

MAANPUOLUSTUSKORKEAKOULU

ENNUSTEVIRHEET. MITÄ NE MEISTÄ KERTOVAT JA MITÄ VOIMME NIISTÄ OPPIA?

Yleisesikuntaupseerikurssin diplomityö

Komentajakapteeni
Juho Vanhatalo

Yleisesikuntaupseerikurssi 60
Rajavartiolinja

Elokuu 2021

Kurssi Yleisesikuntaupseerikurssi 60	Linja Rajavartiolinja
Tekijä Komentajakapteeni Juho Vanhatalo	
Opinnäytetyön nimi Ennustevirheet. Mitä ne meistä kertovat ja mitä voimme niistä oppia?	
Oppiaine, johon työ liittyy Johtaminen	Säilytyspaikka Maanpuolustuskorkeakoulun kirjasto
Aika Elokuu 2021	Tekstisivuja 107 Liitesivuja -
<p>TIIVISTELMÄ</p> <p>Tässä tutkimustyössä analysoitiin Rajavartiolaitoksen toiminnan ja talouden suunnittelua varten laadittujen ennusteiden virheitä. Ennustevirheiden tilastollisen analyysin avulla etsittiin virheistä systematiikkaa ja toistuvuutta, joiden avulla pyrittiin havaitsemaan merkkejä heuristisen ajattelun ja kognitiivisten vääristymien vaikutuksesta ennusteiden laadintaan. Tutkimusta varten saatiin käyttöön aiemmin hyödyntämätön aineisto Rajavartiolaitoksen suoritteista vuosilta 2006–2016 sekä suoritämäriä koskevista, vuosina 2009–2014 laadituista ennusteista.</p> <p>Teoria heuristisesta ajattelusta ja kognitiivisista vääristymistä perustuu käsitykseen siitä, että iso osa ihmisen ajattelusta on nopeaa, kontrolloimatonta intuitiota. Osa ongelmanratkaisusta ihmisen aivoissa tapahtuu nopeasti, yksinkertaistamalla ratkaistavaa ongelmaa ja soveltamalla yksinkertaisia nyrkkisääntöjä. Tätä kutsutaan heuristiseksi ajatteluksi eli heuristiikaksi. Heuristinen ajattelu aiheuttaa ajatusvirheitä, joiden on havaittu toistuvan hyvin samankaltaisina koehenkilöiden iästä, älykkyydestä ja koulutustasosta riippumatta. Koska heuristinen ajattelu ja sen mukanaan tuomat toistuvat virheet ovat kaikille ihmisille yhteisiä, teorialla on relevantteja ja mielenkiintoisia implikaatioita tutkittaessa päätöksenteon ja johtamisen psykologiaa.</p> <p>Se, voidaanko heuristisen ajattelun vaikutuksia havaita monen ihmisen yhteistyönä syntyneissä arvioissa ja päätöksissä, on avoinna oleva empiirinen kysymys, eikä asiaa ole turvallisuusviranomaisten toiminnassa aiemmin tutkittu. Siksi aihealueen tutkimus Rajavartiolaitoksessa on luontevaa aloittaa numeerisista ennusteista, joiden virheet ovat objektiivisesti todettavissa ja mitattavissa. Kognitiivisen psykologian piirissä ilmiötä on pääasiassa tutkittu tutkimuksia varten järjestetyissä testiolosuhteissa. Tämän vuoksi on myös tieteellisestä näkökulmasta arvokasta saada käyttöön aineisto, joka koostuu ihmisten todellisissa työtehtävissään tekemistä ennusteista.</p> <p>Virheellä tarkoitetaan tässä tutkimustyössä ennusteen ja toteuman välistä erotusta. Hyvissäkin ennusteissa on väistämättä virhettä, koska tulevaisuutta koskevaa faktatietoa ei ole. Ennusteiden virheitä aiheuttavat monet eri tekijät, joista suurinta osaa ei käsitellä tässä tutkimuksessa. Tavoitteena oli tuottaa tietoa siitä, voidaanko heuristisella ajattelulla ylipäättään osoittaa olevan merkitystä turvallisuusviranomaisten todellisessa päätöksenteossa. Koska virheiden syitä on monia, tarvitaan suuri otoskoko, jotta heikot signaalit voitaisiin erottaa monien, eri suuntaisten virhemekanismien joukosta.</p> <p>Tutkimuksessa havaittiin, että noin puolia ennusteista ei ollut muutettu verrattuna samasta asiasta aiemmin laadittuun ennusteeseen. Se, oliko ennustetta muutettu aiempaan verrattuna, ei ollut merkitsevästi yhteydessä sen kanssa, kuinka virheellisiä aiemmille vuosille tehdyt ennusteet olivat olleet. Samaa vuotta koskevat ennusteet olivat tarkempia kuin pitemmälle aikavälille laaditut. 1–4 vuoden päähän tehtyjen ennusteiden välillä ei ollut merkitseviä eroja ennustetarkkuudessa.</p> <p>Hallintoyksiköiden välillä oli ennustevirheissä vuodesta toiseen samankaltaisena toistuvia eroja, vaikka eri hallintoyksiköiden ennusteet oli tarkastettu ja hyväksytty samassa yläjohtoportaan, ja tieto aiempien vuosien ennusteiden virheistä olisi ollut ennusteiden laatijoiden käytettävissä.</p>	

Ennusteita oli muutettu vähemmän alas- kuin ylöspäin. Toteuman laskuun oli reagoitu (ennustetta muuttamalla) harvemmin ja vähemmän kuin toteuman kasvuun: Toteuman kasvuun reagoiminen muuttamalla ennustetta ylöspäin oli merkittävästi todennäköisempää kuin toteuman laskuun reagoiminen muuttamalla ennustetta alaspäin. Ennusteisiin tehdyt tarkistukset olivat myös kooltaan pienempiä silloin, jos toteumassa oli laskeva trendi.

Virheet aiempien vuosien ennusteissa ennustivat seuraavien vuosien ennustevirheitä lähes yhden suhde yhteen. Toisin sanoen, jos ennuste oli virheellinen, seuraavan vuoden ennusteesta samasta asiasta oli odotusarvoisesti samankokoinen ja -suuntainen virhe. Yksinkertaisen tilastollisen ennusteen (kolmen edellisen vuoden toteumien keskiarvo) virhe olisi ollut koko aineistossa keskimäärin 4.9 prosenttiyksikköä pienempi kuin aineistona olleiden asiantuntijoiden tekemien ennusteiden.

Asiantuntijaennusteet olivat (suhteessa tilastolliseen ennusteeseen) sitä parempia, mitä suurempi oli kyseisen tulosrivin suoritteiden lukumäärä, ja mitä suurempi oli kyseiselle suoritteelle tuloksellisuuden laskennassa annettu painokerroin. Lisäksi saatiin viitteellistä näyttöä siitä, että asiantuntijaennusteet olisivat olleet keskimäärin yhtä tarkkoja kuin tilastolliset ennusteet, jos aiempien vuosien ennusteiden virheitä ei olisi toistettu.

Tutkimustuloksia voidaan pitää ainakin viitteellisenä näyttönä heuristisen ajattelun vaikutuksista ennusteiden laatimiseen. Aihealueen kirjallisuudessa kuvataan kiinnittyminen aiempiin lukuihin ja reagoimattomuus huonoon ennustettavuuteen. Kirjallisuudessa kuvataan myös, ettei ihminen osaa siirtyä kohti regressiivisempää ennustetta, vaikka ennustetarkkuus heikentyisi, ja että ihminen ottaa helpommin vastaan informaatiota, joka sopii yhteen hänen omien ennakkokäsitystensä kanssa. Näiltä osin tutkimuksen empiiriset havainnot olivat yhdensuuntaisia sen kanssa, mitä teorian perusteella voitiin olettaa.

Valitulla tutkimusmenetelmällä ja -aineistolla ei voida saada tosiasiatietoa havaittujen korrelaatioiden syistä. Vaikka tutkimustulokset ja teorian perusteella tehdyt oletukset eivät ole ristiriidassa, kausaalista suhdetta ennustevirheiden systematiikan ja heuristisen ajattelun välillä ei voida tässä työssä miltään osin todistaa. Regressiomallien selitysasteet olivat hyvin matalia, eli havaitut korrelaatiot selittivät ennustevirheitä vain hyvin pienen osan.

Tutkimustyöllä tuotettiin Rajavartiolaitokselle kattava analyysi ennustevirheistä, joita ei ollut aikaisemmin systemaattisesti tarkasteltu. Tutkimuksen havaintojen perusteella voitiin myös antaa seuraavat suositukset ennusteiden laatimisen kehittämisestä: Ennusteita ei ole suositeltavaa laatia vanhojen ennusteiden perusteella. Ennusteet tulisi laatia joko ilman lähtöarvoja tai siten, että lähtöarvona käytetään esimerkiksi edellisten vuosien toteuman keskiarvoa. Yhdenmukaisuutta aiempiin lukuihin tulisi tarkastella vasta kun ennusteet on laadittu. Virheiden säännöllinen tarkastelu olisi suositeltavaa ottaa säännölliseksi osaksi prosessia.

AVAINSANAT

ennakointi, ennusteet, ennustettavuus, heuristiikka, johtaminen, kognitiivinen psykologia, tilannekuva

SISÄLLYS

1. JOHDANTO	1
1.1. Tutkimuksen taustaa	1
1.2. Virheestä käsitteenä ja ennusteiden laadinnasta	3
1.3. Rajaukset	6
1.4. Heuristinen ajattelu ja kognitiiviset vääristymät	7
1.5. Tutkimuksen tarpeellisuus	8
1.6. Tutkimuskysymykset	9
1.7. Tutkimuksen tavoitteet	10
1.8. Tutkimusmenetelmä	11
2. KOGNITIIVISET VÄÄRISTYMÄT ENNUSTEIDEN LAADINNASSA	12
2.1. Kahden tyyppistä prosessointia	12
2.2. Yhteenveto Dual-Process -teoriasta	13
2.3. Heuristinen ongelmanratkaisu: Yksinkertaisia vastauksia monimutkaisiin ongelmiin	15
2.4. Kontrollin vaikeus	18
2.5. Heuristisen ajattelun aiheuttamat vääristymät	19
2.6. Edustavuus	20
2.7. Tavoitettavuus	23
2.8. Kysymyksenasettelun vaikutus vastauksiin	25
2.9. Ankkuroituminen	27
2.10. Tunnehajaiseen arviointiin liittyvät vääristymät	28
2.11. Vääristymien kompensoiminen ja välttäminen	29
2.12. Älykkyyden ja koulutuksen merkitys	31
2.13. Heuristisen ajattelun merkitys ennusteiden laadinnan kannalta	33
2.14. Ennustettavuuden arvioinnin merkitys	37
2.15. Intuiitiivinen reaktio heikkoon ennustettavuuteen	38
2.16. Heuristisen ajattelun merkitys riskien arvioinnin kannalta	38
3. TUTKIMUSMENETELMÄ	42
3.1. Kvantitatiivinen tutkimusasetelma sosiaalitieteissä	42
3.2. Aineisto	43
3.3. Otoksen kokoaminen	44
3.4. Saatavilla olevat tiedot kustakin tilastorivistä	46
3.5. Tulosluvussa käytetyt suureet:	47
Ennustevirhe ja virheiden muuntaminen yhteismitallisiksi	47
Ennusteen harhaisuus ja tarkkuus	48
Suunnitelmaluvut	50
Toteumaluvut	50

Huono odotettu ennustettavuus	51
Aiemmat virheet.....	53
Regressiivinen ennuste.....	53
Aikavälit.....	54
Painokerroin ja suuruusluokka	54
Organisaation osaa kuvaavat binäärimuuttujat.....	55
3.6. Menetelmät	56
Huomautus termeistä.....	56
Hajontalukujen kuvaaminen ja vertailu ryhmittäin	56
Kahden muuttujan suhde.....	57
Lineaarinen regressio	57
Regressioanalyysin lähtöoletukset	58
Heteroskedastisuus ja sen testaaminen.....	59
Multikollineaarisuus.....	59
Normaalisuus ja sen testaaminen	59
Autokorrelaatio paneelidatassa	61
Autokorrelaation ja heteroskedastisuuden huomioivat keskivirheet.....	61
Muuttujamuunnokset.....	62
4. TULOKSET	64
4.1. Eri aikaväleille laaditut ennusteet.....	64
4.2. Painokerroin ja toteuman suuruusluokka.....	69
4.3. Hallintoyksiköiden erot	73
4.4. Ennusteiden muuttaminen	77
4.5. Heikko etukäteinen (odotettu) ennustettavuus.....	81
4.6. Vertailu tilastolliseen ennusteeseen.....	83
5. DISKUSSIO	91
5.1. Samojen virheiden toistaminen.....	91
5.2. Usko suoritelmäärien kasvuun.....	95
5.3. Ennustettavan tulosrivin merkitys ja tärkeys	96
5.4. Mitä ennusteilla ilmaistaan?	98
5.5. Odotukset, joita ennusteisiin kohdistetaan.....	100
5.6. Tutkimuksen rajoitteet ja luotettavuuden arviointi	103
5.7. Jatkotutkimuksen tarve	105

LÄHTEET

ENNUSTEVIRHEET. MITÄ NE MEISTÄ KERTOVAT JA MITÄ VOIMME NIISTÄ OPPIA?

1. JOHDANTO

1.1. Tutkimuksen taustaa

Muuttuvan toimintaympäristön ennakointi ja riskien arviointiin ja analysointiin perustuva ennakoiva johtaminen ovat keskeisiä teemoja yhteiskunnan turvallisuuden rakentamisessa. Viranomaisten toiminta pyritään mitoittamaan vastaamaan uhkia. Riskit pyritään tunnistamaan ja analysoimaan siten, että riskien ja uhkien taso kyetään huomioimaan suunnittelussa ja toiminnassa, ja kyetään tunnistamaan toimenpiteitä, joilla riskitasoa pyritään alentamaan tai uhkaa torjumaan. Tavoitteena on tilanne, jossa kyetään ennakoimaan ja luomaan toiminnan painopisteitä, ja samaan aikaan kuitenkin kyetään käyttämään resursseja taloudellisesti.

Kun tehdään osittain tuntemattomia olosuhteita tai tulevaisuutta koskevia arvioita, on hyvin epätodennäköistä, että arvio tai ennuste osuisi täsmälleen oikeaan. Niin voi käydä, mutta se on hyvin harvinaista. Yleisempää lienee se, että ennusteet ovat oikean suuntaisia, mutta vähintään yksityiskohdiltaan ne jotenkin eroavat toteutuvasta todellisuudesta. Ennusteet voivat olla niin laveita tai epämääräisiä, että se vähentää niiden hyödyllisyyttä. Vastaavasti ennusteet, joissa on konkreettisia yksityiskohtia, kuten aikoja, paikkoja tai lukumääriä, eivät voi osua täysin oikeaan kuin äärimmäisen harvoin.

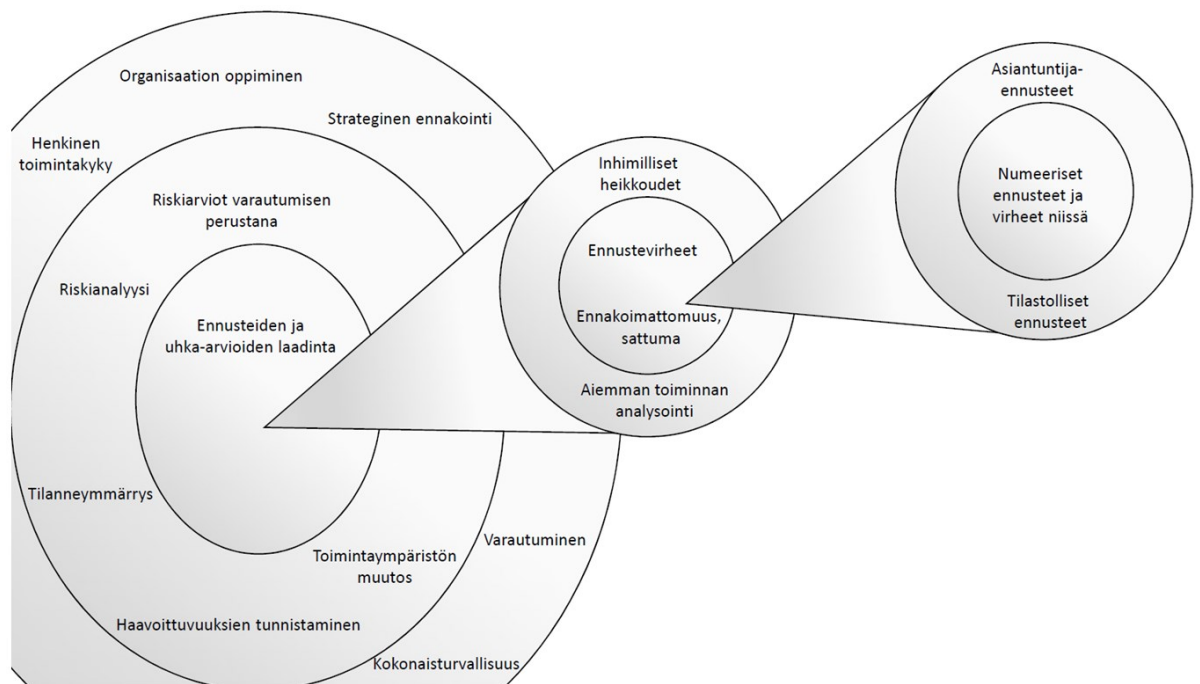
Tässä työssä käsitellään ennusteita ja ennustevirheitä. Ennuste käsitetään tässä työssä laajaksi käsitteeksi. Se voi, päätöksentekotilanteen yksityiskohdista riippuen olla hyvin epäformaali, nopea, lähes tiedostamaton orientaatio siihen, mitä aivan välittömässä lähitulevaisuudessa tapahtuu. Toisessa ääripäässä ennuste voi olla systemaattisen työn tuloksena tietyllä menetelmällä laadittu määrämuotoinen ennuste tai analyysi pitkän aikavälin kehityksestä. Joka tapauksessa ennakoivaa päätöksentekoa edeltää jonkinlaisen ennusteen muodostaminen, tavalla tai toisella. Osa jäljempänä kuvattavasta logiikasta ja ennusteisiin liittyvistä ongelmista pätee myös muun muassa uhka-arvioihin, tilannearvioihin ja riskianalyysiin.

Tehtyjen virheiden jälkikäteinen tarkastelu mahdollistaa organisaation oppimisen ja lisää avoimuutta. Esimerkiksi Suomen Pankki on julkaissut säännöllisesti raportteja ennustevirheistään (ks. esim. Newby & Orjasniemi 2011). On tärkeää pyrkiä luomaan mahdollisuuksia organisaation oppimiselle, jotta samoja virheitä ei toistettaisi (Mälkki 2015, 112). Kaikkia virheitä ei ole mahdollista välttää, mutta voimme pyrkiä tiedostamaan niiden laadun ja laajuuden. Jos tietoa ei ole tai sitä ei hyödynnetä, tilannetietoisuutemme jää rajalliseksi, eikä ole mahdollista kyseenalaistaa vallitsevia oletuksia (Martelius 2019).

