>0,05853</b> | <b>51,1401</b> | <b>1,1E-12</b> |
| t <b>Sitcom</b>                 | <b>0,145261</b> | <b>0,04425</b> | <b>10,7785</b> | <b>0,00104</b> |
| e <b>Sport/betting</b>          | <b>0,247739</b> | <b>0,07453</b> | <b>11,0486</b> | <b>0,0009</b>  |
| g <b>Talkshow</b>               | <b>0,641415</b> | <b>0,08447</b> | <b>57,655</b>  | <b>4,2E-14</b> |
| o <b>Talkshow, kultur</b>       | <b>-0,2548</b>  | <b>0,08784</b> | <b>8,41404</b> | <b>0,00375</b> |
| r <b>Talkshow, diskussion</b>   | <b>0,319876</b> | <b>0,0597</b>  | <b>28,7047</b> | <b>9,1E-08</b> |
| i TV-serie, criminal            | 0,098523        | 0,04571        | 4,64539        | 0,03122        |
| TV-serie, dokumentär            | -0,07336        | 0,04083        | 3,2278         | 0,0725         |
| TV-serie, drama                 | -0,06366        | 0,0491         | 1,68082        | 0,19492        |
| TV-serie, historia              | 0,122796        | 0,0689         | 3,17637        | 0,07482        |
| <b>TV-serie, hjälp</b>          | <b>0,381961</b> | <b>0,06498</b> | <b>34,5506</b> | <b>4,6E-09</b> |
| <b>TV-serie, inredning</b>      | <b>0,164944</b> | <b>0,05005</b> | <b>10,8597</b> | <b>0,001</b>   |
| TV-serie, kultur                | -0,07138        | 0,0448         | 2,53897        | 0,11118        |
| TV-serie, reality               | 0,034485        | 0,1059         | 0,10604        | 0,74473        |
| <b>TV-serie, tävling</b>        | <b>0,399292</b> | <b>0,06604</b> | <b>36,5572</b> | <b>1,7E-09</b> |
| <b>TV-serie, utslagstävling</b> | <b>0,311387</b> | <b>0,05254</b> | <b>35,121</b>  | <b>3,5E-09</b> |
| TV-serie, yrke                  | 0,000483        | 0,05291        | 8,3E-05        | 0,99272        |
| TV-serie, övernaturlig          | -0,07418        | 0,08086        | 0,84157        | 0,35903        |
| <b>Värdet</b>                   | <b>-0,38944</b> | <b>0,05373</b> | <b>52,5311</b> | <b>5,4E-13</b> |

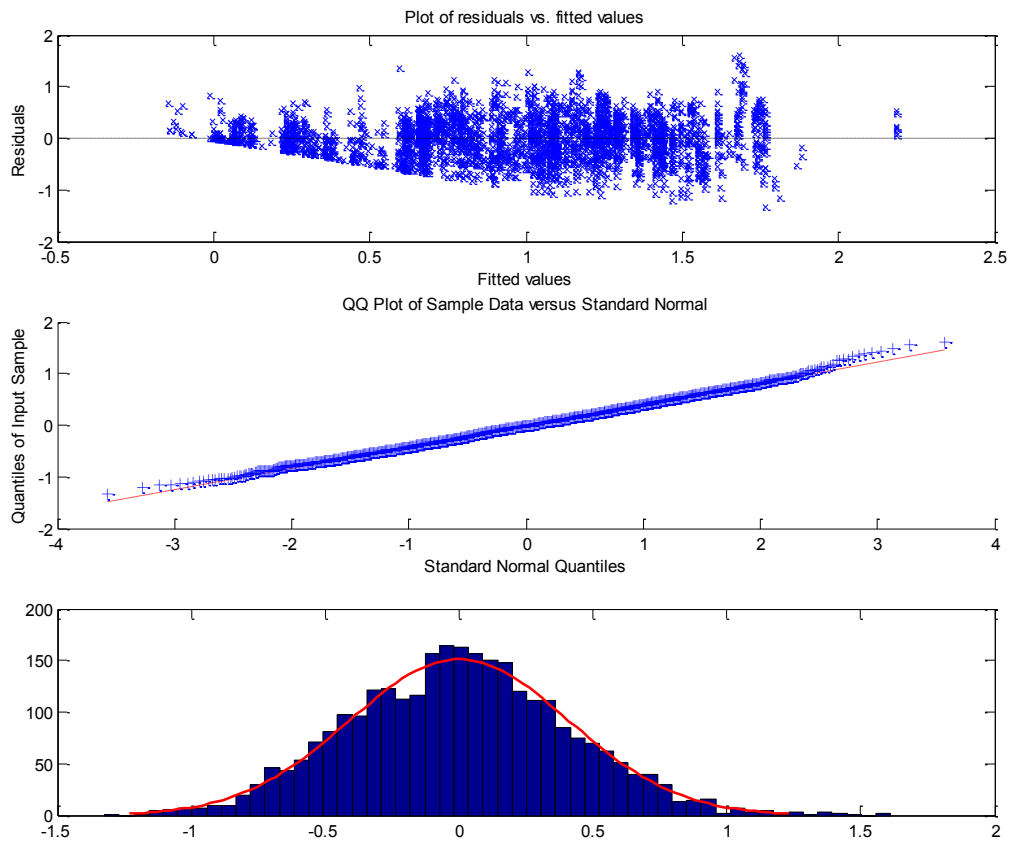
Fortsättning av Regressionen\*

| Män, 15-34 år     | Estimat         | White SE       | F-värde        | P-värde        |
|-------------------|-----------------|----------------|----------------|----------------|
| K <b>Svt1</b>     | <b>0,176799</b> | <b>0,03326</b> | <b>28,2626</b> | <b>1,1E-07</b> |
| a <b>TV3</b>      | <b>0,119293</b> | <b>0,04133</b> | <b>8,33136</b> | <b>0,00393</b> |
| n <b>TV4</b>      | <b>0,755623</b> | <b>0,03032</b> | <b>621,158</b> | <b>0</b>       |
| a <b>Kanal 5</b>  | <b>0,260775</b> | <b>0,04047</b> | <b>41,5162</b> | <b>1,4E-10</b> |
| l <b>TV6</b>      | <b>0,425644</b> | <b>0,04097</b> | <b>107,921</b> | <b>0</b>       |
| <b>TV8</b>        | <b>-0,53733</b> | <b>0,03686</b> | <b>212,542</b> | <b>0</b>       |
| <b>TV10</b>       | <b>-0,59718</b> | <b>0,03853</b> | <b>240,165</b> | <b>0</b>       |
| D Måndag          | 0,036207        | 0,02673        | 1,83424        | 0,17574        |
| a Onsdag          | 0,006824        | 0,02825        | 0,05836        | 0,80912        |
| g Torsdag         | -0,01886        | 0,02845        | 0,43974        | 0,5073         |
| Fredag            | -0,01288        | 0,03189        | 0,16322        | 0,68624        |
| Lördag            | -0,03399        | 0,03237        | 1,10239        | 0,29383        |
| <b>Söndag</b>     | <b>0,105722</b> | <b>0,03232</b> | <b>10,7013</b> | <b>0,00108</b> |
| F Program i följd | 0,053267        | 0,03345        | 2,53566        | 0,11141        |
| L Programlängd    | 0,00031         | 0,0004         | 0,60836        | 0,43547        |

Antalet observationer: 2857, Frihetsgrader: 2814  
 RMSE: 0,413  
 $R^2$ : 0,531, Justerat- $R^2$ -värde: 0,524

\* De som med signifikans skiljer sig från interceptet är fetmarkade.

**Tabell 17.** Regressionsanalys med beräknade  $\beta$ -estimat, Whites standardavvikelse, F-värde och P-värde för alla variabler i X-matrisen. Detta för målgruppen Män, 15-34 år.



**Figur 7.** Översta plotten utgörs av "Residuals Vs. Fitted", mellersta plotten är en QQ-plot och den nedersta plotten visar ett histogram av residualerna mot det röda strecket, vilken utgörs av normalfördelningen.

Se målgruppen Alla, 15-49 år vad gäller programlängd och -följd samt Wald-testet.

| Wald-tester för Män, 15-34 år                            |         |       |                           |
|--|---------|-------|---------------------------|
| Hypoteser  | Utfall  |       | Slutsats                  |
| Programkategori har ingen inverkan                       | W       | 431   | kan förkasta hypotesen    |
|  | P-värde | 0     |                           |
| Dagar i veckan har ingen inverkan                        | W       | 25,4  | kan förkasta hypotesen    |
|  | P-värde | 0     |                           |
| TV-kanalen har ingen inverkan                            | W       | 1850  | kan förkasta hypotesen    |
|  | P-värde | 0     |                           |
| Torsdag, fredag och lördag kan slås ihop med interceptet | W       | 1,180 | kan ej förkasta hypotesen |
|  | P-värde | 0,240 |                           |

**Tabell 18.** Hypotestest vid heteroskedasticitet för målgruppen Män, 15-34 år. Programlängd och -följd har båda en signifikant påverkan.

## 5.6. Män, 25 – 49 år, kl. 20:00-22:59

Tabellerna för regressionsanalysen för målgruppen Män, 25-49 år, följer nedan och de viktigaste resultaten sammanfattas under diskussionsdelen 6.-6.7, varför ingen längre förklaring ges här. Utöver detta återfinns även  $R^2$  och andra mått av relevans för regressionen i tabellen nedan.

Regression för målgruppen Män, 25-49 år\*

| Män, 25-49 år            | Estimat         | White SE       | F-värde        | P-värde        |
|--------------------------|-----------------|----------------|----------------|----------------|
| I (Intercept)            | 1,27            | 3,09E-02       | 1,69E+03       | 0              |
| P Debatt                 | <b>-0,23152</b> | <b>0,04944</b> | <b>21,9283</b> | <b>3E-06</b>   |
| r Dokumentär             | <b>-0,4884</b>  | <b>0,06961</b> | <b>49,2294</b> | <b>2,8E-12</b> |
| o Film                   | -0,02593        | 0,05595        | 0,21478        | 0,64308        |
| g Kriminalmagasin        | <b>0,177138</b> | <b>0,05667</b> | <b>9,7714</b>  | <b>0,00179</b> |
| r Magasin, humor         | 0,044973        | 0,05655        | 0,63246        | 0,42652        |
| a Matlagningsprogram     | -0,02505        | 0,06668        | 0,14109        | 0,70723        |
| m Nyheter, kultur        | -0,11666        | 0,05884        | 3,93084        | 0,04751        |
| k Nyheter, sport         | <b>0,140062</b> | <b>0,04815</b> | <b>8,46062</b> | <b>0,00366</b> |
| a Pausprogram            | <b>-0,81746</b> | <b>0,0754</b>  | <b>117,534</b> | <b>0</b>       |
| t Sitcom                 | -0,10015        | 0,04314        | 5,38839        | 0,02034        |
| e Sport/betting          | 0,117366        | 0,07236        | 2,63102        | 0,10491        |
| g Talkshow               | <b>0,569709</b> | <b>0,08627</b> | <b>43,6062</b> | <b>4,8E-11</b> |
| o Talkshow, kultur       | <b>-0,4943</b>  | <b>0,10151</b> | <b>23,7097</b> | <b>1,2E-06</b> |
| r Talkshow, diskussion   | <b>0,238691</b> | <b>0,06292</b> | <b>14,3934</b> | <b>0,00015</b> |
| i TV-serie, criminal     | 0,023439        | 0,04807        | 0,23772        | 0,6259         |
| TV-serie, dokumentär     | <b>-0,29648</b> | <b>0,04189</b> | <b>50,0877</b> | <b>1,8E-12</b> |
| TV-serie, drama          | <b>-0,17973</b> | <b>0,05469</b> | <b>10,8007</b> | <b>0,00103</b> |
| TV-serie, historia       | 0,001046        | 0,07054        | 0,00022        | 0,98817        |
| TV-serie, hjälp          | <b>0,233019</b> | <b>0,06654</b> | <b>12,2619</b> | <b>0,00047</b> |
| TV-serie, inredning      | 0,092072        | 0,05381        | 2,92806        | 0,08716        |
| TV-serie, kultur         | <b>-0,22885</b> | <b>0,05604</b> | <b>16,6751</b> | <b>4,6E-05</b> |
| TV-serie, reality        | 0,001616        | 0,10216        | 0,00025        | 0,98738        |
| TV-serie, tävling        | <b>0,186055</b> | <b>0,06641</b> | <b>7,8485</b>  | <b>0,00512</b> |
| TV-serie, utslagstävling | <b>0,141487</b> | <b>0,05055</b> | <b>7,83535</b> | <b>0,00516</b> |
| TV-serie, yrke           | <b>-0,16267</b> | <b>0,05795</b> | <b>7,88082</b> | <b>0,00503</b> |
| TV-serie, övernaturlig   | <b>-0,27455</b> | <b>0,08934</b> | <b>9,44501</b> | <b>0,00214</b> |
| Värdet                   | <b>-0,35474</b> | <b>0,04727</b> | <b>56,3097</b> | <b>8,3E-14</b> |

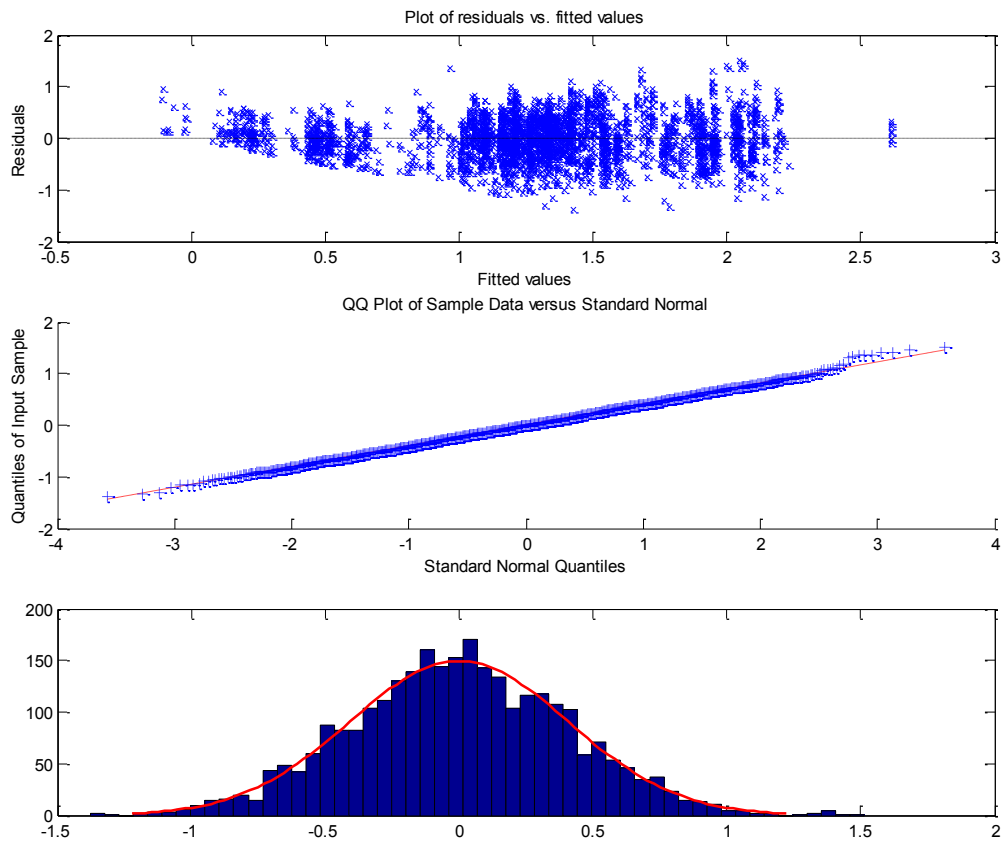
Fortsättning av Regressionen\*

| Män, 25-49 år     | Estimat         | White SE       | F-värde        | P-värde        |
|-------------------|-----------------|----------------|----------------|----------------|
| K Svt1            | <b>0,216976</b> | <b>0,03718</b> | <b>34,048</b>  | <b>6E-09</b>   |
| a TV3             | <b>-0,16116</b> | <b>0,04314</b> | <b>13,9524</b> | <b>0,00019</b> |
| n TV4             | <b>0,616737</b> | <b>0,031</b>   | <b>395,761</b> | <b>0</b>       |
| a Kanal 5         | 0,011016        | 0,04259        | 0,06691        | 0,79591        |
| l TV6             | <b>-0,15635</b> | <b>0,04163</b> | <b>14,1074</b> | <b>0,00018</b> |
| TV8               | <b>-0,8923</b>  | <b>0,04098</b> | <b>474,18</b>  | <b>0</b>       |
| TV10              | <b>-0,91808</b> | <b>0,04403</b> | <b>434,768</b> | <b>0</b>       |
| D Måndag          | 0,056673        | 0,02714        | 4,36084        | 0,03686        |
| a Onsdag          | 0,055124        | 0,0279         | 3,90458        | 0,04825        |
| g Torsdag         | 0,015455        | 0,02867        | 0,2906         | 0,58988        |
| Fredag            | -0,00779        | 0,03228        | 0,05822        | 0,80935        |
| Lördag            | -0,02459        | 0,03364        | 0,53415        | 0,46493        |
| Söndag            | <b>0,128933</b> | <b>0,03152</b> | <b>16,727</b>  | <b>4,4E-05</b> |
| F Program i följd | 0,059403        | 0,0313         | 3,60273        | 0,05779        |
| L Programlängd    | 0,000978        | 0,00042        | 5,36105        | 0,02066        |

Antalet observationer: 2857, Antalet frihetsgrader: 2814  
 RMSE: 0,411  
 $R^2$ : 0,585, Justerat- $R^2$ -värde: 0,579

\* De som med signifikans skiljer sig från interceptet är fetmarkade.

