

Multinomiaalisen logistisen regression sovellus vuoden 2014  
EVAn kansalliseen asennetutkimukseen

Teemu Paajanen

Tilastotieteen pro gradu

Jyväskylän yliopisto  
Matematiikan ja tilastotieteen laitos  
30. toukokuuta 2017

JYVÄSKYLÄN YLIOPISTO

Matematiikan ja tilastotieteen laitos

**Teemu Paajanen:** Multinomiaalisen logistisen regression sovellus vuoden 2014 EVAn kansalliseen asennetutkimukseen

Tilastotieteen pro gradu -tutkielma, 48 sivua + liitteitä 4 sivua, 30. toukokuuta 2017

### **Tiivistelmä**

Kun vastemuuttuja on luokiteltu ja siinä on enemmän kuin kaksi luokkaa, niin mallinusta voidaan tehdä multinomiaalisen logistisen regression avulla. Tämän alaluokka on nominaalinen logistinen regressio ja sitä käytetään, kun vasteen luokkia ei saada perustellusti yhteen tiettyyn järjestykseen. Siinä yksi vasteen luokka valitaan referenssiluokaksi, johon muita vasteen luokkia verrataan. Parametrit estimoidaan suurimman uskottavuuden menetelmällä. Malleja vertaillaan uskottavuusosamäärätestin ja AIC:n avulla. Mallin sopivuutta aineistoon tarkastellaan Pearsonin jäännöksillä.

Nominaalista logistista regressiota sovelletaan Elinkeinoelämän valtuuskunnan (EVA) vuoden 2014 kansalliseen asennetutkimukseen. Tarkasteltavat väittämät koskevat Venäjää, NATOa, USA:ta ja EU:ta. Jokaisen väittämän kohdalla malliin valikoituu prediktoreiden sukupuoli, ikäluokka ja puoluekanta päävaikutukset. Päävaikutusten kertoimien tulkinta tapahtuu niin, että vertaillaan tarkasteltavan prediktorin suhteen erilaisia, mutta muiden prediktoreiden suhteen samanlaisia henkilöitä.

Naiset ovat suhtautuneet kaikkiin neljään väittämään kielteisemmin kuin myönteisemmin verrattuna miehiin. Tosin EU-väittämän kohdalla ero ei ole merkitsevää. Naiset ovat myös vastanneet kaikkiin neljään väittämään todennäköisemmin ”vaikea sanoa” kuin että olisivat olleet jotain mieltä verrattuna miehiin. Ikäluokista nuorin ikäluokka on ollut keskimäärin negatiivisempi kuin positiivisempi verrattuna vanhimpaan ikäluokkaan kaikkissa neljässä väittämässä (USA-väittämässä ero ei ole merkitsevää).

Venäjä-väittämässä perussuomalaisten kannattajat ovat olleet keskimäärin negatiivisia kuin positiivisia verrattuna keskustan, kokoomuksen ja SDP:n kannattajiin. SDP:n äänestäjät ovat olleet todennäköisemmin myönteisempiä kuin kielteisiä verrattuna kristillisdemokraattien, perussuomalaisten, RKP:n, vasemmistoliiton ja vihreiden äänestäjiin. NATO- ja USA-väittämissä kokoomuksen kannattajat ovat olleet todennäköisimmin myönteisin ryhmä ja vasemmistoliiton kannattajat kielteisin ryhmä. Kokoomuksen kannattajat ovat olleet todennäköisimmin myönteisiä myös EU-väittämässä, mutta todennäköisimmin kielteisimmät kannattajat ovat olleet perussuomalaisilla. Ei-äänestäneet ovat vastanneet kaikkiin neljään väittämään todennäköisemmin ”vaikea sanoa” kuin monien puolueiden kannattajat.

### **Avainsanoja**

Multinomiaalinen logistinen regressio, nominaalinen logistinen regressio, multinomijakauma, suurimman uskottavuuden estimointi, EVA, kansallinen asennetutkimus.

# Sisältö

<b>1 Johdanto</b>	<b>1</b>
<b>2 Nominaalivasteinen regressio</b>	<b>3</b>
2.1 Multinomivaste yleistetyissä lineaarisissa malleissa . . . . .	3
2.2 Nominaalinen logistinen regressiomalli ja uskottavuusfunktio . . . . .	3
2.3 Suurimman uskottavuuden estimointi . . . . .	5
2.4 Kertoimien tulkinta ja Pearsonin jäännökset . . . . .	6
2.5 Uskottavuusosamäärätesti ja AIC . . . . .	6
<b>3 Aineisto</b>	<b>8</b>
<b>4 Tuloksia Venäjä-väittämään</b>	<b>16</b>
4.1 Mallinvalinta . . . . .	16
4.2 Sukupuoli ja ikäluokka . . . . .	17
4.3 Puoluekanta . . . . .	18
4.4 Puoluekannan tulosten vertailua muihin tuloksiin ja näkemyksiin . . . . .	23
<b>5 Tuloksia NATO-väittämään</b>	<b>25</b>
5.1 Mallinvalinta . . . . .	25
5.2 Sukupuoli ja ikäluokka . . . . .	26
5.3 Puoluekanta . . . . .	27
5.4 Puoluekannan tulosten vertailua muihin tuloksiin ja näkemyksiin . . . . .	31
<b>6 Tuloksia USA-väittämään</b>	<b>31</b>
6.1 Mallinvalinta . . . . .	32
6.2 Sukupuoli ja ikäluokka . . . . .	33
6.3 Puoluekanta . . . . .	34
<b>7 Tuloksia EU-väittämään</b>	<b>38</b>
7.1 Mallinvalinta . . . . .	39
7.2 Sukupuoli ja ikäluokka . . . . .	39
7.3 Puoluekanta . . . . .	40
7.4 Puoluekannan tulosten vertailua muihin tuloksiin ja näkemyksiin . . . . .	44
<b>8 Yhteenveto</b>	<b>45</b>
<b>Lähteet</b>	<b>47</b>
<b>Liitteet</b>	<b>49</b>

# 1 Johdanto

Elinkeinoelämän valtuuskunta (EVA) teettää Suomessa kansallisia asennetutkimuksia. Tässä tutkielmassa tarkastellaan vuoden 2014 tutkimusta. Tavoitteena on vertailla, millaisia ovat olleet suomalaisten asenteet neljää ulkopoliittista väittämää kohtaan tuona ajankohtana. Väittämät koskevat Venäjää, NATOa, USA:ta ja Euroopan unionia. Aineisto on saatu Tampereen yliopiston Yhteiskuntatieteellisestä tietoarkistosta.

Tutkimusongelmia lähestytään multinomiaalisen logistisen regression avulla. Logistista regressiota käytetään tapauksissa, joissa vasteena on muuttuja, joka voi saada vain kahta eri arvoa. Jos vaste on luokiteltu ja siinä on enemmän kuin kaksi luokkaa, niin yleistetyille lineaarisille malleille on kaksi vaihtoehtoa. Toinen lähestymistapa on multinomiaalinen logistinen regressio ja toinen on log-lineaarinen mallinnus. Log-lineaarisia malleja käytetään, kun mikään muuttujista ei ole vasteen asemassa. Jos yksi muuttuja on selkeästi vasteen asemassa ja muut ovat prediktoreita, niin tällöin voi kokeilla mallinnusta multinomiaalisen logistisen regression avulla. (Dobson, 2002, s. 135.)

Kun muuttuja on luokiteltu ja siinä on enemmän kuin kaksi luokkaa, niin muuttuja voi olla nominaali- tai järjestysasteikollinen. Jos vaste on regressiomallinnuksessa useampi kuin kaksiluokkainen, niin tällöin mallinnuksesta käytetään yleisesti nimitystä multinomiaalinen logistinen regressio. Multinomiaalinen logistinen regressio hajaantuu erikoistapauksiin vasteen mukaan. Tunnetuimmat erikoistapaukset ovat nominaalinen logistinen regressio sekä järjestysasteikollinen logistinen regressio. Nimitykset tulevat yksinkertaisesti siitä, että jos vaste on nominaalinen (vasteen luokkia ei saa yhteen tiettyyn järkevään järjestykseen), niin silloin käytetään nominaalista logistista regressiota, ja jos vasteen luokat voidaan laittaa vain yhteen järkevään järjestykseen, niin silloin on tavallista käyttää järjestysasteikollista logistista regressiota. (Dobson, 2002, s. 135.) Tutkittavien väittämien vastausvaihtoehtoja ei saada yhteen tiettyyn järjestykseen vaikea sanoa -luokan takia, joten tästä syystä syvennytään nominaaliseen logistiseen regressioon.

Tutkielman ensimmäiseksi tutkimusongelmaksi on valikoitunut suomalaisten suhtautuminen Venäjään vuoden 2014 alussa. Elinkeinoelämän valtuuskunnan teettämässä kansallisessa asennetutkimuksessa yhtenä kysymyksenä (tai väittämänä) on ollut tuolloin: "Vaikka Venäjällä on omat ongelmansa, suomalaisilla ei ole nykyisin mitään syytä suhtautua suureen naapuriinsa kielteisesti.". Väittämä on kiinnostava, koska on mielenkiintoista tutkia, millaisia ovat olleet suomalaisten asenteet Venäjää kohtaan juuri ennen Ukrainan kriisiä, joka alkoi kyseisen vuoden keväänä.

Taustamuuttujina EVAn tutkimuksessa ovat olleet muun muassa sukupuoli, syntymävuosi, asuinmaakunta, asuinkunnan koko, peruskoulutus, ammatillinen koulutus, ammattiryhmä, toimiala, puoluekanta, ammatillinen keskusjärjestö ja yhteiskuntaluokka. Prediktoreiksi Venäjä-väittämä -vasteelle valitaan aluksi sukupuoli, ikä jaettuna neljään luokkaan, puoluekanta, asuinalue, ammatillinen koulutus ja ammattiryhmä. Riittävin selit-

tävä malli valitaan lopullisen tarkastelun kohteeksi. Osoittautuu, että riittävin malli on sellainen, jossa prediktoreina ovat sukupuoli, ikäluokka ja puoluekanta.

Mallin mukaan naiset ovat suhtautuneet Venäjään keskimäärin kielteisemmin kuin miehet ja nuoremmat ikäluokat ovat suhtautuneet keskimäärin kielteisemmin kuin vanhemmat ikäluokat. Puoluekantoja vertailtaessa RKP:n kannattajat sekä perussuomalaiset ovat suhtautuneet keskimäärin kielteisemmin kuin varsinkin SDP:n kannattajat. Lisäksi ne, jotka eivät osanneet sanoa, miten olisivat äänestäneet, ovat vastanneet varsinkin joidenkin puolueiden (esimerkiksi kokoomuksen ja perussuomalaisten) äänestäjiä todennäköisemmin vaikea sanoa kuin että olisivat vastanneet “eri mieltä”. Venäjä-kysymykseen vaikea sanoa -vastanneita verrataan lopuksi logistisella regressiolla niihin, jotka ovat olleet jotain mieltä, jotta huomataan ovatko ei-äänestäneet verrattuna äänestäneisiin vastanneet Venäjä-väittämään todennäköisemmin vaikea sanoa kuin että olisivat olleet jotain mieltä.

Venäjä-väittämän innoittamana on tutkittu myös kolmea muuta väittämää. Väittämät koskevat NATOa, USA:ta ja EU:ta. Näiden muiden väittämien kohdalla malli valitaan sukupuolen, ikäluokan ja puoluekannan joukosta. Näin toimitaan siksi, koska halutaan erityisesti tutkia, miten puoluekannan kertoimet vaihtelevat kysymyksittäin ja vastaavatko puoluekantojen näkemuserot suurinpiirtein sitä, mihin on totuttu näiden aiheiden kohdalla.

Tämän kyselyn perusteella osoittautuu, että Venäjä-väittämän lisäksi naiset ovat olleet myös näissä kolmessa muussa väittämässä keskimäärin kielteisiä kuin myönteisiä verrattuna miehiin (tosin tulos ei ole merkitsevä EU-väittämässä). Lisäksi naiset ovat vastanneet kaikkiin neljään väittämään todennäköisemmin “vaikea sanoa” kuin että olisivat olleet jotain mieltä. Kun ikäluokkia vertaillaan, niin nuorin ikäluokka on suhtautunut kaikkiin väittämiin keskimäärin negatiivisemmin kuin positiivisemmin suhteessa vanhimpaan ikäluokkaan (tosin tulos ei ole merkitsevä USA-väittämässä). Puolueiden välillä on sen sijaan enemmän eroja, kun vertaillaan eri kysymyksiä.

NATO- ja USA-väittämissä kokoomusta äänestäneet ovat olleet keskimäärin myönteisiä kuin kielteisiä suhteessa vasemmistoliittoa äänestäneisiin. EU-väittämän kohdalla kokoomuksen kannattajat ovat suhtautuneet keskimäärin positiivisemmin kuin negatiivisemmin verrattuna perussuomalaisten kannattajiin. Näitä tuloksia ja eroja käydään läpi vielä enemmän tuloksissa ja yhteenvedossa. Yksi tulos kuitenkin pysyy puolueiden kohdalla samana. Kun kyseessä on ei-äänestävä, niin hän on vastannut kaikkiin neljään väittämään todennäköisemmin “en osaa sanoa” kuin että olisi ollut jotain mieltä verrattuna monen puolueen äänestäjiin.

## 2 Nominaalivasteinen regressio

Aluksi todetaan, että multinomivasteita voidaan mallintaa yleistettyjen lineaaristen mallien avulla. Sen jälkeen keskitytään nominaalivasteisen regression kertoimien estimointiin sekä tulkintaan ja sopivan mallin toteamiseen. Noudatetaan Dobsonin (2002, s. 135-143) tapaa käydä asiaa läpi.

### 2.1 Multinomivaste yleistetyissä lineaarisissa malleissa

Oletetaan, että satunnaisvektori  $\mathbf{Y}$  noudattaa multinomijakaumaa, jossa on  $J$  luokkaa. Luokkaan  $j$  kuulumisen todennäköisyys on  $\pi_j, j = 1, \dots, J$ , missä  $\pi_1 + \dots + \pi_J = 1$ .

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_J \end{pmatrix}, \quad \text{ja} \quad \mathbf{Y}_j \sim \text{Multinom}(N, \pi_1, \dots, \pi_J).$$

Havaintoa  $\mathbf{Y}$ :stä merkitään

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_J \end{pmatrix}, \quad \text{ja} \quad \sum_{j=1}^J y_j = N, \quad \text{missä } y_j\text{:t ovat kokonaislukuja ja ei-negatiivisia.}$$

Silloin

$$P(\mathbf{Y} = \mathbf{y}) = f(\mathbf{y}|n) = \frac{n!}{y_1! y_2! \dots y_J!} \pi_1^{y_1} \pi_2^{y_2} \dots \pi_J^{y_J}.$$

Jos vasteen jakauma kuuluu eksponentiaaliseen jakaumaperheeseen, niin voidaan tehdä mallinnusta yleistettyjen lineaaristen mallien avulla. Multinomijakauman voi todistaa kuuluvan eksponentiaaliseen jakaumaperheeseen, mutta kuulumisen voidaan esittää myös multinomijakauman ja Poisson-jakauman yhteydellä, kuten Dobson on tehnyt.

### 2.2 Nominaalinen logistinen regressiomalli ja uskottavuusfunktio

Nominaalista logistista regressiota käytetään, kun vaste on luokiteltu, siinä on enemmän kuin kaksi luokkaa ja luokkia ei saada perustellusti vain yhteen tiettyyn luonnolliseen järjestykseen. Lähestymistapa on samankaltainen kuin tavallisessa kaksi luokkaisen vasteen logistisessa regressiossa. Nominaalisessa logistisessa regressiossa yksi luokka valitaan referenssiluokaksi, johon muita luokkia verrataan.

Oletetaan seuraavaksi, että on  $n$  kappaletta riippumattomia multinomiaalisia satunnaisvektoreita  $\mathbf{Y}_i$ , jokainen on  $J$ -luokkainen ja  $\mathbf{Y}_i \sim \text{Multinom}(N_i, \pi_{i1}, \dots, \pi_{iJ})$ . Oletetaan, että vektoreiden  $\mathbf{Y}_i$  havaitut arvot ovat  $\mathbf{y}_i, i = 1, \dots, n$ . Lisäksi  $\sum_j y_{ij} = N_i$ , ja jokaisella havainnolla on oma multinomiaalinen todennäköisyysjakaumansa  $\boldsymbol{\pi}_i = (\pi_{i1}, \dots, \pi_{iJ})$  sekä  $\sum_j \pi_{ij} = 1$ .