Suurin osa Rajavartiolaitoksessa tehtävistä ennusteista on laadullisia ja sanallisia, ja niiden oikeellisuutta saattaa olla vaikea jälkikäteenkään määrittää tai ainakaan mitata tai kvantifioida. Tämä ei tarkoita sitä, ettei laadullisia ja sanallisia ennusteita ja arvioita kannattaisi jälkikäteen tarkastella ja pyrkiä analysoimaan niiden osuvuutta ja sitä, mitkä ovat ne syyt prosesseissa, kommunikaatiossa tai päätöksentekotilanteissa, jotka johtivat niihin onnistumisiin ja epäonnistumisiin, joita ennakoimiseen aina väistämättä liittyy. Numeeristen ennusteiden jälkikäteinen tarkastelu on helpompaa. Niiltä osin kuin ennustetaan esimerkiksi suoritteiden lukumääriä, ja saadaan jälkikäteen toteumatieto käyttöön, voidaan kvantifioida myös se, kuinka kaukana totuudesta ennusteet olivat.

Mitattavuus ei itsessään kuitenkaan tee asiasta tärkeämpää. Turvallisuusviranomaisten toiminnassa tehdään valtava määrä olennaisen tärkeitä laadullisia arvioita operaatioiden toimintalinjoista ja toteuttamiskelpoisuudesta, uhkien vaikuttavuudesta ja todennäköisyyksistä sekä tilanearvioita, riskianalyseja ja suunnitelmia. Ei ole mitään perusteita olettaa, että määrälliset ennusteet olisivat ennusteista tärkeimpiä. Sen sijaan kvantitatiivinen tarkastelu on luonteva ensimmäinen askel siksi, että numeerinen data mahdollistaa objektiivisen ja kiihkottoman lähestymistavan. Numeeristen ennusteiden virheet ja niiden kvantitatiivinen tarkastelu liittyvät kuitenkin edellä kuvattuun laajempaan kokonaisuuteen.

Ennusteiden ja arvioiden laadinta on hyvin monimutkainen ja monipuolinen asia, johon liittyy paljon eri näkökulmia. Tästä kokonaisuudesta pieni osa on ennusteiden osuvuus tai osumattomuus: sen tarkastelu, miten olemme aiemmin onnistuneet ennakoimaan uhkia ja muutoksia, sekä sen arviointi, miten hyvin oletamme onnistuvamme tulevien uhkien ja muutosten ennakoinnissa. Näistä, edelleen hyvin laajoista kysymyksistä, yksi yksityiskohta ovat kvantitatiiviset ennusteet ja aiempien kvantitatiivisten ennusteiden tarkastelu.



Kuva 1. Tutkimuksen viitekehys

1.2. Virheestä käsitteenä ja ennusteiden laadinnasta

Ennusteen virhe tarkoittaa erotusta ennusteen ja toteutuneen todellisuuden välillä. Suomen kielen sana virhe on arkikieliseltä merkitykseltään laajempi, joten termin merkitystä on syytä käsitellä. Hyvissäkin ennusteissa on väistämättömästi virheitä, koska tulevaisuutta koskevaa dataa ei ole. Tämän tutkimustyön aineiston ennusteet ovat nk. piste-ennusteita. Piste-ennusteessa koko todennäköisyysjakaumaa kuvataan yhdellä luvulla (odotusarvolla), joka on mahdolliseksi arvioitujen toteumien todennäköisyyksillä painotettu keskiarvo. Ei voida siis sanoa, että piste-ennuste olisi väärä, jos ennuste ei ole yhtä suuri kuin toteuma. Kuitenkin ennusteen ja toteuman erotuksesta käytetään nimitystä virhe.

Numeerisen ennusteen helpoimmin laskettavat ominaisuudet ovat tarkkuus ja harhattomuus. Epätarkkuus (*error*) on virheiden itseisarvojen keskiarvo, joka kertoo, kuinka kaukana toteumasta ennusteet keskimäärin ovat. Harhaisuus (*bias*) on virheiden keskiarvo, joka kertoo, ovatko ennusteet systemaattisesti jommallakummalla puolella (yli tai alle) toteumaa. Etenkään epätarkkuus ei edellytä ennusteen laatijan väärää menettelyä. Varsinkin suurten muutosten ennustettavuus voi olla erittäin huono, jolloin hyvissäkin ennusteissa on suuria virheitä.

Ennusteen laatijasta aiheutuvat virheet ovat psykologisia virheitä. Ne voidaan karkeasti jakaa, keskustelun yksinkertaistamiseksi, tahattomiin ja tahallisiin virheisiin. Tahattomat virheet ovat virheitä ajatusprosessissa, ja näihin kuuluvat tässä työssä tarkasteltavat heuristiset vääristymät. Kirjallisuudessa on osoitettu, että heuristiset virheet eivät rajoitu naiiveihin subjekteihin, vaan koskevat myös ammattilaisia ja korkeasti koulutettuja asiantuntijoita (ks. esim. Kahneman 2003, 711; Kahneman & Tversky 1982, 68; Yudkowsky 2008, 99). Näiden virheiden syntymekanismit liittyvät siihen, miten kaikkien ihmisten aivot toimivat.

Tahalliset virheet ovat ennusteen (tai toteuman) manipulointia jonkin tavoitteen edistämiseksi. Todellisuudessa tahallisten ja tahattomien virheiden välillä ei ole tarkkaa rajaa. Ennusteen laatija voi turvautua yksinkertaiseen heuristiikkaan myös tietoisesti, jos hän arvioi aikasäästön arvokkaammaksi kuin ennustevirheen välttämisen. Samoin lukujen manipulaatio voi tapahtua tiedostamatta; ihmiset reagoivat kannustimiin myös huomaamattaan.

Ennusteiden laatimisen menetelmät voidaan jakaa tilastollisiin ennusteisiin ja asiantuntija-arvioihin. Tilastolliset ennusteet ovat yleensä numeerisia, ja ne ovat yleensä jonkinlaisia ekstrapolaatioita olemassa olevasta datasta. Myös asiantuntija-arviot voivat olla numeerisia. Menetelmällisesti ne ovat yhdistelmiä intuitiivisesta (heuristisesta) ja loogisesta (analyttisestä) ajattelusta. Lisäksi ennusteita voidaan laatia myös yhdistetyillä menetelmillä, joissa hyödynnetään sekä tilastollista ennustetta että asiantuntija-arviota.

Numeeristen ennusteiden ajatellaan usein olevan laadullisia ennusteita objektiivisempia. Asia ei välttämättä ole näin, silloin kun käsitellään sosiologisia ilmiöitä (ihmisten toimintaa). Muutoin kuin kaikkein teknisimmissä sovelluksissa, tilastollisten ja matemaattisten menetelmien objektiivisuus on sikäli näennäistä, että mallin valinta tai kokoaminen, datan valinta, mahdollisten eksogeenisten suureiden arvot (ja valinta) sisältävät joka tapauksessa subjektiivisen elementin. Jos ennusteen laatimisen prosessi on sellainen, että asiantuntijat voivat halutessaan hylätä tilastollisen ennusteen, on oikeastaan olemassa vain asiantuntija-arvioita ja yhdistetyillä menetelmillä tehtyjä ennusteita.

Numeerinen ennuste voidaan laatia tilastollisin menetelmin, asiantuntija-arviona tai yhdistetyillä menetelmillä. Laadullisen ennusteen laatiminen perustuu aina asiantuntija-arvioon, vaikka laatimisessa voidaan noudattaa jotain tiettyä metodologiaa, jonka avulla ajattelua pyritään strukturoimaan. Laadullisen ennusteen ytimessä ovat asiantuntijan intuitiivinen ja

analyttinen kyvykkyys. Puhtaita tilastollisia tai laskennallisia ennusteita, joita ei laskennan jälkeen altistettaisi asiantuntijoiden subjektiiviselle tarkastelulle, on Rajavartiolaitoksessa laadittavista ennusteista hyvin pieni osa, vaikka tällaisiakin on (esimerkiksi öljypäästöjen leviämisen mallinnukset tai meripelastuksen ajelehtimislaskelmat). Yhdistelmämenetelmien etuna on teoriassa se, että tilastollisten ja heurististen virhelähteiden pitäisi periaatteessa olla toisistaan riippumattomia. Tutkimusaineistossa on asiantuntija-arviona tai yhdistelmämenetelmillä tuotettuja numeerisia ennusteita.

Mitkä seikat aiheuttavat ennusteiden virheitä? Virheitä on paljon erilaisia: eri kokoisia, eri suuntaisia, eri tyyppisiä, eri syistä johtuvia, eri kohdassa prosessia syntyviä, ja niin edelleen. Kaikissa ennustemenetelmissä on omat heikkoutensa. Datassa voi olla virheitä, tilastollinen malli voi olla väärin rakennettu. Asiantuntijoiden arviot voivat olla väärää monesta eri syystä, käsitys vallitsevasta tilanteesta voi olla puutteellinen. Informaatio voi olla ristiriitaista, tai yksinkertaisesti saatetaan arvata väärin, mihin suuntaan tilanne on kehittymässä. Ennustevirheiden taustalla voi olla myös organisatorisia syitä, tai aikapaine. Tietynlaisissa ennusteissa ja arvioissa voi olla operatiivisia syitä tietoisesti pyrkiä välttämään tietynlaisia virheitä (esim. liian vähän pelastusyksiköitä tehtävällä) ja erehtyä mieluummin tiettyyn suuntaan oikeasta vastauksesta (esim. mieluummin liikaa kuin liian vähän pelastusyksiköitä).

Kuten todettu, on myös nk. tahallisia virheitä, ja moninaisia syitä näiden taustalla, erilaisia maine- tai resurssikannustimiin liittyviä asioita. Ennustevirheisiin itsessäänkin voi liittyä kannustimia ennustaa tahallaan väärin (Ehrbeck & Waldmann 1996). Organisaation sisältä voi tulla ohjausta tai painetta ennusteen laadintaan liittyen. Voi esiintyä ristiriitaisia näkemyksiä. Jos ennusteita tehdään useamman henkilön yhteistyönä, kyseessä on myös sosiaalinen prosessi. Siihen, mitä sellaisessa sosiaalisessa vuorovaikutuksessa tapahtuu, liittyy monia kysymyksiä. Tässä tutkimuksessa aineistona olevat Rajavartiolaitoksen toiminnan ja talouden suunnitteluun kuuluvat ennusteet toimivat nekin osaltaan pohjana resurssien käytön suunnittelulle ja resurssien jakamiselle, jolloin niihin voi liittyä kannustimia, joita emme välttämättä kaikkia kykene edes tunnistamaan.

Ennusteisiin voi liittyä myös tarkoituksellista skenaarioiden luontia tai jonkin ajatuksen tai skenaarion myymistä. On mahdollista, että jonkin uhkan kuvaaminen tietynlaisena palvelee jollain tavalla uhkakuvan tai ennusteen laatijaa. Samoin voi esiintyä päämies-agentti - ongelmia, jossa arvion tekijän kannustimet ovat erisuuntaiset kuin prosessin omistajan; esimerkiksi alajohtoportaan kannustimet voivat olla erilaiset kuin yläjohtoportaan.

1.3. Rajaukset

Tulevaisuuden ennakoinnin vaikeutta tai suoranaista mahdottomuutta olisi mahdollista tutkia erittäin monesta näkökulmasta; ennusteiden laadinnan prosesseja, ennusteiden virheitä ja virheiden lähteitä voisi tarkastella hyvin monen eri tieteenalan ja teorian kautta. Ennakointi ja ennusteiden laadinta ovat monia tieteenaloja leikkaavia, poikkitieteellisiä ongelmia. Yksittäisessä tutkimuksessa ei ole mahdollista käsitellä syvällisesti kaikkia niitä mahdollisia syitä, jotka aiheuttavat ennusteen ja toteuman välisen erotuksen, ts. ennustevirheen. Tässä työssä ei voida vastata kaikkiin ilmiöön liittyviin kysymyksiin, eikä tämä työ voi millään selittää kaikkia syitä ennusteiden virheille. Tässä tutkimustyössä ennustevirheiden joukosta etsitään viitteitä ihmisen intuitiivisessa ajattelussa toistuvista ajatusvirheistä, ja työssä on tehty kolme tärkeää rajausta liittyen aineistoon, tutkimusmenetelmään ja valittuun teoriaan.

Tässä tutkimustyössä keskitytään numeerisiin ennusteisiin. Tilanteessa, jossa ennusteiden paikkansapitävyyttä ei ole systemaattisesti selvitetty, numeromuodossa esitetyt ennusteet ovat hyvä ja luonteva aloituspiste. Siksi aineistoksi on valittu ainoastaan numeerisia ennusteita.

Ei ole mitään perustetta olettaa, että määrälliset ennusteet olisivat tärkeämpiä kuin laadulliset ja sanalliset ennusteet, arviot ja riskianalyysit. Numeerisista ennusteista on kuitenkin helpoin lähteä liikkeelle, koska numeeristen ennusteiden virheet ovat kvantifioitavissa, ja tarkasteltavissa objektiivisesti ja riidattomasti. Jos valitun teorian mukaisia systemaattisia virheitä kyetään tunnistamaan numeerisessa aineistossa, voidaan diskussiossa pyrkiä ottamaan kantaa siihen, olisiko uskottavaa olettaa, että samankaltaisia virheitä olisi myös laadullisissa ja sanallisissa ennusteissa.

Tässä työssä ei pyritä tulkitsemaan merkkejä tahallisista virheistä, koska valittu teoria ei pyri selittämään sen kaltaisia ilmiöitä. Kaikista niistä mahdollisista syistä, jotka voivat aiheuttaa tahattomia systemaattisia virheitä, on tässä työssä valittu näkökulmaksi tarkastella virheitä kognitiivisen psykologian kirjallisuudessa kuvattujen heuristisen ajattelun ja kognitiivisten vääristymien kautta. Tämä on kenties keskeisin rajaus, ja liittyy valittuun teoriaan.

Nyt käytössä olevan aineiston avulla on mahdollista yrittää tutkia systemaattisia tai toistuvia ennustevirheitä piste-ennusteissa, joita Rajavartiolaitoksessa on tehty suoritteiden ja

havaintojen määrien kehityksestä, ja joilla osaltaan kuvataan toimintaympäristön kehittymistä. Suunnitelmaluvuista koostetun aineiston pohjalta saatavia tuloksia ei voi yleistää muihin päätöksentekotilanteisiin ilman jatkotutkimusta. Siksi tässä työssä ei käsitellä kaikkea sitä muuta tilanteenarviointia, ennakointia, riskianalyysia ja päätöksentekoa, joista teorian perusteella voisi tehdä hypoteeseja. Tutkimustulosten mahdollisista implikaatioista voidaan kuitenkin keskustella diskussiossa.

1.4. Heuristinen ajattelu ja kognitiiviset vääristymät

Ihminen on rationaalinen toimija, mutta aika- ja resurssirajoitteet ajavat ihmisen käyttämään aivojaan eri tavoin eri tilanteissa. Kognitiivisen psykologian kirjallisuudessa on kuvattu ihmisen rationaalisuuden rajoja, heuristista ajattelua, kognitiivisia vääristymiä, ja ihmisen erilaisia tapoja ratkaista ongelmia. Tätä tutkimusta on julkaistu 1970-luvulta lähtien, ja sillä on nykyään melko vakiintunut asema. Daniel Kahneman sai hänen ja Amos Tverskyn heuristista ajattelua ja vääristymiä koskevasta tutkimuksesta taloustieteen Nobel-palkinnon vuonna 2002. Tämä *Heuristics and Biases* -nimellä tunnettu tutkimusala käsittelee eroja objektiivisesti oikeiden vastausten, ja ihmisten empiirisissä kokeissa valitsemien vastausten välillä.

Tahdonalaiselle ja tahdosta riippumattomalle ajattelulle on kirjallisuudessa annettu monia eri termipareja. Näitä eri malleja kutsutaan yhteisnimellä Dual-Process theories. Nämä ovat käsitteellisiä malleja ihmisten ajatustoiminnasta, ja tarjoavat selityksiä niille virheille, joita heuristisen ajattelun tutkimuksessa on kuvattu. Tässä työssä puhutaan Stanovichin (2011, 19) mukaan ”tyypin I ja tyypin II prosessoinnista”. Kun ajattelua on kehitetty, on kuvattu neljä pääasiallista prosessoinnin tasoa, joista kaksi nopeinta kuuluu tyypin I: Affektiivisiin eli tunnereaktioihin perustuva päättelyminen on kaikkein nopeinta. Tunnereaktioiden lisäksi myös seuraava taso, intuitiivinen ajattelu, on tahdosta riippumatonta ja erittäin nopeaa. Tyypin II prosessointiin kuuluu tietoinen ajattelu, jossa ihminen ei teekään intuitiivista johtopäätöstä tai ei tyydy siihen, vaan käyttää asian ajattelemiseen aikaa ja energiaa. Tyypin II prosessointiin kuuluu myös ylin, reflektiivinen taso; eräänlainen metaprosessointi, jossa ihminen tietoisesti ajattelee ajattelemista, ratkaisee sitä, pitääkö jonkin asian miettimiseen käyttää aikaa ja resursseja, ja kyseenalaistaa omia johtopäätöksiään.

Aihealueen tutkimuksessa kuvataan, millaisia virheitä ihmiset intuitiivisessa ajattelussa tyypillisesti tekevät, sekä yksilöllisiä eroja ja yhtäläisyyksiä ihmisten ajattelussa. Tämän tutkimustyön kannalta olennaisia ovat virheet, joita ihmiset intuitiivisesti tekevät ennusteita ja arvioita laatiessaan.

Kirjallisuus ja teoria esitellään luvussa 2. Teoriaosuudessa käsitellään heuristisen ajattelun ja kognitiivisten vääristymien vaikutusta päätöksentekoon yleisesti, ja ennusteiden laadintaan erityisesti. Lisäksi teoriaosuudessa käydään läpi heuristisen ajattelun vaikutuksia modifioivia tekijöitä, kuten relevantti koulutus, älykkyys ja käytettävissä oleva aika, ja mitä tällaisten tekijöiden vaikutuksesta aiemman tutkimuksen perusteella tiedetään.

Kognitiivisessa psykologiassa ja taloustieteessä heuristisen ajattelun aiheuttamia vääristymiä on tutkittu yleensä testiolosuhteissa, jossa koehenkilöt suorittavat tutkijoiden laatimia tehtäviä. Näin on saatu esiin erittäin mielenkiintoisia, ja huomattavan samankaltaisina toistuvia havaintoja siitä, mihin ihmisen intuitiivinen ajattelu pystyy ja toisaalta ei pysty. Näiden ilmiöiden osoittamiseksi tutkimukset eivät käsittele useamman ihmisen yhteistä päätöksentekoa vaan nimenomaisesti yksittäisiä ihmisiä.

On tärkeää ymmärtää, että kognitiivisissa vääristymissä ei ole kyse motivaation puutteesta tai laiskuudesta. Tätä kuvastaa mm. se, ettei oikeista vastauksista tarjotuilla rahallisilla palkinnoilla ole kyetty vähentämään koehenkilöiden virheitä. Myös motivoituneiden asiantuntijoiden toistuvat epäonnistumiset samoissa testeissä osoittavat, että analyttisen ja intuitiivisen ajattelun välillä ei ole kyse yksikäsitteisesti tietoisesta valinnasta. Intuitiivinen ajattelu on erottamaton osa sitä, miten me ihmiset erilaisiin johtopäätöksiin päädymme. Resurssin allokointi jonkin asian pohdinnalle tietoisesti on jo erittäin edistynyttä rationaalista toimintaa. Sellainenkin ihminen, joka osaa kontrolloida itseään näin, tekee sitä hyvin harvoin. Suuren osan päätöksenteosta on pakko olla intuitiivista, koska käytettävissä oleva aika on rajallinen.