Tabell 19. Regressionsanalys med beräknade  $\beta$ -estimat, Whites standardavvikelse, F-värde och P-värde för alla variabler i X-matrisen. Detta för målgruppen Män, 25-49 år.



**Figur 8.** Översta plotten utgörs av "Residuals Vs. Fitted", mellersta plotten är en QQ-plot och den nedersta plotten visar ett histogram av residualerna mot det röda strecket, vilken utgörs av normalfördelningen.

Se målgruppen Alla, 15-49 år vad gäller programlängd och -följd samt Wald-testet.

| Wald-tester för Män, 25-49 år                            |         |       |                           |
|--|---------|-------|---------------------------|
| Hypoteser  | Utfall  |       | Slutsats                  |
| Programkategori har ingen inverkan                       | W       | 560   | kan förkasta hypotesen    |
|  | P-värde | 0     |                           |
| Dagar i veckan har ingen inverkan                        | W       | 33,2  | kan förkasta hypotesen    |
|  | P-värde | 0     |                           |
| TV-kanalen har ingen inverkan                            | W       | 2240  | kan förkasta hypotesen    |
|  | P-värde | 0     |                           |
| Torsdag, fredag och lördag kan slås ihop med interceptet | W       | 1,550 | kan ej förkasta hypotesen |
|  | P-värde | 0,229 |                           |

**Tabell 20.** Hypotestest vid heteroskedasticitet för målgruppen Män, 25-49 år. Programlängd och -följd har båda en signifikant påverkan.

## 6. Matematisk diskussion

Som nämnts drogs slutsatsen att det enda intervallet som är lämpligt att titta på med den uppställda regressionsmodellen i det här kandidatexamensarbetet är intervallet prime time: 20:00-22:59 (för motivering, se Kapitel 4.3. Heteroskedasticitet och transformering av  $Y$ -variabeln). Det är flest tittare denna tid på dygnet och således fås bättre variation i  $Y$ -värdena. Problemet med övriga tidsintervaller är att det är få som tittar. Det är så få som tittar övriga tider på dygnet att den överhängande mängden observationer består av 0 % tittare. Uppenbarligen blir variationen i den förklarande variabeln väldigt låg och således även svår att transformera och som en följd även svår att analysera.

I det intervall som valts att lägga fokus på, det vill säga prime time, anses modellen vara god för prediktion. Denna slutsats dras utifrån att RMSE-värdena för de olika regressionerna är omkring 2 procentenheter. I prediktionssyfte kan modellen användas för att ungefärligt se vilken tittarprocentnivå som ett visst program i en programkategori kommer att generera utifrån vilken dag och kanal som programmet är tänkt att sändas på.

Själva grundekvationen, se Formel (21), beskriver drivande faktorer bakom TV-tittande väl. Observera den omformulering som gjorts av grundekvationen (se Kapitel 4.3. och Kapitel 4.6.). Detta grundar sig i att  $R^2$ -värdena och justerade- $R^2$ -värdena låg på en nivå kring 0,6 för samtliga regressioner. På så sätt kan regressionsmodellen användas för att dra slutsatser kring vilka program som bidrar till flest TV-tittare och således hur TV-kanalernas tablåer bör se ut för att locka flest tittare inom en viss målgrupp (för vidare diskussion se Kapitel 6.2.-6.7.).

För att ytterligare förbättra denna modell och eventuellt få ännu bättre  $R^2$ -värden skulle exempelvis fenomenet trailing kunna inkluderas i grundekvationen (för fler kompletterande variabler och för beskrivning av trailing se Kapitel 7.). Utöver detta rekommenderas att programlängd samt programföljd utesluts som variabler då de inte har statistisk signifikans i samtliga målgrupper. Vidare gick det att se att framtida modeller kan inkludera torsdag, fredag och lördag i benchmark (se Wald-test-tabellerna under resultatdelen).

Det kan diskuteras huruvida logaritmering av  $(Y + 1)$  är ett bra tillvägagångssätt vid transformeringen av  $Y$ -värdena för att uppnå normalfördelade residualer. Att addera 1 är en linjär transformation, medan logaritmering är en icke-linjär transformation, vilket kan få komplikationer när värdena ska transformeras tillbaka till rätt enhet. Då RMSE:n dock låg på 2 procentenheter är denna transformering godtagbar i denna regressionsanalys.

Även om modellen i sig kan påstås vara bra, har det överhängande arbetet varit att dels formulera grundekvationen helt fristående, och dels anpassa MMS-data till en regressionsvänlig input i modellen. Det mest tidskrävande arbetet har varit kategoriseringen och omskrivningen till dummyvariabler av TV-programmen, som återfinns i bilagorna 5 och 6. Regressionsanalys av MMS-data har dessutom inte gjorts tidigare och har därför varit ett omfattande arbete för författarna, eftersom det saknas riktlinjer och tidigare rekommendationer. Den mest värdefulla slutsatsen är således att det går att tillämpa regressionsanalys på TV-tittarhistorik, och inte modellen i sig. De diskuterade åtgärderna under Kapitel 6.8. om rekommendationer, utgör riktlinjer vid fortsatt analyserande av TV-tittarhistorik med MMS-data som grund och regressionsanalys som verktyg.



## 6.1. Kanalgenerell diskussion

Under resultatdelen framgår det att program bidrar till kanalens tittarsiffror i olika stor utsträckning och att det till viss del skiljer sig mellan målgrupper och till stor del mellan kanaler vad gäller deras tittarsiffror. I den kanalspecifika diskussionen som följer kommer det även att framgå att TV-kanalerna visar väldigt olika programkategorier, vilka alla inte når ut till den definierade målgruppen. Programkategorin som konsekvent har visat sig vara populär utifrån tabellerna i Kapitel 6.2.-6.7. om Resultat är ”Talkshow”, som är den populäraste hos alla målgrupper med bred marginal. Vad programkategorin ”Talkshow” byggs upp av exakt för program finns i Bilaga 5. Bland annat utgörs den av talkshows som Hellenius Hörna, men också svensknorska talkshowen Skavlan.

Vädret är, å andra sidan, sällsynt impopulärt och kan konstateras vara ett förlegat inslag i nyhetssändningen, eftersom alla målgrupper mer eller mindre byter kanal eller slutar titta när det kommer på. Nyheter är däremot desto mer populärt och kan kanske inte i lika stor utsträckning ersättas av applikationer och hemsidor, såsom vädret kan. För att återkoppla till inledningen av det här kandidatexamensarbetet så är vädret väldigt smidigt tillgängligt i applikationer i dagsläget. Varför vädret är mindre populärt på TV än nyheter kan ha sin förklaring i att information om vädret lätt går att få överskådligt via bilder i mobilen, medan nyheter delvis måste läsas och på så sätt kräver större uppmärksamhet av användaren. Det går i varje fall att konstatera att vädret konsekvent är bland de minst populära programmen hos de målgrupperna som undersöks.

Programföljd och -längd är överflödiga variabler i regressionen för de flesta målgrupper, eftersom deras P-värden är höga (se Resultatdelen på sidorna 24-36). När programlängd och -följd däremot har en inverkan på TV-tittandet är den vagt positiv. Det är väntat att programföljd är positiv eftersom TV-tittaren troligtvis tycker om det program som denne tittar på och det är större sannolikhet att personen även gillar nästkommande program om det är samma än om det hade varit ett godtyckligt TV-program efteråt. Att programlängd har vagt positiv inverkan är också relativt väntat eftersom Pausprogram som Anslagstavlan är mest impopulära av alla program och dessa är alla under 20 minuter långa. Det behöver således inte betyda att ju längre, desto mer populärt, utan snarare är det nog som så att korta program generellt sett har negativ påverkan på TV-tittande.

Enstaka dokumentärinslag är generellt sett också väldigt impopulära bland de målgrupper som undersökts. Det är rimligt att anta att det är svårt att lyckas med ett program som sänds en gång när det gäller att fånga TV-tittarnas uppmärksamhet. Vid fortsatt analyserande med regressionsanalys rekommenderas programkategorin ”Dokumentär” delas upp efter innehåll. På så sätt går det att identifiera vilka dokumentärtyper som är populära i vilka målgrupper, för att möjliggöra för TV-bolagen att i framtiden både sända, köpa in och producera rätt dokumentärer för rätt kanal och målgrupp. Generellt sett går det dock att konstatera att dokumentärer attraherar synnerligen få TV-tittare i respektive målgrupp.

Utöver ovan nämnda generella tendenser, går det att utläsa att TV4 lyckats väldigt bra med att nå ut till TV-tittarna. För att återkoppla till inledningen av det här kandidatexamensarbetet har TV4 som kanal historiskt alltid varit konkurrensinriktad och lyckades även med bedriften att bli den första marksända reklamkanalen i Sverige. Detta har legat till kanalens favör och kan förklara varför TV3, trots att kanalen lanserades före TV4, inte har lyckats i lika stor utsträckning. Vad TV3 specifikt kan göra för att förbättra sin position på marknaden går genom under Kapitel 6.4.

För ytterligare diskussion om de olika kanalernas framgång bör tilläggas att det gick att se att i samtliga regressionsstabeller (Tabell 9, 11, 13 15, 17 och 19) så särskiljer sig de olika programkategorierna inte särskilt mycket när det gäller deras popularitet i olika målgrupper. Som en följd är det naturligt att ifrågasätta om målgruppsdefinitionen utifrån ålder och kön är rätt sätt att gå tillväga eller om detta kan anses vara ett förlegat tillvägagångssätt vad det gäller analys av tittarbeteende. Som gick att utläsa ur exempelvis TV10:s målgrupp i Tabell 19. har denna kanal dessvärre ej lyckats särskilt bra utifrån sin definierade målgrupp. TV10 är dock en väldigt nischad kanal som ger bilden av att de sänder intellektuella program. Denna typ av program kanske inte går hem som helhet hos en målgrupp av en viss ålder och ett visst kön, utan snarare en mindre population ur respektive målgrupp och således sträcker sig över målgruppsspannen, så som de är definierade i dagsläget. För TV10 skulle det således vara intressantare att ställa upp en regression utifrån en målgrupp som utformats utifrån specifika intressen istället för ålder och kön. På samma sätt skulle eventuellt en vidare regressionsanalys av vissa typer av kanaler, specifikt nischade kanaler, återspegla de drivande faktorerna bakom den egentliga målgruppens TV-tittande i större utsträckning om dessa målgrupper ställs upp efter andra och/eller ytterligare kriterier än de som använts i detta arbete.

Nedan följer en mer ingående diskussion av resultatet rörande de sex olika målgrupperna var för sig och vad som kan utläsas ur respektive regressionsanalys. Observera att kanalerna består av fri-TV-kanaler och abonnemangskanaler. Således måste detta tas hänsyn till vid jämförelse mellan kanaler. Att till exempel TV3 inte lyckas nå ut bra i en viss målgrupp där SVT1 gör det reflekterar till viss del deras olika popularitet och till viss del tillgången på abonnemangspaket. De resultat från regressionsmodellen som med statistisk signifikans påverkar TV-tittande är sammanfattade i tabeller för varje målgrupp i följande kapitel. Inom varje kategori av dummyvariabler är de ingående programmen eller kanalerna sorterade från störst positiv påverkan till mest negativ påverkan på TV-tittande inom målgruppen. Observera att för att jämföra  $\beta$ -värden mellan olika regressioner och målgrupper måste omskrivningen  $e^{X\beta} - 1$  göras, där  $\beta$  återfinns under rubriken Estimat i följande kapitel.

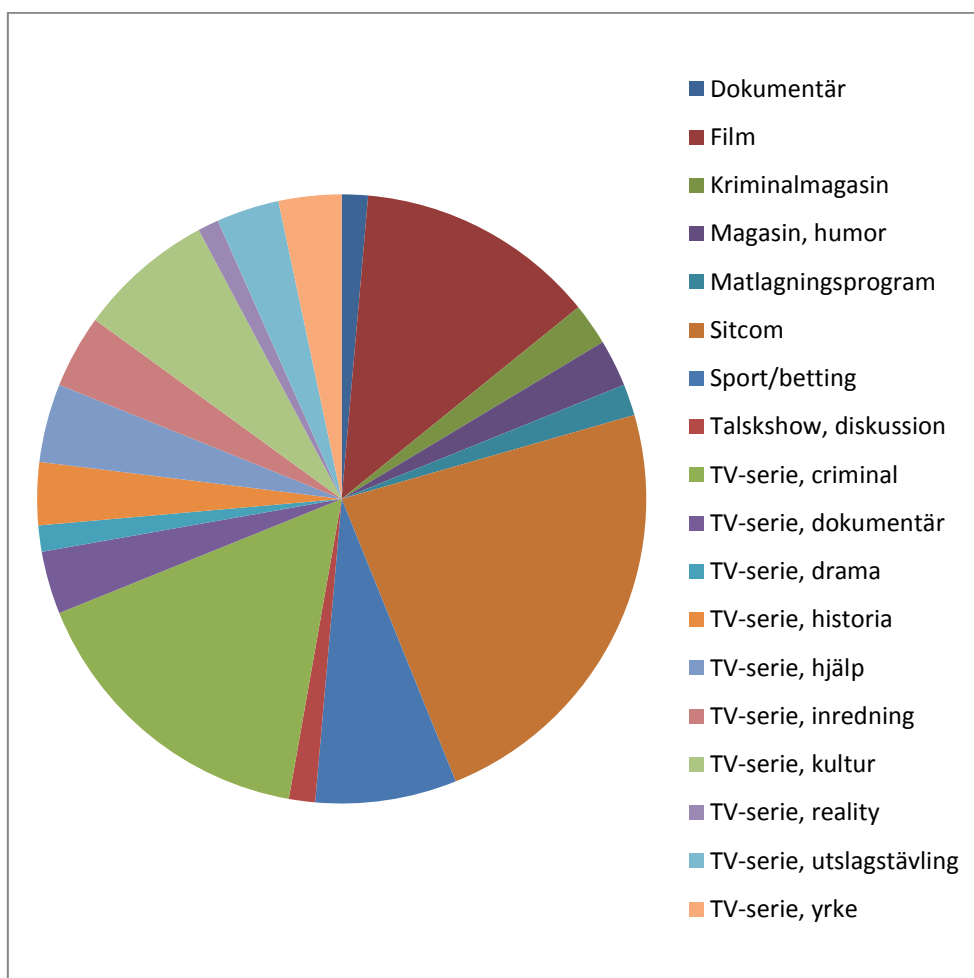
## 6.2. Alla, 3 – 99 år, kl. 20:00-22:59

| Alla, 3-99 år              | Estimat  | White SE | F-värde  | P-värde | Alla, 3-99 år     | Estimat  | White SE | F-värde | P-värde |
|----------------------------|----------|----------|----------|---------|-------------------|----------|----------|---------|---------|
| l (Intercept)              | 1,55     | 2,81E-02 | 3,07E+03 | 0       | K TV4             | 0,445806 | 0,029    | 236,299 | 0       |
| P Talkshow                 | 0,564719 | 0,08728  | 41,8629  | 1,2E-10 | a Svt1            | 0,302633 | 0,04164  | 52,8195 | 4,7E-13 |
| r TV-serie, hjälp          | 0,285926 | 0,06789  | 17,7399  | 2,6E-05 | n Kanal 5         | -0,40232 | 0,04291  | 87,9261 | 0       |
| o Talskshow, diskussion    | 0,214615 | 0,06344  | 11,4428  | 0,00073 | a TV3             | -0,462   | 0,04271  | 117,022 | 0       |
| g TV-serie, tävling        | 0,193108 | 0,07483  | 6,65965  | 0,00991 | l TV6             | -0,61077 | 0,04147  | 216,951 | 0       |
| r Kriminalmagasin          | 0,192638 | 0,0692   | 7,7491   | 0,00541 | TV8               | -1,02106 | 0,0419   | 593,949 | 0       |
| a TV-serie, utslagstävling | 0,149743 | 0,0496   | 9,11396  | 0,00256 | TV10              | -1,14056 | 0,04144  | 757,649 | 0       |
| m Debatt                   | -0,23773 | 0,04295  | 30,6409  | 3,4E-08 | Söndag            | 0,101079 | 0,02963  | 11,6399 | 0,00065 |
| k TV-serie, drama          | -0,25331 | 0,05883  | 18,5376  | 1,7E-05 | F Program i följd | 0,064097 | 0,02324  | 7,60381 | 0,00586 |
| a TV-serie, historia       | -0,29857 | 0,06384  | 21,8743  | 3E-06   | L Programlängd    | 0,001712 | 0,00046  | 13,9739 | 0,00019 |
| t Sitcom                   | -0,29941 | 0,04312  | 48,2164  | 4,7E-12 |                   |          |          |         |         |
| e TV-serie, yrke           | -0,32855 | 0,05901  | 31,0032  | 2,8E-08 |                   |          |          |         |         |
| g Film                     | -0,33132 | 0,05724  | 33,507   | 7,9E-09 |                   |          |          |         |         |
| o TV-serie, kultur         | -0,40949 | 0,06002  | 46,5544  | 1,1E-11 |                   |          |          |         |         |
| r TV-serie, dokumentär     | -0,42676 | 0,04381  | 94,8998  | 0       |                   |          |          |         |         |
| i Vädret                   | -0,45027 | 0,04116  | 119,661  | 0       |                   |          |          |         |         |
| Talkshow, kultur           | -0,56981 | 0,11189  | 25,9353  | 3,8E-07 |                   |          |          |         |         |
| TV-serie, övernaturlig     | -0,57909 | 0,0714   | 65,7797  | 0       |                   |          |          |         |         |
| Dokumentär                 | -0,71709 | 0,06643  | 116,526  | 0       |                   |          |          |         |         |
| Pausprogram                | -0,91187 | 0,06799  | 179,88   | 0       |                   |          |          |         |         |

**Tabell 21.** Resultat från regressionsanalysen för Alla, 3-99 år, sammanfattas ovan, med Nyheter, tisdagar och program i följd som benchmark.