Logistisessa regressiossa, kuten myös nominaalisessa logistisessa regressiossa, linkki-funktio toimii logit-funktio. Merkitään kovariaattivektoria  $\mathbf{x}_i = (1, x_{i1}, \dots, x_{ip})$  ja kerroinvektoria  $\boldsymbol{\beta}_j = (\beta_{j0}, \beta_{j1}, \dots, \beta_{jp})$ , missä  $p$  on estimoitavien parametrien lukumäärä. Jos oletetaan, että nominaalisessa logistisessa regressiossa vasteen luokka 1 on valittu referenssiluokaksi, niin silloin regressioyhtälö saa muodon

$$\text{logit}(\pi_{ij}) = \log\left(\frac{\pi_{ij}}{\pi_{i1}}\right) = \mathbf{x}_i^T \boldsymbol{\beta}_j, \quad \text{kaikilla } j = 2, \dots, J \text{ ja } i = 1, \dots, n.$$

Alaindeksi  $j$  viittaa vasteen luokkaan. Alaindeksi  $i$  voidaan tulkita siten, että se viittaa prediktorien kombinaation tiettyyn ryhmään, esimerkiksi tässä tutkielmassa yksi ryhmä on keskustaa äänestäneet vanhimman ikäluokan naiset. Yllä olevasta kaavasta saadaan

$$\pi_{ij} = \pi_{i1} \exp(\mathbf{x}_i^T \boldsymbol{\beta}_j), \quad j = 2, \dots, J \text{ ja } i = 1, \dots, n.$$

Koska  $\pi_{i1} + \pi_{i2} + \dots + \pi_{iJ} = 1$ , niin

$$(1) \quad \pi_{i1} = \frac{1}{1 + \sum_{j=2}^J \exp(\mathbf{x}_i^T \boldsymbol{\beta}_j)},$$

ja

$$(2) \quad \pi_{ij} = \frac{\exp(\mathbf{x}_i^T \boldsymbol{\beta}_j)}{1 + \sum_{j=2}^J \exp(\mathbf{x}_i^T \boldsymbol{\beta}_j)}, \quad j = 2, \dots, J.$$

Uskottavuusfunktio voidaan kirjoittaa

$$L(\boldsymbol{\pi}_1, \dots, \boldsymbol{\pi}_J) = \prod_{i=1}^n \binom{N_i}{y_{i1}, \dots, y_{iJ}} \times \pi_{i1}^{y_{i1}} \times \dots \times \pi_{iJ}^{y_{iJ}}.$$

Sen logaritmi on

$$l(\boldsymbol{\pi}_1, \dots, \boldsymbol{\pi}_J) = \sum_i (y_{i1} \log(\pi_{i1}) + \dots + y_{iJ} \log(\pi_{iJ})) + \text{vakio}.$$

Kun tähän sijoitetaan kaavat (1) ja (2), niin saadaan

$$l(\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_J) = \sum_i \left( y_{i1} \log\left(\frac{1}{1 + \sum_{j=2}^J \exp(\mathbf{x}_i^T \boldsymbol{\beta}_j)}\right) + y_{i2} \log\left(\frac{\exp(\mathbf{x}_i^T \boldsymbol{\beta}_2)}{1 + \sum_{j=2}^J \exp(\mathbf{x}_i^T \boldsymbol{\beta}_j)}\right) + \dots \right. \\ \left. + y_{iJ} \log\left(\frac{\exp(\mathbf{x}_i^T \boldsymbol{\beta}_J)}{1 + \sum_{j=2}^J \exp(\mathbf{x}_i^T \boldsymbol{\beta}_j)}\right) \right) + \text{vakio}.$$

Tiedetään, että  $\sum_j y_{ij} = N_i$ , joten  $y_{i1} = N_i - (y_{i2} + \dots + y_{iJ})$ . Tätä sekä logaritmien las-  
kusääntöjä käyttämällä, saadaan log-uskottavuus sievenemään yksinkertaisempaan muo-  
toon:

$$l(\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_J) = \sum_i \left( N_i \log\left(1 + \sum_{j=2}^J \exp(\mathbf{x}_i^T \boldsymbol{\beta}_j)\right) + y_{i2} \mathbf{x}_i^T \boldsymbol{\beta}_2 + \dots + y_{iJ} \mathbf{x}_i^T \boldsymbol{\beta}_J \right) + \text{vakio}.$$

## 2.3 Suurimman uskottavuuden estimointi

Nominaalisen logistisen regression kuten myös tavallisen logistisen regression parametrit estimoidaan suurimman uskottavuuden menetelmällä. Suurimman uskottavuuden estimaatti  $\hat{\boldsymbol{\beta}}$  toteuttaa yhtälön  $\max_{\boldsymbol{\beta}} l(\boldsymbol{\beta}) = l(\hat{\boldsymbol{\beta}})$ .

Log-uskottavuusfunktio on konkaavi, joten tavallisesti log-uskottavuusfunktiolle maksimin antava äärellinen piste, jos se on olemassa, on yksikäsitteinen ja se löytyy asettamalla osittaisderivaatat nolliksi. Kun verrataan vasteen  $j$ . luokkaa ensimmäiseen, niin osittaisderivaatat ovat samankaltaisia kuin tavallisessa logistisessä regressiossa, mutta kyseisiä vasteen luokkien vertailuja on nominaalisessa logistisessä regressiossa  $J - 1$  kappaletta, koska muita luokkia verrataan valittuun referenssiluokkaan. Osittaisderivaatat ovat

$$\frac{\partial l(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_j} = \sum_i \left( y_{ij} \mathbf{x}_i - \frac{N_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}_j)}{1 + \exp(\mathbf{x}_i^T \boldsymbol{\beta}_j)} \mathbf{x}_i \right).$$

Kun osittaisderivaatat asetetaan nolliksi, niin yhtälöt voidaan koota vektoryhtälöksi

$$\sum_i y_{ij} \mathbf{x}_i = \sum_i \frac{N_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}_j)}{1 + \exp(\mathbf{x}_i^T \boldsymbol{\beta}_j)} \mathbf{x}_i.$$



Uskottavuusyhtälöt ratkeavat numeerisesti ja se täytyy tehdä iteratiivisesti.

## 2.4 Kertoimien tulkinta ja Pearsonin jäännökset

Kaavojen (1) ja (2) avulla saadaan estimaatit todennäköisyyksille vasteen eri luokille jokaisessa prediktorien ryhmässä, kun uskottavuusyhtälöistä estimoidut  $\hat{\beta}_j$  sijoitetaan  $\beta_j$  paikalle kyseisiin kaavoihin (esim. todennäköisyydet positiivisesti, vaikea sanoa ja negatiivisesti vastanneille nuorimman ikäluokan naisille, jotka kannattavat kokoomusta). Kun tiedetään ryhmän  $i$  koko  $N_i$  ja estimoidut todennäköisyydet  $\hat{\pi}_{ij}$ , niin saadaan odotetut frekvenssit

$$e_{i1} = N_i \hat{\pi}_{i1} = N_i \frac{1}{1 + \frac{\hat{\pi}_{i2}}{\hat{\pi}_{i1}} + \dots + \frac{\hat{\pi}_{iJ}}{\hat{\pi}_{i1}}} = N_i \frac{1}{1 + \sum_{j=2}^J \exp(\mathbf{x}_i^T \hat{\beta}_j)}$$

ja

$$e_{ij} = N_i \hat{\pi}_{ij} = N_i \frac{\frac{\hat{\pi}_{ij}}{\hat{\pi}_{i1}}}{1 + \frac{\hat{\pi}_{i2}}{\hat{\pi}_{i1}} + \dots + \frac{\hat{\pi}_{iJ}}{\hat{\pi}_{i1}}} = N_i \frac{\exp(\mathbf{x}_i^T \hat{\beta}_j)}{1 + \sum_{j=2}^J \exp(\mathbf{x}_i^T \hat{\beta}_j)}$$

Kertoimien tulkinta onnistuu yleensä helpoiten vedonlyöntisuhteiden avulla, kuten tavallisessa logistisessa regressiossa. Nominaalisen logistisen regression tapauksessa vedonlyöntisuhteet muodostetaan vasteen referenssiluokan ja tarkasteltavan luokan välille. Myös tässä tutkielmassa vedonlyöntisuhteita käytetään erojen tulkinnassa ja todennäköisyyksien estimaatteja sekä sovitettuja arvoja käytetään Pearsonin jäännöksiä laskemisessa.

Jos kaikki prediktorit ovat luokiteltuja, niin odotettujen frekvenssien avulla voidaan laskea Pearsonin jäännökset. Mallin sopivuutta aineistoon voidaan tarkastella Pearsonin jäännöksiä avulla. Pearsonin jäännökset lasketaan kaavalla

$$r_{ij} = \frac{y_{ij} - e_{ij}}{\sqrt{e_{ij}}}, \quad i = 1, \dots, n \quad \text{ja} \quad j = 1, \dots, J.$$

## 2.5 Uskottavuusosamäärätesti ja AIC

Sovitetun mallin sopivuutta aineistoon voidaan tarkastella vertaamalla sitä saturoituun malliin (Dobson, 2002, s. 75-77). Saturoitu malli tarkoittaa maksimaalista mallia, jossa kaikki mahdolliset interaktiot ovat mukana. Merkitään  $m$ :llä parametrien suurinta määrää, joka voidaan estimoida. Olkoon  $\beta_{max}$  parametrivektori saturoidulle mallille ja  $\hat{\beta}_{max}$  suurimman uskottavuuden estimaattori  $\beta_{max}$ :lle sekä  $L(\hat{\beta}_{max}) =: \hat{L}_{sat}$  uskottavuusfunktion arvo saturoidulle mallille kohdassa  $\hat{\beta}_{max}$  ja sen logaritmi  $l(\hat{\beta}_{max}) =: \hat{l}_{sat}$ . Olkoon vielä

$p$  sovitetun mallin parametrien lukumäärä ja  $L(\hat{\beta}) =: \hat{L}$  sovitetun mallin uskottavuusfunktion suurin arvo ja  $l(\hat{\beta}) =: \hat{l}$  sen logaritmi. Tällöin uskottavuusosamäärän logaritmi

$$D = 2 \log \left( \frac{L(\hat{\beta}_{max})}{L(\hat{\beta})} \right) = 2[l(\hat{\beta}_{max}) - l(\hat{\beta})].$$

Symbolia  $D$  käytetään siksi, että uskottavuusosamäärää kutsutaan myös devianssiksi. Suurilla  $D$ :n arvoilla sovitettu malli kuvaa aineistoa huonosti verrattuna saturoituun malliin. Dobson myös perustelee, että  $D$  noudattaa  $\chi^2$ -jakaumaa vapausastein  $m - p$ .

$$D \sim \chi^2(m - p).$$

Multinomiaalisen mallin tapauksessa saturoidun mallin estimaatit ovat suhteellisia osuuksia  $\bar{\pi}_{ij} = y_{ij}/N_i$  rajoituksella  $\sum_j \bar{\pi}_{ij} = 1$ . Saturoidun mallin uskottavuusfunktion maksimiarvo on

$$\begin{aligned} \hat{L}_{sat} &= \prod_{i=1}^n \binom{N_i}{y_{i1}, \dots, y_{iJ}} \times \bar{\pi}_{i1}^{y_{i1}} \times \dots \times \bar{\pi}_{iJ}^{y_{iJ}} \\ &= \prod_{i=1}^n \binom{N_i}{y_{i1}, \dots, y_{iJ}} \times \left(\frac{y_{i1}}{N_i}\right)^{y_{i1}} \times \dots \times \left(\frac{y_{iJ}}{N_i}\right)^{y_{iJ}} \end{aligned}$$

ja sen logaritmi

$$\hat{l}_{sat} = \sum_{i=1}^n \log \binom{N_i}{y_{i1}, \dots, y_{iJ}} + y_{i1} \log \left(\frac{y_{i1}}{N_i}\right) + \dots + y_{iJ} \log \left(\frac{y_{iJ}}{N_i}\right).$$

Merkitään sovitetun nominaalisen logistisen regressiomallin suurimman uskottavuuden estimoitua todennäköisyyttä taas  $\hat{\pi}_{ij}$  ja vastaavaa sovitettua arvoa  $e_{ij} = N_i \hat{\pi}_{ij}$ . Silloin uskottavuusfunktion maksimiarvo on

$$\begin{aligned} \hat{L} &= \prod_{i=1}^n \binom{N_i}{y_{i1}, \dots, y_{iJ}} \times \hat{\pi}_{i1}^{y_{i1}} \times \dots \times \hat{\pi}_{iJ}^{y_{iJ}} \\ &= \prod_{i=1}^n \binom{N_i}{y_{i1}, \dots, y_{iJ}} \times \left(\frac{e_{i1}}{N_i}\right)^{y_{i1}} \times \dots \times \left(\frac{e_{iJ}}{N_i}\right)^{y_{iJ}} \end{aligned}$$

ja sen logaritmi

$$\hat{l} = \sum_{i=1}^n \log \binom{N_i}{y_{i1}, \dots, y_{iJ}} + y_{i1} \log \left(\frac{e_{i1}}{N_i}\right) + \dots + y_{iJ} \log \left(\frac{e_{iJ}}{N_i}\right).$$

Tällöin devianssi on

$$D = 2 \log \left( \frac{\hat{L}_{sat}}{\hat{L}} \right) = 2(\hat{l}_{sat} - \hat{l})$$

$$= 2 \sum_{i=1}^n [y_{i1} \log \left( \frac{y_{i1}}{e_{i1}} \right) + \dots + y_{iJ} \log \left( \frac{y_{iJ}}{e_{iJ}} \right)].$$

Uskottavuusosamäärätestillä (UO-testi) voidaan verrata myös laajempaa mallia ja laajemman mallin sisältämää suppeampaa mallia. Kyseisiä malleja kutsutaan myös sisäkkäisiksi malleiksi. Oletetaan, että laajemmassa nominaalisessa logistisessa regressiomallissa on  $p$  prediktorina ja suppeammassa mallissa on  $q$  prediktorina ( $q < p$ ). Nollahypoteesina on, että suppeampi malli on tosi. Testi suoritetaan maksimoimalla sekä  $p$ :n prediktorin malli että suppeampi  $q$ :n prediktorin malli. Testisuure saadaan devianssien erotuksena

$$2 \log \left( \frac{\hat{L}_p}{\hat{L}_q} \right) = 2 \log \left( \frac{\hat{L}_{sat} \hat{L}_p}{\hat{L}_q \hat{L}_{sat}} \right) = D_q - D_p,$$

missä  $D_p$  on laajemman mallin devianssi ja  $D_q$  suppeamman mallin devianssi. Kun  $D_q - D_p$  on suuri, suppeampi  $q$ :n prediktorin malli kuvaa aineistoa huonosti verrattuna laajempaan  $p$ :n prediktorin malliin. Kun nollahypoteesi (suppeampi malli) on tosi, niin

$$D_q - D_p \sim \chi^2(p - q).$$

Tässä tutkielmassa malleja vertaillaan myös AIC:n avulla. AIC:n avulla voidaan vertailla muitakin kuin sisäkkäisiä malleja. Jos merkitään mallin estimoitujen parametrien lukumäärää  $k$ :lla, niin Davisonia (2003, s. 152) mukailten saadaan

$$AIC = 2k - 2 \log(\hat{L}).$$

Jos mallin valinta tehdään informaatiokriteerin perusteella, niin valitaan se malli, jolla on pienin AIC.

Laskenta suoritetaan R-ympäristössä (R Core Team, 2015) multinomiaaliseen logistiseen regressioon tarkoitettulla `nnet`-paketilla (Venables & Ripley, 2002).

### 3 Aineisto

EVA on teettänyt kansallisen asennetutkimuksen vähintään kahden vuoden välein vuodesta 1984 lähtien. Tutkimukset on toteuttanut EVAn toimeksiannosta pääosin Yhdyskuntatutkimus. Tämän tutkielman aineisto on peräisin vuoden 2014 asennetutkimuksesta

ja sen on kerännyt Taloustutkimus. Kyseinen aineisto, ja myös muiden vuosien tutkimusten aineistoja, on ladattavissa Tampereen yliopiston Yhteiskuntatieteellisen tietoarhivon Ailan palveluportaalista. (EVAn kansallinen asennetutkimus, 2014.)

Otos vastaajista on poimittu Taloustutkimuksen internet-paneelista 18–70-vuotiasta Suomen (pl. Ahvenanmaa) väestöä edustavaksi vastaajien iän, sukupuolen, asuinläänin, koulutuksen ja ammatin/aseman mukaan. Vastajat on kutsuttu Internet-paneeliin henkilökohtaisesti perustuen satunnaisotantaan Manner-Suomen 15–79-vuotiaasta väestöstä (Keckman-Koivuniemi 2015). Otoksen poiminnassa panelista on käytetty satunnaisotantaa, joka on muodostettu ohjelmallisesti ja poiminnassa on huomioitu eri vastaajaryhmien ennakoitu vastauskäyttäytyminen. Tutkimuskutsut on lähetetty valitulle kohderyhmälle sähköpostitse. Vastajat täyttivät sähköisen kyselylomakkeen internetissä Taloustutkimuksen suojatussa tutkimusympäristössä. Tiedonkeruun aikana on lähetetty neljä muistutuspostia siihen mennessä vastaamatta tai vastaamisen kesken jättäneille. Tutkimukseen vastasi hyväksytysti 30.1.2014 mennessä 2052 vastaajaa. Vastausprosentti oli 30 %. (EVAn kansallinen asennetutkimus, 2014.)

Vastausprosentti on melko tavanomainen sähköpostin kautta lähetetyille kyselyille, joten puuttuvaa tietoa on paljon. Puuttuvalle tiedolle voidaan määrittää kolme lajia (Rubin & Little, 2002). Seuraavaksi käydään läpi taustamuuttujia ja huomataan, että vanhimman ikäluokan vastausprosentti on pienempi kuin perusjoukon samanikäisten prosentuaalinen osuus, joten puuttuva tieto ei ole ainakaan täysin satunnaista (MCAR). Satunnainen puuttuvuus (MAR) tarkoittaa tämän kyselyn kohdalla sitä, että vastauksien puuttuminen on riippumaton kyselyssä esitetyistä kysymyksistä ja väittämistä. Koska puuttuvasta tiedosta ei ole enempää informaatiota käytettävissä, niin tämän tutkimuksen analyysit tehdään olettaen, että puuttuva tieto on satunnaista.

Prediktoreita otetaan lähempään tarkasteluun kuusi: sukupuoli, ikäluokka, puolue, alue, koulutustaso ja ammattiryhmä. Ikä jaetaan neljään luokkaan: 18–25-vuotiaat (nuori), 26–40-vuotiaat (alempi keski-ikä), 41–60-vuotiaat (ylempi keski-ikä) sekä yli 60-vuotiaat (vanha). Sukupuolen (taulukko 1) ja ikäluokkien (taulukko 3) jakaumat ovat suunnilleen sitä, mitä ne ovat olleet koko Suomessa vuoden 2014 alussa (taulukot 2 ja 4). Ainoastaan yli 60-vuotiailla on muita suurempi ero perusjoukon ja aineiston jakaumien välillä. Tämä johtuu luultavasti siitä, että kysely on lähetetty sähköpostitse, ja tästä syystä yli 60-vuotiaiden tavoittaminen on ollut vaikeampaa.

Taulukko 1: Naisten ja miesten frekvenssit ja prosenttiosuudet aineistossa.

Sukupuoli	n	%
Nainen	997	48.6
Mies	1055	51.4

Taulukko 2: 18–70-vuotiaiden naisten ja miesten frekvenssit ja prosenttiosuudet Suomessa 31.12.2013 (Tilastokeskuksen PX-Web-tietokannat, 2016).

Sukupuoli	n	%
Nainen	1 851 815	49.7
Mies	1 873 199	50.3

Taulukko 3: Ikäluokkien frekvenssit ja prosenttiosuudet aineistossa.

Ikäluokka	n	%
yli 60-vuotiaat	265	12.9
41-60-vuotiaat	867	42.3
26-40-vuotiaat	610	29.7
18-25-vuotiaat	310	15.1

Taulukko 4: Ikäluokkien frekvenssit ja prosenttiosuudet Suomessa 31.12.2013 (Tilastokeskuksen PX-Web-tietokannat, 2016).