1.5. Tutkimuksen tarpeellisuus

Rajavartiolaitos on tehnyt numeerisia ennusteita toimintaympäristötekijöistä ja omasta toiminnastaan, osana toiminnan ja talouden seurantaan ja tuloksellisuuden laskentaa, ainakin 1990-luvun puolivälistä lähtien. Näitä ennusteita käytetään strategisen tason johtamisen

yhtenä osatekijänä toiminnan ja talouden suunnitteluprosessissa. Ennusteiden virheitä ei ole koskaan systemaattisesti tarkasteltu. Tarkastelulle on tarve, vaikka se ei johtaisi mihinkään löydöksiin, koska sekin jo itsessään lisää tietouttamme. Ja, jos systemaattisia vääristymiä löydetään, saatamme oppia laatimaan ennusteita eri tavoin tai suhtautumaan niihin jäsenytyneemmin.

Virheiden analysointi tuottaa arvokasta tietoa organisaatiosta ja toimintaympäristöstä: ”Millainen on kykymme ennakoita toimintaympäristömme muutoksia? Tiedämmekö, mistä virheet johtuvat? Millaisia virheitä yleensä teemme? Toistammeko samoja virheitä tai samantyyppisiä virheitä? Voisiko osaa virheistä jotenkin välttää?” Tulemalla tietoisiksi virheistä, niiden määristä ja tyypeistä, on mahdollista kompensoida eri mekanismein – kuten tarkistamalla sitä, millaisia odotuksia asetetaan millekin ennusteelle, milläkin aikavälillä.

Kognitiivisten vääristymien esiintyvyys hierarkkisessa päätöksenteossa on aihe, jota kirjallisuudessa ei varsinaisesti ole aiemmin käsitelty. Se, voidaanko eri organisaatiotasojen yhteisen työn tuloksena syntyneistä ennusteista luotettavasti havaita merkkejä kirjallisuudessa kuvatusta yksilöiden heuristisesta ajattelusta, on aiemmin vastaamaton empiirinen kysymys.

Heuristisen ajattelun ja kognitiivisten vääristymien tutkimuksen kenties useimmin esitetty kritiikki on, että tulokset on saatu testiolosuhteissa (ks. esim. Gilovich & Griffin 2002, 11). Kognitiivisia vääristymiä on tutkittu hyvin vähän varta vasten järjestettyjen tutkimustilanteiden ulkopuolella, koska tarkoitukseen soveltuvia tietoaaineistoja on vaikea koota. Tämän työn aineisto tarjoaa harvinaisen mahdollisuuden etsiä heuristisen ajattelun merkkejä tosimaailman havainnoista.

Vaikka teoreettisista, menetelmällisistä ja aineistoon liittyvistä syistä käsitellään vain tietynlaisia numeerisia ennusteita, tutkimustuloksia voidaan käyttää yhtenä syötteenä keskusteluun ja herättämään ajatuksia myös laadullisten ennusteiden osalta, ja ennakoivasta johtamisesta laajemminkin.

1.6. Tutkimuskysymykset

Tässä työssä pyritään empiirisesti kuvaamaan pientä osaa siitä, millaisia epätarkkuuksia esiintyy tilanteiden arvioinnissa ja ennakoinnissa. Tutkimusongelma on: Mitä ennustevirheet

kertovat heuristisen ajattelun vaikutuksista ennusteisiin? Koska ennustevirheitä ei ole koskaan systemaattisesti tarkasteltu, on virheiden tilastollinen kuvaaminen (deskriptiiviset tulokset) keskeisessä osassa tutkimusongelmaan vastattaessa.

Lisäksi, osana tutkimusongelman tarkastelua pyritään vastaamaan seuraaviin tutkimuskysymyksiin:

- Voidaanko empiirisesti havaita systemaattisia tai toistuvia ennustevirheitä?
- Ovatko ennustevirheet sellaisia kuin heuristisen ajattelun ja kognitiivisten vääristymien teoria esittää?

1.7. Tutkimuksen tavoitteet

Tulosten perusteella pyritään sekä luomaan ehdotuksia ennusteiden laatimisesta että esittämään hypoteeseja siitä, mitä heurististen virheiden mahdollinen esiintyvyys hierarkkisessa päätöksenteossa tarkoittaa yleisemmin, tilannearvioiden, riskianalyyysien ja ennakoivan päätöksenteon kannalta.

Tiedonintressi on kahtalainen. Rajavartiolaitoksella on organisaation oppimisen ja kehittymisen kannalta tarve deskriptiivisille tuloksille, ja suosituksille ennusteiden laadinnan kehittämisestä. Akateeminen tiedonintressi liittyy heuristisen ajattelun jälkien etsintään usean ihmisen yhteistyön tuloksena, hierarkkisessa organisaatiossa, syntyneestä tosimaailman aineistosta.

Tutkimuksen tavoitteet ovat:

1. Virheiden kunnollinen tilastollinen kuvaaminen.
2. Heuristisen ajattelun ja kognitiivisten vääristymien mahdollisen vaikutuksen ymmärtäminen.
3. Esitykset ennusteiden laatimisen prosessin kehittämiseksi siten, että olemassa oleva informaatio ja tutkimus hyödynnetään, ja mahdollisia psykologisia tekijöitä kyettäisiin kompensoimaan.
4. Jatkotutkimuksen ja keskustelun mahdollistaminen kognitiivisen psykologian kuvaamien ilmiöiden merkityksestä, numeerisia ennusteita laajemmassa kontekstissa.

1.8. Tutkimusmenetelmä

Tutkimusaineistona on Rajavartiolaitoksen toiminnan ja talouden suunnitteluprosessin toimintaympäristön kehitystä kuvaavia suunnitelmalukuja vuosilta 2009-2014, sekä ennusteita vastaavat toteumat. Aineistossa on myös tieto kunkin tilistorivin Rajavartiolaitoksen tuloksellisuuden laskennassa käytetystä painokertoimesta, sekä tieto siitä, mitä hallintoyksikköä suunnitelmaluku koskee. Aineiston on luovuttanut tutkimuskäyttöön Rajavartiolaitoksen esikunnan Suunnittelu- ja talousyksikkö, Rajavartiolaitoksen myöntämän tutkimusluvan (1917167, RVLDno-2019-315) mukaisesti. Tutkimukseen mukaan otettu aineisto on julkista.

Tutkimusmenetelmänä on ennustevirheiden tilastollinen kuvaaminen, sekä aineiston eri muuttujien välisten suhteiden kuvaaminen korrelaatioiden ja lineaarisen regression avulla. Tietokannan kokoaminen ja rakenne, sekä tutkimusmenetelmä on kuvattu tarkemmin luvussa 3.

2. KOGNITIIVISET VÄÄRISTYMÄT ENNUSTEIDEN LAADINNASSA

2.1. Kahden tyyppistä prosessointia

Tässä luvussa kuvataan ihmisen ajattelua ns. Dual-Process -teorian kautta. Tämän mukaan ihmisen ajattelu jakautuu tyyppin I nopeaan, intuitiiviseen ja tunnepohjaiseen ajatteluun, ja tyyppin II analyttiseen, algoritmiseen ja reflektiiviseen ajatteluun. Intuitiivista ja heuristista ajattelua käytetään tarkoittamaan samaa asiaa. Termit eroavat toisistaan alkuperänsä osalta; heuristinen ajattelu on kognitiivisen psykologian termi, ja intuitio hieman epämääräisempi ja yleiskielisempi ilmaisu. Kumpaakin käytetään, jotta säilytetään tarkkuus kirjallisten lähteiden hyödyntämisessä.

Ajatus siitä, että ihmisen aivoissa on erilaisia osioita, joilla on erilaiset tehtävät, on tuttu jo antiikin kirjallisuudesta; jo Platon erotteli sielun refleктоivan osan toisesta osasta, joka kaipaa primitiivisten tarpeiden tyydyttämistä (Stanovich 2011, 16). Sosiaalipsykologiassa, kognitiivisessa psykologiassa ja muillakin tieteenaloilla on viime vuosikymmeninä kehitetty useita kahden prosessin malleja (*Dual-Process theories*) (sama, 19). Viime vuosien aikana arkikieleenkin on vakiintunut kahdenlaista prosessointia kuvaavat termit *systemi I* ja *systemi II* (ensimmäisenä Stanovich & West 2000, 658).

Kahdesta systeemistä puhuminen saattaa luoda harhakäsityksen siitä, että kukin systemi olisi aivoissa yksittäinen, tarkkarajainen kokonaisuus, vaikka näin ei ole (Stanovich 2011, 19). On havainnollisempaa puhua *tyypin I* ja *tyypin II* prosessoinnista. Nimityksiä ja erotteluja eri ajatusprosessien välillä on kirjallisuudessa esitetty useita. Näistä termeistä varhaisimpien ja suosituimpien joukkoon kuuluu termipari heuristinen ajattelu - analyttinen ajattelu (sama). Tyyppin I ja tyyppin II prosessointia voitaisiin kuvata myös esimerkiksi termeillä intuitiivinen - tahdonalainen. Näiden heikkous on tarkan määrittelyn puute, mutta vahvuus vastaavasti termien kattavuus.

Klassisen rationaalisen päätöksenteon malli ei edellytä, ettei ihminen tekisi virheitä. Edellytyksenä on kuitenkin, että virheissä ei ole systematiikkaa, eli että virheet ovat satunnaisia (Gilovich & Griffin 2002, 2; Stanovich & West 2000, 647). 1960- ja 1970-luvuilla alkanut heuristisen ajattelun tutkimus muutti olennaisesti ihmisen päätöksenteon tutkimuksen. Kuvaus ihmisen turvautumisesta ajattelussaan heuristisiin yksinkertaistuksiin tarjosi selityksen sille, että ihmiset tekevät systemaattisia virheitä ajattelussaan, edellyttämättä kuitenkaan rationaalisuuden oletuksen hylkäämistä. (Gilovich & Griffin 2002, 1.)

Heuristinen ajattelu perustuu ongelmanratkaisua helpottavien ajatusmallien hyödyntämiseen. Näitä ovat mm. ankkuroituminen (esimerkiksi numeerisiin lähtöarvoihin), ja edustavuuden (havaintojen yhteensopivuus ennakkokäsitysten kanssa) arviointi todennäköisyyksien arvioinnin asemesta. Näihin ilmiöihin liittyviä toistuvia virheitä ovat mm. kyvyttömyys ymmärtää konjunkttiivisia todennäköisyyksiä, kyvyttömyys reagoida kertaluokan muutoksiin (numeerisessa tiedossa), suhteellisten osuuksien liiallinen painotus päätöksenteossa verrattuna absoluuttisiin lukumääriin, ja liiallinen luottamus omiin johtopäätöksiin. Näillä kaikilla on merkitystä riskien arvioinnin ja ennusteiden laatimisen kannalta.

2.2. Yhteenveto Dual-Process -teoriasta

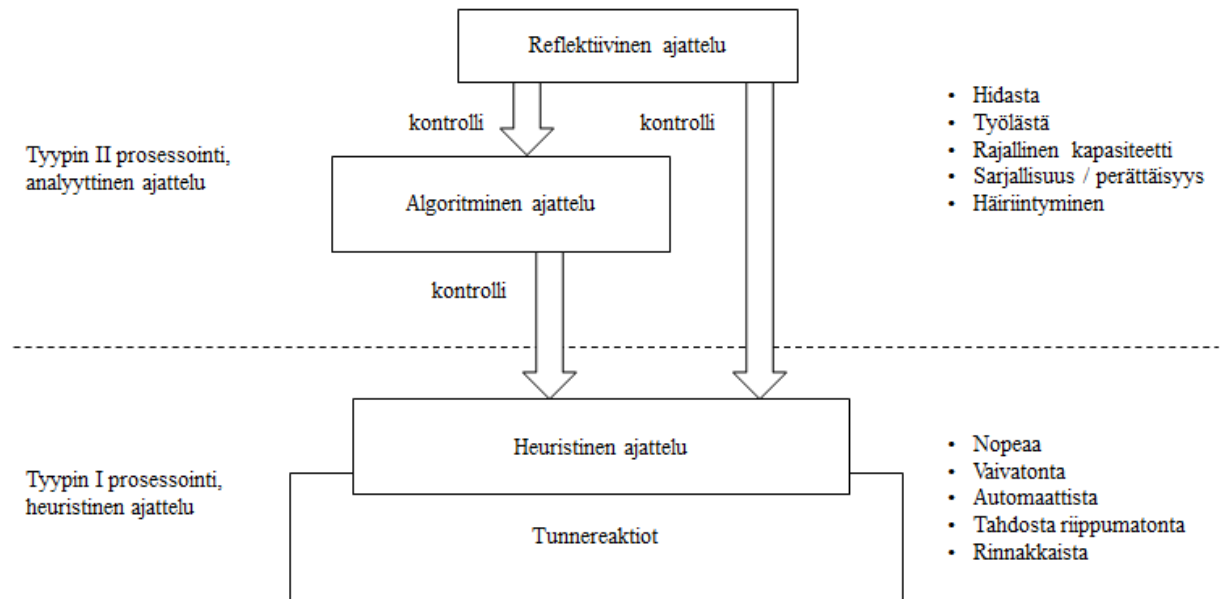
Heuristinen ajattelu (tyypin I prosessointi) on autonomista, tiedostamatonta ja tahdosta riippumatonta (Kahneman 2003, 698; Stanovich & West 2000, 658; Stanovich 2011, 19). Tietty heräte (engl. *stimulus*) käynnistää heuristisen prosessin väistämättömästi. Tämä on tyypin I prosessien merkittävin määrittävä piirre (Stanovich 2011, 19). Heuristisille, autonomisille prosesseille on tyypillistä, että ne eivät edellytä syötteitä korkeammilta ajattelun tasoilta. Heuristinen ajattelu on nopeaa, se ei kuluta merkittävästi aivojen resursseja tai kapasiteettia, eivätkä samanaikaiset heuristiset prosessit häiritse toisiaan tai tyypin II prosessointia (Gilovich & Griffin 2002, 4; Kahneman 2003, 698; Stanovich 2011, 19).

Tyyppin I prosessointiin kuuluvat myös tunnereaktiot. Mielikuvat ja niihin liittyvät positiiviset ja negatiiviset tunteet ohjaavat arviointia ja päätöksentekoa (Slovic ym. 2007, 1335). Asioiden arvottamista tunnepohjaisesti kutsutaan affektiksi, ja sen tiedetään aiheuttavan saman tyyppisiä systemaattisia vääristymiä kuin muiden heurististen prosessien (sama). On ihmiselle tyypillistä tehdä valinta tunnepohjaisesti, ja sitten jälkikäteisesti esittää erilaisia perusteluja valinnoille (Slovic ym. 2007, 1334; Yudkowsky 2008, 99). Tunnereaktiot ovat usein ensimmäiset ja kaikkein nopeimmat, ne syntyvät automaattisesti, ja ohjaavat sen jälkeen informaation prosessointia ja päätöksentekoa (Slovic ym. 2007, 1334). Tunnereaktiot kuuluvat tyyppin I nopeaan, automaattiseen prosessointiin, ja ne ovat osa sitä mekanismeja, jolla ihminen hyödyntää oppimaansa.

Analyyttinen ajattelu, eli tyyppin II prosessointi, on epäautonomista, hidasta ja kallista (sitoo aivokapasiteettia ja kuluttaa energiaa). Analyyttiselle ajattelulle on tyypillistä myös sen sarjallisuus, ts. prosessit pitää toteuttaa yksi kerrallaan, ajallisesti peräkkäin. Analyyttinen ajattelu on usein kieleen perustuvaa, mutta ei kaikissa tapauksissa. (Stanovich 2011, 20.)

Yksi analyttisen ajattelun tärkeimmistä tehtävistä on kumota heuristisen ajattelun tuottamia ratkaisuja (Stanovich 2011, 20; Stanovich & West 2000, 662; Stanovich & West 2002, 439). Tämä on tarpeen, koska heuristinen ajattelu on tarkoitettu tuottamaan *oikean suuntaisia* vastauksia, joiden tarkkuus ei kuitenkaan ole riittävä, jos kyseessä on tärkeä päätös. (Stanovich 2011, 20). Kriittisyys ja omien johtopäätösten epäileminen kuuluvat näihin korkeampiin prosesseihin, analyttiseen ajatteluun (Kahneman 2003, 702).

Analyttinen ajattelu voidaan jakaa edelleen algoritmiseen ajatteluun ja reflektiiviseen ajatteluun (Stanovich 2011, 35-36). Ihmiset eroavat toisistaan älykkyyden osalta, mikä kuvaa eroja algoritmisen ajattelun tehokkuudessa. Ihmiset kuitenkin eroavat toisistaan myös reflektiivisessä ajattelussa. Tähän sisältyvät erot ajattelutavoissa: suhtautumisessa informaatioon, aktiivisuudessa, kriittisyydessä, ymmärtämisen tarpeessa, halussa tehdä itse omat johtopäätöksensä, tieto- ja todellisuuskäsityksissä, halussa miettiä ratkaisujen seurauksia, epävarmuuden sietämisessä, kiinnittymisessä dogmaattisiin käsityksiin, ja niin edelleen (Stanovich 2011, 35).



Kuva 2. Yhteenveto Dual-Process -mallista (mukaan Stanovich 2011, 62; Stanovich & West 2011, 122).

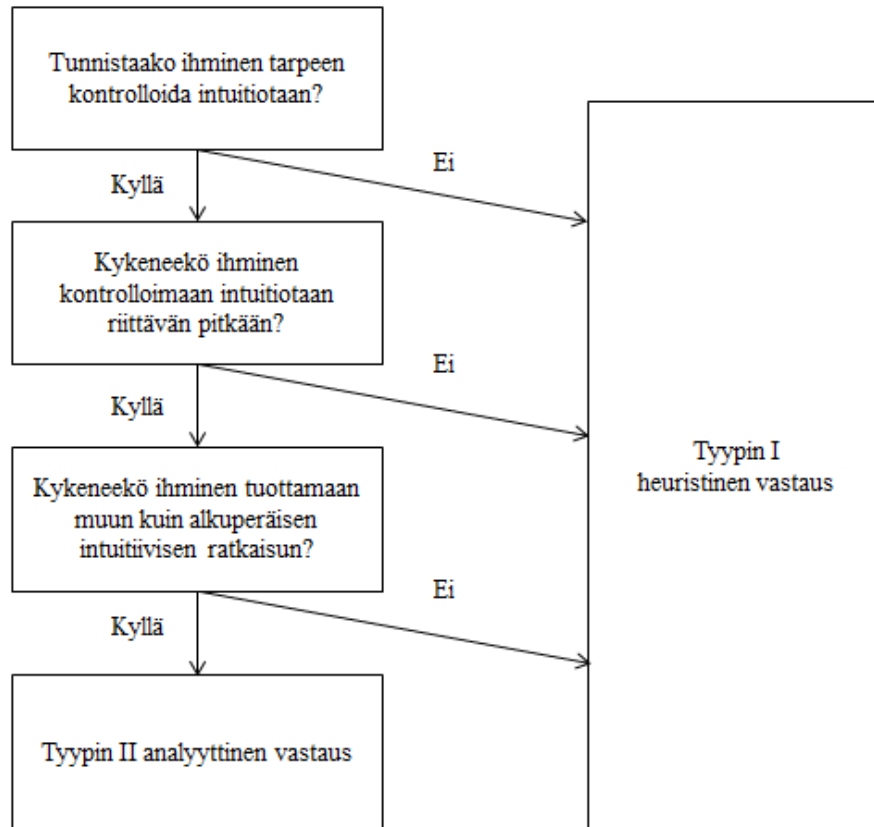
Heuristisen ajattelun ja sen aiheuttamien vääristymien tutkimusta ja *Dual-Process* -teorioita on keskeisesti kritisoitu kuvaileviksi, ts. että ne eivät tarjoa selityksiä havaituille ilmiöille (Stanovich & West 2002, 423; Gilovich & Griffin 2002, 11). Muilta osin kritiikki on kohdistunut normatiivisten oikeiden vastausten liian kapeaan määrittelyyn koehenkilöille esitetyissä tehtävissä (Friedman 2019, 182), ja teorioiden epämääräisyyteen, joka mahdollistaa pahimmillaan kaiken evidenssin selittämisen yhteensopivaksi teorioiden kanssa (Gigerenzer 1996, 593).