Den här målgruppens regressionsresultat är dels intressant att jämföra med övriga målgrupper, dels är det även intressant för de TV-kanaler som inte har en definierad målgrupp, utan vill nå ut till så många som möjligt samtidigt. Som regressionsresultatet i Tabell 21. visar är ”Talkshow”, ”TV-serie, hjälp” och ”Talkshow, diskussion” de populäraste programkategorierna. Programkategorier som bör undvikas är ”TV-serie, Övernaturlig”, ”Dokumentär” och ”Pausprogram”. Program som Skavlan, Lyxfällan, Arga Snickaren, Berg Flyttar In och Så Mycket Bättre är följaktligen rekommenderade i största allmänhet. Program att undvika i den här målgruppen är däremot Anslagstavlan och enstaka dokumentärinslag, men också TV-serier som True Blood och Grimm. Vid nyproduktion av TV-program ger det här en fingervisning angående vad TV-bolagen bör satsa på. Det är även, till viss del, rekommenderat att sända TV-program i följd efter varandra när det finns fler avsnitt eftersom det påverkar tittarsiffrorna positivt. Den kanal som i särklass når ut bäst till Sveriges TV-tittare är TV4.

### 6.3. Alla, 15 – 49 år, kl. 20:00-22:59, kanalgruppens generella målgrupp



**Figur 9.** TV3, TV6, TV8 & TV10:s TV-tablå, fördelad på programkategorier under prime time i november 2013, hämtad ur MMS-data.

| Alla, 15-49 år             | Estimat  | White SE | F-värde  | P-värde |
|----------------------------|----------|----------|----------|---------|
| l (Intercept)              | 1,11     | 2,74E-02 | 1,65E+03 | 0       |
| P Talkshow                 | 0,626637 | 0,08433  | 55,21    | 1,4E-13 |
| r TV-serie, utslagstävling | 0,287175 | 0,04639  | 38,3252  | 6,9E-10 |
| o TV-serie, hjälp          | 0,274581 | 0,06394  | 18,4387  | 1,8E-05 |
| g TV-serie, tävling        | 0,248828 | 0,06737  | 13,6408  | 0,00023 |
| r Talkshow, diskussion     | 0,212878 | 0,06091  | 12,2153  | 0,00048 |
| a Kriminalmagasin          | 0,142663 | 0,05333  | 7,15675  | 0,00751 |
| m TV-serie, yrke           | -0,17151 | 0,051    | 11,3083  | 0,00078 |
| k Debatt                   | -0,24382 | 0,03921  | 38,6645  | 5,8E-10 |
| a TV-serie, kultur         | -0,25124 | 0,04555  | 30,4179  | 3,8E-08 |
| t TV-serie, övernaturlig   | -0,25922 | 0,06988  | 13,7606  | 0,00021 |
| e TV-serie, dokumentär     | -0,26469 | 0,03697  | 51,2498  | 1E-12   |
| g Vädret                   | -0,38904 | 0,04281  | 82,595   | 0       |
| o Dokumentär               | -0,47468 | 0,05906  | 64,6037  | 0       |
| r Talkshow, kultur         | -0,48223 | 0,08565  | 31,6993  | 2E-08   |
| i Pausprogram              | -0,72723 | 0,05743  | 160,35   | 0       |

| Alla, 15-49 år | Estimat  | White SE | F-värde | P-värde |
|----------------|----------|----------|---------|---------|
| K TV4          | 0,735216 | 0,02787  | 695,784 | 0       |
| a Svt1         | 0,249789 | 0,03462  | 52,0722 | 6,8E-13 |
| n Kanal 5      | 0,159527 | 0,04091  | 15,2033 | 9,9E-05 |
| a TV3          | 0,017997 | 0,03986  | 0,20383 | 0,65168 |
| l TV6          | -0,10845 | 0,03852  | 7,92676 | 0,0049  |
| TV8            | -0,74208 | 0,03675  | 407,687 | 0       |
| TV10           | -0,81595 | 0,03664  | 495,795 | 0       |
| Söndag         | 0,103421 | 0,02904  | 12,6852 | 0,00037 |

**Tabell 22.** Resultat från regressionsanalysen för Alla, 15-49 år, sammanfattas ovan, med Nyheter, tisdagar och program i följd som benchmark.

Nepa formulerade den här målgruppen som en generell målgrupp för kanalgruppen, vilken som sagt består av TV3, TV6, TV8 och TV10.

Som synes i Tabell 22. är ”Talkshow”, ”TV-serie, utslagstävling” och ”TV-serie, hjälp” de tre programkategorierna som driver högst TV-tittande inom målgruppen, Alla 15-49 år. I jämförelse med de program som sänds i kanalgruppen; TV3, TV6, TV8 och TV10 ser man att ”TV-serie, utslagstävling” och ”TV-serie, hjälp” sänds men inte ”Talkshow”. Rekommendation till kanalgruppen är således att börja sända denna typ av programkategori under tidsintervallet 20:00-22:59. Vidare sänder kanalgruppen programkategorierna ”Sitcom” och ”Film” i stor utsträckning. Tyvärr gick det inte att säga någonting om just de  $\beta$ -värdena för den här målgruppen. ”Talkshow, diskussion” är högt värderat av tittarna och sänds relativt ofta av kanalgruppen, vilket är fördelaktigt.

Programkategorier som framförallt bör undvikas är ”Pausprogram”, ”Talkshow, kultur” och ”Dokumentär”. Som också gäller för övriga målgrupper är programkategorin ”Vädret” ett överflödigt inslag i nyhetsprogram. Nyheter används som benchmark och trots att vädret är en del av nyhetssändningen byter många kanal just när det kommer på.

Vidare går det att utläsa att kanalgruppen utgörs av de kanaler som driver minst tittarsiffror inom målgruppen. TV6, TV8 och TV10 har dessutom en negativ inverkan på tittarsiffrorna vid jämförelse med benchmark. Således kan slutsatsen dras att kanalgruppen lyckats dåligt med att nå ut till sin gemensamt uppställda målgrupp. Extra tydligt blir detta vid jämförelse med TV6, som är en fri-TV-kanal med Kanal 5, som är en abonnemangskanal, där Kanal 5 trots detta driver högre tittarprocent än TV6 i den här målgruppen, som är en bedrift i sig.

## 6.4. Kvinnor, 25 – 44 år, kl. 20:00-22:59, TV3:s målgrupp

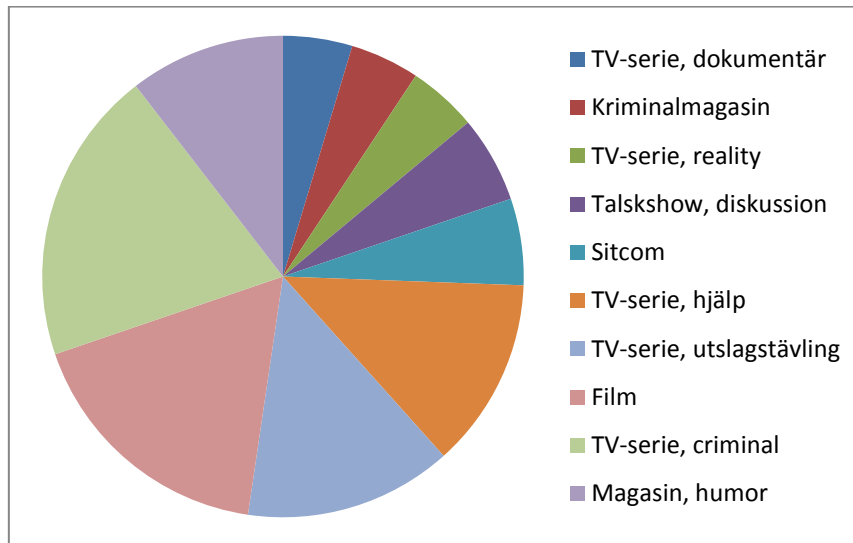
3



Sände oftast programmen 2013: How I Met Your Mother, Jims Värld, Lyxfällan

Populäraste programmen på TV3 2013: Lyxfällan – hur gick det sen, Lyxfällan, Efterlyst

**Tabell 23.** Tabell som beskriver TV3 generellt i november 2013. Populäritet och frekvens är hämtade ur MMS-data.



**Figur 10.** TV3:s TV-tablå, fördelad på programkategorier under prime time i november 2013, hämtad ur MMS-data.

| Kvinnor, 25-44 år          | Estimat  | White SE | F-värde  | P-värde | Kvinnor, 25-44 år | Estimat  | White SE | F-värde | P-värde |
|----------------------------|----------|----------|----------|---------|-------------------|----------|----------|---------|---------|
| l (Intercept)              | 1,18     | 3,32E-02 | 1,27E+03 | 0       | K TV4             | 0,832791 | 0,03091  | 726,007 | 0       |
| P Talkshow                 | 0,685388 | 0,09132  | 56,3345  | 8,1E-14 | a Kanal 5         | 0,361716 | 0,04831  | 56,0558 | 9,4E-14 |
| r TV-serie, utslagstävling | 0,310053 | 0,05121  | 36,663   | 1,6E-09 | n Svt1            | 0,291623 | 0,03983  | 53,6184 | 3,2E-13 |
| o TV-serie, hjälp          | 0,258152 | 0,0727   | 12,608   | 0,00039 | a TV3             | 0,137344 | 0,04545  | 9,13156 | 0,00253 |
| g TV-serie, tävling        | 0,202385 | 0,07627  | 7,04071  | 0,00801 | l TV6             | -0,3093  | 0,04375  | 49,9816 | 1,9E-12 |
| r TV-serie, inredning      | 0,168023 | 0,05663  | 8,80475  | 0,00303 | TV8               | -0,70606 | 0,04337  | 264,988 | 0       |
| a Sitcom                   | -0,12273 | 0,04622  | 7,05034  | 0,00797 | TV10              | -0,82493 | 0,04057  | 413,439 | 0       |
| m Sport/betting            | -0,16764 | 0,06292  | 7,09837  | 0,00776 | Söndag            | 0,114244 | 0,03349  | 11,6363 | 0,00066 |
| k TV-serie, yrke           | -0,24245 | 0,06001  | 16,3238  | 5,5E-05 | Måndag            | 0,076274 | 0,02873  | 7,04908 | 0,00798 |
| a TV-serie, historia       | -0,25031 | 0,08327  | 9,03668  | 0,00267 |                   |          |          |         |         |
| t TV-serie, kultur         | -0,31404 | 0,04967  | 39,9706  | 3E-10   |                   |          |          |         |         |
| e TV-serie, dokumentär     | -0,34066 | 0,04162  | 66,9844  | 0       |                   |          |          |         |         |
| g Debatt                   | -0,34599 | 0,04453  | 60,3778  | 1,1E-14 |                   |          |          |         |         |
| o Vädret                   | -0,48504 | 0,04904  | 97,8083  | 0       |                   |          |          |         |         |
| r Talkshow, kultur         | -0,51343 | 0,10174  | 25,4683  | 4,8E-07 |                   |          |          |         |         |
| i Dokumentär               | -0,58997 | 0,06462  | 83,3661  | 0       |                   |          |          |         |         |
| Pausprogram                | -0,88269 | 0,07095  | 154,769  | 0       |                   |          |          |         |         |

**Tabell 24.** Resultat från regressionsanalysen för Kvinnor, 25-44 år, sammanfattas ovan, med Nyheter, tisdagar och program i följd som benchmark.

Som synes i Tabell 24. är ”Talkshow”, ”TV-serie, utslagstävling” och ”TV-serie, hjälp” de tre programkategorierna som driver högst TV-tittande inom målgruppen Kvinnor, 25-44 år. I Tabell 23. framgår det även att ”Lyxfällan – hur gick det sen” var det allra populäraste programmet i November 2013, vilket stämmer väl överens med regressionsanalysen, då den tillhör programkategorin ”TV-serie, hjälp”. I jämförelse med de program som sänds på TV3 ser man att ”TV-serie, utslagstävling” och ”TV-serie, hjälp” sänds, men inte ”Talkshow”. TV3 sänder inte heller de fjärde och femte mest populära programkategorierna, ”TV-serie, tävling” och ”TV-serie, inredning”. En rekommendation till TV3 är således att börja sända dessa typer av programkategorier under tidsintervallet 20:00-22:59. Programkategorier som framförallt bör undvikas är ”Pausprogram”, ”Dokumentär” och ”Talkshow, kultur”. ”TV-serie, dokumentär” sänds i viss utsträckning av TV3 i dagsläget, men är inte särskilt högt prioriterat av målgruppen enligt regressionsanalysen. Generellt är ”TV-serie, dokumentär” inte att föredra ur ett TV-tittarmaximerande perspektiv. För se exakt vilka programkategorier som utgörs av vilka program hänvisas till Bilaga 5.

Gällande TV3:s popularitet i målgruppen presterar de bäst inom den kanalgrupp de tillhör, det vill säga bland kanalerna TV3, TV6, TV8 och TV10. TV3 och Kanal 5 konkurrerar på liknande villkor eftersom båda är abonnemangskanaler, men trots det lyckas Kanal 5 mycket bättre i den här målgruppen. Kanal 5 lyckas till och med bättre än SVT1, fast Kanal 5 är en abonnemangskanal och SVT1 en frikanal. Det är anmärkningsvärt att TV3 inte lyckats bättre, eftersom det här är deras målgrupp och de har funnits länge på den svenska TV-marknaden.

## 6.5. Kvinnor, 25 – 49 år, kl. 20:00-22:59, TV8:s målgrupp



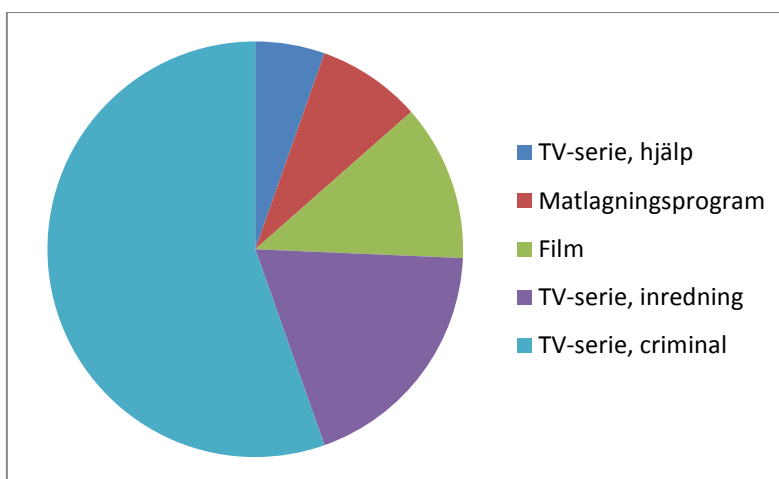
Sände oftast programmen 2013:

En Plats på Landet, Celebrity Masterchef, The Mentalist/Seinfeld/En Plats i Solen Borta Eller Hemma

Populäraste programmen på TV8 2013:

Morden i Midsomer, The Mentalist, Father Brown

**Tabell 25.** *Tabell som beskriver TV8 generellt i november 2013. Populäritet och frekvens är hämtade ur MMS-data.*



**Figur 11.** *TV8:s TV-tablå, fördelad på programkategorier under prime time i november 2013, hämtad ur MMS-data.*

| Kvinnor, 25-49 år          | Estimat  | White SE | F-värde  | P-värde | Kvinnor, 25-49 år | Estimat  | White SE | F-värde | P-värde |
|----------------------------|----------|----------|----------|---------|-------------------|----------|----------|---------|---------|
| I (Intercept)              | 1,27     | 3,03E-02 | 1,76E+03 | 0       | K TV4             | 0,836726 | 0,02933  | 813,603 | 0       |
| P Talkshow                 | 0,666132 | 0,0901   | 54,6638  | 1,9E-13 | a Svt1            | 0,323179 | 0,03982  | 65,869  | 0       |
| r TV-serie, utslagstävling | 0,273482 | 0,05028  | 29,5824  | 5,8E-08 | n Kanal 5         | 0,261155 | 0,04729  | 30,4915 | 3,7E-08 |
| o TV-serie, hjälp          | 0,270382 | 0,07122  | 14,4138  | 0,00015 | a TV6             | -0,38251 | 0,04268  | 80,3205 | 0       |
| g TV-serie, inredning      | 0,15363  | 0,05388  | 8,12902  | 0,00439 | l TV8             | -0,7501  | 0,0425   | 311,441 | 0       |
| r Film                     | -0,15358 | 0,05448  | 7,94551  | 0,00485 | TV10              | -0,88872 | 0,0391   | 516,713 | 0       |
| a Magasin, humor           | -0,15772 | 0,06103  | 6,67863  | 0,00981 | Söndag            | 0,108327 | 0,03189  | 11,5361 | 0,00069 |
| m Sitcom                   | -0,17654 | 0,04478  | 15,5438  | 8,3E-05 |                   |          |          |         |         |
| k Sport/betting            | -0,19392 | 0,06004  | 10,4322  | 0,00125 |                   |          |          |         |         |
| a TV-serie, yrke           | -0,26062 | 0,0571   | 20,8348  | 5,2E-06 |                   |          |          |         |         |
| t TV-serie, historia       | -0,28355 | 0,07882  | 12,9427  | 0,00033 |                   |          |          |         |         |
| e TV-serie, övernaturlig   | -0,29813 | 0,09295  | 10,2888  | 0,00135 |                   |          |          |         |         |
| g TV-serie, kultur         | -0,3521  | 0,04933  | 50,9416  | 1,2E-12 |                   |          |          |         |         |
| o Debatt                   | -0,3524  | 0,04256  | 68,5694  | 0       |                   |          |          |         |         |
| r TV-serie, dokumentär     | -0,36125 | 0,04108  | 77,3475  | 0       |                   |          |          |         |         |
| i Vädret                   | -0,45405 | 0,04203  | 116,729  | 0       |                   |          |          |         |         |
| o Talkshow, kultur         | -0,53812 | 0,09623  | 31,2674  | 2,5E-08 |                   |          |          |         |         |
| o Dokumentär               | -0,6177  | 0,06485  | 90,7336  | 0       |                   |          |          |         |         |
| o Pausprogram              | -0,82687 | 0,06483  | 162,688  | 0       |                   |          |          |         |         |

**Tabell 26.** *Resultat från regressionsanalysen för Kvinnor, 25-49 år, sammanfattas ovan, med Nyheter, tisdagar och program i följd som benchmark.*



TV8:s målgrupp, som detta avsnitt behandlar, skiljer sig inte avsevärt från TV3:s målgrupp, förutom att den också utgörs av upp till fem år äldre kvinnor i åldersspannet, det vill säga kvinnor i åldern 25 till 49 år. Som syns i Tabell 26. är även här ”Talkshow”, ”TV-serie, utslagstävling” och ”TV-serie, hjälp” de mest populära. För att jämföra med vad TV8 sände i november 2013 kan förbättringar göras av TV-tablåns utbud enligt regressionsanalysens resultat. Det går däremot inte att säga någonting om vissa  $\beta$ -värden, exempelvis programkategorin ”TV-serie, criminal”, som TV8 sänder mycket av, men det går att säga desto mer om andra. En viss variation i TV-tablån är givetvis önskvärt ur TV-tittarnas perspektiv, men ”Talkshow”, ”TV-serie, utslagstävling” och ”TV-serie, hjälp” skulle kunna sändas i större utsträckning för att fånga TV-tittarna i den här målgruppen. ”TV-serie, inredning” sänds relativt ofta, vilket är bra för kanalens tittarsiffror, enligt regressionsresultatet. Utöver det prioriteras programkategorin ”Film” till viss del högt av målgruppen, som TV8 sänder relativt ofta, och borde således även göra det framöver.