Ikäluokka	n	%
61-70-vuotiaat	713 522	19.2
41-60-vuotiaat	1 448 697	38.9
26-40-vuotiaat	1 021 643	27.4
18-25-vuotiaat	541 152	14.5

Puoluekanta kyselyssä on kysytty niin, että minkä puolueen edustajaa vastaaja olisi äänestänyt, jos eduskuntavaalit olisi pidetty kyselyhetkellä. Vastausvaihtoehtoja on ollut alunperin 12 (taulukko 5). Prediktorina puolue varten taulukon 5 kolme viimeisintä vastausvaihtoehtoa (ei olisi äänestänyt sekä ei olisi osannut tai halunnut sanoa kantaansa) yhdistetään yksinkertaistamisen vuoksi yhdeksi luokaksi ja nimeksi annetaan ei-äänestäneet.

Taulukko 5: Puoluekantojen frekvenssit aineistossa.

Puolue	n	%
Kokoomus	269	13.1
SDP	203	9.9
Perussuomalaiset	276	13.5
Keskusta	293	14.3
Vasemmistoliitto	132	6.4
Vihreät	228	11.1
RKP	30	1.5
Kristillisdemokraatit	37	1.8
Jokin muu puolue	32	1.6
Ei äänestäisi	70	3.4
Ei osaa sanoa	409	19.9
Ei halua sanoa	73	3.6

Kyselyssä on kysytty asuinmaakuntaa (taulukko 6). Suomi on jaettu asuinmaakuntien mukaan kolmeen osaan: länteen, itään ja pohjoiseen, josta määräytyvät prediktorin alue luokat. Länteen kuuluvat Uusimaa, Varsinais-Suomi, Itä-Uusimaa, Satakunta, Kanta-Häme, Pirkanmaa, Päijät-Häme, Keski-Suomi, Etelä-Pohjanmaa ja Pohjanmaa. Itää edustavat Etelä-Karjala, Pohjois-Karjala, Kainuu, Etelä-Savo, Pohjois-Savo ja Kymenlaakso. Pohjoisessa ovat Lappi ja Pohjois-Pohjanmaa.

Taulukko 6: Asuinmaakuntien frekvenssit aineistossa.

Asuinmaakunta	n	%
Uusimaa	602	29.3
Varsinais-Suomi	218	10.6
Itä-Uusimaa	0	0.0
Satakunta	65	3.2
Kanta-Häme	56	2.7
Pirkanmaa	204	9.9
Päijät-Häme	74	3.6
Kymenlaakso	56	2.7
Etelä-Karjala	44	2.1
Etelä-Savo	41	2.0
Pohjois-Savo	114	5.6
Pohjois-Karjala	59	2.9
Keski-Suomi	116	5.7
Etelä-Pohjanmaa	75	3.7
Pohjanmaa	46	2.2
Keski-Pohjanmaa	29	1.4
Pohjois-Pohjanmaa	158	7.7
Kainuu	30	1.5
Lappi	65	3.2

Yhtenä taustamuuttujana on kysytty myös ammatillista koulutusta (taulukko 7). Prediktorissa koulutustaso luokkina ovat perus-, keski- ja korkea taso. Vain peruskoulun käyneet kuuluvat perustasoon, ammattikoulun tai lukion käyneet kuuluvat keskitasoon ja korkeakoulun käyneet korkeaan tasoon.

Taulukko 7: Ammatillisen koulutuksen frekvenssit aineistossa.

Ammatillinen koulutus	n	%
Ei ammatillista koulutusta	246	12.0
Ammattikurssi, muu lyhyt ammattikoulutus	138	6.7
Ammattikoulu, kouluasteen ammatill. tutkinto	469	22.9
Opistotasoinen ammattikoulutus	417	20.3
Ammattikorkeakoulututkinto	291	14.2
Yliopisto- tai korkeakoulututkinto	491	23.9

Viimeisenä prediktorina on ammattiryhmä. On kysytty, mihin ammattiryhmään vastaaja on katsonut kuuluvansa (taulukko 8). Prediktori ammattiryhmä on luokiteltu viiteen luokkaan: toimihenkilö, työntekijä, opiskelija, eläkeläinen sekä muu tai työtön. Toimihenkilöihin kuuluvat kolme ylintä luokkaa taulukosta 8 ja työntekijöihin työntekijöiden lisäksi yrittäjät, yksityiset ammatinharjoittajat ja maatalousyrittäjät. Muihin ja työttömiin kuuluvat muiden ja työttömien lisäksi kotivanhemmat.

Taulukko 8: Ammattiryhmän frekvenssit aineistossa.

Ammattiryhmä	n	%
Johtavassa asemassa toisen palv.	44	2.1
Ylempi toimihenkilö	250	12.2
Alempi toimihenkilö	273	13.3
Työntekijä	534	26.0
Yrittäjä tai yksityinen ammatinharjoittaja	108	5.3
Maatalousyrittäjä	27	1.3
Opiskelija	212	10.3
Eläkeläinen	350	17.1
Kotiäiti/koti-isä	45	2.2
Työtön	164	8.0
Muu	45	2.2

Tutkimuksessa on kysytty suomalaisten arvoja ja asenteita. Teemoina ovat olleet suomalaisen yhteiskunnan muutostarpeet ja uudistaminen, eriarvoisuus ja tasa-arvo sekä Euroopan unioni. Kiinnostavia väittämiä on esitetty muun muassa NATOsta, EU:sta, maahanmuuttajista, USA:sta ja Venäjältä. Väittäminen Venäjältä on valikoitunut kuitenkin tutkielman ensimmäiseksi tutkimusongelmaksi, koska väittämiä Venäjältä on vain yksi ja se on melko selkeästi esitetty. Lisäksi Ukrainan kriisi käynnistyi kyseisen vuoden keväänä, joten voidaan tarkastella mielipiteiden eroja Venäjältä, kun Ukrainan kriisi ei ollut vielä ehtinyt vaikuttamaan mielipiteisiin.

Väittäminen Venäjältä on muotoiltu seuraavasti: “Vaikka Venäjällä on omat ongelman-  
sa, suomalaisilla ei ole nykyisin mitään syytä suhtautua suureen naapuriinsa kielteisesti.”. Vastausvaihtoehtoja on ollut viisi: täysin samaa mieltä, jokseenkin samaa mieltä, vaikea sanoa, jokseenkin eri mieltä ja täysin eri mieltä. Samaa mieltä väittämisen kanssa olleet voidaan katsoa suhtautuneen Venäjään myönteisemmin tai ainakin neutraalimmin kuin eri mieltä olevat. Ongelman yksinkertaistamiseksi vaste jaetaan kolmeen luokkaan: täysin samaa mieltä ja jokseenkin samaa mieltä yhdistetään positiivisemmaksi kannaksi, vaikea sanoa pysyy omanaan ja täysin eri mieltä sekä jokseenkin eri mieltä yhdistetään negatii-



viseksi suhtautumiseksi. Taulukosta 9 huomataan, että noin 30 prosenttia vastaajista on suhtautunut kielteisesti Venäjään, noin 47 prosenttia on ollut positiivisemmalla kannalla ja 23 prosentille vastaajista on ollut vaikea sanoa.

Taulukko 9: Frekvenssit ja prosenttiosuudet Venäjä-väittämän vastausvaihtoehdoille.

Vastausvaihtoehto	Arvo	n	%
Täysin samaa mieltä	1	222	10,8
Jokseenkin samaa mieltä	2	749	36,5
Vaikea sanoa	3	479	23,3
Jokseenkin eri mieltä	4	473	23,1
Täysin eri mieltä	5	129	6,3

Tämän jälkeen on otettu mukaan myös kolme muuta väittämää. Väittämät koskevat NATOa, USA:ta ja EU:ta. Kaikki valitut kysymykset ovat ulkopoliittisia, erityisesti halutaan saada selville puolueiden kannattajien suhtautumisten eroja näihin kysymyksiin. NATO-väittämä on mukana siitä syystä, että monesti Venäjä-keskustelun ohella käydään myös keskustelua liittymisestä NATOon. Väittämä Yhdysvaltojen ulkopoliitikasta voisi kuvitella noudattavan jossain määrin NATO-väittämän tuloksia, koska Yhdysvallat on NATO:n vahvin maa. Keskustelu EU:sta ja sen toimivuudesta on ollut aina suomalaisille tärkeää ja tällä hetkellä se on erityisen kiinnostava johtuen Iso-Britannian erosta, joten on mielenkiintoista tutkia, millaisia ovat olleet suomalaisten asenteet tätä väittämää kohtaan vuoden 2014 alussa. Vastausvaihtoehdot ovat näissä väittämissä samat kuin Venäjä-väittämän kohdalla. Kaikissa neljässä kysymyksessä samaa mieltä olevat ovat positiivisemmalla kannalla ja eri mieltä olevat ovat negatiivisemmalla kannalla. Myös näistä loppuista kolmesta väittämästä (vasteista) tehdään yksinkertaistamisen vuoksi kolmiluokkaisia ( $J = 3$ ). Yhdistetään luokat täysin samaa mieltä ja jokseenkin samaa mieltä positiiviseksi luokaksi sekä täysin eri mieltä ja jokseenkin eri mieltä negatiiviseksi luokaksi. Vaikea sanoa -luokka pysyy omanaan, ja tässä tutkielmassa siitä käytetään myös nimitystä ei osaa sanoa -luokka.

NATOa koskeva väittämä on esitetty yksinkertaisesti: “Suomen tulisi liittyä NATOon”. Taulukosta 10 nähdään, että noin 51 prosenttia vastaajista on suhtautunut kielteisesti NATOon liittymiseen, noin 18 prosenttia on ollut positiivisemmalla kannalla ja 31 prosentille vastaajista on ollut vaikea sanoa.

Taulukko 10: Frekvenssit ja prosenttiosuudet NATO-väittämän vastausvaihtoehdoille.

Vastausvaihtoehto	Arvo	n	%
Täysin samaa mieltä	1	140	6,8
Jokseenkin samaa mieltä	2	239	11,6
Vaikea sanoa	3	630	30,7
Jokseenkin eri mieltä	4	377	18,4
Täysin eri mieltä	5	666	32,5

Väittäjä Yhdysvalloista on esitetty: “Yhdysvallat toimii maailmanpolitiikassa oikein ja ansaitsee myös suomalaisten tuen”. Taulukosta 11 havaitaan, että noin 62 prosenttia vastaajista on suhtautunut kielteisesti Yhdysvaltojen maailmanpolitiikkaan, noin 8 prosenttia on ollut positiivisemmalla kannalla ja 30 prosentille vastaajista on ollut vaikea sanoa.

Taulukko 11: Frekvenssit ja prosenttiosuudet USA-väittämän vastausvaihtoehdoille.

Vastausvaihtoehto	Arvo	n	%
Täysin samaa mieltä	1	24	1,2
Jokseenkin samaa mieltä	2	143	7,0
Vaikea sanoa	3	605	29,5
Jokseenkin eri mieltä	4	830	40,4
Täysin eri mieltä	5	450	21,9

EU:sta on kysytty seuraavaa: “Suomi on saanut äänensä hyvin kuuluviin EU:n päätöksenteossa”. Taulukosta 12 nähdään, että noin 49 prosentin mielestä Suomi ei ole saanut ääntänsä kuuluviin EU:n päätöksenteossa, noin 17 prosenttia on ollut positiivisemmalla kannalla ja 34 prosentille vastaajista on ollut vaikea sanoa.

Taulukko 12: Frekvenssit ja prosenttiosuudet EU-väittämän vastausvaihtoehdoille.

Vastausvaihtoehto	Arvo	n	%
Täysin samaa mieltä	1	23	1,1
Jokseenkin samaa mieltä	2	330	16,1
Vaikea sanoa	3	693	33,8
Jokseenkin eri mieltä	4	739	36,0
Täysin eri mieltä	5	267	13,0

## 4 Tuloksia Venäjä-väittämään

Kuten aiemmin on todettu, Venäjä-väittäjä on kysytty seuraavasti vuoden 2014 EVAn kansallisessa asennetutkimuksessa: “Vaikka Venäjällä on omat ongelmansa, suomalaisilla ei ole nykyisin mitään syytä suhtautua suureen naapuriinsa kielteisesti.” Väittämää ei ole esitetty täysin suoraan, että suhtautuuko Venäjään kielteisesti. En kuitenkaan näe, että väittämän asettelu ohjaisi vastaamaan suuntaan tai toiseen. Onko vastaaja kenties laitettu miettimään pidemmällä muotoilulla? Samaa mieltä olevat ovat kuitenkin suhtautuneet positiivisemmin kuin eri mieltä olevat.

Oman tutkimukseni alussa oli tarkoitus etsiä kiinnostavin väittäjä ja soveltaa multinomiaalista logistista regressiota siihen. Venäjä-väittäjä on valikoitunut kyseiseksi väittäjäksi Ukrainan kriisin alkamisen ja kyselyhetken ajankohdan lähekkäisyyksien vuoksi. On mielenkiintoista tutkia, minkälaisia eroja eri ryhmien välillä on ollut asennoitumisessa Venäjää kohtaan ennen Ukrainan kriisin alkamista. Kiinnostavaa olisi vertailla näitä tuloksia myöhemmin tehtyjen kyselyiden tuloksiin, ja selvittää, ovatko näkemuserot muuttuneet Ukrainan kriisin alun jälkeen. Nyt keskitytään kuitenkin vain tähän tutkimukseen ja seuraavaksi Venäjä-väittäjälle etsitään sopiva malli, jonka avulla vertaillaan eri ryhmien keskimääräisiä eroja asennoitumisessa Venäjää kohtaan.

### 4.1 Mallinvalinta

Venäjä-väittäjälle etsitään sopiva malli, joka on helposti tulkittavissa, mutta joka on myös riittävän selittävä malli, ja jonka avulla havaitaan mahdolliset ryhmien väliset erot. Riittävin malli etsitään uskottavuusosamäärätestin (UO-testi) ja AIC:n avulla. Taulukosta 13 löytyvät mallien residuaalidevianssit, estimoitujen parametrien lukumäärät, AIC:t ja mallien välisiä UO-testien  $p$ -arvoja. Testisuure UO-testille saadaan taulukon avulla vertailtavien mallien devianssien erotuksena ja vapausasteet vastaavasti mallien estimoitujen parametrien lukumäärien erotuksena. Vertailut aloitetaan mallista, jossa on mukana kaikki kuuden prediktorin parittaiset interaktiot. Tämä malli on niin iso ja monimutkainen

tulkita, että siitä on välttämättä karsittava termejä pois. Lisäksi kyseisen mallin kaikkia parametreja ei saada estimoitua. Aineisto ja prediktorit ovat muutoinkin sellaisia, että parametreja ei saada estimoitua joidenkin interaktioiden tapauksissa, joten tästäkin syystä päädyttäisiin lopulta tarkastelemaan sellaisia malleja, joissa on vain päävaikutuksia. Osoittautuu, että asuinalueen, koulutustason ja ammattiryhmän luokkien välillä ei ole merkitseviä eroja, ja että riittävin malli on sellainen, jossa on sukupuolen, ikäluokan ja puoluekannan päävaikutukset. UO-testin  $p$ -arvo on 0.833, kun vertaillaan sukupuolen, ikäluokan, puoluekannan, asuinalueen, koulutustason ja ammattiryhmän päävaikutusten mallia (malli 2) ja sukupuolen, ikäluokan ja puoluekannan päävaikutusten mallia (malli 4). Lisäksi mallin 4 AIC on pienin vertailtavista malleista, joten tämäkin puoltaa mallin 4 valitsemista.

Taulukko 13: Mallien residuaalidevianssit, estimoitujen parametrien lukumäärät (df), AIC:t ja UO-testien  $p$ -arvoja Venäjä-väittämään. Mallissa 3 kaksoispiste tarkoittaa interaktiota, ts. malli 3 on sukupuolen, ikäluokan ja puolueen täysimalli.

	malli	res. dev	df	AIC	UO-testien $p$ -arvoja
1	kaikki parittaiset interaktiot	3785	366	4517	-
2	kaikki päävaikutukset	4159	44	4247	(1 vs 2) 0.023
3	sukupuoli:ikäluokka:puolue	4026	158	4342	(1 vs 3) 0.059
4	sukupuoli+ikäluokka+puolue	4170	28	4226	(1 vs 4) 0.039
					(2 vs 4) 0.833
					(3 vs 4) 0.184

## 4.2 Sukupuoli ja ikäluokka

Miehille ja vanhemmille ikäluokille kertoimet ovat taulukossa 14. Kertoimia voidaan tulkita, kuten dikotomisen vasteen logistisessa regressiossa. Koska mallissa on vain päävaikutuksia eikä ollenkaan interaktioita, niin tulkittaessa tietyn prediktorin kertoimia, muiden prediktoreiden arvot pysyvät samoina. Esimerkiksi jos tarkastellaan sukupuolten eroja, niin silloin vertaillaan miehiä ja naisia, jotka ovat olleet samasta ikäluokasta ja olisivat äänestäneet samalla tavalla. Tai ikäluokkien kertoimia tulkittaessa vertaillaan niitä, jotka ovat olleet samaa sukupuolta ja olisivat äänestäneet samalla tavalla, mutta poikenneet ikäluokan suhteen. Sukupuolta ja ikäluokkia tarkasteltaessa suurimmat erot löytyvät positiivisen ja negatiivisen suhtautumisen yhtälöstä. Naiset ja nuorin ikäluokka, 18-25-vuotiaat, ovat referenssitasona. Ikäluokan kohdalla tämä tarkoittaa sitä tulkinnan kannalta, että muita ikäluokkia verrataan nuorimpaan ikäluokkaan.

- Positiivisesta kertoimesta (0.458) voidaan päätellä, että miehet ovat suhtautuneet Venäjään positiivisemmin kuin naiset. Mallin mukaan positiivisen mielipiteen to-

dennäköisyys suhteessa negatiiviseen mielipiteeseen on  $\exp(0.458) = 1.580$ -kertainen miesten ryhmässä verrattuna naisiin.

- Ikäluokkien kertoimista voidaan havaita, kun siirrytään ikäluokissa nuorimmasta vanhempaan, niin asennoituminen Venäjään on keskimäärin positiivisempaa.
- Kun vertaillaan nuorinta ja vanhinta ikäluokkaa, niin positiivisen mielipiteen todennäköisyys suhteessa negatiiviseen mielipiteeseen on keskimäärin  $\exp(1.007) = 2.737$ -kertainen vanhimmassa ikäluokassa.

Taulukko 14: Nominaalisen logistisen regression kertoimien estimaatit miehille ja vanhemmille ikäluokille Venäjä-väittämään. Perustasoina naiset ja nuorin ikäluokka. Mallissa on mukana myös puoluekanta.