2.3. Heuristinen ongelmanratkaisu: Yksinkertaisia vastauksia monimutkaisiin ongelmiin

Kahden systeemin mallin mukaisen päätöksen voidaan ajatella syntyvän pääasiallisesti kahdella¹ tavalla: Ensimmäinen tapa on, että syntyy intuitiivinen vastaus ongelmaan tai kysymykseen, jonka jälkeen tahdonalainen ajattelu hyväksyy sen, korjaa arviota (usein liian vähän) tai hylkää sen kokonaan. Toisaalta voi olla, ettei intuitiivista vastausta synny ollenkaan, ja vastaus pitää erikseen tahdonalaisesti päätellä. (Kahneman 2003, 717.)

¹ Kahnemanin (2003, 717) mukaan viidellä eri tavalla, mutta ero syntyy siitä, miten luettelo ryhmitellään.

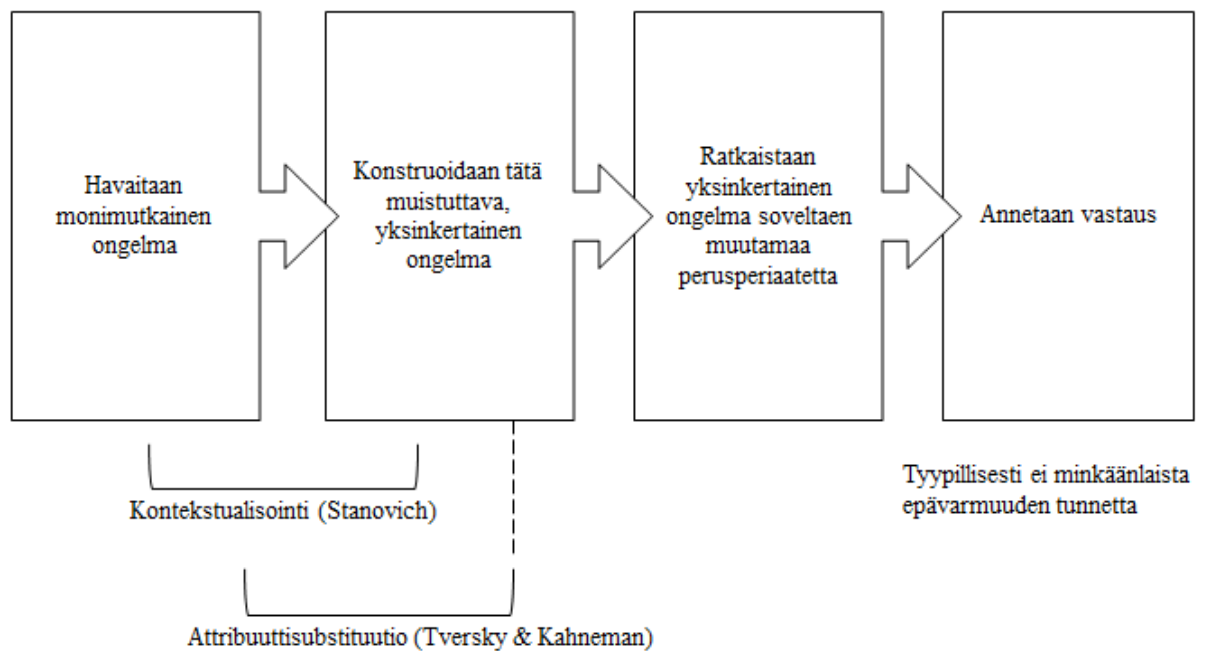
Vaihtoehtoisesti heuristisen tai analyttisen päätöksen syntyminen voidaan ajatella polkuna, jossa kysymyksillä analyttisen ajattelun kontrollin riittävydestä on tietty järjestys (Stanovich & West 2011, 143): onko henkilöllä ajatuksellisia keinoja kontrolliin, tunnistaako hän kontrollin tarpeen, onko kontrollin tarve pitkäkestoinen ja, kykeneekö henkilö jatkamaan kontrollia riittävän pitkään?



Kuva 3. Polku heuristiseen tai analyttiseen vastaukseen (mukaillen Stanovich & West 2011, 143).

Heurististen mallien yhdistävä piirre on monimutkaisten ongelmien korvaaminen yksinkertaisilla ongelmilla, jotka muistuttavat riittävän läheisesti käsillä olevaa monimutkaista ongelmaa. Salamannopeat ja energiaa säästävät heuristiset prosessit tuottavat lähes aina oikean suuntaisia, ns. riittävän hyviä vastauksia (Gilovich & Griffin 2002, 4), ja niiden avulla ihminen kykenee säästämään aikaa ja energiaa. Näin ihminen kykenee tuottamaan verraten edullisesti vastauksia sellaisiinkin kysymyksiin, joiden täysin rationaalinen ratkaiseminen olisi suhteettoman raskasta tai jopa mahdotonta. Koska ajan ja energian säästäminen on ihmiselle edullinen toimintatapa, ihmiselle on kehittynyt tapa pyrkiä aina tuottamaan likimääräinen ratkaisu heuristisen ajattelun kautta (Stanovich 2011, 29).

Olellainen askel heuristisen vastauksen tuottamisessa on monimutkaisen ongelman korvaaminen yksinkertaisella ongelmalla. Tätä voidaan kutsua kontekstualisoinniksi (mm. Stanovich & West 2000) tai attribuuttisubstituutioksi (mm. Kahneman 2003; Kahneman & Frederik 2002, 81). Kontekstualisoinnin ja substituution avulla joko tunnistetaan tai konstruoidaan kohdattua monimutkaista ongelmaa muistuttava yksinkertainen ongelma, ratkaistaan tämä, ja annetaan näin saatu johtopäätös vastauksena alkuperäiseen, monimutkaiseen ongelmaan.



Kuva 4. Heuristinen ongelmanratkaisu.

Koska heuristinen ajattelu on nopeaa ja vaivatonta, se on joissain yhteyksissä yhdistetty laiskuuteen tai motivaation puutteeseen. Tämä on valitettava ja hieman harhaanjohtava miellelyhtymä. Heuristinen ajattelu perustuu erittäin sofistikoituneisiin prosesseihin, kuten muistiin tallentuneen tiedon nopeaan hyödyntämiseen, yhteneväisyyksien tunnistamiseen, jne. Nämä prosessit ovat erittäin keskeinen osa sitä, miten ihminen kykenee tekemään hyviä ja hyödyllisiä päätöksiä hyvinkin nopeasti ja taloudellisesti (Gilovich & Griffin 2002, 3). Kyky heuristiseen ajatteluun ja resurssien tuhannesosasekunneissa tapahtuva allokaatio nopean, heuristisen käsittelyn ja vaativamman, analyttisen käsittelyn välillä on erittäin kehittyntä, korkean tason aivotoimintaa. Kyse ei siis ole laiskuudesta, vaan aika- ja resurssirajoitteiden vallitessa eloonjäämisen kannalta välttämättömästä inhimillisen ajattelun rakenteesta.

Koska heuristinen prosessointi ei ole tahdonalaista, ihminen ei pysty olemaan tekemättä heuristista arviota saamastaan informaatiosta (Gilovich & Griffin 2002, 5). Tämä on tärkeää ymmärtää, kun tarkastellaan heurististen virheiden syntymistä tehtävissä, joihin tehtävän suorittaja keskittyy, ja jotka hän kokee tärkeäksi. Tyypin I prosessointiin liittyy usein myös tunnereaktio ja opitun tavan toistaminen; siksi sitä on erittäin vaikea kontrolloida tai muuttaa (Kahneman 2003, 698). Tyypin II prosessoinnissa on suuriakin yksilöllisiä eroja, mutta autonominen ajattelu on ihmisillä pitkälti samanlaista (Stanovich 2011, 33).

2.4. Kontrollin vaikeus

Voidakseen kontrolloida heuristista ajattelua, analyyttisen ajattelun pitää pystyä toteuttamaan kaksi keskeistä funktiota. Sen tulee kyetä keskeyttämään heuristisen vastauksen (päätöksen) toteuttaminen. Toisekseen sen tulee kyetä hypoteettiseen ajatteluun, jolla tuotetaan heuristisen vastauksen tilalle parempi vastaus. (Stanovich 2011, 21-22.) Analyyttisen ajattelun mahdollisuuksiin havaita ja korjata intuitiivisen ajattelun virheitä vaikuttavat negatiivisesti aikapaine, päällekkäiset kognitiiviset tehtävät, epäedullinen vuorokaudenaika ja - ehkä hieman yllättäen - hyvällä mielellä oleminen (Kahneman 2003, 711).

Ihminen tuntee tyypillisesti erittäin voimakasta luottamusta heuristisen prosessin tuottamiin johtopäätöksiin (Yudkowsky 2008, 107). Kyseenalaistaminen on tyypin II prosessointia; kyky epäilemiseen on metakognitiivinen ominaisuus, joka edellyttää yksilön kykyä konstruoida useita erilaisia, keskenään yhteen sopimattomia ajatuksia samasta asiasta (Kahneman 2003, 702). Analyyttisen ajattelun tehtävänä on kontrolloida heuristista ajattelua, mutta tämä on vaikeaa kahdesta syystä. Ensinnäkin, kontrollin tarve on hyvin vaikea tiedostaa. Toisekseen, intuitiiviset vastaukset ovat hyvin houkuttelevia. Tämä kontrollin vaikeus johtaa siihen, että heuristisen ajattelun tuottamat vastaukset jäävät usein voimaan, ja heuristisen ajattelun tuottamia tyypillisiä virheitä voidaan havaita myös älykkäiden, koulutettujen ja motivoituneiden ihmisten antamista vastauksista.

Heurististen prosessien kontrolloinnin vaikeudesta on todettu myös, että ihmiset luontaisesti soveltavat falsifiointia vain argumentteihin, jotka ovat ristiriidassa heidän ennakkokäsitystensä kanssa (Yudkowsky 2008, 99). On viitteitä siitä, että analyttisessäkin ajattelussa (pitkään harkitessaan) ihmiset pyrkivät etsimään vahvistuksia hypoteeseilleen, eivät falsifioimaan niitä (Stanovich 2011, 63-64; Yudkowsky 2008, 98). Ihmiset valitsevat lähes aina sen vaihtoehdon, jonka he arvioivat todennäköisimmin tulevana valitsemaan. Toisin sanoen, heti kun ihminen kykenee arvioimaan todennäköisyyden sille, että hän tulee tekemään tietyn valinnan tai ratkaisun, hän on käytännössä jo päättänyt asian (Yudkowsky 2008, 101). Kyetäkseen tehokkaaseen kontrolliin analyttisen ajattelun pitäisi toimia ennen kuin ihminen on alkanut edes alustavasti kallistua jonkin tietyn vaihtoehdon kannalle (sama). Kansan kielellä voitaisiin todeta, että kun on kerran saanut jonkin ajatuksen päähänsä, siitä on erittäin vaikea päästä eroon.

2.5. Heuristisen ajattelun aiheuttamat vääristymät

Seuraavaksi käsitellään muutamia parhaiten kirjallisuudessa dokumentoituja heuristisia ratkaisumalleja, ja niiden tyypillisesti aiheuttamia vääristymiä. Tässä yhteydessä keskitytään muutamaa vääristymään, joilla on voimakkaimmat implikaatiot mahdollisiin systemaattisiin virheisiin riskien arvioinnissa ja ennusteiden laadinnassa. Jo aiemmin on käsitelty heuristisen prosessoinnin kontrollointia analyttisen ajattelun tehtävänä, ja heuristisen päätöksen syntymistä. Kuten on jo todettu, tyypin I prosessointi on nopeaa, sen kontrollointi on vaikeaa ja heuristisen ajattelun avulla tuotetut vastaukset ovat joskus vääriä. Sellaisissa päätöksentekotilanteissa, joihin heuristinen ajattelu ei synnyntäisesti sovellu, virheisiin muodostuu systematiikkaa, jolloin syntyy vääristymä.

Tilastotieteessä vääristymän puuttumisesta käytetään suomeksi termiä *harhattomuus*, mutta sen vastakohtana sanan *harha* käyttäminen saattaa aiheuttaa väärinymmärryksiä, koska tässä yhteydessä konteksti ei ole tiukan matemaattinen. Siksi tässä yhteydessä käytetään sanaa *vääristymä* kuvaamaan virheitä, joissa on systematiikkaa eli toistuvuutta. Englannin kielen termi *bias* on merkitykseltään oikeampi, pitäen sisällään ajatuksen virheen puolueellisuudesta, eli virheen taipumuksesta olla samansuuntainen aiempien virheiden kanssa. Englanninkielisessä tieteellisessä kirjallisuudessa tämän aiheen tutkimuksesta käytetään nimitystä *heuristics* (heuristiset ajattelun keinot, joilla yksinkertaiset vastaukset tuotetaan) *and biases* (heuristiseen ajatteluun turvautumisesta aiheutuvat systemaattiset virheet).

Ihmisen intuitio, tyypin I prosessoinnin tuottama arvio, on yleensä hyvä, mutta tietyt todennäköisyyden ja tilastotieteen peruseriaatteet eivät kuulu ihmisen luontaiseen, intuitiiviseen työkalupakkiin, ja näin syntyy virheitä, jotka toistuvat samankaltaisina; syntyy vääristymiä.

2.6. Edustavuus

Edustavuus (*Representativeness*) on kenties keskeisin ja tutkituin (Nisbett ym. 2002, 511) tyypin I prosessointiin kuuluvista heuristiikoista, eli niistä heuristisista ajatusmalleista, joita ihminen ajattelussaan hyödyntää. Edustavuusheuristiikka on prosessi, jossa todennäköisyyden arviointia koskevaan kysymykseen vastataan arvioimalla sitä, miten hyvin asiaan liittyvät mielikuvat vastaavat arvioijan odotuksia.

Klassinen demonstraatio koehenkilöiden turvautumisesta edustavuusheuristiikkaan on nk. Linda-kysymys. Tässä koehenkilöille annetaan lyhyt kuvaus keksitystä henkilöstä nimeltään Linda, ja pyydetään sitten arvioimaan, millä todennäköisyydellä Lindaa koskevat vastausvaihtoehdot pitävät paikkansa. Lindasta annettu kuvaus ei sisällä informaatiota, jonka perusteella vastausvaihtoehtojen todennäköisyyttä voisi objektiivisesti arvioida. Lindasta kerrotaan mm. hänen ikänsä, luonteenpiirteitä sekä hänen huolensa yhteiskunnallisista asioista ja ydinvoimasta:

"Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student she was deeply concerned with issues of discrimination and social justice and also participated in antinuclear demonstrations." (Kahneman 2003, 708-709; Tversky & Kahneman 1982b, 92)

Vastausvaihtoehtojen joukossa on kaksi olennaista: "Linda työskentelee pankissa" ja "Linda työskentelee pankissa ja on aktiivinen feministi". Koehenkilöt arvioivat jälkimmäisen vastausvaihtoehdon todennäköisemmäksi (Tversky & Kahneman 1982b, 93), vaikka todennäköisyyslaskennan lainalaisuuksien mukaisesti jälkimmäisen vaihtoehdon todennäköisyyden on (konjunkttiivisen todennäköisyyden vuoksi) pakko olla pienempi kuin ensimmäisen.

Lindasta annetun lyhyen kuvauksen perusteella koehenkilöistä ilmeisesti tuntui, että pelkkä pankissa työskenteleminen ei vastannut Lindasta annettua kuvausta. Tämä osoitettiin myös kysymällä toiselta ryhmältä koehenkilöitä, miten hyvin Lindasta annettu kuvaus vastaa tietyn vastausvaihtoehdon tyypillistä edustajaa. Tähän kysymykseen annetut vastaukset olivat lähes identtiset todennäköisyyttä koskevaan kysymykseen vastanneen ryhmän kanssa (Tversky & Kahneman 1982b, 92-93).

Hyvin ilmeisesti vastausvaihtoehto, jossa mainittiin feminismi, sopi paremmin yhteen sen mielikuvan kanssa, jonka Lindasta annettu kuvaus oli koehenkilöiden mielessä luonut. Tämä arkiselta vaikuttava toteamus sisältää vuosikymmeniä kestäneen tieteellisen keskustelun herättäneen väitteen: koehenkilöt eivät, tehtävänannosta huolimatta, arvioineetkaan todennäköisyyksiä vaan mielikuviansa keskinäistä yhteensopivuutta - edustavuutta.

Ihmiset eivät spontaanisti ymmärrä konjunktiiivisia todennäköisyyksiä, vaan heuristinen ajattelu tuottaa vastauksen hyödyntäen nopeasti arvioitavissa olevaa mielikuvien samankaltaisuutta eli edustavuusheuristiikkaa. Tästä aiheutuu samankaltaisena toistuvia heikkouksia ihmisen ajattelussa, ja tällä on erittäin tärkeitä implikaatioita sekä ennusteiden laadinnan että riskien arvioinnin kannalta. Yksityiskohtainen uhkaskenaario saattaa tuntua ihmisestä intuitiivisesti uskottavammalta, vaikka tosiasiasa spesifien yksityiskohtien lisääminen uhkakuvaan vähentää sen toteutumisen todennäköisyyttä (Yudkowsky 2008, 95).