Programkategorier som bör undvikas är ”Pausprogram”, ”Dokumentär” och ”Talkshow, kultur”. Alla programkategorier med de respektive programmen återfinns i Bilaga 5.

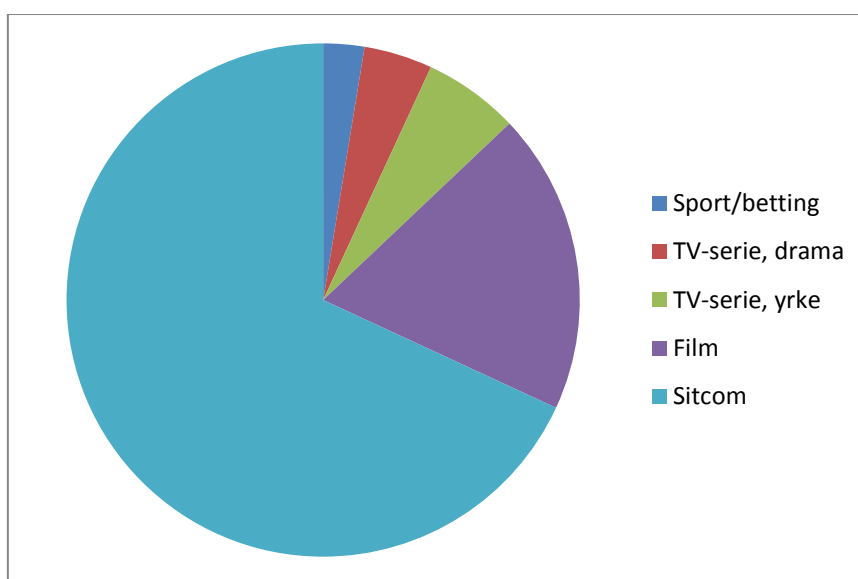
Gällande TV8:s popularitet i målgruppen i fråga så framgår det i Tabell 26. att TV8 presterar bättre än TV10. I övrigt presterar dock samtliga av de andra kanalerna i regressionsanalysen bättre. TV8 är en nystartad kanal i jämförelse med övriga TV-kanaler, förutom TV10, och det kan delvis vara en förklaring till varför de presterar så dåligt vid en jämförelse.

## 6.6. Män, 15 – 34 år, kl. 20:00-22:59, TV6:s målgrupp



|                                     |   |
|-------------------------------------|---|
| Sände oftast programmen 2013:       | Simpsons, Scrubs, Family Guy  |
| Populäraste programmen på TV6 2013: | (Fotboll CL Psg-And), Star Wars Episod 6 Jedins Återkomst, 2 ½ Män, Unstoppable |

**Tabell 27.** Tabell som beskriver TV6 generellt i november 2013. Populäritet och frekvens är hämtade ur MMS-data.



**Figur 12.** TV8:s TV-tablå, fördelad på programkategorier under prime time i november 2013, hämtad ur MMS-data.

| Män, 15-34 år              | Estimat  | White SE | F-värde | P-värde |
|----------------------------|----------|----------|---------|---------|
| l (Intercept)              | 0,675    | 2,91E-02 | 539     | 0       |
| P Talkshow                 | 0,641415 | 0,08447  | 57,655  | 4,2E-14 |
| r TV-serie, tävling        | 0,399292 | 0,06604  | 36,5572 | 1,7E-09 |
| o TV-serie, hjälp          | 0,381961 | 0,06498  | 34,5506 | 4,6E-09 |
| g Kriminalmagasin          | 0,321085 | 0,05996  | 28,6759 | 9,2E-08 |
| r Talkshow, diskussion     | 0,319876 | 0,0597   | 28,7047 | 9,1E-08 |
| a TV-serie, utslagstävling | 0,311387 | 0,05254  | 35,121  | 3,5E-09 |
| m Sport/betting            | 0,247739 | 0,07453  | 11,0486 | 0,0009  |
| k Magasin, humor           | 0,186522 | 0,05688  | 10,7537 | 0,00105 |
| a TV-serie, inredning      | 0,164944 | 0,05005  | 10,8597 | 0,001   |
| t Sitcom                   | 0,145261 | 0,04425  | 10,7785 | 0,00104 |
| e Film                     | 0,124236 | 0,05222  | 5,66001 | 0,01742 |
| g Dokumentär               | -0,2029  | 0,06095  | 11,0807 | 0,00088 |
| o Talkshow, kultur         | -0,2548  | 0,08784  | 8,41404 | 0,00375 |
| r Vädret                   | -0,38944 | 0,05373  | 52,5311 | 5,4E-13 |
| i Pausprogram              | -0,41858 | 0,05853  | 51,1401 | 1,1E-12 |

| Män, 15-34 år | Estimat  | White SE | F-värde | P-värde |
|---------------|----------|----------|---------|---------|
| K TV4         | 0,755623 | 0,03032  | 621,158 | 0       |
| a TV6         | 0,425644 | 0,04097  | 107,921 | 0       |
| n Kanal 5     | 0,260775 | 0,04047  | 41,5162 | 1,4E-10 |
| a Svt1        | 0,176799 | 0,03326  | 28,2626 | 1,1E-07 |
| l TV3         | 0,119293 | 0,04133  | 8,33136 | 0,00393 |
| TV8           | -0,53733 | 0,03686  | 212,542 | 0       |
| TV10          | -0,59718 | 0,03853  | 240,165 | 0       |
| Söndag        | 0,105722 | 0,03232  | 10,7013 | 0,00108 |

**Tabell 28.** Resultat från regressionsanalysen för Män, 15-34 år, sammanfattas ovan, med Nyheter, tisdagar och program i följd som benchmark.

Som synes i Tabell 28. över regressionsresultaten når TV6 ut till sin målgrupp väl. Det är den näst mest populära kanalen efter TV4, vilket inte är fallet bland de andra målgrupperna. TV6 sänder främst sitcom-serier, som synes i Tabell 27. Sitcom-serier hamnar dock inte särskilt högt upp i tabellen över de program som bidrar mest till TV-tittande för den här målgruppen. Av den anledningen kan TV6 med fördel sända programmen ovanför sitcom i tabellen i större utsträckning istället och på så sätt troligtvis öka sina tittarsiffror i det här tidsspannet. Observera att allt underlag baseras på tidsspannet 20:00-22:59 och följaktligen kan programkategorin ”Sitcom” vara populär andra tider på dygnet, men just under prime time är det rekommenderat att sända programkategorier högst upp i tabellen för att nå ut till målgruppen Män, 15-34 år, i så stor utsträckning som möjligt.

I dagsläget sänder TV6 även filmer flitigt, vilket syns i Figur 11., och som skulle kunna prioriteras ned och ge plats för de programkategorier med högre estimat i Tabell 28. som till exempel ”Talkshow”. Det är också intressant att nämna att varken programföljd eller –längd bidrar till tittarsiffrorna. Utöver det är söndag den enda dagen, som enskilt med signifikans skiljer sig från benchmark. Som synes i  $\beta$ -estimatet bidrar söndag positivt till tittarsiffror. Vidare bör TV6 undvika att sända programkategorierna längst ned i tabellen. Angående programkategorin Sport/betting gick det inte att säga någonting för de andra målgrupperna, men för Män, 15-34 år, ger den en positiv påverkan på TV-tittandet och som synes i Tabell 27. är även det populäraste programmet på TV6 under november 2013 en fotbollsmatch. För exakt vilka program som utgör de olika programkategorierna hänvisas till Bilaga 5.

## 6.7. Män, 25 – 49 år, kl. 20:00-22:59, TV10:s målgrupp



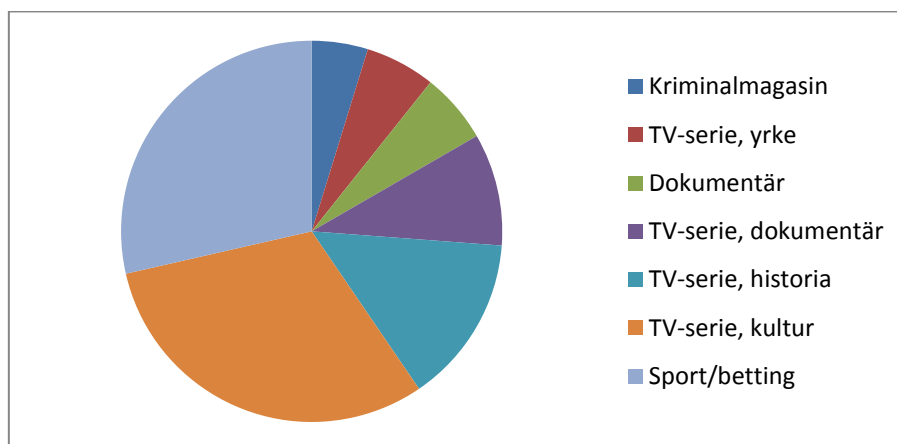
Sände oftast programmen på TV10 2013:

Rick's Antikhandel, Zachrissons Osminne, Träskjägarna

Populäraste programmen bland alla 2013:

(Ishockey KC Sve-Tje), Jordan Runt med Ewan McGregor, Världens Märkligaste Vapen, (NHL Nyr-Van), Band of Brothers

**Tabell 29.** Tabell som beskriver TV10 generellt i november 2013. Populäritet och frekvens är hämtade ur MMS-data.



**Figur 13.** TV8:s TV-tablå, fördelad på programkategorier under prime time i november 2013, hämtad ur MMS-data.

| Män, 25-49 år              | Estimat  | White SE | F-värde  | P-värde | Män, 25-49 år | Estimat  | White SE | F-värde | P-värde |
|----------------------------|----------|----------|----------|---------|---------------|----------|----------|---------|---------|
| l (Intercept)              | 1,27     | 3,09E-02 | 1,69E+03 | 0       | K TV4         | 0,616737 | 0,031    | 395,761 | 0       |
| P Talkshow                 | 0,569709 | 0,08627  | 43,6062  | 4,8E-11 | a Svt1        | 0,216976 | 0,03718  | 34,048  | 6E-09   |
| r Talskshow, diskussion    | 0,238691 | 0,06292  | 14,3934  | 0,00015 | n TV6         | -0,15635 | 0,04163  | 14,1074 | 0,00018 |
| o TV-serie, hjälp          | 0,233019 | 0,06654  | 12,2619  | 0,00047 | a TV3         | -0,16116 | 0,04314  | 13,9524 | 0,00019 |
| g TV-serie, tävling        | 0,186055 | 0,06641  | 7,8485   | 0,00512 | l TV8         | -0,8923  | 0,04098  | 474,18  | 0       |
| r Kriminalmagasin          | 0,177138 | 0,05667  | 9,7714   | 0,00179 | TV10          | -0,91808 | 0,04403  | 434,768 | 0       |
| a TV-serie, utslagstävling | 0,141487 | 0,05055  | 7,83535  | 0,00516 | Söndag        | 0,128933 | 0,03152  | 16,727  | 4,4E-05 |
| m Nyheter, sport           | 0,140062 | 0,04815  | 8,46062  | 0,00366 |               |          |          |         |         |
| k TV-serie, yrke           | -0,16267 | 0,05795  | 7,88082  | 0,00503 |               |          |          |         |         |
| a TV-serie, drama          | -0,17973 | 0,05469  | 10,8007  | 0,00103 |               |          |          |         |         |
| t TV-serie, kultur         | -0,22885 | 0,05604  | 16,6751  | 4,6E-05 |               |          |          |         |         |
| e Debatt                   | -0,23152 | 0,04944  | 21,9283  | 3E-06   |               |          |          |         |         |
| g TV-serie, övernaturlig   | -0,27455 | 0,08934  | 9,44501  | 0,00214 |               |          |          |         |         |
| o TV-serie, dokumentär     | -0,29648 | 0,04189  | 50,0877  | 1,8E-12 |               |          |          |         |         |
| r Vädret                   | -0,35474 | 0,04727  | 56,3097  | 8,3E-14 |               |          |          |         |         |
| i Dokumentär               | -0,4884  | 0,06961  | 49,2294  | 2,8E-12 |               |          |          |         |         |
| o Talkshow, kultur         | -0,4943  | 0,10151  | 23,7097  | 1,2E-06 |               |          |          |         |         |
| o Pausprogram              | -0,81746 | 0,0754   | 117,534  | 0       |               |          |          |         |         |

**Tabell 30.** Resultat från regressionsanalysen för Män, 25-49 år, sammanfattas ovan, med Nyheter, tisdagar och program i följd som benchmark.

TV10:s målgrupp, liksom alla övriga målgrupper, tittar mest på programkategorierna överst i Tabell 30., vilka går att utläsa som ”Talkshow”, ”Talkshow, diskussion” och ”TV-serie, hjälp”. TV10:s programmix går att ifrågasätta i dagsläget, eftersom de sänder sällsynt impopulära program bland den målgrupp de ämnar nå ut bäst till. Vissa programkategorier går det däremot inte att säga någonting om, men när det gäller ”TV-serie, kultur” och ”TV-serie, dokumentär” hamnar dessa ganska långt ned i tabellen, och trots detta sänds de ofta på TV10. Programkategorin ”Dokumentär”, som består av enskilda dokumentäraavsnitt som inte tillhör någon serie, hamnar synnerligen lågt ner i tabellen och prioriteras följaktligen inte alls av den här, eller någon annan av målgrupperna som går att undersöka. Det är uppenbarligen inte en bra kombination. Måhända profilerar sig TV10 som en lite mer intellektuell kanal med tanke på vilka program de sänder, men om målet är få så höga tittarsiffror som möjligt inom den definierade målgruppen, prioriterar de fel i dagsläget. Att komma in på någon diskussion rörande marknadssegment eller nischmarknadsföring är utanför det här kandidatexamensarbetets ramar. I sammanhanget bör också påpekas att TV10 är den mest nystartade TV-kanalen bland dem som undersöks, varför det inte är konstigt om TV-tittarna inte hittat dit än. För exakta program som sänts under prime time i de respektive programkategorierna, hänvisas till Bilaga 5.

## 6.8. Rekommendationer

Även om slutsatsen i det här kandidatexamensarbetet är att det går att tillämpa regressionsanalys på MMS-data för att analysera TV-tittande, har framtida rekommendationer formulerats av författarna för att förenkla eventuella fortsatta studier inom området. MMS-data är exempelvis omständligt att skriva om manuellt till ett regressionsvänligt underlag. Förutsatt att multipel linjär regression blir ett vedertaget verktyg för analys av TV-tittarhistorik, skulle det vara fördelaktigt om det gick att dra ut underlag för TV-tittarhistoriken i form av dummyvariabler direkt ur MMS:s program HotProg. I nuläget består TV-tittarhistorikutdraget ur HotProg av förklarande text, medan regressionsvänlig historik är uppdelad efter kategori med 1:or och 0:or, enligt Kapitel 2.2. som behandlar Dummyvariabler.

MMS är även inkonsekvent i sin kategorisering av TV-program. Exempelvis har troligtvis ekonominyheterna slagits ihop med nyhetssändningen 2012 och framåt eftersom programmet dessförinnan återfinns i datan, medan det hädanefter inte går att hitta, även om det inslaget fortfarande är en del av nyhetssändningen på TV. Det är en felkälla som leder till problem i regressionsanalysen. MMS:s programkategorisering är även bristfällig ur ett regressionsanalysperspektiv, eftersom den i många fall avviker från andra källor som exempelvis [www.tv.nu](http://www.tv.nu) och [www.imdb.com](http://www.imdb.com), som sinsemellan också är inkonsekventa. Det är således långt ifrån självklart var ett TV-program hör hemma. Vidare betecknar MMS många program som "Övrigt", vilket inte är önskvärt ur ett regressionsperspektiv. Ett försök till att lägga en grund för hur en regressionsvänlig programindelning ska se ut har formulerats i det här arbetet, men även den har många förbättringsmöjligheter. Beroende på vad som är intressant i regressionsanalysen kan kategorierna delas upp ytterligare. Om analysen ämnar fokusera på programkategorin "Film" eller "Dokumentär" är det motiverat att dela in dem ytterligare i komedi, thriller och så vidare respektive natur- och sportdokumentärer exempelvis.