<b>kerroin (neg vs pos)</b>	estimaatti	keskivirhe	$\exp(\text{estimaatti})$
mies	0.458	0.110	1.580
61jaYli	1.007	0.212	2.737
41-60v	0.432	0.156	1.540
26-40v	0.191	0.162	1.210
<b>kerroin (neg vs eos)</b>	estimaatti	keskivirhe	$\exp(\text{estimaatti})$
mies	-0.060	0.129	0.942
61jaYli	0.633	0.258	1.884
41-60v	0.385	0.184	1.470
26-40v	0.292	0.189	1.339

### 4.3 Puoluekanta

Puoluekantoihin liittyvät kertoimet löytyvät taulukosta 15. Puoluekannan kertoimet tulkitaan vastaavasti siten, että vertaillaan niitä, jotka ovat samaa sukupuolta ja samaa ikäluokkaa, mutta joiden näkemykset ovat eronneet puoluekannassa. Referenssitasoksi on valittu RKP, koska RKP:n kerroin on pienin positiivisen ja negatiivisen suhtautumisen yhtälössä. Myös perussuomalaisilla on pieni kerroin, ja kun huomioidaan myös luottamuvälit, niin näiden puolueiden kannattajat ovat suhtautuneet keskimäärin negatiivisemmin Venäjään kuin Venäjä-väittämän positiivisempien puolueiden kannattajat. Positiivisimminkin Venäjään ovat suhtautuneet SDP:n kannattajat.

- Positiivisesta kertoimesta (1.007) voidaan päätellä, että SDP:n kannattajat ovat suhtautuneet Venäjään positiivisemmin kuin RKP:n kannattajat. Positiivisen mielipiteen todennäköisyys suhteessa negatiiviseen mielipiteeseen on keskimäärin

$\exp(1.007) = 2.737$ -kertainen SDP:n kannattajien keskuudessa verrattuna RKP:n kannattajiin.

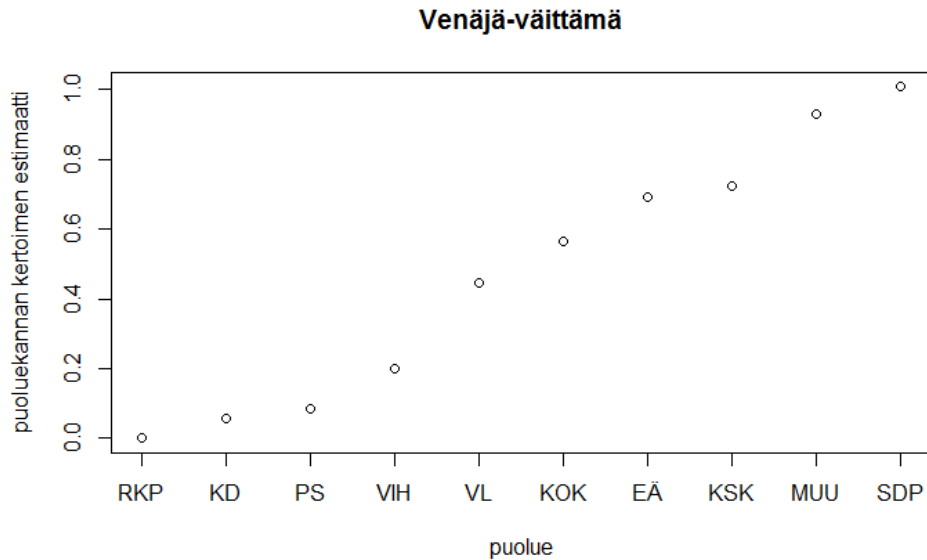
Taulukko 15: Nominaalisen logistisen regression kertoimien estimaatit puoluekannoille Venäjä-väittämään. Perustasona on RKP. Mallissa on mukana myös sukupuolen ja ikäluokan päävaikutukset.

<b>kerroin (neg vs pos)</b>	estimaatti	keskivirhe	exp(estimaatti)
keskusta	0.722	0.454	2.059
kokoomus	0.567	0.454	1.762
kristillisdemokraatit	0.056	0.593	1.057
perussuomalaiset	0.084	0.453	1.088
SDP	1.007	0.470	2.738
vasemmistoliitto	0.445	0.476	1.561
vihreät	0.200	0.461	1.221
ei äänestäneet	0.693	0.446	2.000
muut puolueet	0.929	0.590	2.532
<b>kerroin (neg vs eos)</b>	estimaatti	keskivirhe	exp(estimaatti)
keskusta	-0.141	0.500	0.869
kokoomus	-0.446	0.505	0.640
kristillisdemokraatit	0.232	0.616	1.262
perussuomalaiset	-0.268	0.495	0.765
SDP	0.394	0.514	1.483
vasemmistoliitto	-0.225	0.528	0.799
vihreät	-0.054	0.498	0.948
ei äänestäneet	0.603	0.480	1.827
muut puolueet	-1.122	0.911	0.325

Seuraavaksi tarkastellaan negatiivisen ja positiivisen suhtautumisen yhtälön kertoimia. Puolueet laitetaan järjestykseen Venäjä-väittämän osalta siten, miten heidän kannattajansa ovat suhtautuneet väittämään tässä kyselyssä. Todennäköisimmin kielteisimmin suhtautuneet ovat kuvan alkuosassa ja myönteisimmin suhtautuneet loppuosassa. Lisäksi järjestykset vielä ryhmitellään kolmeen tai neljään luokkaan kertoimien samankaltaisuuksien mukaan. Yhteenvedon yhteydessä tarkastellaan, ovatko joidenkin puolueiden kannattajat suhtautuneet samalla tavalla tarkasteltuihin väittämiin.

Puolueet ovat järjestyksessä Venäjä-väittämän kertoimien osalta kuvassa 1. Kun puolueet lajitellaan kertoimien samankaltaisuuksien mukaan, niin RKP:n, kristillisdemokraattien, perussuomalaisen ja vihreiden kannattajat kuuluvat kielteisempään luokkaan tässä

kyselyssä. Ei-äänestäneet sekä keskustaa, muita puolueita ja SDP:tä äänestäneet kuuluvat myönteisempään luokkaan. Kielteisen ja myönteisen luokkien väliin jäävät vasemmistolii-  
ton ja kokoomuksen kannattajat.



Kuva 1: Venäjä-väittämä ja puolueet järjestettynä kertoimien mukaan pienimmästä suurimpaan.

Alla on vielä taulukoitu luottamusvälit huomioon ottaen kaikki parittaiset erot. Vertailut saataisiin taulukosta 15 kertoimien osalta siten, että puolueiden kertoimet vähennetään toisistaan. Keskivirheitä ei kuitenkaan saada laskettua samaa taulukkoa käyttäen, joten vertailut on tehty siten, että jokainen puoluekannan luokka on ollut referenssiluokkana ja muita luokkia verrattu vuorossa olevaan referenssiluokkaan. Tällä tavoin saadaan selville, minkä puolueiden kannattajat ovat suhtautuneet Venäjään todennäköisemmin negatiivisemmin kuin positiivisemmin suhteessa jonkin toisen puolueen kannattajiin. Samalla tietysti saadaan selville, ketkä ovat suhtautuneet todennäköisemmin positiivisesti kuin negatiivisesti suhteessa toiseen puolueeseen. Tärkein ero verrattuna kuvaan 1 on siinä, että puoluekantojen parittaiset erot voidaan yleistää koko perusjoukkoon.

Taulukosta 16 nähdään, että perussuomalaisten kannattajat ovat olleet todennäköisemmin negatiivisia kuin positiivisia Venäjää kohtaan, kun verrataan SDP:n, keskustan ja kokoomuksen kannattajiin sekä ei-äänestäneisiin. Toisena negatiivisena ryhmänä voidaan pitää vihreiden kannattajia, jotka ovat olleet todennäköisemmin negatiivisempia kuin SDP:tä tai keskustaa äänestäneet. Vaikka RKP:n kannattajien kerroin on pienin,

niin he ovat olleet merkitsevästi negatiivisia kuin positiivisia vain SDP:n kannattajiin nähden. Tämä johtuu siitä, että RKP:n kannattajia on ollut tässä kyselyssä vähän ja siksi luottamusväli on suuri. SDP:n kannattajat ovat olleet selkeästi positiivisimpia ja heidän voidaan katsoa olleen todennäköisesti positiivisempia kuin negatiivisia Venäjää kohtaan kuin RKP:n, kristillisdemokraattien, perussuomalaisen, vasemmistoliittolaisten ja vihreiden kannattajien. Toisena hieman positiivisempina ovat olleet keskustan kannattajat, jotka ovat olleet todennäköisesti positiivisempia kuin vihreitä tai perussuomalaisia kannattaneet.

Taulukko 16: Aiheena Venäjä-väittämä ja taulukoituna puolueet sekä minkä puolueiden kannattajiin nähden kyseisen puolueen kannattajat suhtautuvat negatiivisesti ja mihin positiivisesti, kun ikä ja sukupuoli vakioitu. Vertailut perustuvat parittaisiin luottamusväleihin. KSK = keskusta, KOK = kokoomus, KD=kristillisdemokraatit, PS = perussuomalaiset, VL = vasemmistoliitto, VIHR = vihreät, MUU = jokin muu puolue, EÄ = ei äänestäneet.

Puolue	Negatiivisempi kuin	Positiivisempi kuin
KSK	-	PS, VIHR
KOK	-	PS
KD	SDP	-
PS	KSK, KOK, SDP, EÄ	-
RKP	SDP	-
SDP	-	KD, PS, RKP, VL, VIHR
VL	SDP	-
VIHR	KSK, SDP, EÄ	-
MUU	-	-
EÄ	-	PS, VIHR

Kun tarkastellaan vaikea sanoa -mielipidettä ja negatiivista suhtautumista, niin havaitaan yksi mielenkiintoinen tulos. Ne, jotka eivät olisi äänestäneet silloisissa seuraavissa eduskuntavaaleissa, ovat vastanneet todennäköisemmin “vaikea sanoa” kuin että olisivat vastanneet “eri mieltä” Venäjä-väittämään verrattuna monien puolueiden äänestäjiin. Isoista puolueista ero näkyy selvimmin Kokoomusta äänestäviin nähden.

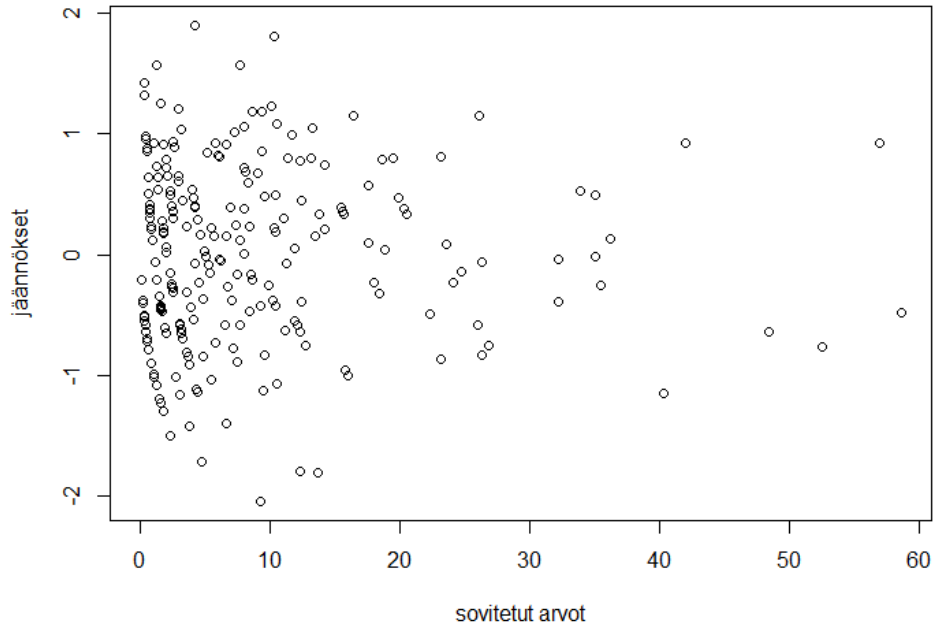
- Ei osaa sanoa -mielipiteen todennäköisyys suhteessa negatiiviseen mielipiteeseen on mallin mukaan  $\exp(0.603) = 1.828$ -kertainen ei äänestävien ryhmässä verrattuna RKP:n kannattajiin, kun vertaillaan samasta ikäluokasta ja samaa sukupuolta ole-



via. Tämä ero ei ole kuitenkaan merkitsevää, toisin kuin esimerkiksi ero kokoomuksen kannattajiin nähden.

Lopuksi vielä tarkastellaan kyseistä ei-äänestäneiden eroa äänestäneisiin tavallisen logistisen regression avulla. Ne, jotka ovat vastanneet Venäjä-väittämään “vaikea sanoa”, ovat yksi ryhmä ja kaikki muut, toisin sanoen jotain mieltä olleet, ovat toinen ryhmä. Logistisessa regressiossa additiiviset prediktorit ovat samat kuin nominaalivasteisen regression tapauksessa. Nyt prediktorin puolue perustasona on ei-äänestäneet, koska kyseinen ryhmä on mielenkiinnon kohteena. Liitteestä 1 huomataan, kun vertaillaan samaa sukupuolta ja samaa ikäluokkaa edustavia, että ei-äänestäneet ovat vastanneet todennäköisemmin “vaikea sanoa” kuin että olisivat olleet jotain mieltä Venäjä-väittämän kanssa, kun vertaillaan keskustaa, kokoomusta, perussuomalaisia, SDP:tä, vasemmistoliittoa, vihreitä tai jotain muuta puoluetta äänestäneisiin. Ero kristillisdemokraattien ja RKP:n kannattajiin ei ole merkitsevää luultavasti siitä syystä, että näiden kahden puolueen kannattajia on ollut vähän tutkimuksessa.

Pearsonin residuaalit edellä tulkitulle nominaalivasteisen regression mallille ovat kuvassa 2. Teoriaosuuden kaavoilla voidaan laskea sovitettut arvot tämän mallin tapauksessa. Sovitettujen ja havaittujen arvojen avulla voidaan laskea Pearsonin jäännökset, ja Pearsonin jäännöksien avulla voidaan tarkastella mallin sopivuutta aineistoon. Kuvasta 2 huomataan, että kuvio ei osoita, etteikö malli sopisi aineistoon. Nähdään, että jäännökset jakautuvat suunnilleen tasan nollan molemmille puolille. Lisäksi alkupään suuremmat vaihtelut nollan molemmin puolin verrattuna loppupään vaihteluihin selittyvät sillä, että sovitettuja (kuten myös havaittuja) arvoja on paljon enemmän nollan lähetyillä kuin esimerkiksi arvosta 20 eteenpäin.



Kuva 2: Pearsonin jäännökset sovitetulle nominaalivasteisen regression mallille Venäjäväittämään.

#### 4.4 Puoluekannan tulosten vertailua muihin tuloksiin ja näkemyksiin

Jotkin tulokset ovat hieman yllättäviä internetistä etsittyihin artikkeleihin ja yleiseen käsitykseen verrattuna. Löydettyjen artikkeleiden, haastatteluiden ja tutkimusten mukaan vihreät ja keskusta ovat olleet ennen ja jälkeen tämän kyselyn suopeimpia Venäjää kohtaan. Kokoomus ja perussuomalaiset ovat sen sijaan olleet kielteisimpiä isoista puolueista. Artikkeleissa kerrotaan suoraan suhtautumisesta Venäjään, mutta myös suhtautumisesta Venäjän ja Suomen väliseen mahdolliseen viisumivapauteen sekä Venäjältä mahdollisesti kohdistuvaan uhkaan.

Taloustutkimuksen selvityksen mukaan vihreiden kannattajat ovat olleet ainoa enemmän myönteisesti kuin kielteisesti suhtautuva kannattajaryhmä Venäjä-matkailun mahdolliseen viisumivapauteen vuonna 2014, kun taas perussuomalaisten kannattajat ovat olleet muita kielteisempiä (Taloustutkimus, 2014). Vaikka Suomen ja Venäjän välinen viisumivapaus ei olekaan sama kysymys kuin suomalaisten suhtautuminen Venäjään, niin vihreiden kannattajien positiivisuus viisumivapautta kohtaan Taloustutkimuksen kyse-

lyssä, ja joidenkin puolueiden kannattajia negatiivisempi asenne Venäjää kohtaan EVAn kansallisessa asennetutkimuksessa, ovat ehkä yllättäviä tuloksia. Se on kuitenkin yhdenmukaista näissä tutkimuksissa ja Venäjää koskevissa kysymyksissä, että perussuomalaiset ja heidän kannattajansa ovat olleet keskimäärin muita negatiivisempia.

Ulkopolitiikan internet-sivuilla Markku Jokisipilä on arvioinut ennen vuoden 2015 eduskuntavaaleja, että keskusta on aina ollut ymmärtäväisempi Venäjää kohtaan kuin SDP sekä kokoomus (Pörsti, 2015). Lisäksi Helsingin Sanomien vuoden 2014 lopulla ilmestyneessä artikkelissa kerrotaan, miten kokoomuksen Alexander Stubb on syyttänyt keskustan Venäjä-linjaa liian myönteiseksi (Teittinen, 2014). Vaikka nämä molemmat artikkelit ja näkemykset ovat Ukrainan kriisin jälkeen, niin ne antavat ymmärtää, että keskusta olisi ollut jo pidemmän aikaa ymmärtäväisempi Venäjää kohtaan kuin SDP ja varsinkin kokoomus. Tämän tutkimuksen mukaan keskustan kannattajat ovat olleet positiivisempia Venäjää kohtaan kuin kokoomuksen kannattajat, mutta eivät merkitsevästi, että tuloksen voisi yleistää. Sen sijaan SDP:n kannattajat ovat olleet tässä tutkimuksessa positiivisempia kuin keskustan kannattajat, mitä voi pitää yllättävänä Markku Jokisipilän lausuntoihin verrattuna.