Kyvyyttömyys ymmärtää konjunktiiivisia todennäköisyyksiä ei ole ainoa edustavuusheuristiikan soveltamisesta aiheutuva ongelma. Ironisesti nimetty, niin kutsuttu "pienten numeroiden laki" (*Law of small numbers*) on osa tätä samaa. Tällä viitataan siihen, että ihmiset luulevat populaation ominaisuuksien tulevan esiin pienissäkin otoksissa. Toisin sanoen, ihminen ei osaa intuitiivisesti huomioida pieneen otoskokoön liittyvää epävarmuutta. Eräässä kokeessa koehenkilöt arvioivat kolikolla heitettävien sarjojen todennäköisyyksiä. Sarjaa H-T-H-T-T-H² pidettiin todennäköisempänä kuin sarjaa H-H-H-H-T-H, koska jälkimmäinen edustaa huonommin populaatiota, jossa puolet on kruunaa ja puolet klaavaa; samoin ensimmäistä sarjaa pidettiin todennäköisempänä kuin sarjaa H-H-H-T-T-T, koska viimeksi mainittu sarja ei ihmisten mielissä edusta satunnaisuutta (Tversky & Kahneman 1982a, 7). Tosiasiasa kaikki sarjat ovat yhtä todennäköisiä. Tämä on hyvä esimerkki siitä, ettei edustavuusheuristiikkaa soveltava tyypin I prosessointi kykene huomioimaan itsenäisten

² Merkinässä H merkitsee kruunaa ja T klaavaa.

tapahtumien itsenäisyyttä. Ihmisen on vaikeaa spontaanisti ymmärtää satunnaisuutta käsitteenä.

Heuristinen ajattelu edustavuuden perusteella ei mahdollista myöskään taustatodennäköisyyksien huomioimista. Klassisen demonstraatio tästä on testi, jossa koehenkilöitä pyydetään arvioimaan, onko Jack insinööri vai juristi. Ennen vastaamista koehenkilöt saivat tiedon, että Jack kuuluu 100 ihmisen joukkoon, joista 30 on insinöörejä ja 70 juristeja, sekä kuvauksen Jackista ihmisenä. Jackista annettu kuvaus oli täynnä stereotyyppisesti insinööreihin liitettäviä luonteenpiirteitä, joista ei kuitenkaan pitäisi voida tehdä pitkälle vietyjä johtopäätöksiä. Kuitenkin todennäköisyysarviot siitä, onko Jack insinööri, olivat koehenkilöillä samankaltaiset kuin kontrolliryhmässä, jossa tehtävä oli muuten sama, mutta 100 ihmisen joukkoon kerrottiin kuuluvan 70 insinööriä ja 30 juristia. (Tversky & Kahneman 1982a, 5.)

Kun edellä kuvattua koetta toistettiin ja kehitettiin edelleen, havaittiin, että vain silloin, jos lisäinformaatiota ei ole lainkaan, ihmiset hyödyntävät taustatodennäköisyyksiä (Kahneman & Tversky 1982, 53). Jos taustatodennäköisyyksien lisäksi annetaan lisäinformaatiota, se vaikuttaa ihmisten arvioihin, vaikka ihmiset kokisivat lisäinformaation hyödyttömäksi. Klassisesta 70/30 kokeesta toistettiin versio, jossa annetusta kuvauksesta ei voinut päätellä yhtään mitään siitä, onko kyseessä insinööri vai juristi:

"Dick is a 30 year old man. He is married with no children. A man of high ability and high motivation, he promises to be quite successful in his field. He is well liked by his colleagues."
(Tversky & Kahneman 1982a, 5)

Hyödytön kuvaus sai ihmiset arvioimaan, että esimerkkihenkilön insinööriyden todennäköisyys on 50% (Kahneman & Tversky 1982, 56; Tversky & Kahneman 1982a, 5). Sellaisissakaan tutkimuksissa, joissa lisäinformaation hyödyttömyys tai epäluotettavuus on eksplisiittisesti tuotu esiin tehtävänannossa, koehenkilöt eivät osaa huomioida taustatodennäköisyyttä (Kahneman & Tversky 1982, 52).

On huomionarvoinen löydös, että ihmiset käyttäytyvät eri tavalla silloin, jos informaatio on hyödytöntä ja epäluotettavaa, kuin silloin, jos sitä ei ole lainkaan (Kahneman & Tversky 1982, 56). Ihmiset tukeutuvat helpommin kokemuksiin ja kertomuksiin kuin tilastollisiin perusteluihin (Stanovich 2011, 43).

2.7. Tavoitettavuus

Tavoitettavuudella, heuristisena ajatusmallina, tarkoitetaan sen arviointia, miten helposti jokin asia tulee mieleen. Tätä heuristiikkaa hyödyntäessään ihminen korvaa mielessään todennäköisyyttä koskevan kysymyksen kysymyksellä siitä, miten nopeaa tai vaivatonta vastausvaihtoehto oli keksiä tai muistaa. Tähän heuristiikkaan turvautuminen aiheuttaa kontekstiriippuvuutta sekä skaalaan liittyviä ongelmia. Kontekstiriippuvuuteen keskeisesti liittyvä ilmiö on kysymyksenasettelun vaikutus vastauksiin (*framing effect*). Kysymyksen ja vastauksen skaalaan liittyy mm. yhteismitallisten tietojen merkityksen ylikorostumista. Käytettävissä olevan informaation jotkut aspektit ovat helpommin intuitiivisesti käsitettävissä. Tähän liittyy tutkimuksissa osoitettu suhteellisten määrien liian suuri painotus absoluuttisiin määriin nähden (*proportion dominance*) sekä tunnottomuus suuruusluokkia kohtaan (*scope neglect*). Heuristisiin malleihin liittyvien ongelmien rajat eivät ole tarkkarajaisia, ja tavoitettavuusheuristiikan lisäksi vaikeus hahmottaa suuruusluokkaa liittyy tunnereaktioihin (affektiin).

Se, mitä tietoa vaihtoehtoista saadaan, ja miten ne esitetään (konteksti, ei välttämättä kenenkään tietoisesti rakentama) nostavat tietyt aspektit helposti tavoitettaviksi ja jättävät toiset aspektit piiloon (Kahneman 2003, 702). Saatavan informaation ja tuotetun vastauksen välillä on voimakkaampi yhteys, jos ne ovat yhteismitallisia. Esimerkiksi koehenkilöt saattavat samassa kysymyksessä valita mieluummin uhkapelin A, jossa on suurempi todennäköisyys, vaikka he arvioivat arvokkaammaksi uhkapelin B, jossa on suurempi palkinto (Slovic ym. 2007, 1337). Tämä rikkoo suoraan rationaalisen päätöksenteon teoriaa, ja on selitettävissä vain kognitiivisen psykologian kautta.

Ihminen on yllättävän huono hahmottamaan absoluuttisia lukumääriä; lukumäärät eivät tehokkaasti herätä spontaaneja reaktioita, ellei niillä ole voimakasta liityntää ihmisen aiempaan kokemukseen. Ilman kontekstia lukumäärät voivat olla vaikeasti hahmotettavia (Slovic ym. 2007, 1342). Lisäksi tunnereaktiot eivät skaalaudu lukumäärien suhteen. Esimerkiksi halu pelastaa x ihmishenkeä saattaa pohjimmiltaan syntyä helposti mieleen tulevasta mielikuvasta, joka liittyy yhden ihmisen pelastamiseen. Tätä reaktiota - ja siitä syntyvää motivaatiota - ei voi kertoa 200:lla tai 1000:lla (Yudkowsky 2008, 106). Ihmisen aivot eivät sovellu massamaisten tapahtumien kuvittelemiseen tai arvottamiseen.

Jos ihminen yrittää keksiä esimerkkejä lukumäärien arvioinnista, usein ensimmäisenä mieleen tulevat esimerkit liittyvät rahaan. Rahamääriin liittyvät esimerkit eivät välttämättä demonstroi parhaalla mahdollisella tavalla lukumäärien ymmärtämisen vaikeutta, koska rahamäärät yhdistyvät voimakkaasti ihmisen aiempaan elämäkokemukseen, ja niihin liittyy voimakkaita tunnereaktioita. Tässä mielessä mielenkiintoisempia ovat tutkimukset, joissa arvioitaviin lukumääriin ei arkikokemuksen perusteella ole muodostunut samanlaista valmista arviointiasteikkoa.

Eräässä tutkimuksessa koehenkilöitä pyydettiin kertomaan, kuinka kannatettavana he pitäisivät tehtävänannossa esitettyä parannusta lentokenttien turvallisuuteen. Kolmelle eri ryhmälle kerrottiin, että muutos estäisi "150 kuolonuhria", "85% 150:stä kuolonuhrista", ja "98% 150:stä kuolonuhrista". Ihmisten halukkuus pelastaa ihmishenkiä oli enemmän riippuvainen pelastettavien suhteellisesta osuudesta kuin absoluuttisesta määrästä. "98%" sai voimakkaimman kannatuksen, ja myös mahdollisuus "estää 85% 150:stä kuolonuhrista" oli suositumpi kuin "150:n kuolonuhrin" estäminen. (Slovic ym. 2007, 1341.) Tunnettomuus kertaluokalle vain pahenee, kun luvut suurenevat. Ihmisen on käytännössä hyvin vaikea erottaa suurten lukujen suuriakaan eroja. Miljoona pakolaista tai kymmenen miljoonaa pakolaista eivät oikeastaan tunnu erilaisilta. (Yudkowsky 2008, 106.)

Toisessa tutkimuksessa koehenkilöt arvioivat kahta sanakirjaa: uuden veroista pienempää sanakirjaa, ja laajempaa sanakirjaa, jossa oli revennyt kansi. Ryhmä, joka arvioi sanakirjoja rinnakkain, arvioi laajemman sanakirjan arvokkaammaksi. Verrokkiryhmät, jotka arvioivat vain jompaakumpaa, arvioivat paremmassa kunnossa olevan kirjan keskimäärin arvokkaammaksi. (Slovic ym. 2007, 1339.) Arviointi on riippuvainen kontekstista. Sitä, onko 10000 sanaa sanakirjassa paljon vai vähän, on vaikea arvioida ilman kontekstia.

2.8. Kysymyksenasettelun vaikutus vastauksiin

Se, että kysymyksen asettelulla ja näkökulman valinnalla voidaan ohjata vastausta haluttuun suuntaan, ei sinänsä kaipaa pitkällistä todistelua. Tämä on yleisesti hyväksytty totuus, ja ihmisille yleisen elämäkokemuksen myötä intuitiivisesti selvää, jopa siinä määrin, että sitä hyödynnetään arkisessa sosiaalisessa kanssakäymisessäkin täysin rutiininomaisesti. Oikeastaan on intuitiivisesti vaikeampi asia tavoitella objektiivista vastausta, jos halutaan, että kysymyksenasettelu ei vaikuttaisi vastaukseen. Tutkimuksessa on kuitenkin pyritty kuvaamaan yksityiskohtaisemmin sitä, missä määrin konteksti saa ihmisen rikkomaan rationaalisuutta ja, miten tämä yksityiskohtaisesti tapahtuu.

Ihminen ei spontaanisti tai automaattisesti muokkaa saamaansa informaatiota toiseen muotoon, vaan hyväksyy passiivisesti sen muotoilun (*framing*), jonka hän kohtaa; ihminen ei esimerkiksi vierekkäisistä palikoista automaattisesti laske, kuinka korkean tornin niistä saisi (Kahneman 2003, 703). Lisäksi tietyn vaihtoehdon hylkääminen vaikuttaa olevan eri ajatusprosessi kuin tietyn vaihtoehdon hyväksyminen. Se, miten koehenkilöt arvioivat oman preferenssinsä, on hyvin herkkä sille, miten vaihtoehdot kuvataan heille. Ihminen hyödyntää positiivisena pitämäänsä informaatiota päättäessään valita tietyn vaihtoehdon, ja negatiivisena pitämäänsä informaatiota päättäessään hylätä tietyn vaihtoehdon. (Shafir 1993, 546-547.) Arvioinnin helppouteen vaikuttaa suuresti myös sisään tulevan informaation (vastausvaihtoehdon kuvaus) ja ulos annettavan informaation (missä yksikössä vastaus pitää antaa) yhteismitallisuus (sama, 546).

Kuuluisassa esimerkissä koehenkilöille kerrotaan, että Yhdysvaltoihin on leviämässä harvinainen, aasialainen tauti, jonka oletetaan vaativan jopa 600 ihmisen hengen. Epidemian torjumiseen on kaksi vaihtoehtoista suunnitelmaa, "A" ja "B". Ensimmäiselle ryhmälle kerrotaan, että suunnitelma A johtaa 200 ihmishengen pelastumiseen, ja epävarmempi vaihtoehto B tarjoaa 1/3 mahdollisuuden pelastaa kaikki 600 ihmistä ja 2/3 mahdollisuuden siihen, että kukaan 600:sta potilaasta ei pelastu. Toiselle ryhmälle koehenkilöitä vaihtoehtojen muotoilu muutetaan niin, että suunnitelma A johtaa 400 ihmisen kuolemaan, kun taas suunnitelma B voi 1/3 todennäköisyydellä estää kaikki kuolonuhrit, mutta pitää sisällään 2/3 mahdollisuuden sille, että kaikki 600 ihmistä kuolevat. (Tversky & Kahneman 1981, 453.)

Kun molemmat muotoilut esitetään rinnakkain, on helppo huomata, että vaihtoehdot "A" ja "B" ovat molemmilla ryhmillä täsmälleen samanlaiset. 400:n ihmisen varma kuolema ei kuitenkaan tunnu ihmisestä hyväksyttävältä, toisin kuin 200 ihmisen varma pelastuminen. Alkuperäisessä tutkimuksessa ensimmäinen vastaajaryhmä valitsi 72-prosenttisesti suunnitelman A, kun taas toisesta vastaajaryhmästä 78% kannatti suunnitelmaa B (Tversky & Kahneman 1981, 453). Tutkimustuloksessa on huomionarvoista se, että vastausten muuttumiseen liittyy vastaajien siirtyminen riskin välttämisestä riskin ottamiseen, joka on vastoin rationaalisen päätöksenteon malleja. On tyypillistä, että päätöksissä, joihin liittyy jotain saavutettavaa, ihminen välttää riskiä; päätöksissä, joihin liittyy potentiaalisia menetyksiä, ihmiset ovat alttiimpia ottamaan riskejä (Tversky & Kahneman 1981, 453; LeBoeuf & Shafir 2003, 72).

Eräässä tutkimuksessa pyydettiin puolta koehenkilöistä valitsemaan yksi kolmesta arvonnasta, ja toista puolta koehenkilöistä hylkäämään kaksi kolmesta. Vaihtoehdot olivat:

1. 50% mahdollisuus voittaa \$50,
2. 60% mahdollisuus voittaa \$40, ja
3. 80% mahdollisuus voittaa \$150 dollaria ja 20% mahdollisuus hävitä \$20.

Ensimmäisestä ryhmästä 61% valitsi arvonnasta numero 3. Toisesta ryhmästä 56% hylkäsi arvonnasta numero 3 (ja jommankumman toisista vaihtoehdoista); sama vaihtoehto oli siis "sekä paras että huonoin" (Shafir 1993, 552). Arvonnasta numero 3 valitsi enemmistö niistä, joita pyydettiin valitsemaan, ja hylkäsi enemmistö niistä, joita pyydettiin hylkäämään. (sama.)

Samankaltaisia tutkimustuloksia on paljon. Esimerkiksi mielipidemittausten tulosten on osoitettu olevan täysin erilaiset, jos osallistujilta kysytään: "Kumpaa äänestäisit?" kuin, jos heiltä kysytään: "Kumpaa et äänestäisi?" (Shafir 1993, 552). Ensimmäinen kysymys saa vastaajat vertailemaan ehdokkaiden hyviä puolia, kun taas jälkimmäinen saa heidät vertailemaan ehdokkaiden huonoja puolia.

2.9. Ankkuroituminen

Referenssipisteeseen kiinnittyminen ja halu välttää tappioita (saavutetun aseman heikkenemistä) selittävät monia ihmisen tekemiä valintoja (Kahneman 2003, 705). Edustavuuden ja tavoitettavuuden ohella kolmas alkuperäisistä, jo 1970-luvulla kuvatuista heuristisista päätöksentekomalleista on nk. ankkuroituminen ja korjaus (*Anchoring and adjustment*). Tämä heuristinen ratkaisumalli toimii siten, että ihminen aloittaa vastauksen arvioinnin jostain alkuarvosta, arvioi vastauksensa uskottavuutta, ja alkuarvosta käsin lähtee korjaamaan epäuskottavana pitämäänsä vastausta, kunnes vastaus tuntuu mahdolliselta. Näin menetellen mahdollinen vastaus on tyypillisesti ensimmäinen vastaus, joka tuntuu jollain tasolla mahdolliselta. Tämän ajattelumallin seurauksena lähtöarvosta (ankkurista) tehty korjaus on tyypillisesti liian pieni.

Ankkuroituminen on osa ilmiötä, jota voidaan laajemmin nimittää kontaminaatioksi. Tutkimustulokset tukevat käsitystä, jonka mukaan lähes mikä tahansa informaatio voi vaikuttaa kognitiiviseen päätöksentekoon (Yudkowsky 2008, 102). Kuten jo aiemmin todettu, ihmiset reagoivat eri lailla silloin kun käytössä ei ole mitään dataa, kuin silloin, jos käytössä on ilmeisenkin hyödyttömiä ja epävalidia dataa (Kahneman & Tversky 1982, 56).

Eräässä tutkimuksessa koehenkilöitä pyydettiin arvioimaan, montako Afrikan valtiota on YK:n jäseniä. Ennen kysymystä koehenkilön nähden arvottiin onnenpyörästä luku 1:n ja 100:n välillä. Koehenkilöiden vastaukset korreloivat voimakkaasti onnenpyörästä arvottujen lukujen kanssa. Esimerkiksi arvotun numeron 10 nähneiden koehenkilöiden mediaanivastaus oli 25, kun taas onnenpyörästä luvun 65 nähneiden mediaanivastaus oli 45. (Tversky & Kahneman 1982a, 14.)

Koehenkilöt ankkuroituivat arvottuun numeroon, vaikka he tiesivät sen satunnaiseksi ja täysin asiaan liittymättömäksi. Kun onnenpyöräkin vaikuttaa, ei liene yllätys, että myös fiktiivinen aineisto vaikuttaa ihmisten arvioihin, vaikka fiktiivisyys olisi arvioijan tiedossa. Jos fiktiivinen aineisto on relevanttia ja aihepiiriin liittyvää, sillä on vaikutusta (Yudkowsky 2008, 103). Riskianalyysin kannalta tällaista relevanttia aineistoa on paitsi uhkakuvista kertovat kirjat ja elokuvat, myös mm. aiemmin samoista aiheista tehdyt riskianalyysit. Ennusteiden laadinnassa ankkuroituminen on niin ikään huomioon otettava tekijä, jos esimerkiksi ennustettavalle ajanjaksolle toteumasta on jo aiemmin tehtyjä ennusteita, ja uusi ennuste tehdään päivittämällä aiempaa.

2.10. Tunnepohjaiseen arviointiin liittyvät vääristymät

Se, että tunnetilat vaikuttavat päätöksiin, on populaaristi tunnettu asia. Yleisessä käsityksessä tunteiden vaikutuksesta on kuitenkin merkittävä ero siihen asemaan, joka tunteille kognitiivisen psykologian piirissä, *Dual-Process* -teorioissa annetaan. Ensin mainitussa yhteydessä tunteet nähdään rationaalisuuden ulkopuolisena asiana, ja niiden laajasti popularisoitu vaikutus inhimillisenä heikkoutena; jonain sellaisena, joka ei saisi vaikuttaa objektiiviseen päätöksentekoon silloin, kun se tehdään oikein. *Dual-Process* -teoriassa tunnereaktiot hyväksytään rationaalisen päätöksenteon osana, osana tyypin I autonomista prosessointia, ja niiden osuutta pyritään kuvaamaan yksityiskohtaisemmin.