Regressionsmodellen går att förbättra så till vida att exempelvis torsdag, fredag och lördag kan slås ihop med benchmarket tisdagar, förutsatt att regressionsanalysen utförs på tidsspannet 20:00-22:59. Utöver den åtgärden skulle fenomenet Trailing med fördel också kunna tas med som en variabel i modellen för att öka dess förklaringsgrad. För förklaring kring Trailing och ytterligare variabler att föra in i regressionsmodellen, se Kapitel 7. När det å andra sidan gäller det ekonomiska perspektivet, att olika program kostar olika mycket att köpa in, har inte det tagits med i det här arbetet då det uteslutande antagits att TV-kanaler vill maximera tittaranantalet. Avvägningen mellan föreslagna åtgärder och kostnaden att implementera dem har lagts på TV-kanalen. Författarna är väl införstådda med att en åtgärd som genererar dubbelt så många tittare, kan vara exempelvis tre gånger så dyr som den i dagsläget, vilket eventuellt inte är försvarbart ur ledningens synvinkel. Det skulle kunna tas med i vidare analyser på det här ämnet eftersom det är en avvägning TV-kanalerna måste göra, mellan optimala investeringar ur ekonomiska respektive tittargenererande perspektiv.

När det gäller transformeringen av  $Y$ -variabeln i den uppställda regressionsmodellen finns andra tillvägagångssätt än de som tagits upp i det här kandidatexamensarbetet. Som beskrivs under Kapitel 4.3. är  $Y$ -variabeln i regressionsmodellen begränsad underifrån vid värdet 0 och begränsad ovanifrån vid värdet 100, även om den övre begränsningen aldrig nås, eftersom 100% av andelen TV-tittare på ett program kan anses vara praktiskt omöjligt. När  $Y$ -variabeln ser ut som den gör i det här kandidatexamensarbetet lämpar sig transformeringsmetoder i olika stor utsträckning. En annan transformeringsmetod, som inte har tagits upp i arbetet hittills, är en regressionsmetod kallad TOBIT,

vilken tar hänsyn till just denna typ av observationsmängd<sup>53</sup>. Således rekommenderas det utredas i vilken mån regressionanalys av MMS-data, med grundekvationen som Formel (21), kan förbättras med tillämpning av TOBIT. TOBIT skulle möjligtvis även möjliggöra analys i de övriga tidsspannen.

---

<sup>53</sup> Peter Kennedy, *A Guide to Econometrics*, 6th edn (United States of America: Blackwell Publishing Ltd, 2008), pp. 263-65.

## 7. Kritik riktad mot regressionsmodellen med stöd ur marknadsföringsteori

### 7.1. Inledning

Regressionsmodellen i det här kandidatexamensarbetet är uppställd enligt Formel (21) på sida 17. Författarna inser dess begränsningar eftersom modellen är en förenkling av verkligheten. Dock hade den varit begränsat praktiskt tillämpbar utan viss reducering. Exempelvis är det svårt att få tag i underlag för alla faktorer som påverkar kunden, men det är också svårt att direkt mäta vissa av dessa faktorer. Som kritik och komplement definieras därför en ny regressionsmodell här, med stöd ur marknadsföringsteori, och som ämnar fånga upp *alla* ur kundens perspektiv värdeskapande faktorer, som TV-bolagen påverkar själva.

### 7.2. Teori

För att definiera och utvärdera en produkt, i det här fallet TV-kanalen, är det viktigt att förstå dess ingående delar och framförallt dess konkurrensfördelar, sätt på vilka den kan utmärka sig på marknaden och på så sätt få ett försprång i jämförelse med sina konkurrenter (exempelvis hur en TV-kanal kan prioritera olika för att nå fram till olika kundsegment). Produkter och tjänster kan analyseras med en rad olika metoder. Ett klassiskt, vedertaget instrument för att kombinera marknadsföringsmixen för en produkt är De Fyra P:na.<sup>54</sup>

De Fyra P:na i marknadsföringsmixen för produkter består av:

1. *Produkt* – någonting som kan erbjudas marknaden för att uppnå uppmärksamhet, förvärv, användning eller konsumtion och tillgodoser en vilja eller ett behov. Produkten kan vara ett fysiskt objekt, personer, upplevelser, organisationer och/eller idéer.
2. *Pris* – den summa pengar som debiteras i utbyte mot en produkt eller en tjänst. Pris återspeglar summan av det ekonomiska värde konsumenten väger mot de fördelar innehavet eller användandet av produkten medför.
3. *Plats* – alla företagsaktiviteter som leder till att produkten görs tillgänglig för målgruppen.
4. *Reklam (Promotion)* – aktiviteter som kommunicerar produkten eller tjänsten och dess fördelar till målgruppen och övertalar dem om att de behöver den.

För tjänster kompletteras dessa med:<sup>55</sup>

5. *Paket/Fysiskt bevis/Process* – allting kunden upplever från första gången denne kommer i kontakt med företaget och hela vägen genom köpprocessen, inklusive någon form av fysiskt bevis som ett kvitto på att ett värdeutbyte ägt rum. Exempel på ett sådant fysiskt bevis av en utförd tjänst skulle kunna vara en hårklippning hos frisören.

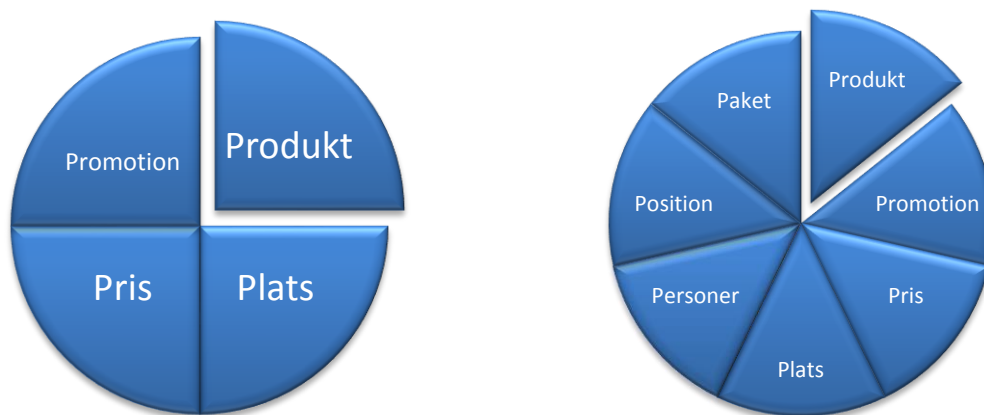
---

<sup>54</sup> Philip Kotler, *Principles of Marketing*, 2nd edn (New Jersey: Prentice Hall Europe, 1999), pp. 109-11.

<sup>55</sup> Brian Tracy, *Million Dollar Habits* (United States: Entrepreneur Media Incorporated, 2004).



6. *Personer* – ansvariga inom och utom affärsverksamheten för varje delelement av försäljningen och marknadsföringen. I boken *Good to Great* av Jim Collins (2001) görs en metafor där företag liknas vid en buss. De bästa företagen ser inte bara till så att rätt personer är på bussen, men också att de sitter på rätt plats, utifrån uppställd målformulering.
7. *Positionering* – utgörs av vad kunder och potentiella kunder känner inför företaget rent känslomässigt, men också hur de uttrycker sig om företaget utan dess närvaro. Ett slags fördoms- eller förväntansbaserat P för företaget och som indirekt skapar mervärde för kunden. Om företaget lyckas pränta in en positiv förväntan hos kunderna, kan de ta ut ett högre pris och följaktligen större marginal på sina produkter, som exempelvis Apple lyckats med.



**Figur 14.** De fyra P:na för produkter respektive de sju P:na för tjänster.

### 7.3. Regressionsmodellen – ur ett marknadsföringsperspektiv

Kanalpaketet är en benämning som författarna själva myntat och som ämnar fånga upp alla aspekter av det kanalen har att erbjuda kunden, eller tittaren. Kanalpaketet är inte en självklar produkt, eftersom det inte är en fysisk produkt med ett definierat pris. Det är å andra sidan inte heller en självklar tjänst, eftersom den inte lämnar några fysiska bevis, i enlighet med tidigare presenterad litteratur på ämnet. På grund av att kanalpaketet inte passar in i någon tidigare formulerad modell görs ett försök att formulera de sju P:na för kanalen med utgångspunkten att återge en fullskalig bild av vad kanalen kan erbjuda sina tittare. För att därigenom kunna komplettera den förenklade regressionsmodellen uppställd under den matematiska delen av det här kandidatsexamensarbetet.

De Fyra P:na i marknadsföringsmixen för kanalpaketet består av:

1. *Produkt* – utgörs av kanalen och dess programmix, där olika slags program lockar målgrupps uppmärksamhet i olika stor utsträckning och på så sätt tillgodoser ett tittarbehov hos kunden.
2. *Pris* – är mer svårdefinierat eftersom det i kanalernas fall inte handlar om en direkt transaktion för kunden, dels eftersom vissa kanaler är så kallade fri-TV-kanaler, dels för att flera

betalkanaler bakas in i en abonnemangsavgift och det inte går att utröna hur TV-tittarna skulle värdera respektive kanal i pengar. Hur tittarna värderar sin egna tid – det vill säga den mängd reklam som de är villiga att se innan de byter kanal – går det dock att sätta ett slags pris på. Detta kan ses som det pris tittarna är villiga att betala i termer av sin egen tid för att få se ett program. Det påstås här utgöra den indirekta kostnaden för kunden och den direkta intäkten för TV-bolaget. Reklamtidsuppoffring går också att likställa med ekonomisk uppoffring så till vida att efterfrågan närmar sig noll när reklam och pris går mot oändligheten.

3. *Plats* – sändningsplattformen utgörs av sändningstiden. Sändningstiden, utgörs i sin tur av tid och veckodag och är en tvådimensionell variabel i sammanhanget.
4. *Reklam (Promotion)* – marknadsföring av TV-kanalen kan göras för enstaka program, så kallad trailing av programmet, det vill säga att det pushas innan TV-sändningen.

De Fyra P:na för kanalpaketet kompletteras även med:

5. *Paket/Fysiskt bevis/Process* – TV-kanalen efterlämnar inte ett fysiskt bevis hos tittaren, eftersom inget egentligt utbyte av fysisk karaktär äger rum mellan kunden och producenten av tjänsten i det här fallet. Däremot omfattas TV-kanalen av ett abonnemangspaket (i betalkanalernas bemärkelse) eller ett marknadsnät (i frikanalernas bemärkelse). Givet att kunden redan har marknadsnätet, och eventuellt abonnemangspaketet, sker fri konkurrens inom paketet och mellan kanalerna, varför detta inte utgör ett paket i direkt anslutning till någon kanal, utan snarare en teknisk lösning som erbjuder kunderna flera kanaler. Utan något slags abonnemangspaket konkurrerar endast fri-TV-kanalerna SVT1, SVT2, TV4 och TV6. Med abonnemangspaket konkurrerar alla kanaler som berörs av det här kandidatexamensarbetet. Av den anledningen argumenteras det för att TV-kanalerna<sup>56</sup> endast omfattas av det fysiska beviset abonnemangspaket och att variabeln antar värdena noll eller ett. Dummyvariabeln representerar innehav eller avsaknad av abonnemangspaket.
6. *Personer* – människor involverade i försäljningsprocessen vad gäller kanalpaketet. TV-personligheterna och programledarna utgör främst den här delen av marknadsföringsmixen. Det är dessa personer som ökar programmets slagkraft bland tittarna genom sin närvaro i programmen och är ansvariga för hur kanalpaketet upplevs.

Som understöd till det argumentet vill författarna lägga fram helomvändningen efter programledarbytet av "På Spåret" i SVT, då TV-tittarna blev både yngre och fler i antal.<sup>57</sup> Enligt Tabell 30. nedan visas hur enstaka program av Sommarpratarna, som förvisso är väldigt fokuserade på personer, men som icke desto mindre har väldigt varierande tittarsiffror, varierar i sin tittarprofil. Avsnittet 2012-11-05 attraherade Kvinnor, 25-49 år, mest, medan män 15-34 tittade i omkring medelutsträckning. Avsnittet 2013-11-25 tittade däremot Män, 15-34 i nästan dubbelt så stor utsträckning på som på avsnittet 2012-11-05 och kvinnor, 25-49 år, tittade mindre på det än på något annat avsnitt av dem som undersöks.

---

<sup>56</sup> Av de kanalerna som är intressanta här, det vill säga, Svt1, Svt2, TV3, TV4, Kanal 5, TV6, TV8 och TV10.

<sup>57</sup> Jan-Olov Andersson, 'Årets säsong av "På Spåret" slår rekord', in *aftonbladet.se*

<<http://www.aftonbladet.se/nojesbladet/kronikorer/janolovandersson/article16250596.ab>> [accessed 15 February 2013]

Både programledare och TV-personligheter har således stor påverkan på såväl antalet tittare som vilken målgrupp som attraheras. Av den anledningen är även produktmixens Personer en tvådimensionell variabel för att fånga upp de här två sätten som TV-programmet kan göras mer populärt på.

| Datum    | Starttid | Sluttid | Alla,<br>3-99 år | Alla,<br>15-49 år | Män,<br>15-34 år | Män,<br>25-49 år | Kvinnor,<br>25-49 år | Kvinnor,<br>25-44 år |
|----------|----------|---------|------------------|-------------------|------------------|------------------|----------------------|----------------------|
| 20121105 | 20:01    | 20:58   | 10,68            | 4,62              | 1,36             | 4,10             | 7,11                 | 6,63                 |
| 20121112 | 20:01    | 20:58   | 7,32             | 2,76              | 1,28             | 2,55             | 4,49                 | 4,17                 |
| 20121119 | 20:01    | 20:58   | 9,86             | 4,54              | 3,12             | 5,37             | 5,94                 | 5,64                 |
| 20121126 | 20:01    | 20:58   | 10,54            | 4,48              | 2,04             | 5,33             | 6,65                 | 5,70                 |
| 20131111 | 20:01    | 20:58   | 8,95             | 3,17              | 0,91             | 2,07             | 5,87                 | 5,54                 |
| 20131118 | 20:01    | 20:58   | 11,33            | 4,64              | 2,75             | 4,88             | 6,64                 | 5,11                 |
| 20131125 | 20:01    | 20:58   | 9,49             | 3,36              | 2,26             | 3,86             | 3,89                 | 2,83                 |

**Tabell 30.** *Andelen tittare i procent i varje målgrupp som tittade på Sommarpratarna under olika avsnitt som alla sändes på måndagar i november på SVT1 och som tillhör programkategorin "Talkshow, diskussion".<sup>58</sup>*

7. *Positionering* – vad väljer TV-tittarna att titta på allra först, utan eftertanke eller insikt i vad för program som går. Med andra ord: vilken bild har kanalerna sedan tidigare förmedlat till kunden och hur pass väl attraherar den här bilden den målgrupp kanalen har? Vad är förväntningarna på respektive kanal, i enlighet med teorin punkt 7 i marknadsföringsmixen? Positionering av kanalen är starkt kopplat till hur kanalen marknadsför sig själv.

Genom att definiera attribut med vilka TV-kanalen kan göra sig mer konkurrenskraftig, det vill säga ta fram den individuella marknadsföringsmixen för kanalpaketet som har gjorts ovan, kan en kompletterande ekvation till den matematiska delen i det här kandidatexamensarbetet ställas upp. Generellt ser den ut som Formel (22) och specifikt som Formel (23).

$$Y = \beta_0 + \text{Produkt} * \beta_1 + \text{Pris} * \beta_2 + \text{Plats} * \beta_3 + \text{Reklam} * \beta_4 + \text{Paket} * \beta_5 + \text{Personer} * \beta_6 + \text{Positionering} * \beta_7 + e \quad (22)$$

$$Y = \beta_0 + \text{Programkategori} * \beta_1 + \text{Reklamtid} * \beta_2 + \text{Tid på dygnet} * \beta_{3.1} + \text{Dag i veckan} * \beta_{3.2} + \text{Trailing} * \beta_4 + \text{Abonnemangspaket} * \beta_5 + \text{Programledare} * \beta_{6.1} + \text{Mediaprofiler} * \beta_{6.2} + \text{Marknadsföring} * \beta_7 + e \quad (23)$$

Däremot erbjuds inte någon lösning kring hur den här modellen skulle implementeras i regressionsmodellen, eftersom exempelvis variabeln Position är svår att mäta. Kanalernas alla programledare skulle kunna införas som dummyvariabler i regressionsanalysen för att se vilken programledare som verkligen bidrar till tittarsiffrorna, samt vilken som inte gör det. Om underlag för hur mycket trailing TV-kanalerna gör för respektive program fanns tillgänglig skulle exempelvis den summa pengar det kostade, eller den reklamtid i TV som trailingen utgjorts av, kunna tas med i modellen för att även här se om det var särskilt givande. Observera att modellen och alla dummyvariabler kan bli väldigt lång.

Det förväntas att  $\beta$ -värdet som estimeras för Pris i Formel (22) och Reklamtid i Formel (23) har negativ inverkan. Ju mer reklamtid, desto oftare kommer troligtvis TV-tittaren att byta kanal, vilket

<sup>58</sup> MMS, 'Våra tjänster - Mediemätning' <<http://www.mms.se/start/vara-tjanster/mediematning/>> [accessed 06 April 2014]

följaktligen bidrar negativt till tittarsiffrorna. Programledare och Mediaprofiler skulle exempelvis kunna ha både negativa och positiva estimerade  $\beta$ -värden i jämförelse med benchmark. Som nämnts innan finns tydliga indikeringar på att det finns ett samband mellan vilken programledare och vilka TV-gäster, eller mediaprofiler, som står för programmet och dess slutgiltiga tittarsiffror.