Sen sijaan Verkkouutisten artikkelin mukaan SDP:n kansanedustajista yli 90 prosenttia ei ole pitänyt Venäjää sotilaallisena uhkana, mikä on selvinnyt MTV:n vaalikoneesta keväältä 2015. Vasemmistoliiton kansanedustajien joukossa tilanne on ollut samanlainen. Keskustan, kristillisdemokraattien ja vihreiden edustajista noin 40 prosenttia on pitänyt Venäjää sotilaallisena uhkana. Kokoomuksen kansanedustajista noin 70 prosenttia ja perussuomalaisten edustajista noin 60 prosenttia on pitänyt Venäjää uhkana. Pakotelinjauksessa keskustan kansanedustajat ovat olleet muiden puolueiden edustajia hieman suopeampia. (Ahtokivi, 2015.) Vaikka nämäkään eivät kerro välttämättä suoraan puolueiden suhtautumisesta Venäjään vuoden 2014 alussa, niin jotain näkemystä nämä kuitenkin antavat. Tämän tutkimuksen tulos, että SDP:n kannattajat ovat olleet positiivisempia kuin monien muiden puolueiden kannattajat, sopii paremmin yhteen Verkkouutisten artikkelin kanssa, vaikkakin artikkeli on peräisin vuodelta 2015 ja siinä on käytetty kansanedustajaehdokkaiden mielipiteitä ja on kysytty hieman eri asiaa. Myös keskustan suhtautuminen verrattuna muihin puolueisiin voidaan katsoa samankaltaiseksi tämän tutkimuksen ja Verkkouutisten MTV:n vaalikoneesta kaivamien tietojen kanssa. Perussuomalaiset ovat olleet Verkkouutistenkin tapauksessa negatiivisimpia puolueita, mikä sopii jälleen yhteen tämän tutkimuksen kanssa. Sen sijaan vasemmistoliiton kansanedustajaehdokkaiden kaltainen positiivisuus ei ole tullut esiin heidän kannattajissaan EVAn kansallisessa asennetutkimuksessa. Kristillisdemokraattien kansanedustajaehdokkaat ovat olleet Verkkouutisten mukaan positiivisimpia puolueita, mutta tämän tutkimuksen perusteella kannattajat ovat olleet negatiivisimpia. Verkkouutisten selvittämä kokoomuksen kansanedustajaehdokkaiden kielteisyyttä ei tule näkyviin kannattajien mielipiteissä tässä EVAn kyselyn pohjalta tehdyssä tutkimuksessa. (Poimitujen artikkeleiden internet-sivut ovat

lähteissä omana kappaleenaan.)

## 5 Tuloksia NATO-väittämään

Varsinkin mediassa kysymys liittymisestä NATOon nostetaan usein esiin, kun on ensin kirjoitettu ja puhuttu Venäjästä. Joten tässäkin tutkielmassa Venäjä-kysymystä seuraa kysymys NATOsta.

Jos Venäjä-väittämää ei olla kysytty täysin suoraan, niin NATO-väittämän kohdalla tilanne on toinen. Väite on esitetty yksinkertaisesti: “Suomen tulisi liittyä NATOon.”.

### 5.1 Mallinvalinta

Muiden kuin Venäjä-väittämän kohdalla ei karsita prediktoreita pois, vaan riittävin malli etsitään suoraan sukupuolen, ikäluokan ja puoluekannan yhdistelmien joukosta. Näin toimitaan siksi, koska jatkossa ollaan kiinnostuneita vain näistä kolmesta prediktorista ja erityisesti puolueiden kannattajien eroavaisuuksista valittujen väittämien kohdalla.

Mallien vertailuja on taulukossa 17 samaan tapaan kuin Venäjä-väittämän tapauksessa. Huomataan, että varsinkin AIC:tä vertailemalla päädyttäisiin samaan malliin, vaikka otettaisiin huomioon kaikki samat prediktorit ja mallien vertailut kuin Venäjä-kysymyksessä. Kun verrataan sukupuolen, ikäluokan ja puoluekannan kyllästettyä mallia päävaikutusten malliin uskottavuusosamäärätestillä,  $p$ -arvoksi saadaan noin 0.335, joten päävaikutusten malli on riittävä. Myös tämän väittämän kohdalla tehdään sama huomautus, että aineisto ja prediktorit ovat sellaisia, että parametreja ei saada estimoitua joidenkin interaktioiden tapauksissa, joten tästäkin syystä päädyttäisiin lopulta tarkastelemaan päävaikutusten mallia.

Taulukko 17: Mallien residuaalidevianssit, estimoitujen parametrien lukumäärät (df), AIC:t ja UO-testien  $p$ -arvoja NATO-väittämään. Mallissa 3 kaksoispiste tarkoittaa interaktiota, ts. malli 3 on sukupuolen, ikäluokan ja puolueen täysimalli.

	malli	res. dev	df	AIC	UO-testien $p$ -arvoja
1	kaikki parittaiset interaktiot	3379	366	4111	-
2	kaikki päävaikutukset	3768	44	3856	(1 vs 2) 0.006
3	sukupuoli:ikäluokka:puolue	3646	158	3962	(1 vs 3) 0.004
4	sukupuoli+ikäluokka+puolue	3782	28	3838	(1 vs 4) 0.009
					(2 vs 4) 0.607
					(3 vs 4) 0.335

## 5.2 Sukupuoli ja ikäluokka

Taulukossa 18 ovat kertoimet prediktoreille sukupuoli ja ikäluokka NATO-väittämän osalta. Koska mallissa on taas vain päävaikutuksia, tarkastellaan jälleen heitä, jotka eroavat tarkasteltavan prediktorin suhteen toisistaan, mutta ovat muutoin samanlaisia. Referenssiluokkina ovat jälleen naiset ja nuorin ikäluokka. Ikäluokkien kohdalla suurimmat erot löytyvät positiivisen ja negatiivisen suhtautumisen yhtälöstä.

- Kun verrataan nuorinta ikäluokkaa vanhempiin ikäluokkiin, niin asennoituminen NATOon liittymiseen on todennäköisemmin negatiivisimpaa nuorimmassa ikäluokassa.
- Kun vertaillaan nuorinta ja vanhinta ikäluokkaa, niin positiivisen mielipiteen todennäköisyys suhteessa negatiiviseen mielipiteeseen on keskimäärin  $\exp(0.991) = 2.693$ -kertainen vanhimmassa ikäluokassa. (Vanhin ikäluokka suhtautuu positiivisemmin kuin nuorin ikäluokka.)

Naiset ovat vastanneet todennäköisemmin “vaikea sanoa” kuin että olisivat vastanneet “eri mieltä” verrattuna miehiin. Naiset ovat olleet myös todennäköisemmin negatiivisempia kuin positiivisempia verrattuna miehiin, mikä selviää negatiivisen ja positiivisen suhtautumisen yhtälöstä.

- Positiivisesta kertoimesta (0.753) voidaan päätellä, että miehet ovat suhtautuneet NATOon liittymiseen positiivisemmin kuin naiset. Positiivisen mielipiteen todennäköisyys suhteessa negatiiviseen mielipiteeseen on keskimäärin  $\exp(0.753) = 2.122$ -kertainen miesten ryhmässä verrattuna naisten ryhmään.

Taulukko 18: Nominaalisen logistisen regression kertoimien estimaatit miehille ja vanhemmille ikäluokille NATO-väittämään. Perustasoina ovat naiset ja nuorin ikäluokka. Mallissa on mukana myös puoluekanta.

<b>kerroin (neg vs pos)</b>	estimaatti	keskivirhe	$\exp(\text{estimaatti})$
mies	0.753	0.141	2.122
61jaYli	0.991	0.259	2.693
41-60v	0.765	0.224	2.148
26-40v	0.680	0.235	1.975
<b>kerroin (neg vs eos)</b>	estimaatti	keskivirhe	$\exp(\text{estimaatti})$
mies	-0.584	0.109	0.557
61jaYli	-0.163	0.213	0.849
41-60v	0.161	0.155	1.175
26-40v	0.256	0.161	1.292

### 5.3 Puoluekanta

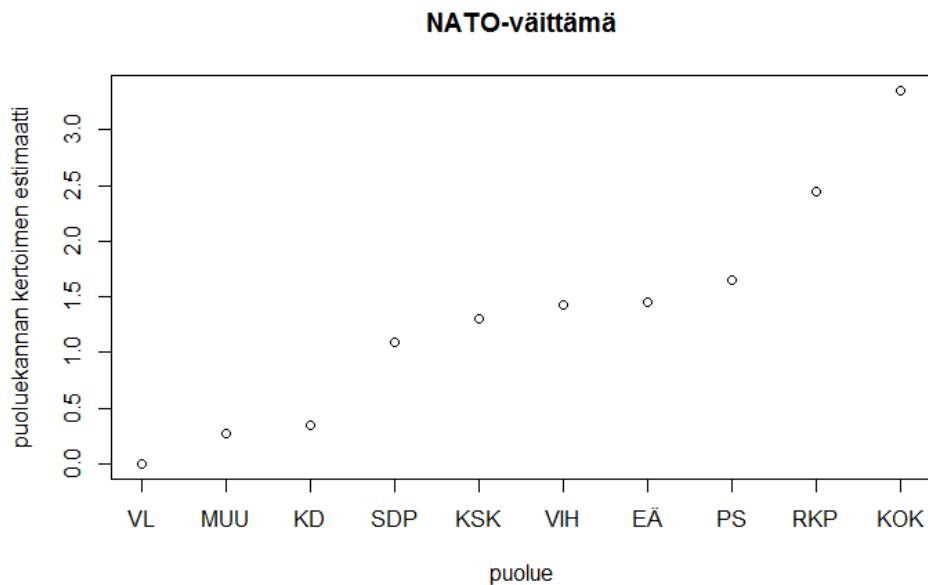
Puoluekantaan liittyvä kerroin (taulukko 19) on pienin vasemmistoliiton kannattajilla positiivisen ja negatiivisen suhtautumisen yhtälössä, joten vasemmistoliitto on valittu referenssitasoksi. Kun huomioidaan luottamusvälit, niin myös SDP:n kannattajat ovat suhtautuneet keskimäärin negatiivisemmin NATOon liittymiseen kuin positiivisempien puolueiden äänestäjät. Positiivisimmin NATOon liittymiseen ovat suhtautuneet kokoomusta ja RKP:tä kannattaneet.

- Positiivisesta kertoimesta (3.352) voidaan päätellä, että kokoomuksen kannattajat ovat suhtautuneet NATOon liittymiseen keskimäärin positiivisemmin kuin vasemmistoliiton kannattajat. Positiivisen mielipiteen todennäköisyys suhteessa negatiiviseen mielipiteeseen on keskimäärin  $\exp(3.352) = 28.560$ -kertainen kokoomuksen ryhmässä verrattuna vasemmistoliittoa äänestäneisiin.

Taulukko 19: Nominaalisen logistisen regression kertoimien estimaatit puoluekannoille NATO-väittämään. Perustasona on vasemmistoliitto. Mallissa on mukana myös sukupuolen ja ikäluokan päävaikutukset.

<b>kerroin (neg vs pos)</b>	estimaatti	keskivirhe	exp(estimaatti)
keskusta	1.299	0.426	3.666
kokoomus	3.352	0.426	28.560
kristillisdemokraatit	0.341	0.841	1.406
perussuomalaiset	1.647	0.424	5.191
RKP	2.442	0.587	11.496
SDP	1.093	0.446	2.983
vihreät	1.422	0.448	4.145
ei äänestäneet	1.453	0.416	4.276
muut puolueet	0.265	0.843	1.303
<b>kerroin (neg vs eos)</b>	estimaatti	keskivirhe	exp(estimaatti)
keskusta	0.832	0.292	2.298
kokoomus	2.349	0.305	10.475
kristillisdemokraatit	1.194	0.433	3.300
perussuomalaiset	1.167	0.294	3.212
RKP	1.443	0.525	4.233
SDP	0.718	0.308	2.050
vihreät	1.087	0.292	2.965
ei äänestäneet	1.726	0.268	5.618
muut puolueet	1.225	0.467	3.404

Kuvassa 3 puolueet ovat järjestyksessä NATO-väittämän kertoimien suhteen. Puolueet on lajiteltu kertoimien samankaltaisuuksien mukaan neljään luokkaan. Vasemmistoliiton, muiden puolueiden ja kristillisdemokraattien kannattajat kuuluvat kielteisimpään luokkaan. Myönteisimpiä ovat olleet RKP:n ja Kokoomuksen äänestäjät. Perussuomalaiset on oma luokkansa, koska heidän kannattajat ovat olleet todennäköisemmin myönteisiä, mutta eivät niin myönteisiä kuin RKP:n ja kokoomuksen kannattajat. Kielteisimpien ja myönteisimpien suhtautumisien väliin jäävät SDP:n, keskustan ja vihreiden äänestäjät sekä ei-äänestäneet.



Kuva 3: NATO-väittäjä ja puolueet järjestettynä kertoimien mukaan pienimmästä suurimpaan

Alla on taulukoitu luottamusvälit huomioon ottaen, minkä puolueiden kannattajat ovat suhtautuneet NATOon liittymiseen todennäköisemmin negatiivisesti kuin positiivisesti ja vastaavasti todennäköisemmin positiivisesti kuin negatiivisesti suhteessa toisten puolueiden kannattajiin, kun vertaillaan saman sukupuolen ja ikäluokan edustajia. Taulukosta 20 nähdään, että vasemmistoliiton kannattajat ovat olleet todennäköisemmin negatiivisia kuin positiivisia liittymään Natoon, kun verrataan ei-äänestäneisiin sekä keskustaa, kokoomusta, perussuomalaisia, RKP:tä, SDP:tä ja vihreitä äänestäneisiin. Toisena negatiivisempaa ryhmänä on ollut SDP:n kannattajat. SDP:n kannattajat ovat olleet todennäköisemmin negatiivisempia kuin positiivisia verrattuna kokoomusta, perus-

suomalaisia tai RKP:tä kannattaneisiin. Kokoomuksen ja RKP:n kannattajat ovat olleet selkeästi positiivisimpia NATO-väittämän kohdalla kuin muiden puolueiden kannattajat. Kokoomuksen äänestäjät ovat olleet todennäköisemmin positiivisia kuin negatiivisia muiden puolueiden äänestäjiin paitsi RKP:tä äänestäneisiin verrattuna. RKP:n kannattajat ovat sen sijaan olleet todennäköisemmin positiivisia kuin negatiivisia kaikkien muiden paitsi kokoomuksen ja perussuomalaisten kannattajiin verrattuna.

Taulukko 20: NATO-väittäjä ja taulukoituna puolueet sekä minkä puolueiden kannattajiin nähden kyseisen puolueen kannattajat suhtautuvat negatiivisesti ja mihin positiivisesti, kun ikä ja sukupuoli vakioitu. Vertailut perustuvat parittaisiin luottamusväleihin. KSK = keskusta, KOK = kokoomus, KD=kristillisdemokraatit, PS = perussuomalaiset, VL = vasemmistoliitto, VIHR = vihreät, MUU = jokin muu puolue, EÄ = ei äänestäneet.

Puolue	Negatiivisempi kuin	Positiivisempi kuin
KSK	KOK, RKP	VL
KOK	-	KSK, KD, PS, SDP, VL, VIHR, MUU, EÄ
KD	RKP, KOK	-
PS	KOK	SDP, VL
RKP	-	KSK, KD, SDP, VL, VIHR, MUU, EÄ
SDP	KOK, PS, RKP	VL
VL	KSK, KOK, PS, RKP, SDP, VIHR, EÄ	-
VIHR	KOK, RKP	VL
MUU	KOK, RKP	-
EÄ	KOK, RKP	VL

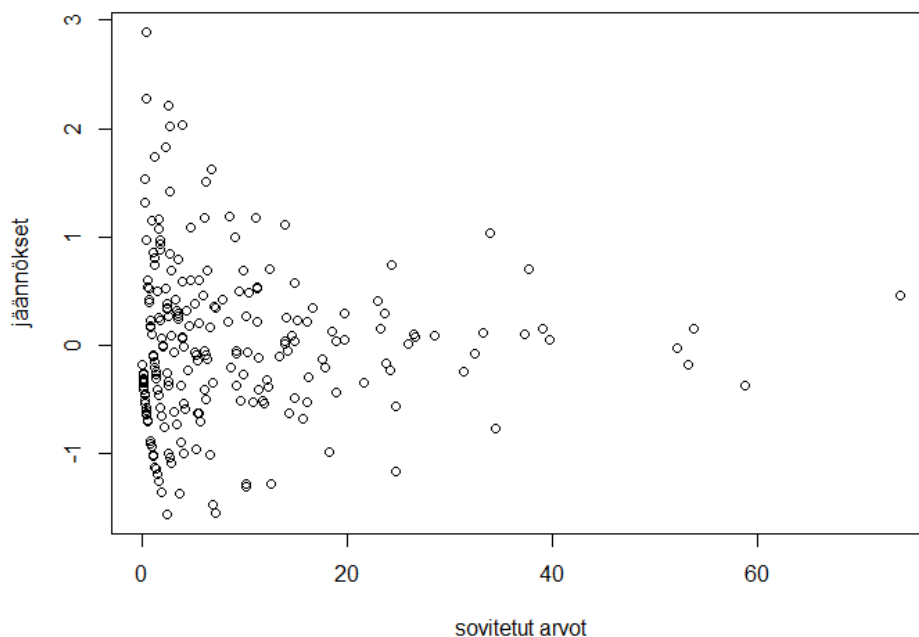
Kun tarkastellaan vaikea sanoa -mielipidettä ja negatiivista suhtautumista, niin havaitaan, että kokoomuksen kannattajat ovat vastanneet todennäköisemmin “vaikea sanoa” kuin että olisivat vastanneet “eri mieltä” verrattuna vasemmistoliiton kannattajiin. Tämä johtunee siitä, että kokoomuksen kannattajat ovat olleet niin paljon selkeämmin NATOon liittymisen puolella kuin vasemmistoliiton kannattajat, että kokoomusta äänestäneet ovat vastanneet mieluummin “vaikea sanoa”, kuin että olisivat olleet eri mieltä verrattuna vasemmistoliittoa äänestäneisiin. Tästä syystä niitä, jotka eivät ole osanneet kertoa kantaansa NATO-väittämään, verrataan heihin, joilla on jokin näkemys (puolesta tai vastaan) logistisella regressiolla.

Liitteestä 2 voidaan havaita, että ei-äänestäneet ovat vastanneet tässäkin väittämässä todennäköisemmin “vaikea sanoa” kuin että olisivat olleet jotain mieltä verrattuna monen puolueen äänestäjiin. Ne, jotka ovat vastanneet NATO-väittämään “vaikea sanoa”, ovat



logistisessa regressiossa yksi ryhmä ja kaikki muut kuuluvat toiseen ryhmään. Prediktorin puolue perustasona on jälleen ei-äänestäneet, koska kyseinen ryhmä on mielenkiinnon kohteena. Kun vertaillaan samaa sukupuolta ja samaa ikäluokkaa edustaneita, niin ei-äänestäneet ovat vastanneet todennäköisemmin “vaikea sanoa” kuin että olisivat olleet jotain mieltä liittymisestä NATOon, kun verrataan keskustaa, perussuomalaisia, SDP:tä, vasemmistoliittoa ja vihreitä äänestäneisiin. Ero kokoomuksen kannattajiin ei ole merkitsevä ja kokoomuksen kannattajat ovat olleet toinen ryhmä, joka on vastannut muita todennäköisemmin “vaikea sanoa” kuin että on ollut jotain mieltä.