Tunnereaktioiden vaikutus päätöksentekoon on pääosin positiivinen. Kuten kuvattu edellä, on jotain näyttöä siitä, että tunnereaktioiden avulla ihminen kykenee nopeasti hyödyntämään aiempaa kokemustaan. Tunnereaktiot, kuten muukin autonominen prosessointi, saattavat kuitenkin joskus johtaa ihmisiä harhaan. (Slovic ym. 2007, 1347.) Tavoitettavuusheuristiikan syyksi pistetyt vääristymät todennäköisyyksien ja frekvenssien arvioinnissa voivat osittain selittyä tunnereaktioilla (sama, 1345).

Tunnereaktioihin perustuvien arvioiden heikkouksista yhtenä erimerkkinä on jo aiemmissa luvuissa mainittu kyvyttömyys huomioida kertaluokkia (*scope neglect*, joissain yhteyksissä *extension neglect*). Varovasti voidaan arvioida, että lukujen kasvaessa tunnereaktio voimistuu vain hyvin suuren kokoluokan muutoksen seurauksena, jos lainkaan. Eräässä tutkimuksessa haastateltavat olivat valmiita maksamaan keskimäärin saman suuruisen summan yhden, nimetyn, Ontarion osavaltion järven vesien puhdistuksesta, kuin kaikkien Ontarion osavaltion järvien vesien puhdistuksesta (Kahneman 2003, 713).

Sellaisen toiminnan arvottamiseen, jossa kyseessä ovat ihmishenget (kuten meripelastuksen suorituskykyjen rakentaminen tai käyttö), tämä sopii moraalisesti. Yleisesti katsotaan, että ihmishengelle ei pidäkään voida asettaa arvoa, ja yhdenkin ihmishengen pelastamiseen pitää suunnata kaikki käytettävissä olevat resurssit. Kuitenkin on päätöksenteon psykologian kannalta hyödyllistä ymmärtää, että ihminen kuvittelee mielessään yhden hukkuvan tai palavan uhrin, ja arvottaa tätä mielikuvaa. Ihminen ei osaa kertoa tätä tuskaa viidellä tai kahdellakymmenellä tai varsinkaan tuhansilla.

Tunnereaktioilla on myös muita, vaikeammin hahmotettavia ja huonommin tunnettuja vaikutuksia. Esimerkiksi eräässä tutkimuksessa koehenkilöt kokivat paljon houkuttelevammaksi uhkapelin, jossa oli mahdollisuus hävitä vähän rahaa, kuin muuten identtisen pelin ilman tappion mahdollisuutta (Slovic ym. 2007, 1338). Tiedetään lisäksi, että pelkästään saman asian näkeminen monta kertaa riittää luomaan positiivisen miellelyhtymän (sama, 1336).

Tunnereaktiot edustavat kaikkein nopeinta prosessointia. Myös tunnereaktioiden oppiminen ja liittäminen asioihin on nopeaa. Tästä on esimerkkinä tutkimus, jossa kiinaa osaamattomille koehenkilöille näytettiin kiinalaista kirjoitusmerkkiä, jota edelsi 1/250 sekunnin ajan heidän tiedostamattaan näytetty hymyilevä tai irvistävä kasvokuva. Koehenkilöt arvioivat pitävänsä hymyileviä kasvoja seuranneista kirjoitusmerkeistä enemmän. Vaikka myöhemmin järjestystä vaihdettiin siten, että merkit saivat eteensä vastakkaista tunnetta ilmaisevat kasvot, tällä ei enää ollut merkitystä. (Slovic ym. 2007, 1336.) Toisin sanoen miellelyhtymä syntyi salamannopeasti ja oli pysyvä.

2.11. Vääristymien kompensoiminen ja välttäminen

Heuristisen ajattelun aiheuttamat kognitiiviset vääristymät vaikuttavat ihmisissä universaalisti. Vääristymien ilmenemisessä on kuitenkin eroja, koska joskus ihmiset kykenevät hylkäämään virheellisen heuristisen vastauksen ja tuottamaan oikean vastauksen (Stanovich & West 2011, 122). Heuristisen ongelmanratkaisun ensimmäistä askelta, kontekstualisointia, vastaan yksilöllä suojaa taipumus riisua turha konteksti ratkaistavan ongelman ympäriltä (Stanovich 2011, 44).

Jos ihminen tulee tietoiseksi heuristisesta ajattelusta, hän on valmiimpi kompensoimaan (Kahneman 2003, 711). Kontaminaatiota on pyritty tutkimuksissa poistamaan maksamalla oikeista vastauksista, kertomalla heuristisesta ajattelusta, neuvomalla koehenkilöitä olemaan ankkuroitumatta ja käsittelemällä tutkimustehtävissä todellisen maailman oikeita ongelmia (Yudkowsky 2008, 102). Muotoilun ja kontekstin vaikutusta on esitetty vähentävän vaatimus käyttää aikaa tehtävän ratkaisemiseen sekä esittää perustelut valinnalle (LeBoeuf & Shafir 2003, 79). Aika ja perustelutarve saavat tämän hypoteesin mukaan ihmisen etsimään vaihtoehtoisia vertailukohtia, mikä vähentää esitetyn muotoilun vaikutusta.

Taloudellisten kannustimien on todettu parantavan ihmisten keskittymistä, ja saamaan koehenkilöt yrittämään kovemmin. Kuitenkaan ei ole kyetty saamaan yhtään toistettavissa olevaa tulosta, jossa heuristisen ajattelun aiheuttamat vääristymät saataisiin häviämään tarjoamalla koehenkilöille palkintoja. (Gilovich & Griffin 2002, 4; Camerer & Hogarth 1999, 7) Vastaamiseen käytettävän ajan on huomattu pitävän olla merkittävästi pidempi, jotta vaikutus vastausten laatuun olisi tilastollisesti merkitsevä (LeBouef & Shafir 2003, 79). LeBouefin ja Shafirin (2003, 80) mukaan vääristymien kompensoimisesta kirjallisuudessa esitetty näyttö on ristiriitaista. Heidän mukaansa (sama, 85) ei ole mitään näyttöä siitä, että pitempi harkinta tai vaatimus perusteluista vähentäisivät kysymyksenasettelun vaikutusta. Kahnemanin (2003, 711) mukaan tutkimuksissa esillä olleet korjausmekanismit liittyvät tavoitettavuuteen, eivätkä toimiessaankaan tarjoaisi keinoa välttää edustavuusheuristiikan aiheuttamia ongelmia. Yudkowskyn (2008, 102) mukaan mistään ei ole ollut juuri apua.

Ongelma erityisesti rahallisen palkitsemisen suhteen on, etteivät heuristiset vääristymät aiheudu ahkeruuden, keskittymisen tai kiinnostuksen puutteesta, vaan siitä, miten ihmisen aivot ja ajattelu toimivat. Lisäksi voisi argumentoida, että koetehtävät, joissa tarjotaan suhteessa korkeita palkintoja, ovat siitä näkökulmasta epärealistisia tai ainakin epäedustavia, että ihminen joutuu tekemään joka päivä satoja ja taas satoja pieniä arvioita ja päätöksiä, joita ei mitenkään voi kaikkia priorisoida. Vastausajan suhteen vaikuttaa olevan riidatonta vain se, että aikapaine lisää ongelmia ainakin kontaminaation (Yudkowsky 2008, 103) ja affektin (Finucane ym. 2000, 8) osalta.

Tutkimuksiin osallistuvat henkilöt tyypillisesti itse arvioivat, ettei heidän ajattelunsa ole vääristynyttä. Koehenkilöiden oman arvion mukaan esimerkiksi kontaminoiva tieto (kuten onnenpyörä) ei ole vaikuttanut heihin (Yudkowsky 2008, 103). Ilmeisesti odotettavissa olevaan epävarmuuteen liittyy intuitioita, joita ei voi muuttaa parilla tilastotieteen kurssilla (Kahneman & Tversky 1982, 68). Kuitenkin kehotukset ottaa tilastolliset tekijät huomioon, relevantti koulutus ja monitorointijärjestelmien herättely oikeanlaisilla vihjeillä parantavat tyypin II prosessoinnin mahdollisuuksia korjata heuristisia virheitä, mikä parantaa koehenkilöiden laatimien ennusteiden tarkkuutta (Kahneman 2003, 711; Mellers ym. 2014, 1109).

Tyyppin I prosessointiin liittyvistä tutkimuksissa dokumentoiduista virheistä ja vääristymistä ei voi päätellä, etteikö tyyppin II prosessointiin niin ikään liittyisi virheitä ja vääristymiä. Tutkimuksissa on tehty ajoittain sellaisiakin havaintoja, että kokeneet ammattilaiset suoriutuvat paremmalla tasolla silloin kun he luottavat intuitioonsa (Kahneman 2003, 699). Vaikka koehenkilön tutkimustehtävässä antama vastaus ei olisi tiukan määritelmän mukaan rationaalinen, voi heuristinen ajattelu silti olla optimaalinen tapa prosessoida käsillä olevaa tilannetta, kun huomioidaan aika- ja resurssirajoitteet (Stanovich 2011, 8).

2.12. Älykkyyden ja koulutuksen merkitys

Keinoiksi välttää ajattelun vääristymiä on yleisimmin esitetty älykkyyttä, koulutusta, ja erilaisia yksityiskohtaisempia tekniikoita kuhunkin ongelmaan erikseen. Vaikka heuristisen ajattelun tutkimuksissa suurin osa ihmisistä poikkeaa teorian olettamalla tavalla rationaalisista vastauksista, silti yksilöllinen vaihtelu on suurta, ja usein osa koehenkilöistä kykenee antamaan normatiivisen rationaalisen vastauksen (Stanovich 2011, 13). Yksilöllisiä eroavaisuuksia heurististen virheiden välttämiseksi eivät selitä älykkyydestein mitattavat erot; kyse on reflektiivisen tason eroista, eroista ajattelutavoissa (Stanovich 2011, 44; Stanovich & West 2011, 121-122). On esitetty, että älykkyyys tai taipumus huolelliseen, analyyttiseen ajatteluun ja harkintaan vähentäisivät alttiutta mm. muotoilun ja kysymyksenasettelun vaikutuksille (LeBouef & Shafir 2003, 79). Tällainen hypoteesi on aivan ilmeinen, ja voidaan esittää koskien intuitiivisen ajattelun vääristymiä yleisemminkin. Tutkimustulokset ovat kuitenkin ristiriitaisia.

Älykkyydestein mitattavat erot ovat eroja algoritmisen tason kognitiivisessa kapasiteetissa, eivätkä yleisesti selitä yksilöiden eroja reflektiivisessä ajattelussa (Stanovich 2011, 36, 38). Toisin sanoen, älyllisen kapasiteetin omaaminen ja alttius sen hyödyntämiseen ovat kaksi eri asiaa. Älykkyydelle on olemassa myös laajoja määritelmiä, jolloin termi laajenee merkitsemään myös muita rationaalisuuden osa-alueita kuin algoritmista suorituskkyä, mutta nämä laajat määritelmät eivät vastaa sitä tapaa, jolla älykkyyttä mitataan kokeissa ja testeissä (Stanovich & West 2011, 121). Tämä selittää osaltaan sitä, miksi älykkyyys ei suojaa ihmisiä heuristisen ajattelun virheiltä. Testein mitattava älykkyyys ei korreloi uteliaisuuden, huolellisuuden, avoimuuden, velvollisuudentunnon tai muiden keskeisesti ihmisen rationaaliin ratkaisuihin vaikuttavien ajattelutapojen erojen kanssa (Stanovich 2011, 38). Reflektiivisten ajattelutapojen lisäksi älykkyyys ei pidä sisällään monia muitakaan

rationaaliseen toimintaan tarvittavia ei-kognitiivisia, sosiaalisia ja emotionaalisia kykyjä (Stanovich & West 2011, 122).

Osa vääristymistä älykkyys kuitenkin poistaa (Stanovich & West 2011, 139). Vaikuttaa siltä, että älykkyydestä on hyötyä silloin, kun tehtävät eivät ole liian vaikeita; jos tehtävät ovat melkein kaikille osallistujille liian vaikeita, älykkyydestä voi olla jopa haittaa (Kahneman & Frederik 2002, 68). Älykkäät ihmiset suoriutuvat koetehtävistä muita paremmin, jos heille kerrotaan (tai koejärjestely sisältää viitteitä), mikä vääristymä on kyseessä ja, miten sitä voi välttää. Ilman tällaisia apuja erot kuitenkin häviävät. (Stanovich & West 2011, 153.) Joissain olosuhteissa älykkyys saattaa heikentää heuristisen ajattelun kontrolloimisen tarpeen tunnistamista - älykkäät ihmiset saattavat *luulla* ajatustensa olevan vähemmän vääristyneitä (sama, 134-135).

Reflektiivisen tason yksilöllisiä eroja on pyritty mittaamaan erilaisilla kriittisen ajattelun testeillä, ja näillä on saatu esiin merkitseviä korrelaatioita rationaalisen ajattelun tehtävissä havaittavien virheiden kanssa. Esimerkiksi ihmiset, jotka kriittisen ajattelun testeissä ovat mieleltään avoimempia uudelle informaatiolle, hyödyntävät tilastollista, populaatiotason tietoa paremmin kuin muut, vaikka älykkyystestissä mitattua älykkyyttä käytettäisiin kontrollimuuttujana (Stanovich 2011, 44). Kyvyttömyys sietää viivytystä (*intolerance to delay*) korreloi merkitsevästi kognitiivisissa koetehtävissä epäonnistumisen kanssa (Kahneman 2003, 699).

Reflektiivisen tason kyvytkään eivät täysin suojaa vääristymiltä. Yksilön taipumusta huolelliseen, analyttiseen ajatteluun on mitattu *Need for Cognition* (NC) -asteikolla (LeBouef & Shafir 2003, 79). Eräässä tutkimuksessa samoille koehenkilöille esitettiin samoista *framing effect* -tutkimuksessa käytetyistä testikysymyksistä molemmat versiot, osana tunnin kestänyttä tutkimuskyselyä. Korkeammat NC-pisteet saaneet koehenkilöt vastasivat *johdonmukaisemmin*, mutta osoittivat *framing*-vaikutuksen sen mukaan, kumman muotoilun olivat nähneet ensin. (sama, 86-87, 89.) Huolelliseen, analyttiseen ajatteluun taipuvaiset ihmiset kykenivät siis paremmin tunnistamaan kysymysten yhtäläisyyden, ja pyrkivät vastaamaan johdonmukaisesti, mutta tämä ei suojannut heitä kysymyksenasettelun vaikutukselta.

Koulutuksen kautta ihminen voi oppia myös tilastollisia nyrkkisääntöjä, kuten "älä luota pieneen otokseen", mutta ilman oikeanlaisia herätteitä ja muistutuksia nämä tilastolliset säännöt eivät ole helposti tavoitettavissa (Kahneman 2003, 711). 1970-luvulta alkaen on toistuvaa näyttöä siitä, että tilastotieteellisen koulutuksen saaneet tutkijat eivät kykene intuitiivisesti noudattamaan tilastollisen todennäköisyyden peruseriaatteita, jotka he hyvin tuntevat (sama, 697). Koulutetuilla ihmisillä on enemmän työkaluja argumenttien kumoamiseen, mutta vain silloin kun heillä on siihen motivaatiota, minkä vuoksi koulutetut, kriittiset ihmiset voivat olla jopa keskimääräistä huonompia vaihtamaan mielipidettään (Yudkowsky 2008, 99).

Yhteenvedon voidaan todeta, että hyödylliset ajattelutavat suojaavat ihmistä heuristisilta vääristymiltä paremmin kuin älykkyys tai koulutus. Tällaisia hyödyllisiä reflektiivisen tason ominaisuuksia ovat mm. huolellisuus, uteliaisuus, kriittisyys, avoimuus uutta informaatiota kohtaan, kyky sietää epävarmuutta ja viivytystä, taipumus analysoida asioita ja halu ajatella itse. Kukaan ihminen ei kuitenkaan ole täysin immuuni heuristisen ajattelun aiheuttamille vääristymille, varsinkaan jos tehtävä on vaikea, aikaa vähän ja häiriötekijöitä paljon.

2.13. Heuristisen ajattelun merkitys ennusteiden laadinnan kannalta

Jos edellä esitetyt väitteet heuristisesta ajattelusta pitävät paikkansa, ja ihminen ei spontaanisti kykene ymmärtämään satunnaisuutta, pieneen otokseen liittyvää epävarmuutta, konjunkttiivisia todennäköisyyksiä, eikä käytössä olevan informaation epäluotettavuutta, tällä on voimakkaita implikaatioita ennusteiden laatimisen suhteen. Ennusteiden laadinnan näkökulmasta nämä tyypin I prosessoinnin heikkoudet aiheuttavat sen, että toteutunutta todellisuutta ei osata nähdä satunnaisvaihtelun tuloksena edes osittain, käytössä olevan informaation huonoa ennustearvoa ja datan vähäistä määrää ei osata huomioida, eikä näin ollen osata siirtyä kohti regressiivisempää ennustetta silloin, kun ennustettavuus on huono.

Yksinkertaisia vastauksia monimutkaisiin ongelmiin tuottavia heuristisia ajattelumalleja on kirjallisuudessa kuvattu useita. Jo ensimmäisenä kirjallisuudessa kuvatut, alkuperäiset kolme heuristiikkaa - edustavuus, tavoitettavuus ja ankkuroituminen (Kahneman 2003, 707) - selittävät useita ennusteisiin ja riskien arviointiin liittyviä vääristymiä, kuten: epäregressiiviset ennusteet, taustatodennäköisyyksien huomioimatta jättämisen, liiallisen luottamuksen, ja helposti mieleen tulevien tai hyvin yksityiskohtaisten skenaarioiden todennäköisyyden yliarvioinnin. Seuraavassa pyritään avaamaan tätä problematiikkaa hieman lisää.

Ihmisellä on primitiivinen tarve yrittää hallita ympäristöään, ja tähän liittyvä alitajuinen usko siihen, että se on mahdollista. Vaikka ihmiset puhuessaan osoittaisivat tiedostavansa satunnaisuuden, he useissa tilanteissa silti käyttäytyvät kuin voisivat kontrolloida kaikkia lopputulokseen vaikuttavia tekijöitä (Langer 1982, 231). Ihmiset eivät tyypillisesti ymmärrä ääritulosten epätodennäköisyyttä, ja lisäksi ihminen ennustaa tulevaa kuin sisään tuleva informaatio olisi täysin virheistä vapaata (Kahneman & Tversky 1982, 67).

Eräs tilastollisten todennäköisyyksien periaate, joka ei ole ihmiselle intuitiivisesti selvää (eli tyypin I prosessointi ei kykene sitä ymmärtämään) on paluu kohti todennäköisyysmassan keskiarvoa (*regression towards the mean*). Eräässä tutkimuksessa lennonopettajat kertoivat huomanneensa, että kun oppilasta kehuu poikkeuksellisen hyvin onnistuneen harjoitteen jälkeen, hänen seuraava suorituksensa on huonompi (Kahneman & Tversky 1982, 67-68). Tämän havainnon vuoksi opettajat olivat ottaneet tavaksi olla kehumatta hyviä suorituksia. Tutkimuksen seuraavassa vaiheessa lennonopetuksen kehujen problematiikkaa esiteltiin koehenkilöille ja pyydettiin heitä arvioimaan, mikä voisi selittää lennonopettajien havainnot kehujen vaikutuksista. Yksikään tutkimuksen koehenkilöistä, jotka olivat tilastotieteellistä koulutusta saaneita yliopiston jatko-opiskelijoita, ei osannut ehdottaa, että selitys voisi olla paluu kohti todennäköisyysmassan keskiarvoa (sama, 68); eli, että suoriuduttuaan omaan tasoonsa nähden poikkeuksellisen hyvin, oppilaan seuraava suoritus on jo pelkästään tilastollisen todennäköisyyden vuoksi hyvin todennäköisesti lähempänä oppilaan normaalia suoritustasoa - eli huonompi kuin huippusuoritus.