Som även har gjorts i den matematiska delen så bör programföljd, -längd och alla dagar utom söndagar slås ihop med respektive intercept för de flesta målgrupper eftersom de inte har en avsevärd påverkan på tittarantalet under prime time. Det bör följaktligen undersökas i fortsatta studier om alla införda variabler verkligen är relevanta alla tider på dygnet, eftersom den presenterade regressionsanalysen visade att så inte var fallet på tidsintervallet 20:00-22:59.

Tack vare att författarna till detta kandidatexamensarbete, utöver fördjupande kurser i Matematik, även gått kurser som Marknadsföring, med kurskod ME1307, men också Verksamhetsledning, med kurskod ME1308, på Kungliga Tekniska högskolan, har den rent matematiskt framtagna och förenklade modellen kunnat kompletteras med ytterligare variabler för att återge en fullständig bild av vad som påverkar TV-tittarna, enligt Formel (23)

Den nya modellen enligt Formel (23) är mer till för att återge en fullständig bild av vad som påverkar TV-konsumenten, än praktiskt tillämpbar. Det finns begränsningar vad gäller inhämtningen av den typen av data, eftersom TV-kanalerna exempelvis inte gör trailing- och andra marknadsföringsrelaterade investeringar tillgängliga för allmänheten. Det här tjänar mer som ett komplement till den matematiska delen eftersom det inte går att förenkla modellen till Formel (21), utan den borde snarare se ut som Formel (23), med stöd ur marknadsföringsteori.

Författarna menar att den här modellen fullständigt beskriver kanalpaketets marknadsföringsmix och skulle eliminera de relativt stora feltermerna i den tidigare uppställda modellen, utan att på något sätt adressera svårigheten med att tillämpa modellen. Vidare studier inom regressionsanalys och TV-tittande rekommenderas att fortsätta undersökningen kring vilka variabler som påverkar tittarna. Nyckelpersoner inom verksamhetsledning på de olika TV-kanalerna rekommenderas även använda resultatet av regressionsanalysen för att på så sätt få högre tittarsiffror.

## Källhänvisning

- Andersson, Jan-Olov/*Aftonbladet*, ”Årets säsong av "På Spåret" slår rekord”, tillgänglig: <http://www.aftonbladet.se/nojesbladet/kronikor/janolovandersson/article16250596.ab>, 2013, (hämtad: 2014-05-01)
- Ender, Phil, ”Linear Statistical Models: Regression”, tillgänglig: <http://www.philender.com/courses/linearmodels/notes1/trans1.html>, 1999, (hämtad: 2014-05-24)
- Faraway, Julian J, ”Practical Regression and Anova using R”, tillgänglig: <http://cran.r-project.org/doc/contrib/Faraway-PRA.pdf>, 2002, (hämtad: 2014-04-27)
- Fox, John, *Applied Regression Analysis and Generalized Linear Models*. upplaga 2., Washington: SAGE Publications, Inc, 2008.
- Jamal, Åsa/*Spotlight TV4-gruppen*, ”Framtidens TV”, tillgänglig: <https://spotlight.tv4.se/Nyheter/Arkiv/2013/Maj/Framtidens-tv/>, 2013 (hämtad: 2014-04-06)
- Kaggle Inc., ”Root Mean Squared Error (RMSE)”, tillgänglig: <https://www.kaggle.com/wiki/RootMeanSquaredError>, 2014, (hämtad: 2014-05-19)
- Kennedy, Peter, *A Guide to Econometrics*. 6., United States of America: Blackwell Publishing Ltd, 2008.
- Konkurrensverket/Myndigheten för Radio och TV, *Rörlig Bild - En Analys av Marknaden för Rörlig Bild*, Stockholm: Davidssons Tyckeri AB, 2012.
- Kotler, Philip/Armstrong, Gary, *Principles of Marketing*. Europeisk upplaga 2., New Jersey: Pearson Education, Inc., 2012.
- Lang, Harald, *Topics on Applied Mathematical Statistics*. version 0.97., Stockholm: Kungliga Tekniska Högskolan, 2013.
- Milld, Jan/*BGF historia*, ”Television”, tillgänglig: <http://www.bgf.nu/historia/6/tv.html>, 2008, (hämtad: 2014-04-05)
- MMS, ”Glossary - Nielsen Audience Measurement”, tillgänglig: <http://mms.se/glossary/nielsen-audience-measurement/>, 2013, (hämtad: 2014-04-06)
- MMS, ”Historik”, tillgänglig: <http://www.mms.se/start/om-mms/historik/>, 2013, (hämtad: 2014-04-06)
- MMS, ”Om MMS”, tillgänglig: <http://www.mms.se/start/om-mms/>, 2013, (hämtad: 2014-04-06)
- MMS, ”Våra tjänster - Mediemätning”, tillgänglig: <http://www.mms.se/start/vara-tjanster/mediematning/>, 2013, (hämtad: 2014-05-01)
- MMS, ”Våra tjänster - Mediemätning - Så mäts TV-tittandet”, tillgänglig: <http://mms.se/start/vara-tjanster/mediematning/sa-mats-tv-tittandet/>, 2013, (hämtad: 2014-04-06)
- Nepa, ”Om oss”, tillgänglig: <http://www.nepa.se/sidor/om-oss/>, 2011, (hämtad: 2014-04-06)
- Nordfors, Andreas, intervju av Anna Leijon och Sara Holmsäter, 2014, Nepa AB.
- Nilsson, Björn, *Boken om TV*, Malmö: Bengt Forsbergs Förlag, 1961.

Osborne, Jason W, "Improving your data transformations: Applying the Box-Cox transformation", *Practical Assessment, Research & Evaluation*, Vol. 15, 2010: ss. 2-9.

Oxford Journals, "Regression Diagnostics", *Our Journals*, vol. 5, 2014, ss. 49-51.

*Stat Trek*, "Transformations to Achieve Linearity", tillgänglig: <http://stattrek.com/regression/linear-transformation.aspx>, 2014, (hämtad: 2014-05-19)

*The Analysis Factor*, "Assessing the Fit of Regression Models", tillgänglig: <http://www.theanalysisfactor.com/assessing-the-fit-of-regression-models>, 2014, (hämtad: 2014-05-20)

*The MathWorks, Inc*, "histfit", tillgänglig: <http://www.mathworks.se/help/stats/histfit.html>, 2014, (hämtad: 2014-05-22)

*The MathWorks, Inc*, "qqplot", tillgänglig: <http://www.mathworks.se/help/stats/qqplot.html>, 2014, (hämtad: 2014-05-22)

*The Pennsylvania State University*, "A residuals vs. fits plot. Online Courses", tillgänglig: <https://onlinecourses.science.psu.edu/stat501/node/36>, (hämtad: 2014-05-04)

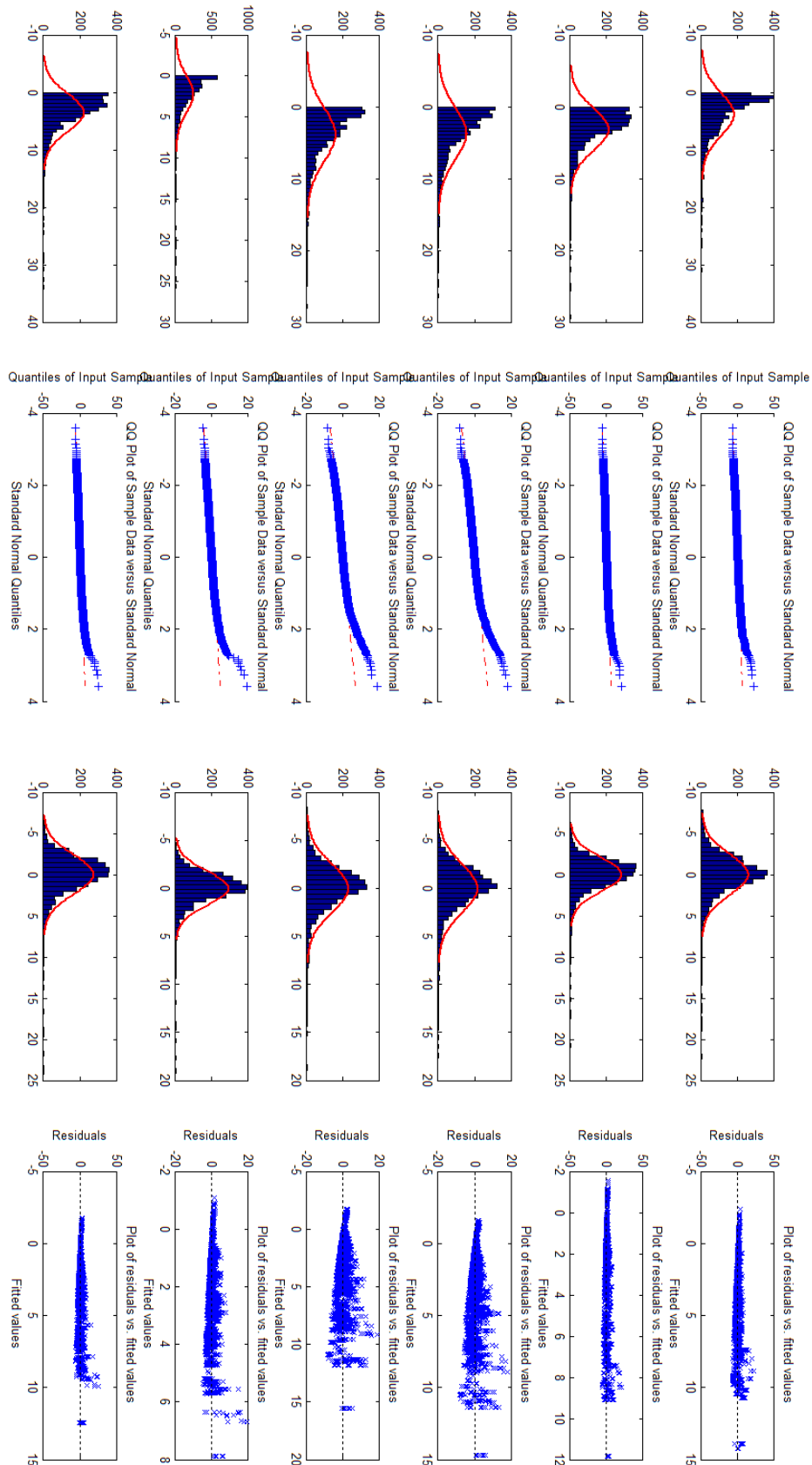
Thelander, Joakim, "En titt på TV", *Populär Historia*, nr. 8, 2002, ss. 10-14. Tillgänglig: <http://www.popularhistoria.se/artiklar/en-titt-pa-tv/> (hämtad 2014-03-04)

Tracy, Brian, *Million Dollar Habits*, United States: Entrepreneur Media Incorporate, 2004.

Östlund, Madeleine/*Dagens Media*, "Mediavision: Tv omsätter 21 miljarder", tillgänglig: <http://www.dagensmedia.se/nyheter/tv/article3626272.ece>, 2013 (hämtad: 2014-04-06)

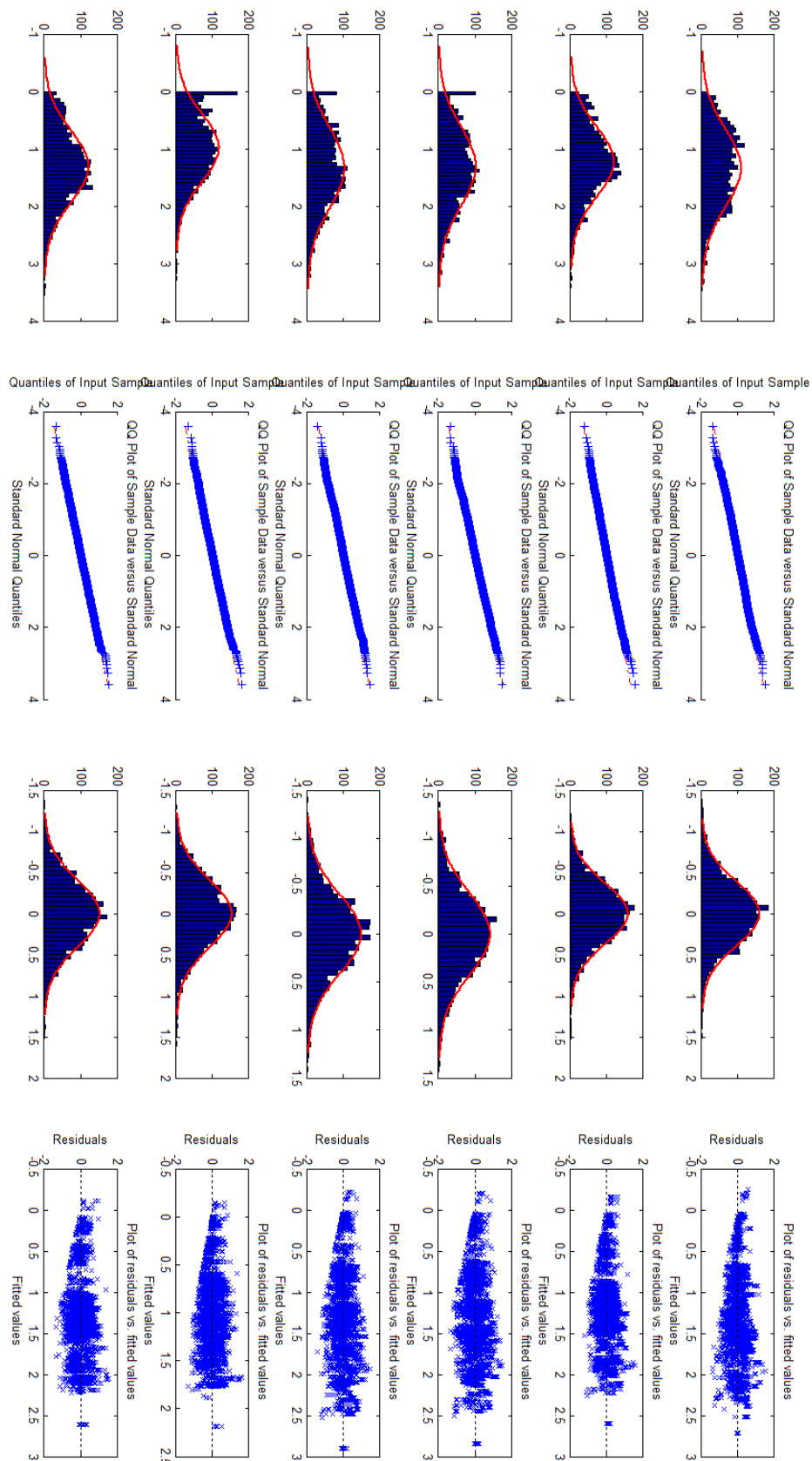
# Bilaga 1. Ursrungsplottar

Ursrungsplottar utan åtgärder vare sig det gäller datan eller modellen. Utifrån ett vykortsperspektiv beskriver raderna sett uppifrån plottar för Alla 3-99 år, Alla 15-49 år, Kvinnor 25-44 år, Kvinnor 25-49 år, Män 15-34 år och sist Män 25-49 år.



## Bilaga 2. Plottar med transformeringsmetod $\log(Y+1)$

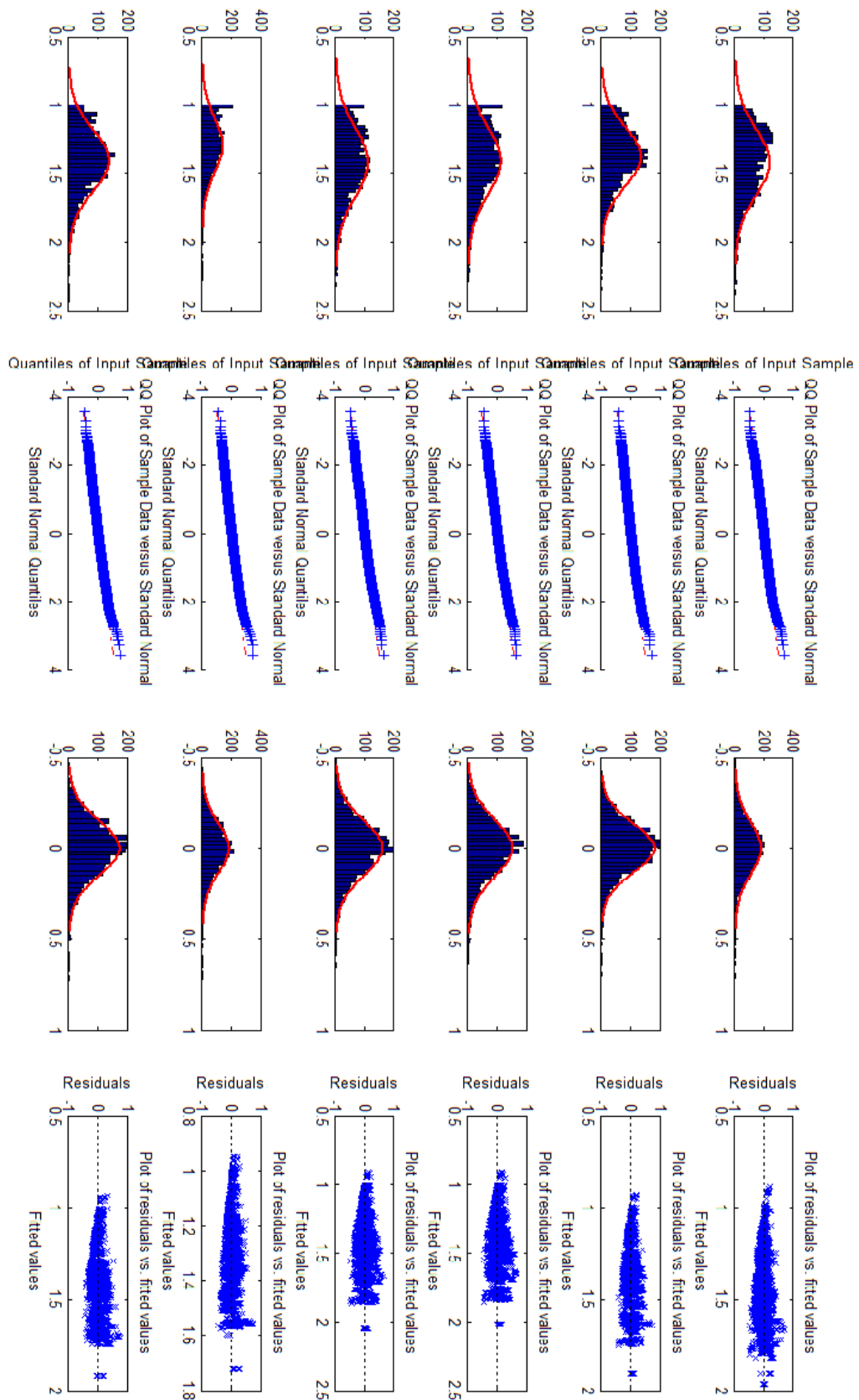
Plottar med åtgärden  $\log(Y + 1)$ . Utifrån ett vykortsperspektiv beskriver raderna sett uppifrån plottar för Alla 3-99 år, Alla 15-49 år, Kvinnor 25-44 år, Kvinnor 25-49 år, Män 15-34 år och sist Män 25-49 år.