Kuvasta 4 huomataan, että jäännöskuvio on melkein samanlainen kuin Venäjä-väittämän kohdalla. Jäännökset jakautuvat suunnilleen tasan nollan molemmille puolille. Alkupään suurempi vaihtelu nollan molemmiin puolin verrattuna loppupään vaihteluun selittyy osin sillä, että sovitettuja ja havaittuja arvoja on paljon enemmän nollan lähetyillä kuin suuremmilla sovitetuilla arvoilla. Loppupään vaihtelu pienenee alun jälkeen kuitenkin niin paljon, että jotain parannuksia olisi varmasti tehtävissä.



Kuva 4: Pearsonin jäännökset sovitetulle nominaalivasteisen regression mallille NATO-väittämään

## 5.4 Puoluekannan tulosten vertailua muihin tuloksiin ja näkemyksiin

Internetistä on etsitty puolueiden näkemyksiä tähänkin väittämään liittyen. Etsittyjen artikkeleiden ja haastatteluiden mukaan varsinkin kokoomus on ollut ennen ja jälkeen tämän kyselyn halukkain liittämään Suomea NATOon. Myös RKP:n kansanedustajaehdokkaat ovat olleet myönteisiä NATOon liittymisen suhteen. Muut puolueet, isoimmista puolueista varsinkin vasemmistoliitto, ovat olleet kielteisimpiä liittymään NATOon. Löydetyt tulokset ja näkemykset ovat varsin samankaltaisia tämän tutkimuksen tuloksien kanssa.

Vuoden 2015 eduskuntavaalien Helsingin Sanomien vaalikoneen mukaan noin 91 prosenttia kokoomuksen ja 64 prosenttia RKP:n kansanedustajaehdokkaista on ollut myönteisiä NATO-jäsenyyteen liittyen. Muiden puolueiden kansanedustajaehdokkaista enemmistö on ollut NATO-jäsenyyttä vastaan. Vasemmistoliiton ehdokkaat ovat olleet noin 97 prosenttisesti kielteisiä. (Eloranta, 2015.) Tässä tutkimuksessa on tullut esille kokoomuksen ja RKP:n kannattajien todennäköisempi halukkuus liittää Suomi NATOon vuoden 2014 alussa verrattuna muiden puolueiden kannattajiin, joten HS:n vaalikoneen ja tämän EVAn kyselyn mukaan kokoomuksen ja RKP:n kannattajat ja kansanedustajaehdokkaat ovat olleet samoilla linjoilla suhteessa muihin puolueisiin. Samanlainen yhtäläisyys pätee vasemmistoliittoon, mutta vasemmistoliitto kuuluu kielteiseen kantaan. Ainoa hieman ristiriitainen tulos on kristillisdemokraattien kannattajien ja kansanedustajaehdokkaiden välillä. Kristillisdemokraattien kannattajien kerroin on tässä tutkimuksessa pienimmästä päästä positiivisen ja negatiivisen suhtautumuksen yhtälössä, mutta HS:n vaalikoneen mukaan kansanedustajaehdokkaat ovat olleet myönteisimmästä päästä ennen vuoden 2015 eduskuntavaaleja. EVAn kyselyn tuloksiin vaikuttaa tosin paljon se, ettei kristillisdemokraattien kannattajia ollut kovinkaan paljon kyselyssä.

Ulkopoliittisen instituutin sivuilta löydetyn artikkelin mukaan kokoomuksen Natotuska pääsi irti eduskunnassa vuoden 2013 alussa (Viljanmaa, 2013). Artikkelin mukaan kokoomuksen kansanedustajat ovat olleet turhautuneita SDP:n, vasemmiston ja vihreiden kielteiseen NATO-linjaan. Tämäkin näkemys on yhdenmukainen tässä tutkimuksessa saadun tuloksen kanssa, että kokoomuksen kannattajat ovat olleet todennäköisemmin myönteisiä kuin kielteisiä liittämään Suomea NATOon vuoden 2014 alussa verrattuna muun muassa SDP:n, vihreiden ja vasemmistoliiton kannattajiin.

## 6 Tuloksia USA-väittämään

Väittäjä Yhdysvalloista on kolmas tutkittava väittäjä. Yhdysvallat on NATO:n vahvin maa, joten NATO-väittämän jälkeen on kiinnostava tutkia, miten paljon yhtäläisyyksiä ja eroavaisuuksia tuloksissa on näiden väittäjien osalta. USA-väittäjä on kysytty seu-

raavasti: “Yhdysvallat toimii maailmanpolitiikassa oikein ja ansaitsee myös suomalaisten tuen”.

## 6.1 Mallinvalinta

USA-kysymys on ainoa, jossa riittävimpiin malliin ei tulisi samat prediktorit kuin kolmen muun kysymyksen kohdalla, jos kaikki kuusi alkuperäistä prediktorista olisivat mukana mallinvalinnassa. Taulukosta 21 havaitaan, että riittävimmissä mallissa olisi sukupuolen ja puolueen päävaikutusten lisäksi ammattiryhmän päävaikutus, jos kaikki kuusi alkuperäistä prediktorista otettaisiin mukaan tarkasteluihin. Sukupuolen ja varsinkaan puoluekannan kertoimet eivät kuitenkaan ole erilaisia, vaikka ikäluokka otetaan ammattiryhmän tilalle. Suurin ero on siinä, että ikäluokkien suhteen ei ole eroja, kun taas ammattiryhmissä opiskelijat ja työttömät sekä muut ovat olleet todennäköisemmin kielteisempiä kuin työntekijät, toimihenkilöt ja eläkeläiset. Varsinkin opiskelijat ovat olleet kielteisempiä kuin eläkeläiset, joten epäsuorasti voidaan päätellä, että nuoremmat ovat olleet kielteisempiä kuin vanhemmat myös tämän väittämän kohdalla, vaikka ikäluokkien kohdalla ero ei olekaan merkitsevä.

Nyt ollaan kuitenkin kiinnostuneita vain prediktoreista sukupuoli, ikäluokka ja varsinkin puolue. Kun verrataan sukupuolen, ikäluokan ja puoluekannan täyttä mallia päävaikutusten malliin UO-testillä,  $p$ -arvoksi saadaan noin 0.184, joten päävaikutusten malli on riittävä. Lisäksi päävaikutusten mallin AIC on pienempi kuin täyden mallin, ja taaskaan kaikkia parametreja ei saada estimoitua joidenkin interaktioiden tapauksissa, joten tarkasteltavaksi valitaan päävaikutusten malli.

Taulukko 21: Mallien residuaalidevianssit, estimoitujen parametrien lukumäärät (df), AIC:t ja UO-testien  $p$ -arvoja USA-väittämään. Mallissa 3 kaksoispiste tarkoittaa interaktiota, ts. malli 3 on sukupuolen, ikäluokan ja puolueen täysimalli.

	malli	res. dev	df	AIC	UO-testien $p$ -arvoja
1	kaikki parittaiset interaktiot	2922	366	3654	-
2	kaikki päävaikutukset	3301	44	3389	(1 vs 2) 0.016
3	sukupuoli:ikäluokka:puolue	3183	158	3499	(1 vs 3) 0.008
4	sukupuoli+ikäluokka+puolue	3327	28	3383	(1 vs 4) 0.007
					(2 vs 4) 0.047
					(3 vs 4) 0.184
5	sukup.+ikäluo.+puolue+am.ryh	3309	36	3381	(1 vs 5) 0.017
					(2 vs 5) 0.375
					(4 vs 5) 0.022
6	sukupuoli+puolue+am.ryh	3316	30	3376	(1 vs 6) 0.017
					(2 vs 6) 0.383
					(5 vs 6) 0.391

## 6.2 Sukupuoli ja ikäluokka

Taulukossa 22 ovat kertoimet miehille ja vanhemmille ikäluokille USA-väittämän osalta. Ikäluokkien välillä ei ole merkitseviä eroja kummankaan yhtälön kohdalla. Miesten ja naisten kesken sen sijaan on eroja.

Jos henkilöt ovat äänestäneet samalla tavalla ja edustaneet samaa ikäluokkaa, niin naiset ovat vastanneet todennäköisemmin “vaikea sanoa” kuin että olisivat vastanneet “eri mieltä” verrattuna miehiin. Naiset ovat olleet taas tässäkin väittämässä todennäköisemmin negatiivisempia kuin positiivempia verrattuna miehiin, mikä selviää negatiivisen ja positiivisen suhtautumisen yhtälöstä.

- Positiivisesta kertoimesta (0.493) voidaan päätellä, että miehet ovat suhtautuneet Yhdysvaltain maailmanpolitiikkaan positiivisemmin kuin naiset. Positiivisen mielipiteen todennäköisyys suhteessa negatiiviseen mielipiteeseen on keskimäärin  $\exp(0.493) = 1.638$ -kertainen miesten ryhmässä verrattuna naisiin, kun vertaillaan naisia ja miehiä, jotka ovat olleet samasta ikäluokasta ja olisivat äänestäneet samalla tavalla.

Taulukko 22: Nominaalisen logistisen regression kertoimien estimaatit miehille ja vanhemmille ikäluokille USA-väittämään. Perustasoina ovat naiset ja nuorin ikäluokka. Mallissa on mukana myös puoluekanta.

<b>kerroin (neg vs pos)</b>	estimaatti	keskivirhe	exp(estimaatti)
mies	0.493	0.188	1.638
61jaYli	0.462	0.323	1.588
41-60v	0.201	0.277	1.222
26-40v	0.222	0.294	1.248
<b>kerroin (neg vs eos)</b>	estimaatti	keskivirhe	exp(estimaatti)
mies	-0.695	0.109	0.499
61jaYli	0.277	0.195	1.319
41-60v	-0.033	0.155	0.968
26-40v	0.227	0.160	1.254

### 6.3 Puoluekanta

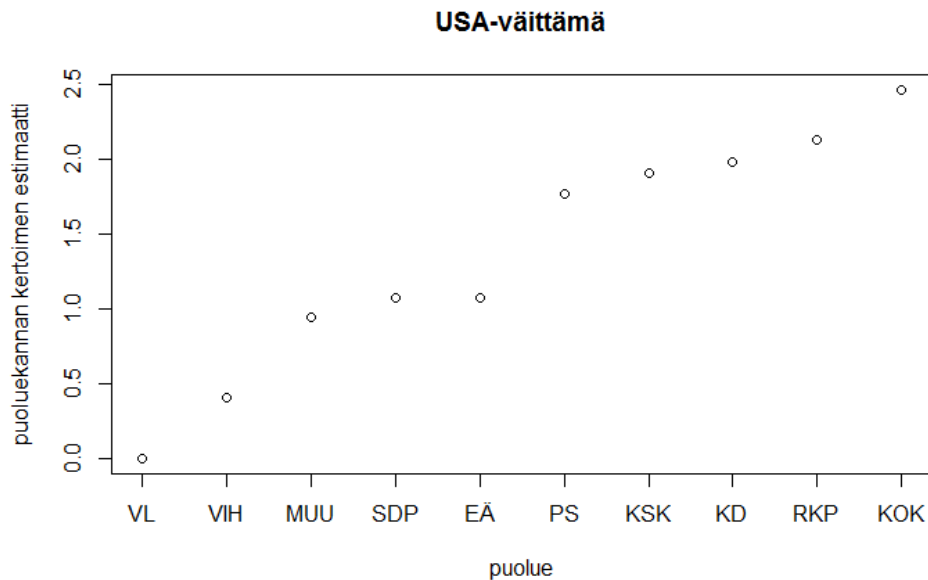
Puoluekantaan liittyvät kertoimet (taulukko 23) ovat pienimpiä vasemmistoliiton (referenssitaso) ja vihreiden kannattajilla positiivisen ja negatiivisen suhtautumisen yhtälössä, joten näiden puolueiden kannattajat ovat suhtautuneet keskimäärin negatiivisemmin Yhdysvaltain maailmanpolitiikkaan kuin USA-väittämän positiivisimpien puolueiden kannattajat. Positiivisimmin Yhdysvaltain maailmanpolitiikkaan ovat suhtautuneet kokoomuksen äänestäjät.

- Positiivisesta kertoimesta (2.465) voidaan päätellä, että kokoomuksen kannattajat ovat suhtautuneet Yhdysvaltain maailmanpolitiikkaan keskimäärin positiivisemmin kuin vasemmistoliiton kannattajat. Positiivisen mielipiteen todennäköisyys suhteessa negatiiviseen mielipiteeseen on keskimäärin  $\exp(2.465) = 11.759$ -kertainen kokoomuksen kannattajien ryhmässä verrattuna vasemmistoliiton kannattajiin.

Taulukko 23: Nominaalisen logistisen regression kertoimien estimaatit puoluekannoille USA-väittämään. Perustasona on vasemmistoliitto. Mallissa on mukana myös sukupuolen ja ikäluokan päävaikutukset.

<b>kerroin (neg vs pos)</b>	estimaatti	keskivirhe	exp(estimaatti)
keskusta	1.909	0.617	6.746
kokoomus	2.465	0.613	11.763
kristillisdemokraatit	1.983	0.802	7.265
perussuomalaiset	1.769	0.619	5.865
RKP	2.131	0.812	8.423
SDP	1.073	0.661	2.924
vihreät	0.408	0.720	1.504
ei äänestäneet	1.075	0.621	2.930
muut puolueet	0.944	0.942	2.570
<b>kerroin (neg vs eos)</b>	estimaatti	keskivirhe	exp(estimaatti)
keskusta	1.443	0.295	4.233
kokoomus	1.799	0.298	6.044
kristillisdemokraatit	1.320	0.451	3.743
perussuomalaiset	1.149	0.303	3.155
RKP	1.597	0.486	4.938
SDP	0.912	0.313	2.489
vihreät	0.743	0.305	2.102
ei äänestäneet	1.474	0.278	4.367
muut puolueet	0.254	0.602	1.289

USA-väittämän osalta puolueet on laitettu järjestykseen kuvaan 5. Puolueet on lajiteltu kertoimien samankaltaisuuksien mukaan neljään luokkaan. Vasemmistoliiton ja vihreiden kannattajat kuuluvat kielteisimpään luokkaan. Myönteisempiä ovat olleet perussuomalaisten, keskustan, kristillisdemokraattien ja RKP:n äänestäjät. Kokoomus on oma luokkansa, koska heidän kannattajat ovat olleet selkeästi todennäköisimmin myönteisiä. Kielteisimpien ja myönteisimpien suhtautumisien välissä ovat SDP:n ja muiden puolueiden kannattajat sekä ei-äänestäneet.



Kuva 5: USA-väittäjä ja puolueet järjestettynä kertoimien mukaan pienimmästä suurimpaan

Alla on taulukoitu, minkä puolueiden kannattajat ovat suhtautuneet Yhdysvaltain maailmanpolitiikkaan todennäköisemmin negatiivisesti kuin positiivisesti ja vastaavasti todennäköisemmin positiivisesti kuin negatiivisesti suhteessa toisten puolueiden kannattajiin. Taulukosta 24 nähdään, että vasemmistoliiton ja vihreiden kannattajat ovat olleet keskimäärin todennäköisemmin negatiivisia kuin positiivisia Yhdysvaltain maailmanpolitiikkaa kohtaan, kun verrataan keskustaa, kokoomusta, kristillisdemokraatteja, perussuomalaisia ja RKP:tä kannattaneisiin. Kokoomuksen, keskustan ja perussuomalaisien kannattajat ovat olleet positiivisimpia USA-väittämän kohdalla kuin monen muun puolueen kannattajat. Kokoomusta äänestäneet ovat olleet todennäköisemmin positiivisia kuin negatiivisia kaikkiin muihin äänestäjäryhmiin paitsi RKP:n ja kristillisdemokraattien kannattajiin verrattuna. Se, että ero ei ole merkitsevä RKP:läisiin ja kristillisdemokraatteihin nähden, voi johtua taas siitä, että näiden kahden puolueen kannattajia on ollut vähän vastaamassa kyselyyn. Keskustan kannattajat ovat olleet todennäköisemmin positiivisia kuin negatiivisia verrattuna ei-äänestäneisiin sekä SDP:n, vasemmistoliiton ja vihreiden kannattajiin.

Taulukko 24: USA-väittämä ja taulukoituna puolueet sekä minkä puolueiden kannattajiin nähden kyseisen puolueen kannattajat suhtautuvat negatiivisesti ja mihin positiivisesti, kun ikä ja sukupuoli vakioitu. Vertailut perustuvat parittaisiin luottamusväleihin. KSK = keskusta, KOK = kokoomus, KD=kristillisdemokraatit, PS = perussuomalaiset, VL = vasemmistoliitto, VIHR = vihreät, MUU = jokin muu puolue, EÄ = ei äänestäneet.

Puolue	Negatiivisempi kuin	Positiivisempi kuin
KSK	KOK	SDP, VL, VIHR, EÄ
KOK	-	KSK, PS, SDP, VL, VIHR, MUU, EÄ
KD	-	VL, VIHR
PS	KOK	VL, VIHR, EÄ
RKP	-	VL, VIHR
SDP	KSK, KOK	-
VL	KSK, KOK, KD, PS, RKP	-
VIHR	KSK, KOK, KD, PS, RKP	-
MUU	KOK	-
EÄ	KSK, KOK, PS	-

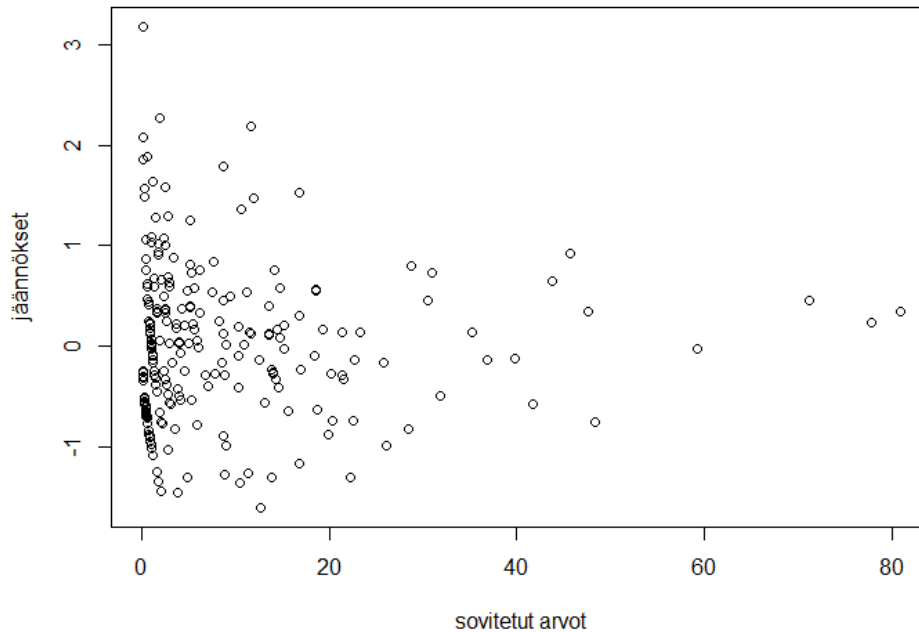
Kun tarkastellaan vaikea sanoa -mielipidettä ja negatiivista suhtautumista, niin havaitaan samankaltainen tilanne NATO-väittämän kanssa. Kokoomusta kannattaneet ovat vastanneet todennäköisemmin “vaikea sanoa” kuin että olisivat vastanneet “eri mieltä” USA-väittämään verrattuna vasemmistoliiton kannattajiin. Suoritetaan taas logistinen regressio, kun vasteen toisena luokkana on “vaikea sanovien -ryhmä” ja toisena ryhmänä kaikki, jotka ovat olleet jotakin mieltä Yhdysvaltain maailmanpolitiikasta. Prediktorin puolue perustasona on jälleen ei-äänestäneet, koska kyseinen ryhmä on mielenkiinnon kohteena.