Toisessa tutkimuksessa kerrottiin henkilön X saaneen älykkyysosamäärää mittaavasta testistä tuloksen 140. Koehenkilöitä pyydettiin arvioimaan 95% luottamusrajat henkilön todelliselle älykkyysosamäärälle. Vastaukset olivat symmetrisiä 140:n ympärillä. (Kahneman & Tversky 1982, 57.) On siis ilmeisesti niin, että tulevaisuuteen ymmärretään liittyvän epävarmuutta, mutta toteutunutta todellisuutta on erittäin vaikea ymmärtää satunnaisvaihtelun tulokseksi.

Satunnaisuus on ihmiselle intuitiivisesti erittäin vaikeaa hahmottaa. Erityisesti tämä koskee jo toteutuneen todellisuuden satunnaisuutta. Silloinkin, kun ilmiöön ymmärretään liittyvän epävarmuutta, ihminen ilmeisesti spontaanisti kuvittelee epävarmuuden sijoittuvan symmetrisesti toteutuneen todellisuuden ympärille. Ihminen saattaa käsittää, että menneisyyteenkin liittyi epävarmuutta, mutta toteutunut todellisuus vaikuttaa usein silti jälkikäteen arvioiden todennäköisimmältä vaihtoehdolta. Ajatus siitä, että toteutunut todellisuus olisi ollut satunnaisvaihtelun tulos, joka poikkesi – ehkä merkittävästikin – *ex ante* oletusarvosta (todennäköisyysmassan keskipisteestä), on ihmiselle äärimmäisen vaikea. Tämän kaltaisen käsityksen muodostaminen on täysin tyypin II prosessoinnin varassa; heuristisen ajattelun avulla ihminen ei kykene lainkaan hahmottamaan informaatiota tällä tavoin.

Ihmisen on vaikeaa nähdä vaihtoehtoisia todellisuuksia (Kahneman 2003, 702). "Koska ihmiset luulevat, perusteetta, että menneisyys olisi ollut ennustettavissa, he päättävät virheellisesti, että tulevaisuuskin on ennustettavissa" (Yudkowsky 2008, 94). Taustatodennäköisyydet voi perustellusti sivuuttaa, jos saatu lisäinformaatio on tarkkaa ja diagnostista (eli sillä on ennustearvoa) (Kahneman & Tversky 1982, 51). Mutta, jos toteutunutta todellisuutta koskeva informaatio on muodostunut osittain sattuman tuloksena, se ei ole tarkkaa, eikä diagnostista.

Ennusteen laadintaan tarvitaan kolmenlaista informaatiota: aiempi taustalla oleva informaatio, käsillä olevaa tapausta nimenomaisesti koskeva lisäinformaatio ja odotusarvo ennusteen tarkkuudesta. Mitä heikompi ennustettavuus on, sitä pienempi paino pitäisi antaa spesifille lisäinformaatiolle. Ihmisen intuitio ei kuitenkaan toimi näin. (Kahneman & Tversky 1982, 51.) Ongelma on kahtalainen: ennen toteumaa vaihteluvälit arvioidaan liian pieneksi (liiallinen luotto arvioiden, ennusteiden tai diagnoosien osuvuuteen), ja toteuman jälkeen ennustettavuus jälkikäteisesti yliarvioidaan (Yudkowsky 2008, 93).

Tämä aiheutuu siitä, että ihminen tyypillisesti ei *laske* ennusteita päässään kuin kone, vaan muodostaa ennusteet heuristisesti. Ennusteen laadintaan vaikuttaa ehkä tärkeimpänä edustavuusheuristiikka. Ihminen valitsee ennusteen, joka hänen mielestään *parhaiten edustaa* käytössä olevia tietoja, eli on samankaltainen ja yhteensopiva niiden kanssa. (Kahneman & Tversky 1982, 48.)

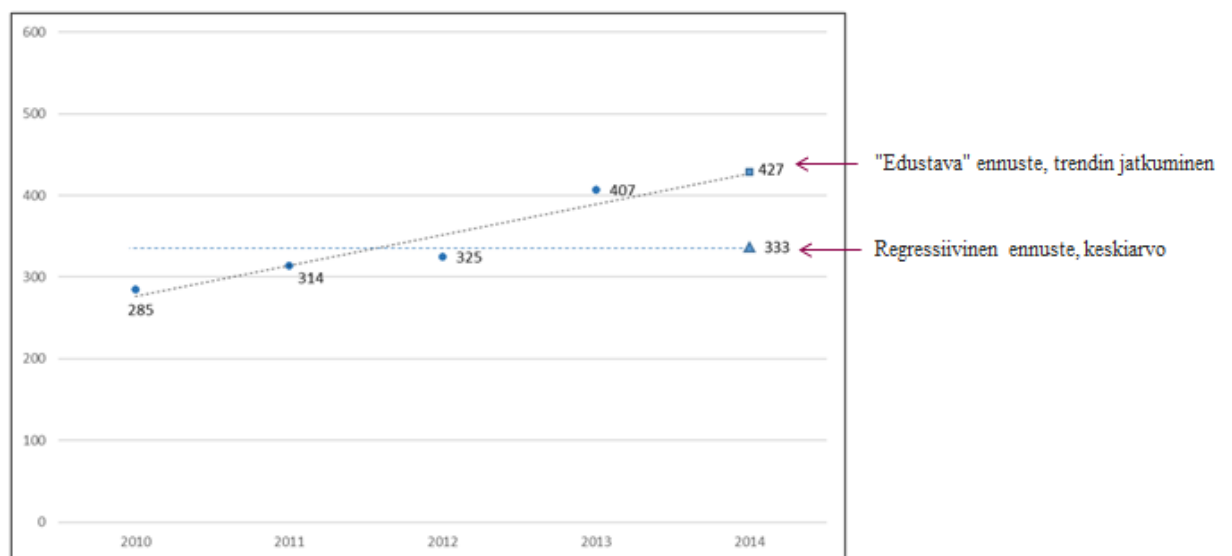
Ennusteiden laadinnassa taustatodennäköisyyksien huomioimatta jättäminen on merkittävin heuristisen ajattelun aiheuttama poikkeama normatiivisista todennäköisyyssäännöistä (Kahneman & Tversky 1982, 57). Mitä suurempi on ennusteen oletettava epävarmuus, sitä pienempi pitäisi olla ennusteiden varianssi taustatodennäköisyyteen nähden (sama), sitä todennäköisemmin toteuma on lähellä jotain keskilukua (Nisbett ym. 2002, 511) ja sitä perustellumpaa olisi ennustaa regressiivisesti. Intuitiiviset käsitykset todennäköisyyksistä syntyvät kuitenkin samankaltaisuuksien ja assosiaation nopeuden kautta, jotka eivät kumpikaan liity epävarmuuden käsittämiseen (Kahneman 2003, 701). Lisäksi ihmiset yhdistävät epävarmuuden *keskimääräisiin* ennusteisiin. Ihmiset kokevat suurempaa varmuutta ennustaessaan ääri lopputuloksia; toisin sanoen, ihmiset luottavat ennusteisiin eniten tapauksissa, joissa niihin pitäisi luottaa vähiten (Kahneman & Tversky 1982, 66).

Analyttisen ajattelun mahdollisuudet kontrolloida heuristista ajattelua heikkenevät, jos päätöksentekoon lisätään aikapaine, kuten mainittu jo aiemmin. Tämän lisäksi reflektiivisen ajattelun yksilölliset erot vaikuttavat siihen, miten ihminen hyödyntää käytettävissä olevaa aikaa. Ihmiset eroavat toisistaan siinä, miten suuresti he kokevat tarvetta saada asia mielessään päätökseen ja, miten hyvin he sietävät päätöksen tekemistä edeltävää epävarmuuden tilaa tai päätöksen tekemisen viivästymistä (Yudkowsky 2008, 109; Stanovich 2011, 36). Nämä yksilölliset erot kuuluvat korkeaan, tyypin II prosessointiin, reflektiivisiin prosesseihin, jotka tuottavat vastauksia metakysymyksiin, kuten: "kelpaako tämä vastaus, vai pitääkö minun miettiä tätä asiaa vielä lisää?" ja ohjaavat näin alemman tason prosessointia.

Yksilöiden välillä on merkittäviä eroja ajattelutavoissa ja suhtautumisessa käsillä oleviin ongelmiin. Tällaisten reflektiivisten prosessien erojen vuoksi syntyviä ratkaisuja rationalisoidaan usein jälkikäteisesti, esimerkiksi korottamalla nopeaa päätöksentekokykyä tilanteesta riippumattomaan arvoon. Jos päätös tehdään nopeasti, tällöin siirrytään tyypillisesti heuristisen ajattelun varaan, ja kognitiivisten vääristymien merkitys korostuu. Riski sille, että ihminen tuntee halua tehdä päätöksen nopeasti, on suuri, koska nopea päätöksenteko tuntuu hyvältä. Varsinkin ihminen, joka ei siedä viivytystä tai epävarmuutta, ja haluaa saada asiat mielessään nopeasti päätökseen, ei halua kokea pitkälliseen tyypin II prosessointiin liittyvää vaivalloisuutta ja päätöksentekoa edeltävää epävarmuutta. Toisin sanoen, ihminen ei halua "vaivata päätään", kuten vanha sanonta kuuluu.

2.14. Ennustettavuuden arvioinnin merkitys

Esitellyllä teorialla on selkeitä implikaatioita ennusteiden laadinnan kannalta. Edellä esitetyn mukaan olennaisia intuitiivisen ajattelun heikkouksia ovat mm. menneisyyden pitäminen väistämättömänä, datan pitäminen virheettömänä, ja pientenkin otosten pitäminen luotettavina. Etenkin kun data visualisoidaan, ihminen näkee hyvin pienissäkin otoksissa trendejä, ja heuristinen ennuste, jota aivot lähes automaattisesti ehdottavat, on trendin jatkuminen. Regressiivinen ennuste (esimerkiksi keskiarvo) voi tuntua intuitiivisesti oudolta.



Kuva 5. Otosta edustava ennuste ja regressiivinen ennuste. Kuvitteellinen esimerkki.

Heuristinen ajattelu ei tarjoa mitään eksaktia tapaa arvioida, kumpi ennuste on parempi. Kuten todettu, ennustettavuuden arviointi (jälkikäteisen tai etukäteisen) on tyypin I prosessoinnille vaikeaa. Intuitiivisesti, heuristisesti, spontaanisti ihminen yliarvioi ennustettavuuden ja ennusteen luotettavuuden jopa siinä määrin, että ihminen ei osaa kaivata tietoa oletusarvoisesta ennustetarkkuudesta. Kuitenkin juuri tieto ennustettavuudesta auttaisi arvioimaan, miten regressiivinen ennusteen tulisi olla.

Esimerkiksi $N = 4$ on erittäin pieni otos. sen vuoksi kuvan 5 kuvitteellisessa esimerkissä on laskennallisesti 90% luottamusvälin rajoissa, että datassa on itse asiassa laskeva trendi, jonka satunnaisvaihtelu vain kätkee alleen. Koska heuristinen ajattelu ei ole kehittynyt tunnistamaan otoskoon tai datan satunnaisuuteen liittyvää problematiikkaa, ajatus siitä, ettei esimerkikuvassa olisikaan nousevaa trendiä, on ihmisen intuition vastainen.

Mikäli kuvan 5 esimerkkitilanteessa viides toteumaluku olisi lähellä otoksen keskiarvoa, se tulkittaisiin helposti laskuksi, jota rationalisoitaisiin ja jolle etsittäisiin syytä. Aiemmin tässä luvussa kuvatun tutkimuksen perusteella (Kahneman & Tversky 1982, 68) voidaan olettaa, että vain hyvin harva osaisi ehdottaa syyksi satunnaisvaihtelusta johtuvaa paluuta kohti keskiarvoa.

2.15. Intuitiivinen reaktio heikkoon ennustettavuuteen

Väite siitä, ettei ihminen osaa intuitiivisesti huomioida heikkoa ennustettavuutta, ei ole tyhjentävä kuvaus asiasta. Monissa heuristisen ajattelun kokeissa koehenkilöt ovat saattaneet tiedostaa prosessiin liittyvän satunnaisuuden. Mutta jos ajatuksellisia työkaluja satunnaisuuden huomioimiseen ei ole, ihminen joutuu tyytymään heuristiseen arviointiin (Nisbett ym. 2002, 512). Alan kirjallisuudessa on kuvattu esimerkkejä, joissa ihmiset pyrkivät spontaanisti kompensoimaan ennusteen epäluotettavuutta. Ihmiset eivät siirry kohti regressiivisempää ennustetta, mutta sen sijaan ennusteen odotettu epätarkkuus saatetaan huomioida esimerkiksi armollisempana suhtautumisena arvioinnin kohteena oleviin henkilöihin (Kahneman & Tversky 1982, 63).

Eräs tutkimus koski tapaa, jolla Israelin armeijan upseerikoulutuksen valintamenettelyssä arvioidaan hakijoita. Tutkimuksessa pyydettiin arvioitsijoita ennustamaan hakijoiden opintomenestys pääsykoearvion perusteella. Ennusteiden jakauma oli muuten täsmälleen samankaltainen kuin opintojen loppuarvosanojen todelliset jakaumat (ts. yhtä suuri varianssi), sillä poikkeuksella, että hylättyjä arvosanoja ei ennustettu hakijoille läheskään niin paljon kuin niitä koulutuksessa annetaan. (Kahneman & Tversky 1982, 63.) Voi siis olla, että ihmisillä on joissain tapauksissa intuitiivinen käsitys ennusteen epätarkkuudesta, ja jopa intuitiivisia keinoja kompensoida epävarmuutta. Nämä keinot eivät kuitenkaan ole tiukassa merkityksessä rationaalisia tai objektiivisesti oikeita.

2.16. Heuristisen ajattelun merkitys riskien arvioinnin kannalta

Edellä on käsitelty heuristisen ajattelun vaikutuksia pääosin ennusteiden ja ennustevirheiden kannalta, koska tämän tutkimustyön empiirinen osuus käsittelee ennusteita. Numeeriset

ennusteet on valittu empiiriseksi aineistoksi siksi, että niiden perusteella on mahdollista suorittaa hypoteesien tilastollista testausta. Se, että jokin asia voidaan kvantifioida ei koskaan ole relevanssin tae. Tiedonintressin kannalta on huomioitava, että Rajavartiolaitoksessa tehdään paljon myös laadullisia ennusteita ja arvioita. Vaikka näiden yhtäläinen tilastollinen arviointi ei ole mahdollista, monilla laadullisilla arvioilla on välitön merkitys Rajavartiolaitoksen toiminnalle. Tämän vuoksi seuraavassa käsitellään sitä, mikä merkitys heuristisella ajattelulla on erimerkiksi riskienhallinnan ja riskianalyysin kannalta, jos kirjallisuudessa esitetyt väitteet ihmisen ajattelusta pitävät paikkansa. Jäljempänä puhutaan tarkoituksellisen epäspesifisesti riskien arvioinnista, koska riskienhallinta ja riskianalyysi viittaavat Rajavartiolaitoksessa tiettyihin melko tarkkarajaisiin menettelyihin, mutta kysymys heuristisista vääristymistä riskejä arvioitaessa eivät rajoitus näihin menettelyihin, vaan koskevat kaikkia tehtäviä.

Edustavuusheuristiikan osalta on jo todettu, että toteutunutta todellisuutta on vaikea nähdä sattuman aiheuttamana ääri lopputuloksena. Pikemminkin ihminen yleensä suhtautuu menneisyyteen kuin muita vaihtoehtoja ei olisi ollut. Tosin, poikkeuksena tästä, hyvin läheiset vaihtoehtoiset todellisuudet on helppo nähdä; ihminen ymmärtää helposti, että kilpailun hiuksen hienosti hävinnyt hevonen *melkein voitti* (Kahneman 2003, 702).

Tämä menneisyyden väistämättömyyden harha saa ihmiset luulemaan, että menneisyys olisi ollut ennakoitavissa, ja että tulevaisuus olisi yhtäläisesti ennakoitavissa. Jälkiviisaus tekee oppimisesta "liian spesifistä" (Yudkowsky 2008, 95). Silloinkin, kun menneestä pyritään oppimaan, on vaarana oppia välttämään vain täsmälleen saman onnettomuuden toistuminen. Onnettomuuden tapahduttua onnettomuuteen johtaneita syitä tutkitaan, ja sitten tarkastellaan, miksi juuri näitä nimenomaisia seikkoja ei oltu hoidettu kuntoon. Kuitenkin on ymmärrettävä se periaate, että esimerkiksi avaruussukkula Challengerin 1986 onnettomuuden estäminen olisi edellyttänyt - ei vain pettäneen tiivisterenkaan uusimista - vaan kaikkien samantasoisten mahdollisten puutteiden korjaamista, läpi vuosien, halki koko organisaation (sama, 94).

Myös vaikeus hahmottaa konjuktiivisia todennäköisyyksiä on riskien arvioinnin kannalta olennaista, ja aiheuttaa hieman samankaltaista ylispesifistä varautumista. Vastoin konjuktiivisen todennäköisyyden todellista matemaattista muodostumista, yksityiskohtien lisääminen riskin tai uhan kuvaukseen lisää sen subjektiivisesti arvioitua uskottavuutta (Yudkowsky 2008, 97). Ihmisten on todettu mm. todettu olevan keskimäärin halukkaita maksamaan enemmän vakuutuksesta, joka korvaisi kuolemantapauksen lentomatkan aikana

tapahtuneen terroriteon johdosta, kuin vakuutuksesta, joka korvaisi kuolemantapauksen lentomatkan aikana mistä tahansa syystä (Johnson ym. 1993, 39-40).

Ihmiset raskaasti yliarvioivat harvinaisten, mutta äärimmäisten onnettomuuksien (kuten väkivaltarikokset, villieläinten hyökkäykset) lukumäärät ja todennäköisyyden (Yudkowsky 2008, 92; Loewenstein ym. 2001, 276). Harvinaiset riskit arvioidaan sitä yleisemmiksi (Yudkowsky 2008, 92; Johnson ym. 1993, 38), ja herättävät sitä enemmän pelkoa (Loewenstein ym. 2001, 279), mitä enemmän ne ovat olleet esillä mediassa. Tavoitettavuusheuristiikan käyttö todennäköisyyksien ja frekvenssien arvioinnissa tarjoaa selityksen tämän kaltaisille havainnoille.