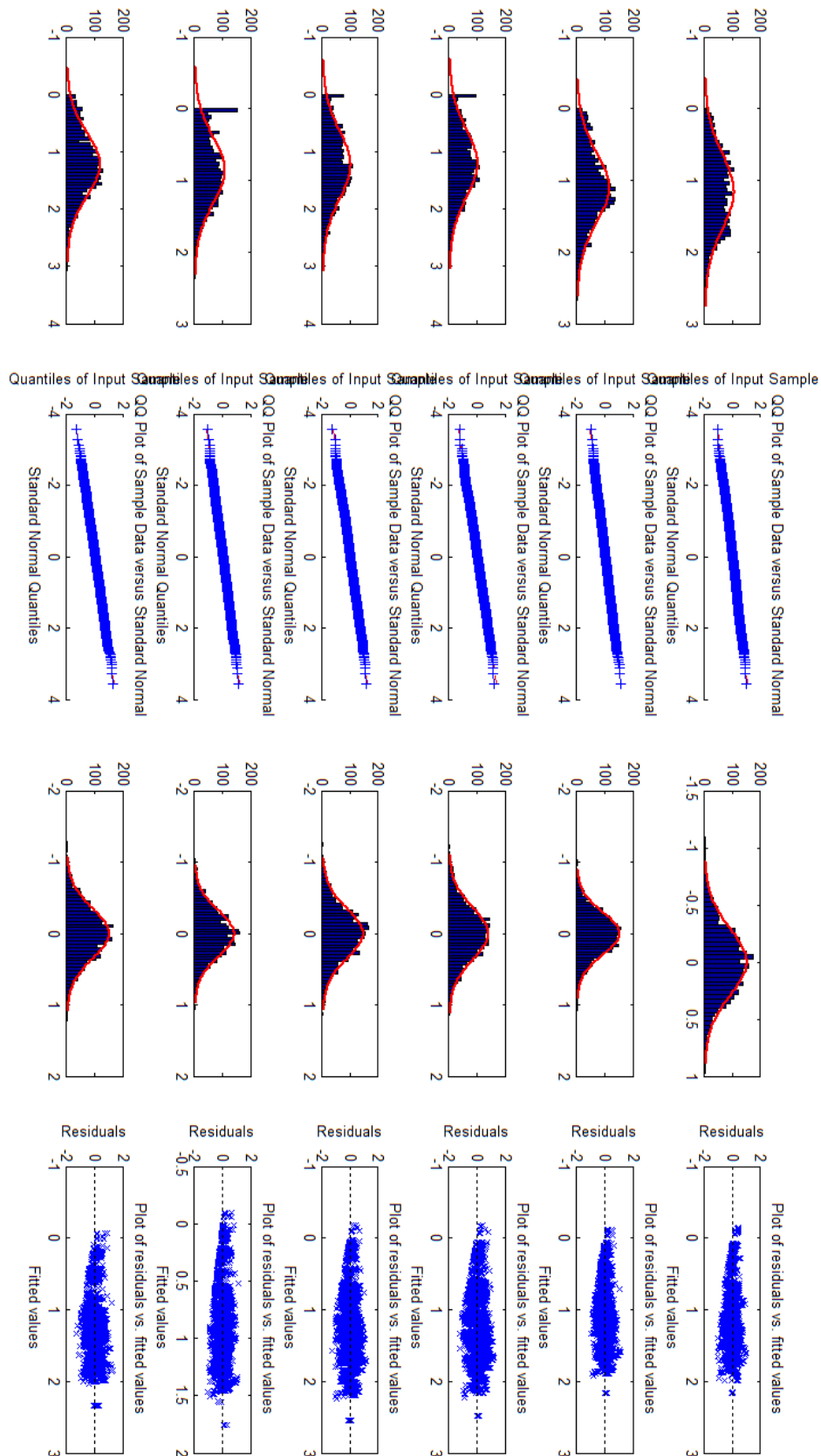
### Bilaga 3. Plottar med transformeringsmetod $(Y+1)^{0,25}$

Plottar med åtgärden  $(Y + 1)^{0,25}$ . Utifrån ett vykortsperspektiv beskriver raderna sett uppifrån plottar för Alla 3-99 år, Alla 15-49 år, Kvinnor 25-44 år, Kvinnor 25-49 år, Män 15-34 år och sist Män 25-49 år.



## Bilaga 4. Plottar med transformeringsmetod BoxCox(Y+1)

Plottar med åtgärden  $BoxCox(Y + 1)$ . Utifrån ett vykortsperspektiv beskriver raderna sett uppifrån plottar för Alla 3-99 år, Alla 15-49 år, Kvinnor 25-44 år, Kvinnor 25-49 år, Män 15-34 år och sist Män 25-49 år.



## Bilaga 5. Alla kategorier med tillhörande program kl. 20:00-22:59

Alla program under prime time i november för år 2009, 2012 och 2013.

|                      |   |
|----------------------|---|
| Debatt               | ADAKTUSSON, AGENDA, ANNAS EVIGA, ASCHBERG, DEBATT, INFÖR PRESIDENTVALET I US, KVÄLLSÖPPET MED EKDAL, .  |
| Dokumentär           | ADRESSLAPP OKÄND, AFR, ANTON CORBIJN INSIDE OUT, ATT BYGGA EN SUPERBIL, BARN UTAN PAPPER, BLACK NATION, BLYBARNEN, CLOSE QUARTER BATTLE, DEN STORA KÄRLEKEN, DOM ANDRA, DÖDSSYNDERNA, EFTER SPRUTAN, ENGLANDS FÖRTÄCKTA BORDE, ESCAPE FACTORY, ETT LIV I TVÅNGSTRÖJA, FALLET, FELIX BAUMGARTNER VÄRLDEN, GAMMAL OCH DÖRR, GREAT GRANDMOTHERS, INGEN SVENSSON LÄNGRE, KALLE MED K, KAPNING, KKK UNDER DEN VITA HUVAN, KOENIGSEGG, KOLPORTÖRER, KVINNORS SEKLUST, LIV MED AUTISM, LUNCH 250 METER OVANFÖR M, LÄRANDETS IDÉHISTORIA, MAMMA TILL EN MÖRDARE, MARK LOMBARDI KONSTEN SOM, MORMONERNAS PARNINGSLEK, MÄNNISKANS RESA, OFF ROAD BAJA 1000, OUTLAW BIKERS, PEDOFILERNAS NATT, SAS FRÅN SUCCÉ TILL KRI, SKINHEADS 25 ÅR SENARE, STRINDBERG 100 ÅR SPÖKSON, STRIPPAN, TANJA, TATUERINGAR JORDEN RUNT, TERROR PÅ FLYGPLATSEN, TESTAMENTET DOKUMENTÄR, THE ACT OF KILLING, UNDERDOG DOKUMENTÄR, UNGDOMSFÄNGELSET, UTÖYA FEL TID FEL PLATS, VI SOM ÄLSKADE, VÄRLDEN FRAMFÖR HENNE, ÅTER TILL PRAG, Grand Total   |
| Film                 | 13 SNART 30, 1408, 2 FAST 2 FURIOUS, 2012, 3000 MILES TO GRACELAND, 6:E DAGEN, ACCEPTED, AIR FORCE ONE, AMERICAN PIE BAND CAMP, AMERICAN PIE THE NAKED MI, ANALYSERA ÄNNU MERA, APORNAS PLANET, ARMORED, AVATAR, BATMAN BEGINS, BECK ADVOKATEN, BECK DEN JAPANSKA SHUNGAM, BECK DEN SVAGA LÄNKEN, BECK GAMEN, BENKNÄCKARGÄNGET KROSSA D, BLADES OF GLORY, BLAST, BLIND SIDE, BRAZIL, BREACH, BROARNA I MADISON COUNTY, BRÖLLOPSPROVET, BURLESQUE, CASSANDRAS DREAM, CECILIE, CHAOS, CIDERHUSREGLERNA, CLIFFHANGER, CLOCKWORK ORANGE, DAY AFTER TOMORROW, DE 12 APORNAS ARME, DE ANDRAS LIV, DEFINITELY MAYBE, DEN TUNNA RÖDA LINJEN, DESPERATE HOURS, DET RÖDA FÄLTET, DEUCE BIGALOW EUROPEAN GI, DIE HARD 4.0, DINNER FOR SCHMUCKS, DJÄVULEN BÄR PRADA, DOA DEAD OR ALIVE, DONNIE BRASCO, DOOM, DRAGONHEART, DRAKARNAS RIKE, DUPLICITY, EDEN A LOUEST, ELIZABETH, EN DAG, EN MAN KOMMER HEM, ERAGON, ETT PÄRON TILL FARSA I L, ETT RUM MED UTSIKT, EVAN DEN ALLSMÄKTIGE, FARSAN, FIRED UP FILM, FIREWALL, FOOLPROOF, FYRA BRÖLLOP OCH EN BEGR, FYRA ÅR TILL, FÖRSONING FILM, GENERALEN, GET HIM TO THE GREEK, GOOD BYE LENIN, GUARDIAN FILM, HAR DU HÖRT RYKTET OM MOR, HAROLD AND KUMAR GO TO W, HEAD OF STATE, HEARTBREAK HOTEL, HUVUDJÄGARNA, HÄRMED FÖRKLARAR JAG ER, I SPEL OCH KÄRLEK, INSIDE MAN, INTOLERABLE CRUELTY, JIDDRA INTE MED ZOHAN, JOHAN FALK ALLA RÅNS MODE, JOHAN FALK DE 107 PATRI, JOHAN FALK SPELETS REGL, JOHN TUCKER MUST DIE, KAIRO 678, KANIN A LA BERLIN, KARATE KID, KILLER ELITE, KNOWING, KULREGN ÖVER BROADWAY, KÄRLEK PÅ MENYN, LAND OF THE LOST, LITTLE MAN, LIZ & DICK, LÖFTET, MADE IN YUGOSLAVIA, MATRIX, MATRIX RELOADED, MEN IN BLACK, MEN IN BLACK 2, MIAMI VICE FILM, MISSION IMPOSSIBLE 4, MITT STORA FETA GREKISKA, MR BEANS SEMESTER, MR DEEDS, MURDER AT 1600, NEVER LET ME GO, NOT ANOTHER TEEN MOVIE, OBSESSED, OCEANS ELEVEN, OCH SÅ KOM POLLY, OM ÖDET FÅR BESTÄMMA, OUT OF SIGHT, PAPPA PÅ BURK, PARADISE NOW, PAYBACK, PITCH BLACK, POLISSKOLAN 3, POSEIDON FILM, PÅ SMÄLLEN, PÅ SPANING MED BRIDGET J, RED, RESIDENT EVIL AFTERLIFE, ROCKY 1, ROCKY 2, ROCKY 3, RYKTET GÅR, SCORPION KING, SHADOW MAN, SHANGHAI KNIGHTS, SHOOTER, SHUTTER ISLAND, SKY CAPTAIN AND THE WORL, SKYLINE, SMEKMÅNADEN, SMÅ OCH STORA BROTT, SNUTEN I VARUHUSET, SOCIAL NETWORK, SOMMAREN MED GÖRAN, SPACE COWBOYS, SPIDERMAN 3, STAR TREK, STAR WARS EPISOD 6 JEDINS, STARSKY & HUTCH, STATE OF PLAY FILM, SWEET HOME ALABAMA, SÖMNLÖS I SEATTLE, TERMINATOR 3, THE BATTLE OF RED CLIFF, THE BIG BOUNCE, THE BOAT THAT ROCKED, THE CELL, THE CLEARING, THE IMAGINARIUM OF DOCTOR, THE KIDS ARE ALL RIGHT, THE LEAGUE FILM, THE LIBRARIAN 3 THE CURS, THE LIBRARIAN RETURN TO K, THE REPLACEMENTS, THE TOWN, THE WEDDING SINGER, TRANSFORMERS, TROJA, TRUE GRIT, TRUE LIES, TWILIGHT, TWILIGHT SAGA ECLIPSE, TWILIGHT SAGA NEW MOON, ULTRAVIOLET, UNSTOPPABLE, V FOR VENDETTA, WALK THE LINE, WANTED FILM, WEDDING DAZE, WILD HOGS, WILL HUNTING, VÄLKOMMEN TILL FAMILJEN, YES MAN, YOU AGAIN, ZACK AND MIRI MAKE A PORN, ÄNGLARNAS STAD |
| Kriminalmagasin      | ARNE DAHL DE STÖRSTA VATT, ARNE DAHL UPP TILL TOPPEN, EFTERLYST, EFTERLYST SPECIAL, KALLA FAKTA, ONDSKA, ROSS KEMP VÄRLDENS FARLIG, ROSS KEMPS EXTREMA VÄRLD, UPPDRAG GRANSKNING, VECKANS BROTT   |
| Matlagningsprogram   | EDWARD BLOMS GÄSTABUD, HEMLAGAT MED GORDON RAMSA, JORDEN RUNT PÅ 80 RÄTTER, KINAS MAT, KINGS RANSOM KÖPET AV WA, NIKLAS MAT, TINAS COOKALONG  |
| Nyheter, kultur      | HYPE, KULTURNYHETERNA   |
| Nyheter, sport       | SPORTEN, SPORTNYTT, SPORTSPEGELN  |
| Nyheter, väder       | LOKALT VÄDER, VÄDRET  |
| Pausprogram          | ANSLAGSTAVLAN, CARPE DIEM KORTFILM, DANSKVÄLL, DJUR JAG DÖDADE FÖRRA SOM, ELVAKAFFE, EN EGEN MUR, FILMKLUBBEN INTRO, FRÖT, FÅGEL DÄRUPPE, HOEDOWN FROM RODEO, IN SWEDEN EVERYBODY CAN, JAKTENS LARM, K 141, KÖREN MED RÖSTEN SOM INS, OFFERROLLSRETORIK, ON SUFFOCATION, OVÄNTAT BESÖK, PIXELS, SIGGE MODIGHS SVANSJÖN  |
| Sitcom               | 2 1/2 MÅN, ALLA ÄLSKAR RAYMOND, ANGER MANAGEMENT SERIE, BIG BANG THEORY, CHRISTINE SERIE, CHUCK, COMMUNITY, COUGAR TOWN, ENTOURAGE, FAMILY GUY, FRIDAY NIGHT DINNER, GARY SERIE, GAVIN OCH STACEY, GLEE, GUSTAFSSON 3 TR, HALLÅ MUMBAI, HAPPY ENDINGS SERIE, HOW I MET YOUR MOTHER, MEN AT WORK, MIDDLE, MIKE AND MOLLY, MODERN FAMILY, MR BEAN SERIE, MY NAME IS EARL, NEW GIRL, PLAYA DEL SOL, SCRUBS, SEINFELD, SIMPSONS, SOLSIDAN, SUBURGATORY, THE BIG C, THE EXES, TRIST HERR MINISTER, TROPHY WIFE, TVÅ PANKA TJEJER, UGLY BETTY, WEEDS, VÄNNER  |
| Sport/betting        | ALPINT VÄRLDSCUPEN, BOXNING, DEN OLYMPISKA DROGSKANDAL, DREAMHACK OPEN, FORMEL1 GP 12 USA, FORMEL1 GP HIGH, FORMEL1 GP HIGH BRA, FORMEL12 GP 12 HIGH BRA, FOTBOLL CHA IPS-BAR, FOTBOLL CL 09 DEB-LIV 1, FOTBOLL CL 09 DEB-LIV 2, FOTBOLL CL 09 DEB-LIV E, FOTBOLL CL 09 DEB-LIV F, FOTBOLL CL 09 DEB-LIV M, FOTBOLL CL 12 PSG-DIZ, FOTBOLL CL ARS-MAR, FOTBOLL CL PSG- AND, FOTBOLL CL09 ATL-CHE 1, FOTBOLL CL09 ATL-CHE 2, FOTBOLL CL09 ATL-CHE E, FOTBOLL CL09 ATL-CHE F, FOTBOLL CL09 ATL-CHE M, FOTBOLL FA WIM-COV, FOTBOLL LANDSKAMP 12, FOTBOLL LANDSKAMP STU 12, FOTBOLL LK ENG-CHI, FOTBOLL LK TYS-ENG, FOTBOLL VM PLAYOFF, FOTBOLL VM PLAYOFF STUDIO, FOTBOLLSGALAN, HISTORIEN OM GARY SPEED, HOCKEYKVÄLL, I ZLATANS FOTSPÅR, ISHOCKEY ALL 12 LEK-MAL M, ISHOCKEY ALL 12 MAGASIN, ISHOCKEY ALL 12 MAL-SÖD, ISHOCKEY ALL 12 OSK-DJU 3, ISHOCKEY ALL 12 OSK-DJU M, ISHOCKEY ALL 12 TRO-TIN 3, ISHOCKEY ALL 12 TRO-TIN M, ISHOCKEY ALL KAR-DJU M, ISHOCKEY ALL TRO-MAL, ISHOCKEY KARJALA CUP 09, KAMPSPORT MMA, MILJONLOTTERIETS PENGARNA, NBA ACTION, NFL DAL-OAK, NFL GAME DAY, NFL NE-PIT, NFL NYG-DAL, NFL SF-CAR, PREMIER LEAGUE EN FOTBOLL, RIDSPORT STOCKHOLM HORSE, VM RALLY, ZACHRISSONS OSMINNE   |
| Talkshow             | BERGS BÄRBARA TALKSHOW, FRÅGA OLLE, HELLENIUS HÖRNA, SKAVLAN  |
| Talkshow, kultur     | BABEL, KRÖNIKA MELODIFESTIVALEN, LONDON LIVE, SVERIGEQUIZEN   |
| Talkshow, diskussion | ALLA ÅR FOTOGRAFER, ANSIKTE MOT ANSIKTE HENN, ATT VARA, BERG FLYTTAR IN, HÜBINETTE, I HUVUDET PÅ GUNDE SVAN, NUGAMMALT 10 SIDOR AV SVE, SETTMAN PÅ PLATS, SOMMARPRATARNA, SÅ MYCKET BÄTTRE  |