Liitteestä 3 voidaan havaita, että ei-äänestäneet ovat vastanneet jälleen todennäköisemmin “vaikea sanoa” kuin että olisivat olleet jotain mieltä verrattuna monen puolueen äänestäjiin. Ei-äänestäneiden kanssa samankaltaiset kertoimet luottamusväleinen on kokoomuksen kannattajien ryhmällä. Kun vertaillaan samaa sukupuolta ja samaa ikäluokkaa edustaneita, niin ei-äänestäneet ja kokoomuslaiset ovat vastanneet todennäköisemmin “vaikea sanoa” kuin että olisivat olleet jotain mieltä Yhdysvaltain maailmanpolitiikasta, kun vertaillaan perussuomalaisia, SDP:tä, vasemmistoliittoa, vihreitä ja jotain muuta puoluetta äänestäneisiin.

Kuvan 6 jäännöskuvio on melkein samanlainen kuin kahden edellisen väittämän kohdalla. Kuvio ei osoita, että mallia tarvitsisi vaihtaa tai suuremmin muokata. Jäännökset jakautuvat jälleen suunnilleen tasan nollan molemmille puolille. Lisäksi alkupään suurempi vaihtelu nollan molemmin puolin verrattuna loppupään vaihteluun selittyy sillä, että



sovitettuja ja havaittuja arvoja on paljon enemmän nollan lähetyvillä kuin esimerkiksi arvosta 20 eteenpäin.



Kuva 6: Pearsonin jäännökset sovitetulle nominaalivasteisen regression mallille USA-väittämään

## 7 Tuloksia EU-väittämään

Tasaisin väliajoin Suomessa nousee keskustelua EU:sta ja sen toimivuudesta. Lisäksi usein mietitään, saako Suomi ääntänsä kuuluviin päätöksenteossa ja ovatko päätökset hyviä Suomen kannalta. Viime aikoina EU:ta koskevat aiheet ovat olleet erityisen kiinnostavia johtuen Iso-Britannian erosta.

Viimeiseksi tarkasteltavaksi väittämäksi on valikoitunut väittämä EU:sta, koska on mielenkiintoista tutkia, millaisia ovat olleet suomalaisten asenteet tätä väittämää kohtaan vuoden 2014 alussa. EU-väittämä on muotoiltu: “Suomi on saanut äänensä hyvin kuuluviin EU:n päätöksenteossa”.

## 7.1 Mallinvalinta

Taulukosta 25 huomataan, että mallia valittaessa päädyttäisiin samoihin additiivisiin prediktoreihin myös EU-väittämässä kuin Venäjä- ja NATO-väittämissä, vaikka otettaisiin huomioon kaikki kuusi prediktorin ja riittävin malli valittaisiin niiden muodostamien mallien joukosta.

Kun verrataan sukupuolen, ikäluokan ja puoluekannan täyttä mallia päävaikutusten malliin uskottavuusosamäärä-testillä,  $p$ -arvoksi saadaan noin 0.062, joten päävaikutusten mallia voidaan pitää riittävänä, jos pitäydytään tarkasti rajassa 0.05. Lisäksi täyden mallin AIC on noin 4117 ja päävaikutusten mallin noin 4013, ja jälleen kaikkia parametreja ei saada estimoitua joidenkin interaktioiden tapauksissa, joten taas päädytään tarkastelemaan päävaikutusten mallia.

Taulukko 25: Mallien residuaalidevianssit, estimoitujen parametrien lukumäärät (df), AIC:t ja UO-testien  $p$ -arvoja EU-väittämään. Mallissa 3 kaksoispiste tarkoittaa interaktiota, ts. malli 3 on sukupuolen, ikäluokan ja puolueen täysimalli.

	malli	res. dev	df	AIC	UO-testien $p$ -arvoja
1	kaikki parittaiset interaktiot	3614	366	4346	-
2	kaikki päävaikutukset	3941	44	4029	(1 vs 2) 0.414
3	sukupuoli:ikäluokka:puolue	3801	158	4117	(1 vs 3) 0.852
4	sukupuoli+ikäluokka+puolue	3957	28	4013	(1 vs 4) 0.422
					(2 vs 4) 0.485
					(3 vs 4) 0.062

## 7.2 Sukupuoli ja ikäluokka

Tulkinnat muodostetaan tälläkin kertaa siten, että vertailtavat ryhmät poikkeavat vain tarkasteltavan prediktorin suhteen ja muutoin ovat samanlaisia. Taulukossa 26 ovat kerroimet miehille ja vanhemmille ikäluokille EU-väittämän osalta. Ikäluokkien välillä nuorin ikäluokka on eronnut hieman muista.

- Kun verrataan nuorinta ikäluokkaa vanhimpaan ikäluokkaan, niin asennoituminen Suomen äänen kuulumiseen EU:ssa on todennäköisemmin positiivisempaa vanhimmassa ikäluokassa ja positiivisen mielipiteen todennäköisyys suhteessa negatiiviseen mielipiteeseen on keskimäärin  $\exp(0.978) = 2.659$ -kertainen vanhimmassa ikäluokassa verrattuna nuorimpaan ikäluokkaan.
- Kun verrataan nuorinta ja toiseksi vanhinta (41-60-vuotiaat) ikäluokkaa, niin nuorin ikäluokka on vastannut todennäköisemmin “vaikea sanoa” kuin että olisi vastannut “eri mieltä” verrattuna 41-60-vuotiaisiin.

Sukupuolten eroja vertailtaessa naiset ovat vastanneet todennäköisemmin “vaikea sanoa” kuin että olisivat vastanneet “eri mieltä” verrattuna miehiin. Tämän väittämän kohdalla naisten ei voida katsoa olleen merkitsevästi todennäköisemmin negatiivisempia kuin positiivisempia verrattuna miehiin.

Taulukko 26: Nominaalisen logistisen regression kertoimien estimaatit miehille ja vanhemmille ikäluokille EU-väittämään. Perustasoina ovat naiset ja nuorin ikäluokka. Mallissa on mukana myös puoluekanta.

<b>kerroin (neg vs pos)</b>	estimaatti	keskivirhe	exp(estimaatti)
mies	0.110	0.133	1.116
61jaYli	0.978	0.251	2.659
41-60v	0.370	0.216	1.447
26-40v	0.098	0.229	1.103
<b>kerroin (neg vs eos)</b>	estimaatti	keskivirhe	exp(estimaatti)
mies	-0.217	0.104	0.805
61jaYli	-0.064	0.198	0.938
41-60v	-0.350	0.149	0.705
26-40v	-0.243	0.153	0.784

### 7.3 Puoluekanta

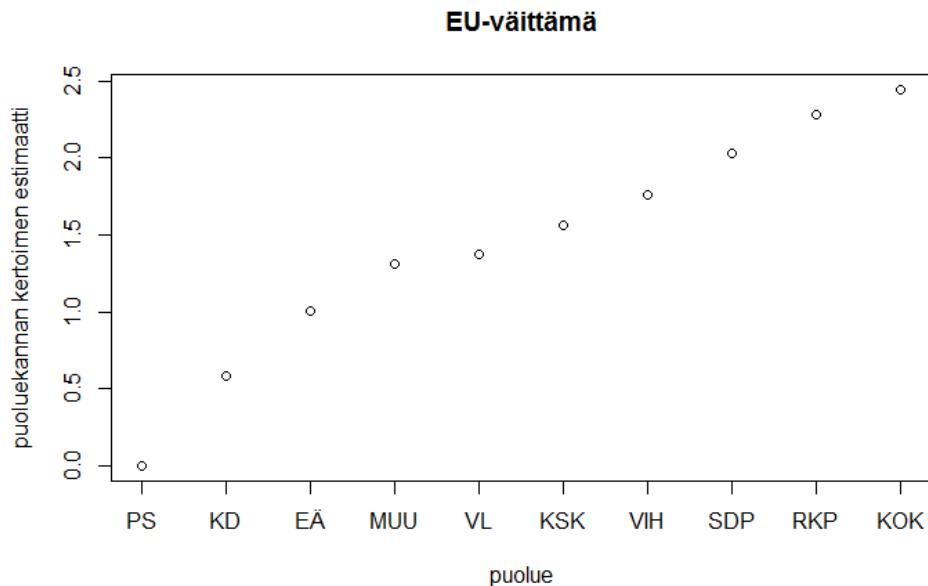
Tämän väittämän kohdalla perussuomalaiset on valittu prediktorin puolue referenssitaksiksi. Kertoimet ovat taulukossa 27. Puoluekantaan liittyvä kerroin on nolla perussuomalaisien kannattajilla positiivisen ja negatiivisen suhtautumisen yhtälössä, kun muilla se on suurempaa, joten perussuomalaiset ovat suhtautuneet negatiivisemmin Suomen äänen kuulumiseen EU:ssa kuin myönteisimpien puolueiden äänestäjät. Kun huomioidaan luottamusvälit, niin positiivisimmin ovat suhtautuneet kokoomuksen ja SDP:n kannattajat. Tämän kyselyn osalta myös RKP:llä on iso kerroin, joten RKP:n kannattajat ovat suhtautuneet todennäköisemmin myönteisesti kuin kielteisesti varsinkin verrattuna perussuomalaisien kannattajiin.

- Positiivisesta kertoimesta (2.444) voidaan päätellä, että kokoomuksen kannattajat ovat suhtautuneet Suomen äänen kuulumiseen EU:ssa keskimäärin positiivisemmin kuin perussuomalaisien kannattajat. Positiivisen mielipiteen todennäköisyys suhteessa negatiiviseen mielipiteeseen on keskimäärin  $\exp(2.444) = 11.524$ -kertainen kokoomuksen kannattajien ryhmässä verrattuna perussuomalaisien kannattajiin.

Taulukko 27: Nominaalisen logistisen regression kertoimien estimaatit puoluekannoille EU-väittämään. Perustasona on perussuomalaiset. Mallissa on mukana myös sukupuolen ja ikäluokan päävaikutukset.

<b>kerroin (neg vs pos)</b>	estimaatti	keskivirhe	exp(estimaatti)
keskusta	1.560	0.297	4.759
kokoomus	2.444	0.295	11.519
kristillisdemokraatit	0.579	0.669	1.784
RKP	2.282	0.495	9.796
SDP	2.030	0.309	7.614
vasemmistoliitto	1.370	0.373	3.935
vihreät	1.764	0.324	5.836
ei äänestäneet	1.009	0.295	2.743
muut puolueet	1.310	0.566	3.706
<b>kerroin (neg vs eos)</b>	estimaatti	keskivirhe	exp(estimaatti)
keskusta	0.699	0.207	2.012
kokoomus	1.276	0.214	3.582
kristillisdemokraatit	0.644	0.393	1.904
RKP	0.681	0.494	1.976
SDP	1.007	0.228	2.737
vasemmistoliitto	1.148	0.245	3.152
vihreät	1.133	0.217	3.105
ei äänestäneet	1.162	0.180	3.196
muut puolueet	0.455	0.451	1.576

Kuvaan 7 on laitettu puolueet järjestykseen EU-väittämän kertoimien osalta taulukkoa 27 apuna käyttäen. Puolueet on lajiteltu taas kertoimien samankaltaisuuksien mukaan neljään luokkaan. Perussuomalaisten ja kristillisdemokraattien kannattajat kuuluvat kielteisimpään luokkaan. Ei-äänestäneet sekä muita puolueita ja vasemmistoliittoa äänestäneet ovat lievästi kielteisessä luokassa. Hieman myönteisemmässä luokassa ovat keskustan ja vihreiden kannattajat. Myönteisimpiä ovat olleet SDP:n, RKP:n ja kokoomuksen kannattajat.



Kuva 7: EU-väittämä ja puolueet järjestettynä kertoimien mukaan pienimmästä suurimpaan

Alla on taulukoitu myös EU-väittämän osalta, minkä puolueiden kannattajat ovat suhtautuneet todennäköisemmin negatiivisesti kuin positiivisesti ja vastaavasti todennäköisemmin positiivisesti kuin negatiivisesti suhteessa toisten puolueiden kannattajiin, kun vertaillaan saman sukupuolen ja ikäluokan edustajia. Taulukosta 28 nähdään, että perussuomalaisten kannattajat ovat olleet todennäköisemmin negatiivisia kuin positiivisia Suomen vaikutusvaltaa kohtaan EU:ssa, kun verrataan ei-äänestäneisiin sekä keskustaa, kokoomusta, RKP:tä, SDP:tä, vasemmistoliittoa, vihreitä ja muita puolueita äänestäneisiin. Toisena negatiivisempänä ryhmänä on ollut ei-äänestäneiden ryhmä. Kokoomuksen kannattajat ovat olleet todennäköisemmin positiivisia kuin negatiivisia kaikkien muiden puolueiden paitsi RKP:n ja SDP:n kannattajiin verrattuna. Myös SDP:tä äänestäneet ovat olleet todennäköisemmin positiivisia kuin negatiivisia verrattuna monien kielteisempien puolueiden kannattajiin.

Taulukko 28: EU-väittämä ja taulukoituna puolueet sekä minkä puolueiden kannattajiin nähden kyseisen puolueen kannattajat suhtautuvat negatiivisesti ja mihin positiivisesti, kun ikä ja sukupuoli vakioitu. Vertailut perustuvat parittaisiin luottamusväleihin. KSK = keskusta, KOK = kokoomus, KD=kristillisdemokraatit, PS = perussuomalaiset, VL = vasemmistoliitto, VIHR = vihreät, MUU = jokin muu puolue, EÄ = ei äänestäneet.

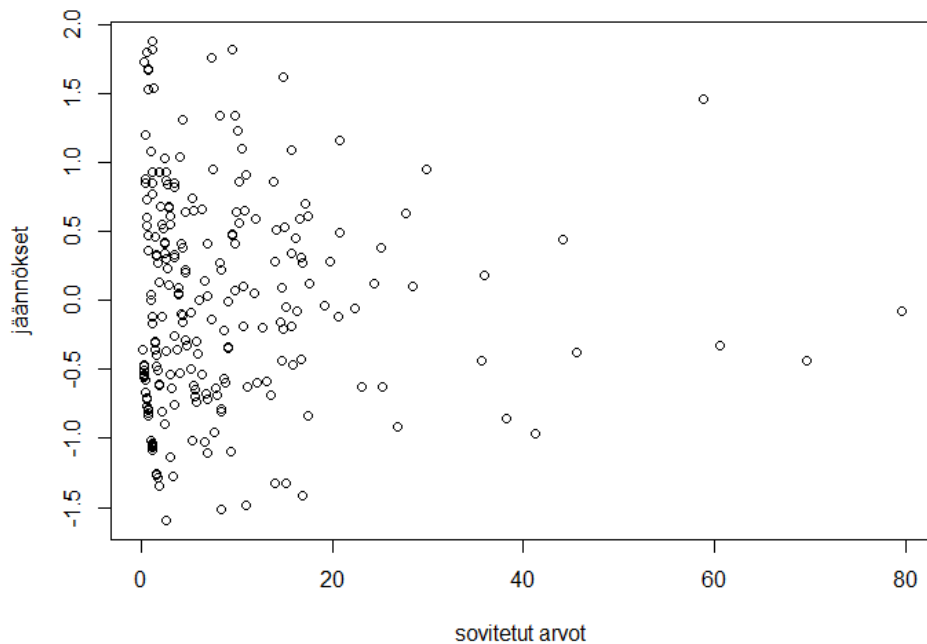
Puolue	Negatiivisempi kuin	Positiivisempi kuin
KSK	KOK, SDP	PS, EÄ
KOK	-	KSK, KD, PS, VL, VIHR, MUU, EÄ
KD	KOK, RKP, SDP	-
PS	KSK, KOK, RKP, SDP, VL, VIHR, MUU, EÄ	-
RKP	-	PS, KD, EÄ
SDP	-	KSK, KD, PS, VL, EÄ
VL	KOK, SDP	PS
VIHR	KOK	PS, EÄ
MUU	KOK	PS
EÄ	KSK, KOK, RKP, SDP, VIHR	PS

Kun tarkastellaan vaikea sanoa -mielipidettä ja negatiivista mielipidettä, niin havaitaan myös tämän väittämän osalta, että kokoomuksen kannattajat ovat vastanneet todennäköisemmin “vaikea sanoa” kuin että olisivat vastanneet “eri mieltä” verrattuna perustason eli tällä kertaa perussuomalaisten kannattajiin nähden. Tämä johtunee siitä, että kokoomuksen kannattajat ovat olleet reilusti selkeämmin EU-väittämän puolella kuin perussuomalaisten kannattajat, joten kokoomuksen kannattajat ovat vastanneet mieluummin “vaikea sanoa”, kuin että olisivat olleet eri mieltä. Tästä syystä EU-väittämän “vaikea sanovien -ryhmää” vertaillaan taas heihin, jotka ovat olleet jotakin mieltä. He, jotka ovat vastanneet EU-väittämään “vaikea sanoa”, ovat logistisessa regressiossa yksi luokka ja kaikki muut kuuluvat toiseen luokkaan. Prediktorin puolue perustasona on jälleen ei-äänestäneet, koska kyseinen ryhmä on mielenkiinnon kohteena.