|                          |   |
|--------------------------|---|
| TV-serie, dokumentär     | 69 SAKER DU VILL VETA OM, BEFRIA SINNET, BEING LIVERPOOL, BILDEN AV VILDEN, DEN STORA UTPLÄNINGEN, DET SVARTA GULDET, DOKUMENT INIFRÅN, DOKUMENT UTIFRÅN, DÖDEN DÖDEN DÖDEN, FATTA KATASTROFEN, FEM ÅR PÅ MARS, FRÅGA OLLE DOKUMENTÄREN, GENERATION FETT, HEMLIGA SVENSKA RUM, HEMLÖS, INLÅST, JAKTEN PÅ DET PERFEKTA LI, JORDEN RUNT MED EWAN MCGR, K SPECIAL, KINA OM KINA, KORRESPONDENTERNA, KUNSKAPSDOKUMENTÄR DE KÄR, KUNSKAPSDOKUMENTÄR PRIVAT, KUNSKAPSDOKUMENTÄR ÖL ÄR, KYSST AV SPRITEN, MITT LIV SOM ROM, PLUS, SEMESTER I FAROZONEN LÅNG, SIR BOBBY MINNS KATASTROF, SOS UPPDRAGET, STAMLIV MED BRUCE PARRY, TEKNIKENS VÄRLD ALLT OM, TORTYR, TV3 DOKUMENTÄR, VETENSKAPENS VÄRLD, WHY POVERTY, WHY POVERTY HIT MED PENGA, WHY POVERTY KAMPEN OM JOR, WHY POVERTY KOPPARKUPPEN, WHY POVERTY LJUSETS MÖDRA, WHY POVERTY MYRORNAS KRIG, WHY POVERTY VÄLKOMMEN TIL, VÄRLDENS KONFLIKTER |
| TV-serie, drama          | ARROW, BETWEEN THE LINES SERIE, BOSS SERIE, BROTHERHOOD, BROTHERS & SISTERS, DAMAGES, DEFIANCE, DOWNTON ABBEY, ENLIGHTENED, GREYS ANATOMY, HART OF DIXIE, HELL ON WHEELS, HJEM, HOMELAND, HOTELLET SERIE, HOUSE, LIVET I FAGERVIK, LYKKE, MOLLY SERIE, NIP/TUCK, NURSE JACKIE, SCANDAL, TRE ÅR MED ELLE, VIKINGS  |
| TV-serie, historia       | BAND OF BROTHERS, HISTORIEÄTARNA, JOHN ADAMS, NAZISTERNAS VAPENBRÖDER, SVENSKA HÄNDELSER, THE KENNEDYS, THE PACIFIC, VÄRLDENS HÄNDELSER, VÄRLDENS MÄRKLIGASTE VAPE, ÖKENKRIGET  |
| TV-serie, hjälp          | ALLT FÖR SVERIGE, ARGA SNICKAREN, DINA FRÅGOR OM PENGAR, ETT FALL FÖR LOUISE, GORDONS HOTEL HELL, GRANNFEJDEN, HJÄLP VI SKA FÖDA, KAN DU SLÅ EN PENSIONÄR, KNIVEN MOT STRUPEN, KOMMANDO SNUSK, LYXFÄLLAN, LYXFÄLLAN HUR GICK DET SE, MARY QUEEN OF THE HIGH ST, MATAKUTEN, NYBYGGARNA, STALKERS, SVERIGES FULASTE HEM, THE FIXER REALITY, TONÅRSBOSSEN, VINNA ELLER FÖRSVINNA SV  |
| TV-serie, humor          | 99 SAKER MED ERIK OCH MAC, AMERICAS FUNNIEST HOME V, AMERICAS FUNNIEST HOME VI, CAFÉ BÄRS, CIRKUS MÖLLER, COMEDY FIGHT CLUB, MITT LIV ÄR ETT SKÄMT, PARLAMENTET, PARTAJ, PENSIONÄRSJÄVLAR AUSTRALI, PENSIONÄRSJÄVLAR USA MED, PULVER, RAW COMEDY CLUB, ROBINS, RUDE TUBE, SNN NEWS, SOMMARTIDER, SVENSK HUMOR, TACK FÖR MUSIKEN, TALA UT, TORSK PA TUBEN  |
| TV-serie, inredning      | DESIGN SIMON & TOMAS, DRÖMHEM VID HAVET, DRÖMKÅKEN, HUSTOPPEN, LYXHOTELLET'S HEMLIGHETER, PROPERTY LADDER, RESTORATION HOME, ÄNTLIGEN HEMMA   |
| TV-serie, kriminal       | BLACKLIST, BOARDWALK EMPIRE, BONES, BRON, BROTTET OCH STRAFFET, BROTTSKOD FÖRSVUNNEN, BURN NOTICE, CASTLE, CHICAGO FIRE, CRIMINAL MINDS, CSI, CSI MIAMI, CSI NY, ELEMENTARY, KOMMISSARIE LYNLEY, LIFE SERIE, MORDEN, MORDEN I MIDSOMER, MÖRDARE OKÄND, NAVY CIS, NCIS LOS ANGELES, OSKYLDIGT DÖMD SERIE, PERCEPTION, SOUTHLAND, THE AMERICANS, THE CLOSER, THE EVENT, THE MENTALIST, WHITECHAPEL, WITHOUT A TRACE   |
| TV-serie, kultur         | ANTIJKAKTEN, ARTIFICIAL BODY VOICES, DOM KALLAR OSS ARTISTER, EXISTENS, FASHION MODEVECKAN I LON, FASHION MODEVECKAN I MIL, FASHION MODEVECKAN I NEW, FLICKAN FRÅN VÄSTERN, K MÄRKTA ORD, KOBRA, RICKS ANTIKHANDEL, SKARPA STRECK SATIRTECKNI, VON SVENSSONS KLÄDER   |
| TV-serie, reality        | DATE & THE CITY, GÖTA KANAL SERIE, MONTAZAMIS MED VÄNNER, OSBOURNES RELOADED, SMÅSTADSDROTTNINGAR, VEM TROR DU ATT DU ÄR  |
| TV-serie, tävling        | DOOBIDOO, GÄSTER MED GESTER, HÅL I VÄGGEN, I SURVIVED A JAPANESE GA, NÖJESPOKALEN, PENGARNA PÅ BORDET, POLISSKOLAN SERIE, SHARK TANK, SILENT LIBRARY, SINGING BEE CELEB, VEM KAN SLÅ FILIP OCH FR, VEM KAN SLÅ FILIP OCH FRE  |
| TV-serie, utslagstävling | ALDRIG MERA FET, AMERICA'S GOT TALENT, BONDE SÖKER FRU, DANSBANDSKAMPEN, DESSERTMÄSTARNA, DET STORA MATSLAGET, ENSAM MAMMA SÖKER, HELLS KITCHEN, HELLS KITCHEN USA, IDOL 2009, IDOL 2013, IDOL EXTRA, INK MASTER SVERIGE, KÖRSLAGET, MAMMAS POJKAR REALITY, MASTERCHEF, PARADISE HOTEL SVENSKT, PARADISE HOTEL XL, PROJECT RUNWAY, PROJECT RUNWAY SVERIGE, PÅ SPÅRET, ROBINSON KARIBIEN, SO YOU THINK YOU CAN DAN, X FACTOR, X FACTOR RÖSTNINGSPROGRAM  |
| TV-serie, yrke           | 112 AINA, AKUTEN, BARNMÖRSKORNA NORGE, COPS, ICE ROAD TRUCKERS, JAMES MAY PÅ GRÄNSEN TILL, JÄMNA PLÅGOR, KIT BYGG DIN DRÖMBIL, MONSTER MOVES, PROJEKT ÖVERLEVA MED CHRI, SOS GUTE, SOS SOMMAR, STOCKHOLM ARLANDA, UNDERCOVER BOSS, UPPFINNARNA  |
| TV-serie, övernaturlig   | GRIMM, MARVELS AGENTS OF SHIELD, SUPERNATURAL SERIE, TRUE BLOOD   |

## Bilaga 6. Alla programkategorier med förklaring

Kompletterande kategoriförklaring till alla 43 kategorier (inte bara dem under prime time) kopplad till Bilaga 5.

|                          |   |
|--------------------------|---|
| Barnprogram              | Program som riktar sig till barn  |
| Debatt                   | Olika typer av debattprogram  |
| Dokumentär               | Dokumentärer i alla dess former exklusive historiedokumentärer (endast ett avsnitt)   |
| Dokumentär, historia     | Dokumentärer som berör historia (endast ett avsnitt)  |
| Film                     | Långfilmer, biograffilmer   |
| Forum                    | Program av olika slag som går mitt på dagen på SVT, se mer info under bilaga  |
| Gala                     | Exempelvis Fotbollsgalan och andra program som inte hör till normal-TV  |
| Handikapp                | Program riktade till hörselskadade  |
| Insamlingsprogram        | Program som är till för att samla in pengar till välgörande ändamål (exklusive galor)   |
| Kriminalmagasin          | Programledarledda i studio-setting som berör brott och liknande   |
| Magasin                  | Exempelvis morgon- och kvällsmagasin, ledda av programledare i en studio-setting  |
| Magasin, humor           | Programledarledda i studio-setting som avser humoristiska inslag  |
| Matlagningsprogram       | Programledarledda recept- och matlagningsstips  |
| Nyheter                  | Allmänna Nyheter, inklusive lokala nyheter  |
| Nyheter, ekonomi         | Nyheter som berör ekonomi   |
| Nyheter, kultur          | Nyheter som berör kultur  |
| Nyheter, sport           | Nyheter som berör sport   |
| Nyheter, vädret          | Nyheter som berör vädret  |
| Pausprogram              | Program som är kortare än 20 minuter (exklusive Nyheter)  |
| Sitcom                   | En humoristisk dramaserie   |
| Sport/betting            | Allt från V75 till hockeymatcher och Keno   |
| Talkshow                 | Renodlad Talkshow där programledaren bjuder in olika typer av gäster i en studio-setting  |
| Talkshow, diskussion     | Programledarledd med mer fokus på gästen/gästerna och deras livserfarenheter eller -situation   |
| Talkshow, hjälp          | Talkshow där programledarens funktion är att hjälpa besökarna i någon form, till exempel agera psykolog, dietist eller liknande   |
| Talkshow, kultur         | Programledarledd studio-setting som berör kulturämnen   |
| TV-serie, dokumentär     | Dokumentärer som är fler än ett program (exklusive historiedokumentärserier)  |
| TV-serie, dokusåpa       | TV-serier där fokus främst ligger på relationer och särskild drama, eventuellt humoristiska inslag  |
| TV-serie, drama          | Drama-TV-serier av det seriösare slaget   |
| TV-serie, fritid         | TV-serier som berör djur och natur, dels renodlade naturprogramserier, men också fritidsrelaterat som fiske och liknande  |
| TV-serie, futuristisk    | TV-serier med särskild rymdsetting, inte bara övernaturlig, utan överjordisk  |
| TV-serie, historia       | TV-serier som handlar om historiska händelser   |
| TV-serie, hjälp          | TV-serier där programledarna hjälper de personer som är med, exempelvis make overs, men också inredning om särskild betoning ligger i att hjälpa individen  |
| TV-serie, inredning      | TV-serier om inredning, där hus och hem ligger i fokus, och inte individerna  |
| TV-serie, kriminal       | TV-serier där fokus ligger på brott eller att lösa brott  |
| TV-serie, kultur         | TV-serier som har med kulturämnen att göra  |
| TV-serie, reality        | TV-serier som exempelvis följer människor i deras vardag, av mer dokumentärslag   |
| TV-serie, resor          | TV-serier som berör resmål och själva resandet  |
| TV-serie, tävling        | TV-serier där tävling är i fokus, men där det inte är nödvändigt att följa varje program, utan där vart och ett är en tävling i sig och det inte är återkommande tävlande i efterföljande avsnitt   |
| TV-serie, utslagstävling | TV-serier med en slutlig segrare där tävlingen löper under en längre period och över flera avsnitt samt är av mer långdragen karaktär där man som tittare är mer benägen att titta flera avsnitt i följd för att följa de tävlande i deras utveckling |
| TV-serie, yrke           | TV-serier där man får följa en eller flera personer i deras dagliga arbete, som exempelvis på sjukhus eller en brandstation   |
| TV-serie, övernaturlig   | TV-serier inriktade på science fiction och övernaturliga karaktärer och händelser   |
| TV-serier, udda          | TV-serier med särskilt reality- och vardagsinslag i en humoristisk miljö som exempelvis Ullared och campingprogram  |
| Undervisningsprogram     | Program som är till för att undervisa tittarna inom klassiska läroämnen som exempelvis språk  |

## Bilaga 7. Matlab-kod

```
%Test för multicollinearitet
[m,n] = size(X)
X1 = [ones(m, 1) X]
collintest(X1)

%Linjär regression med Whites Consistent Variance Estimator
X = "namn"; %Istället för namn står namnet på förklaringsmatrisen
A = "namn"; %Istället för namn står namnet på Y-variabeln
for k=1:length(A)
    A(k)=A(k)+1;
end

A = A';
Y = log(A);
mdl = LinearModel.fit(X,Y);
%Plottar residualerna nedan
subplot(3,1,1), plotResiduals(mdl,'fitted')
residual = mdl.Residuals.raw;
subplot(3,1,2), qqplot(residual)
subplot(3,1,3), histfit(residual)
%Testade att hac räknar Whites Consistent Variance Estimator på samma sätt
som i "Topics on Applied Mathematical Statistics" av Harald Lang, vilket
den gjorde.
%residual = mdl.Residuals.raw;
%Res2 = residual.*residual;
%D = diag(Res2);
%v = ones([2857,1])
%X1 = [v X]
%CovW = ((X1'*X1)^(-1))*(X1'*D*X1)*((X1'*X1)^(-1))
[WhiteCov,WhiteSe,coeff] = hac(mdl,'type','HC','weights',...
    'HC0','display','off'); % White's Estimates
WhiteSe % Skriver ut de korrekta standardavvikelserna för Beta:na
Här följer test för linear restrictions med heteroskedacitet
%De sista r-raderna och r-columnerna i X motsvarar de betan man tillsammans
%vill testa för nollhypotesen
r = "nummer"; %Istället för nummer står numret för antalet B:an som ingår i den linjära restriktionen
tot = "nummer"; %Istället för nummer står numret för antalet B:an som ingår i den linjära modellen
sista = tot-r+1;

v = WhiteCov(sista:tot,sista:tot);
B2 = coeff(sista:tot);
% W = B2*(V^(-1))*B2 har en asymptitisk fördelning som chi2(r) under nollhypotesen
W = B2*(v^(-1))*B2
pchi = chi2pdf(W,r)

%RMSE
Ytest1 = "namn"; %Där "namn" är namnet på Y-vektorn från outhuset, d.v.s. observationerna från 2011
Ytest2 = "namn"; %Där "namn" är namnet på Y-vektorn från insamlet, d.v.s. observationsmängden som
regressionen gjorts på
matris1 = [ones(length(X2011),1) X2011];
y1estimerat = matris1*beta; %Beta är beta-värdena för regressionen
y1 = exp(y1estimerat)-1;
r1 = sqrt( sum( (y1(:)-Ytest1(:)).^2) / numel(y1) ) %RMSE outhuset
X1 = [Ytest2 X]; %Där "X" är namnet på X-matrisen från insamlet, d.v.s. observationsmängden som
regressionen gjorts på
mnew = X1(randperm(size(X1, 1), round(0.05*length(X1))), :); %slumpar fram 5% av observationerna från
insamlet
```

```
RatLrandom = mnew(:,1);  
mat = mnew(:,2:end);  
matrix2 = [ones(length(mat),1) mat];  
y2estimerat = matrix2*beta;  
y2 = exp(y2estimerat)-1;  
r2 = sqrt( sum( (y2(:)-RatLrandom(:)).^2) / numel(y2) ) %RMSE insample
```







TRITA -MAT-K 2014:07  
ISRN -KTH/MAT/K--14/07-SE