Liitteestä 4 havaitaan, että ei-äänestäneet ovat vastanneet EU-väittämänkin kohdalla todennäköisemmin “vaikea sanoa” kuin että olisivat jotain mieltä verrattuna monen puolueen äänestäjiin. Kun vertaillaan samaa sukupuolta ja samaa ikäluokkaa edustavia, niin ei-äänestäneet ovat vastanneet todennäköisemmin “vaikea sanoa” kuin että olisivat olleet jotain mieltä Suomen vaikutusvallasta EU:ssa verrattuna keskustaa, kokoomusta, perussuomalaisia ja SDP:tä äänestäneisiin.

Myös kuvan 8 jäännöskuvio on samankaltainen kuin muiden väittämien kohdalla on nähty. Jäännökset jakautuvat jälleen suunnilleen tasan nollan molemmille puolille. Myös nyt havaitaan, että alkupäässä vaihtelu on suurempaa kuin loppupäässä, mutta ei ai-

van niin selkeästi kuin muiden väittämien kohdalla. Tästä voidaan päätellä, että kuvio ei osoita, etteikö malli sopisi aineistoon. Suurempi vaihtelu nollan molemmin puolin verrattuna loppupään vaihteluun selittyy sillä, että sovitettuja ja havaittuja arvoja on paljon enemmän nollan lähetyvillä kuin esimerkiksi arvosta 20 eteenpäin.



Kuva 8: Pearsonin jäännökset sovitetulle nominaalivasteisen regression mallille EU-väittämään

## 7.4 Puoluekannan tulosten vertailua muihin tuloksiin ja näkemyksiin

EU-väittämän tulokset ovat varsin samanlaisia kuin löydetyissä artikkeleissa on kerrottu. Nyt on etsitty, millaisia ovat olleet puolueiden mielipiteet EU:ta ja Suomen vaikutusvaltaa kohtaan EU:ssa.

Perussuomalaiset ja kristillisdemokraatit ovat olleet kielteisimpiä EU:ta kohtaan, kun on keskusteltu Suomen liittymisestä EU:hun. Positiivisimpia puolueita ovat olleet kokoomus, RKP ja SDP. Neutraaleja puolueita ovat olleet keskusta, vasemmistoliitto ja vihreät (Wikipedia, 2016). Turun Sanomien vuoden 2015 alkupuolella kirjoitetun artikkelin mukaan kansanedustajaehdokkaat perussuomalaisia lukuun ottamatta ovat nähneet, et-

tä EU on ollut Suomelle hyödyksi. Vain perussuomalaiset ja vasemmistoliiton ehdokkaat ovat nähneet vuoden 2015 alun tilanteessa selvää parantamisen varaa EU:ssa. (Ikäheimo, 2015). Tämän tutkimuksen tulokset ovat varsin samanlaisia kuin edellä mainituissa lähteissä on kerrottu, vaikkakin lähteissä on kerrottu kansanedustajista ja ehdokkaista, ja tässä tutkimuksessa on tarkasteltu puolueiden kannattajia.

## 8 Yhteenveto

Tässä tutkielmassa on johdateltu multinomivasteiseen ja erityisesti nominaalivasteiseen logistiseen regressioon. On käyty läpi, että multinomivasteita voidaan mallintaa yleistettyjen lineaaristen mallien avulla, ja miten nominaalisen logistisen regression tapauksessa estimoidaan kertoimet sekä tulkitaan päävaikutusten tapauksessa estimoituja kertoimia. Lisäksi on kiinnitetty huomiota mallinvalintaan ja mallin sopivuuteen.

Kun pohditaan EVAn kansallisesta asennetutkimuksesta (2014) poimittua Venäjäväittämää, niin yksinkertainen ja riittävä malli on sellainen, jossa on sukupuolen, iän ja puoluekannan päävaikutukset. Ammattiryhmän, alueen tai koulutuksen luokilla ei ole merkitseviä eroja Venäjäväittämässä. Tämän jälkeen sukupuolta, ikäluokkaa ja puoluekanta on tutkittu myös kolmen muun väittämän osalta, koska on haluttu tietää varsinkin puoluekantojen eroavaisuuksista näiden väittämien kohdalla. Lopulta kaikissa väittämässä on päädytty tarkastelemaan päävaikutusten mallia.

Perussuomalaisten ja vihreiden kannattajat ovat suhtautuneet Venäjään todennäköisemmin negatiivisesti kuin positiivisesti verrattuna SDP:n ja keskustan kannattajiin. NATOon liittymisen kohdalla kokoomusta ja RKP:tä äänestäneet ovat olleet todennäköisemmin positiivisempia varsinkin verrattuna vasemmistoliittoa ja SDP:tä äänestäneisiin. USA-väittämän tulos on hieman samankaltainen NATO-väittämän kanssa, koska varsinkin vasemmistoliiton kannattajat ovat olleet kielteisempiä verrattuna kokoomuksen kannattajiin. EU-väittämän kohdalla perussuomalaisia äänestäneet ovat olleet selkeästi negatiivisin ryhmä ja kokoomusta tai SDP:tä äänestäneet todennäköisemmin positiivisia kuin negatiivisia verrattuna monien puolueiden kannattajiin.

Millään kahden eri puolueen kannattajilla ei ole täysin samankaltaisia kertoimia kaikissa neljässä väittämässä. Lähimpänä samankaltaisuutta ovat olleet keskustan ja SDP:n kannattajat. Molempien puolueiden kannattajat löytyvät myönteisemmästä päästä Venäjä- ja EU-väittämässä sekä keskivaiheilta NATO- ja USA-väittämässä.

Samankaltaisuuksia löytyy enemmän kertoimien ja sijoitusten osalta kolmessa väittämässä. Tässä kyselyssä kokoomuksen ja RKP:n kannattajat ovat olleet positiivisimpia NATOa, USA:ta ja EU:ta koskevissa kysymyksissä. Venäjäväittämän kohdalla RKP:tä äänestäneet ovat olleet kielteisin ryhmä, kun taas kokoomusta äänestäneet ovat olleet kielteisimpien ja myönteisimpien puolueiden välissä.



Perussuomalaisten ja kristillisdemokraattien kannattajilla on samansuuruisia kertoimia Venäjää, USA:ta ja EU:ta käsittelevissä kysymyksissä. NATO-väittämässä kristillisdemokraattien äänestäjät ovat olleet kielteisiä, kun taas perussuomalaisten äänestäjät myönteisempiä. Vasemmistoliiton ja vihreiden kannattajien kertoimet ovat samankaltaisia Venäjää, USA:ta ja EU:ta koskevissa kysymyksissä, mutta poikkeavat enemmän NATO-väittämässä. Ei-äänestäneiden ja muiden puolueiden kannattajien kertoimet ovat myös samansuuruisia Venäjää, USA:ta ja EU:ta koskevissa kysymyksissä, mutta poikkeavat näitä kysymyksiä enemmän NATO-kysymyksessä.

Multinomivasteisella logistisella regressiolla on voitu tarkastella toisena tapauksena vasteen luokkia vaikea sanoa ja negatiivinen suhtautuminen. Se ei kuitenkaan ole järin kiinnostavaa, miten juuri negatiivisen kannan omaavat eroavat niistä, joilla ei ole näkemystä. Kiinnostavampaa on tutkia, miten vaikea sanoa -luokka eroaa sellaisesta luokasta, jolla on jokin näkemys (negatiivinen tai positiivinen). Tätä on tutkittu tavallisella logistisella regressiolla.

Tällä tavoin selviää, että niiden, jotka eivät olisi äänestäneet tai kertoneet puoluekantaansa, näyttää olleen vaikeampi sanoa, miten suhtautua väittämiin, kun heitä verrataan monien puolueiden kannattajiin.

Ikäluokkia vertailtaessa nuorimman ikäluokan keskuudessa on suhtauduttu jokaiseen neljään kysymykseen todennäköisemmin negatiivisemmin kuin positiivisemmin verrattuna vanhimpaan ikäluokkaan. Tosin ero ei ole merkitsevä USA-väittämän kohdalla, vaikka siinä väittämässä opiskelijat ovat olleet kielteisempiä kuin eläkeläiset. Jokaisen väittämän kohdalla myös naiset ovat olleet negatiivisempia kuin miehet, tosin EU-väittämän kohdalla ero ei ole ollut merkitsevä. Lisäksi naisille on ollut todennäköisemmin vaikea sanoa kuin olla jotain mieltä verrattuna miehiin kaikkien neljän väittämän kohdalla.

Kaikkien väittämien Pearsonin jäännöksiä tarkasteltaessa huomataan, että sovitettujen arvojen alkupään arvot heittelevät enemmän nollan ympärillä kuin loppupään arvot. Tämä johtunee siitä, että sovitettuja, kuten myös havaittuja, arvoja on runsaasti nollan lähetyvillä ja isoja sovitettuja arvoja on vähemmän. Suurempaa ihmettelyn aihetta kuvioissa ei ole, joten voidaan katsoa, että mallit sopivat hyvin aineistoon kaikkien väittämien osalta. Jos malleja kuitenkin haluaisi muuttaa, niin yksi idea voisi olla esimerkiksi puoluekannan luokkien muokkaaminen. Voi muun muassa miettiä, onko muita puolueita äänestäviä järkevää pitää tarkasteluissa mukana, koska heitä on niin vähän, ja lisäksi muita puolueita äänestävien voisi kuvitella olevan heterogeenisempi ryhmä kuin yhtä tiettyä puoluetta äänestävien ryhmät.

## Lähteet

Dobson, A.J. (2002). An introduction to Generalized Linear Models. Second edition: Chapman & Hall/CRC, Florida.

Davison, A.C. (2003). Statistical models. Cambridge University Press, Cambridge.

EVAn kansallinen asennetutkimus (2014) [elektroninen aineisto], FSD2933, versio 2.0 (2014-11-03). Helsinki: Taloustutkimus [aineistonkeruu], (2014). Helsinki: Elinkeinoelämän valtuuskunta [tuottaja], (2014). Tampere: Yhteiskuntatieteellinen tietoaarkisto [jakaaja], (2014).

Rubin, D.B. & Little, R.J. A. (2002). Statistical analysis with missing data (2nd ed.). Wiley, New York.

Kalbfleisch, J.G. (1985). Probability and Statistical Inference, Volume 2: Statistical Inference. Springer-Verlag, New York.

Keckman-Koivuniemi, H. (2015). Henkilökohtainen tiedonanto 2.11.2015. Sähköposti.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Wien. URL <https://www.R-project.org/>. Viitattu 30.11.2015.

Venables, W. N. & Ripley, B. D. (2002). Modern Applied Statistics with S. Fourth Edition. Springer, New York.

Tilastokeskuksen PX-Web-tietokannat (2013). Tilastokeskus:  
[http://pxnet2.stat.fi/PXWeb/pxweb/fi/StatFin/StatFin\\_\\_vrm\\_\\_vaerak/058\\_vaerak\\_tau\\_104.px/?rxid=d62897c0-217a-4979-85a1-7b47ca1dac7b](http://pxnet2.stat.fi/PXWeb/pxweb/fi/StatFin/StatFin__vrm__vaerak/058_vaerak_tau_104.px/?rxid=d62897c0-217a-4979-85a1-7b47ca1dac7b). Viitattu 10.11.2016.

### Venäjä-väittämän internet-lähteet

Venäjän viisumivapauden kannatus romahti (2014) Taloustutkimus:  
<http://www.taloustutkimus.fi/ajankohtaista/uutiskirje/uutiskirje-9-2014/venajan-viisumivapauden-kannatus/>. Viitattu 12.7.2016.

Joonas Pörsti (2015) Keskusta ottaa ohjat. Ulkopoliittinen instituutti.  
[http://www.ulkopoliitikka.fi/artikkeli/1388/keskusta\\_ottaa\\_ohjat/](http://www.ulkopoliitikka.fi/artikkeli/1388/keskusta_ottaa_ohjat/). Viitattu 12.7.2016.

Paavo Teittinen (2014) Suomen turvallisuuspolitiikkaa kirkastettiin Niinistön johdolla: "Selkeä linja luo jo itsessään turvallisuutta". Helsingin Sanomat.  
<http://www.hs.fi/politiikka/a1418963932057>. Viitattu 12.7.2016.

Ilkka Ahtokivi (2015) MTV: Suhtautuminen Venäjään jakaa uuden eduskunnan. Verkkouutiset.  
[http://www.verkkouutiset.fi/politiikka/suhtautuminen\\_rus\\_uusi\\_eduskunta-35101](http://www.verkkouutiset.fi/politiikka/suhtautuminen_rus_uusi_eduskunta-35101). Viitattu 12.7.2016.

### **NATO-väittämän internet-lähteet**

Ville Eloranta (2015). HS:n Vaalikone: Puolueiden puheenjohtajat ovat Nato-jäsenyydestä kieli keskellä suuta. Helsingin Sanomat.  
<http://www.hs.fi/politiikka/a1424146510699>. Viitattu 12.7.2016.

Aki Laurokari (2014). Tutkija keskustan ja kokoomuksen Nato-linjauksista: Ei muuta nykytilannetta. Keski-Pohjanmaan Kirjapaino Oyj:n verkkopalvelu.  
<http://www.kp24.fi/uutiset/lehti/414278/7/Tutkija-keskustan-ja-kokoomuksen-Nato-linjauksista-Ei-muuta-nykytilannetta>. Viitattu 12.7.2016.

Toni Viljanmaa (2013). Kokoomuksen Nato-tuska pääsi irti eduskunnassa. Ulkopoliittinen Instituutti ja Aamulehti.  
[http://www.fia.fi/fi/news/1805/kokoomuksen\\_nato-tuska\\_paasi\\_irti\\_eduskunnassa/](http://www.fia.fi/fi/news/1805/kokoomuksen_nato-tuska_paasi_irti_eduskunnassa/). Viitattu 12.7.2016.

### **EU-väittämän internet-lähteet**

Hannu-Pekka Ikäheimo (2015). Puolueiden EU-kannoissa on selviä eroja. Turun Sanomat.  
<http://www.ts.fi/mielipiteet/lukijoilta/749105/Puolueiden+EUkannoissa+on+selvia+eroja>. Viitattu 12.7.2016.

Suomen EU-kansanäänestys (2016). Wikipedia.  
[https://fi.wikipedia.org/wiki/Suomen\\_EU-kansanäänestys](https://fi.wikipedia.org/wiki/Suomen_EU-kansanäänestys). Viitattu 12.7.2016.

EU-kriittisyys (2016). Wikipedia.  
<https://fi.wikipedia.org/wiki/EU-kriittisyys>.

## Liitteet

**Liite 1** Logistisen regression kertoimien estimaatit Venäjä-väittämään. Vasteen perustasona on jokin mielipide ja sitä verrataan mielipiteeseen vaikea sana. Perustasoina prediktoreissa ovat sukupuoli, ikäluokka ja puoluekannassa ei-äänestävät.

kerroin (mielipide vs eos)	estimaatti	keskivirhe	exp(estimaatti)
vakio	-1.056	0.176	0.348
nainen	0.342	0.110	1.408
61jaYli	-0.016	0.214	0.984
41-60v	0.144	0.164	1.155
26-40v	0.193	0.169	1.213
keskusta	-0.765	0.181	0.465
kokoomus	-0.972	0.197	0.378
kristillidemokraatit	-0.019	0.364	0.981
perussuomalaiset	-0.510	0.177	0.600
RKP	-0.205	0.425	0.815
SDP	-0.432	0.192	0.649
vasemmistoliitto	-0.681	0.238	0.506
vihreät	-0.386	0.181	0.680
muut puolueet	-1.881	0.738	0.152

**Liite 2** Logistisen regression kertoimien estimaatit NATO-väittämään. Vasteen perustasona on jokin mielipide ja sitä verrataan mielipiteeseen vaikea sanoa. Perustasoina prediktoreissa ovat sukupuoli mies, ikäluokassa nuorin ikäluokka ja puoluekannassa ei-äänestävät.

<b>kerroin (mielipide vs eos)</b>	estimaatti	keskivirhe	exp(estimaatti)
vakio	-1.148	0.185	0.317
nainen	0.789	0.104	2.201
61jaYli	0.409	0.174	1.505
41-60v	0.527	0.182	1.694
26-40v	0.413	0.203	1.511
keskusta	-0.857	0.169	0.424
kokoomus	-0.192	0.159	0.825
kristillisdemokraatit	-0.396	0.362	0.673
perussuomalaiset	-0.601	0.168	0.548
RKP	-0.590	0.430	0.554
SDP	-0.937	0.195	0.392
vasemmistoliitto	-1.577	0.266	0.207
vihreät	-0.649	0.172	0.523
muut puolueet	-0.313	0.399	0.731

**Liite 3** Logistisen regression kertoimien estimaatit USA-väittämään. Vasteen perustasona on jokin mielipide ja sitä verrataan mielipiteeseen vaikea sanoa. Perustasoina prediktoreissa ovat sukupuoli mies, ikäluokassa nuorin ikäluokka ja puoluekannassa ei-äänestävät.

<b>kerroin (mielipide vs eos)</b>	estimaatti	keskivirhe	exp(estimaatti)
vakio	-1.062	0.167	0.346
nainen	0.755	0.104	2.128
61jaYli	0.215	0.191	1.240
41-60v	-0.053	0.153	0.948
26-40v	0.203	0.157	1.225
keskusta	-0.121	0.157	0.886
kokoomus	0.119	0.159	1.126
kristillisdemokraatit	-0.251	0.368	0.778
perussuomalaiset	-0.395	0.171	0.674
RKP	-0.007	0.405	0.993
SDP	-0.559	0.190	0.572
vasemmistoliitto	-1.431	0.278	0.239
vihreät	-0.706	0.182	0.494
muut puolueet	-1.207	0.548	0.299

**Liite 4** Logistisen regression kertoimien estimaatit EU-väittämään. Vasteen perustasona on jokin mielipide ja sitä verrataan mielipiteeseen vaikea sana. Perustasoina prediktoreissa ovat sukupuoli mies, ikäluokassa nuorin ikäluokka ja puoluekannassa ei-äänestävät.

<b>kerroin (mielipide vs eos)</b>	estimaatti	keskivirhe	exp(estimaatti)
vakio	-0.511	0.166	0.600
nainen	0.248	0.098	1.281
61jaYli	-0.068	0.156	0.934
41-60v	0.103	0.163	1.108
26-40v	0.367	0.182	1.443
keskusta	-0.582	0.158	0.559
kokoomus	-0.338	0.158	0.713
kristillidemokraatit	-0.459	0.364	0.632
perussuomalaiset	-1.033	0.177	0.356
RKP	-0.848	0.442	0.428
SDP	-0.435	0.176	0.647
vasemmistoliitto	-0.087	0.199	0.917
vihreät	-0.201	0.164	0.818
muut puolueet	-0.768	0.421	0.464