Valitettavasti ei ole mitään loogista perustetta olettaa, että jonkin tilanteen helpoiten saavutettavissa olevat piirteet olisivat päätöksenteon kannalta olennaisimmat (Kahneman 2003, 703). Riskejä arvioitaessa voi merkitystä olla myös sillä, miten kysymys asetetaan³ (Friedman 2019, 182). Muutos positiivisen ja negatiivisen kuvauksen välillä johtuu vastaajien muuttuneesta vertailupisteestä, kuten tässäkin luvussa kuvattu *Asian Disease Problem* osoittaa (ks. s. 26). Ihmiset ovat keskimäärin riskiä välttäviä valitessaan saavutettavien hyötyjen välillä, ja riskinottajia valitessaan väistämättömien tappioiden välillä (Tversky & Kahneman 1981, 453; LeBoeuf & Shafir 2003, 72).

Riskien arviointiin vaikuttaa myös ihmisen intuitiivinen kyvyttömyys reagoida kertaluokan muutoksiin. Ihminen ei kykene luotettavasti ekstrapoloimaan kokemistaan pienistä riskeistä suurten ja vakavien riskien mahdollisuutta tai vaikutuksia (Yudkowsky 2008, 92). Lukumäärien arvioinnin vaikeus saa aikaan sen, että pelastettujen tai menetettyjen ihmishenkien määrä, ilman kontekstia, on vaikeasti arvioitava asia (Slovic ym. 2007, 1341).

Myös tunnereaktioilla on vaikutusta, ja todennäköisyyden lisäksi ne vaikuttavat mielikuvaan uhkan vakavuudesta. Jos ennakoituun lopputulokseen liittyy voimakas tunnereaktio, todennäköisyysvaihtelulle ei anneta niin paljon painoarvoa (Slovic ym. 2007, 1342; Loewenstein ym. 2001, 271). Todennäköisyyksien sijaan ihminen saattaa tosiasiasa arvottaa riskin toteutumiseen liittyvää mielikuvaa (Slovic ym. 2007, 1345) ja sen herättämää

³ Friedman (2019) esittää tämän myös kritiikkinä aihealueen tutkimusta kohtaan. Hänen mukaansa ihmiset eivät niinkään arvioi terrorismin todennäköisyyttä kuin sen epäoikeudenmukaisuutta, mikä vaikuttaa siihen, miten paljon resursseja sen torjuntaan tulee osoittaa. Tämä kritiikki ei kuitenkaan täysin voi selittää esimerkiksi Johnsonin ym. (1993) vakuutus-esimerkissä havaittua epäloogisuutta.

tunnereaktiota. Käsitukset riskin vakavuudesta saattavat selittyä sillä, miten paljon pelkoa riskiin liittyvät mielikuvat arvioitsijassa herättävät (sama, 1342).

Tunteita saattaa liittyä sekä kuviteltuihin lopputuloksiin, että päätöksentekotilanteeseen itseensä (Loewenstein ym. 2001, 271). Tunnereaktioihin perustuva heuristinen ajattelu aiheuttaa myös riskien ja hyötyjen sekoittumista päätöksenteossa. Jos tietyn ratkaisun riskit arvioidaan pieniksi, syntyy positiivinen mielikuva ja affekti, jolloin hyödyt arvioidaan suuriksi. Jos taas riskit arvioidaan suuriksi, syntyy negatiivinen affekti, ja hyödyt arvioidaan pieniksi. Jos arviointiprosessi alkaa (ajallisesti) hyötyjen arvioinnista, pieniksi arvioidut potentiaaliset hyödyt synnyttävät negatiivisen affektin, joka saa ratkaisuun liittyvät riskit vaikuttamaan suurilta. (Finucane ym. 2000, 9, 13; Slovic ym. 2002, 418)

Riskien arviointiin liittyviä vääristymiä ei voida nyt käytössä olevan aineiston kautta tarkastella empiirisesti. Kysymys riskien arviointiin vaikuttavista psykologisista tekijöistä on kuitenkin relevantti, ja ansaitsee tulla huomioiduksi pohdittaessa tämän tutkimuksen tulosten implikaatioita ja jatkotutkimustarvetta.

3. TUTKIMUSMENETELMÄ

3.1. Kvantitatiivinen tutkimusasetelma sosiaalitieteissä

Kvantitatiivinen tutkimus voi tuottaa faktatietoa vain siitä aineistosta, joka on tutkimuksen kohteena. Jokainen aineisto on yksittäistapaus. Kaikki yleistyksen aineiston ulkopuoliseen todellisuuteen ovat vain pohdintaa ja hypoteeseja. Tutkimustuloksina voidaan raportoida mm. lukumääriä, keskilukuja, hajontalukuja ja muuttujien välisiä korrelaatioita. Tilastollisin testein voidaan saada laskennallisia todennäköisyyksiä sille, että havaittu korrelaatio on satunnaisvaihtelun aiheuttamaa. Jos tämä todennäköisyys on pieni, voidaan nollahypoteesi hylätä, ja tietyllä laskennallisella luotettavuustasolla sanoa, ettei havaittu korrelaatio ole satunnaisvaihtelua. Positivistisessa mielessä tämäkin tarkoittaa vain, että otoksessa todella vaikuttaa olevan korrelaatio kahden muuttujan välillä.

Teoriaosuudessa esitetään väitteitä ilmiöiden syistä ja selityksistä. Otoksesta havaittujen korrelaatioiden ja muuttujien lukuarvojen syistä ei kuitenkaan voida tilastollisin menetelmin saada tosiasiatietoa. Korrelaatiosta ei voida päätellä kausaalisuutta. Lisäksi tulos saattaa olla väärä. Otos saattaa olla vääristynyt, laskutapa saattaa olla väärä, ja tilastollisten menetelmien edellyttämien lähtöoletusten täyttymättä jääminen saattaa aiheuttaa vääristymää tuloksiin. Edes satunnaisvaihtelua ei havainnon tosiasiallisena syynä voida koskaan täysin sulkea pois. Tämä sopii fallibilistiseen tietokäsitykseen, jossa varmuutta faktoista ei voida koskaan täysin saavuttaa.

Kuten johdannossa on todettu, toimintaympäristön muutoksen ennakointi, ja ennusteiden laadinta, on monialainen, poikkitieteellinen ongelma. Jos ennustevirheitä tutkittaisiin kvalitatiivisesti, esimerkiksi haastattelututkimuksella, voitaisiin saada tietoa vain konstruktivistisesta todellisuudesta. Esimerkiksi kognitiivisten vääristymien kirjallisuudessa koehenkilöt järjestäen arvioivat, etteivät vääristymät ole vaikuttaneet heihin.

Sosiaalitieteissä tutkimusaineisto on havaintoja ihmisten käyttäytymisestä tai sen seurauksista. Aineisto voi olla kerätty tai valmis. Aineisto voi olla in vitro, eli testiolosuhteissa kerätty, tai in situ, todellisista tilanteista kerätty aineisto. Kaikissa tapauksissa suurempi otoskoko lisää tulosten reliabiliteettia, mutta kaiken tyyppisillä aineistoilla on heikkoutensa.

Tässä työssä käytetään in situ aineistoa: todellisia ennusteita, joita ihmiset ovat laatineet todellisessa ympäristössään, ilman, että lukuja oli tarkoitettu tutkimusaineistoksi. Tällaisella aineistolla on ongelmansa: Mitään ennustevirheitä aiheuttavia seikkoja ei voida sulkea pois, eikä tutkittavaa ilmiötä voida eristää muista tekijöistä. Data ei myöskään ole luotu tutkijan tarpeita varten, eikä koskaan voi olla juuri sellaista, kun tutkimus edellyttäisi. Kognitiivisten vääristymien osalta tällaisen aineiston käyttämiselle on tunnustettu tarve. Teorian kenties keskeisin kritiikki on kohdistunut juuri siihen, että ilmiöt on saatu esiin testiolosuhteissa, ja että tosielämässä ihmisten päätöksenteko on erilaista (Gilovich & Griffin 2002, 11). Esitetyn kritiikin mukaan tulokset selittyvät sillä, että tutkijat ovat tahallaan rakentaneet kompakysymyksiä, jotka palvelevat tutkijoiden omia tarkoituksia.

3.2. Aineisto

Tutkimusaineistona on Rajavartiolaitoksen tulossuunnitelman ja toiminta- ja taloussuunnitelman (TS-TTS) toimintaympäristön kehitystä kuvaavia suunnitelmalukuja vuosilta 2009–2014, sekä ennusteita vastaavat toteumat. Aineistosta koottiin otokseen myös tieto kunkin tilastorivin Rajavartiolaitoksen tuloksellisuuden laskennassa käytetystä painokertoimesta, sekä tieto siitä, mitä hallintoyksikköä suunnitelmaluku koskee.

Suunnitelmaluvuista osa käsitteli mm. hallinnollisia asioita, kuten henkilötyövuosia, tai Rajavartiolaitoksen omaan johtamiseen liittyviä lukuja, kuten partiotunteja. Tällaisissa asioissa kyse on enemmän suunnitelmasta kuin ennusteesta, koska lopullinen toteuma on pääosin viraston omassa vallassa. Osa suunnitelmaluvuista kuitenkin käsitteli organisaation kannalta ulkoisten asioiden lukumäärien kehitystä. Yksinkertaistaen voidaan sanoa, että operatiivisessa toiminnassa kertyvien havaintojen ja suoritteiden lukumäärät ovat tällaisia ulkoisia muuttujia, ja niistä laaditut suunnitelmaluvut ovat ennusteita toimintaympäristön kehittymisestä.

Tässä yhteydessä ulkoisella ts. eksogeenisellä tarkoitetaan suuretta, jonka lukumäärään virasto ei voi suoraan lyhyellä aikavälillä vaikuttaa (esim. rajatarkastukseen saapuvien rattijuopumuksen tunnusmerkistön täyttävien kuljettajien lukumäärä tai viranomaisten apua edellyttävien hätätilanteiden määrä merialueella). Rajavartiolaitoksessa hallintoyksiköiden ja viraston tuloksellisuuden laskennalla ei ole vaikutusta virkamiehen palkkaukseen. Useisiin suoritteista liittyy hallintovallan käyttöä, eikä suoritemääriä tai tuloksellisuuslukuja huomioida asiakkaan asiaa koskevassa päätöksenteossa. Virkamiehen henkilökohtainen

palkanosa riippuu esimiehen tekemästä suoritusarvioinnista, jossa arvioidaan mm. työn laatua ja huolellisuutta. Tämä vahvistaa suoritelmäärien luotettavuutta ja riippumattomuutta. Tämän tutkimuksen tarvitsema otos koottiin jäljempänä kuvatut ehdot täyttäviä suoritteita ja havaintoja koskevista suunnitelmaluvuista.

Vuosina 2009–2014 hallintoyksiköiden esikunnat valmistelivat muun TS-TTS -aineiston osana nk. suunnitelmaluvut seurattavien tilastorivien kehityksestä viiden vuoden suunnittelukaudella. Hallintoyksiköiden laatimat suunnitelmaluvut toimitettiin Rajavartiolaitoksen esikunnalle, joka vastasi lopullisista luvuista. Vuodesta 2015 alkaen prosessia muutettiin siten, että Rajavartiolaitoksen esikunta laati ensimmäiset luonnokset luvuista, ja hallintoyksiköillä oli tämän jälkeen mahdollisuus lausua asiaan. Vuoteen 2008 asti käytössä oli aiempi tuloksellisuuden laskennan järjestelmä, ja 2015 alkaen tuloksellisuuden laskentaa muutettiin joiltain osin siten, että luvut eivät välttämättä olisi täysin vertailukelpoisia. Vertailtavuuden takaamiseksi tämän tutkimustyön otokseen otettiin mukaan suunnitelmalukuja vain aikaväliltä 2009-2014, jolloin käytössä oli, joittenkin yksityiskohtien päivittämistä lukuun ottamatta sama tuloksellisuuden laskutapa, ja samat seurattavat tilastorivit.

Toteumatiedot on alun perin koottu mm. Rajavartioidin tietojärjestelmästä (RVT), Poliisiasian tietojärjestelmästä (PATJA) sekä tehtävän suorittaneista yksiköistä. Tähän työhön toteutuneita lukumääriä koskeva tieto on otettu tutkimusaineistona olevasta TS-TTS -materiaalista. Toteumatiedot on otettu otokseen mukaan vuosilta 2006–2016. Osalla riveistä tilastointi muuttui 2014, ja aikasarjan loppu piti rakentaa kolmen viimeisen vuoden osalta laskemalla yhteen eri tulosrivejä. Tutkimusaineisto on saatu keväällä 2019 Rajavartiolaitoksen esikunnan Suunnittelu- ja talousyksiköstä.

3.3. Otoksen kokoaminen

Tutkimusaineistosta koottiin otos, joka koostuu tietyt ehdot täyttävistä tilastoriveistä. Ehtoina oli, että otokseen mukaan otettava tilastorivi on:

- jossain määrin yksikäsitteinen tilastoitava,
- jossain määrin eksogeeninen, organisaation näkökulmasta ulkopuolinen, ei suoraan viraston hallittavissa, (esim. poliisin virka-apupyynnöt, rajatarkastukseen saapuvien viisumivelvollisten määrä, etsintätehtävät merialueella),
- tilastoitu 2006–2016,

- suunniteltu 2009–2014, ts. toteuman kehityksestä on laadittu tänä aikana suunnitelmaluvut l. ennusteet, ja
- pääosin toteumaltaan eri suuri kuin 0.

Osa otokseen liitetyistä tulosriveistä piti poistaa osasta analyyseja (jos joukossa oli 0-toteumia), n on ilmoitettu aina erikseen. Otokseen ei otettu mukaan salassa pidettävää tietoa, eli esim. sotilaallisen maanpuolustuksen tai muun kriisivalmiuden ylläpitoon liittyviä tietoja.

Otoksessa on mukana tiedot kuudesta Rajavartiolaitoksen hallintoyksiköstä (vartiostosta), jotka ovat alueellaan vastuussa Rajavartiolaitoksen tehtävien suorittamisesta. Mukana ei ole Rajavartiolaitoksen esikunnan, Vartiolentolaivueen ja Raja- ja merivartiokoulun suunnittelua, joka ei ole tämän tutkimustyön näkökulmasta vertailukelpoista (ei riittävää määrää ulkoisia suoritteita). Analyysiin sopivista tiedoista koostui otos, jossa on 779 toteumalukua, ja 1873 näihin liittyvää eripituisen ajan päähän laadittua ennustetta, jotka jakautuivat seuraavasti:

Taulukko 1.

<i>Aikaväli, jolle ennusteet on laadittu.</i>	<i>n</i>
samalle vuodelle	422
seuraavalle vuodelle	415
kahden vuoden päähän	415
kolmen vuoden päähän	345
neljän vuoden päähän	276
yhteensä	1873

3.4. Saatavilla olevat tiedot kustakin tilastorivistä

Otokseen kootuista tilastoriveistä on käytettävissä tieto siitä, mikä tilastoitava suorite tai havainto on kyseessä, toteumatiedot vuosilta 2006-2016, vuosina 2009-2014 laaditut ennusteet (suunnitelmaluvut), tilastorivin painokerroin tuloksellisuuden laskennassa, sekä tieto siitä, mitä hallintoyksikköä luvut koskevat. Tiedot koottiin nk. paneelimuotoon, jossa yhden paneelin tiedot koskevat yhden hallintoyksikön yhtä tilastoriviä (esim. Suomenlahden merivartiostossa rajatarkastetut viisumivelvolliset muodostavat yhden paneelin, ja Kaakkois-Suomen rajavartiostossa rajatarkastetut viisumivelvolliset oman paneelinsa).

Tilastorivien nimet peitettiin koodilla, ja hallintoyksiköt ilmaistiin numeroilla ”Hallintoyksikkö 1 – Hallintoyksikkö 6”. Hallintoyksikköä koskeva tieto on ilmaistu viidellä organisaation osaa kuvaavalla binäärimuuttujalla. Esimerkki paneelista kuvassa 6. Samalla rivillä toteumaluvun kanssa on vuosiluku, ja kyseisen vuoden toteumaa koskevia, eri-ikäisiä ennusteita. Esimerkiksi, kuvassa 6, vuotta 2013 koskeva, vuonna 2009 tehty ennuste (suunnitelmaluku) on vuoden 2013 rivillä kohdassa ”SMA 4v”, joka merkitsee neljän vuoden päähän tehtyä ennustetta.

RIVIKOODI	PAINO	VUOSI	TOTEUMA	SMA 0v	SMA 1v	SMA 2v	SMA 3v	SMA 4v	HallYks 1	HallYks 2	HallYks 3	HallYks 4	HallYks 5
STR03	5	2006	136713						1	0	0	0	0
STR03	5	2007	138358						1	0	0	0	0
STR03	5	2008	137392						1	0	0	0	0
STR03	5	2009	192556	68000					1	0	0	0	0
STR03	5	2010	241097	192000	69000				1	0	0	0	0
STR03	5	2011	326072	240000	194000	70000			1	0	0	0	0
STR03	5	2012	290573	340000	240000	196000	71000		1	0	0	0	0
STR03	5	2013	421598	310000	350000	240000	198000	72000	1	0	0	0	0
STR03	5	2014	317003	360000	360000	360000	240000	200000	1	0	0	0	0
STR03	5	2015	283469		370000	370000	370000	240000	1	0	0	0	0
STR03	5	2016	355128			380000	380000	380000	1	0	0	0	0

Kuva 6. Esimerkki yksittäisestä paneelista.

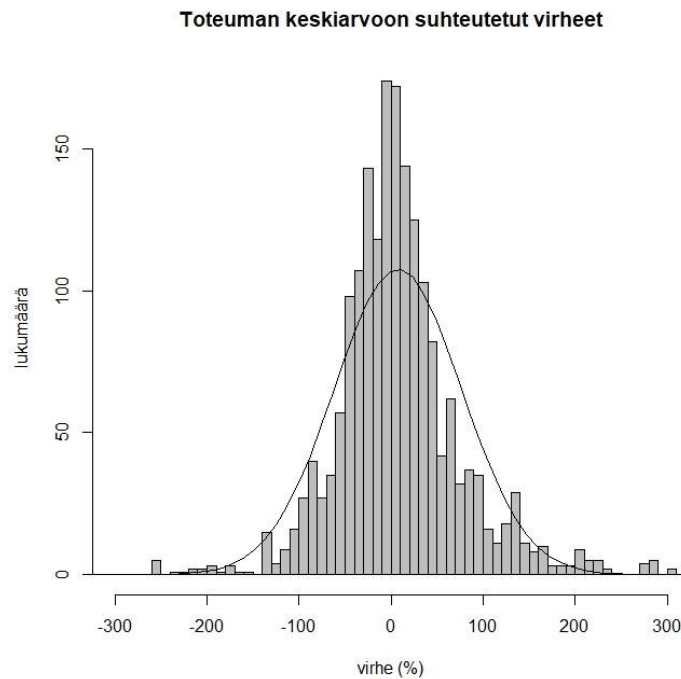
3.5. Tulosluvussa käytetyt suureet:

Ennustevirhe ja virheiden muuntaminen yhteismitallisiksi

Kuten edellä on kerrottu, ennustevirheellä tarkoitetaan ennusteen ja toteuman erotusta. Eri tulosrivejä koskevien ennusteiden virheistä on tehty keskenään vertailukelpoisia suhteuttamalla ne toteumaan. Tämä on tehty jakamalla ennusteen ja toteuman erotus tulosrivin toteuman keskiarvolla. Näin saatava luku on negatiivinen silloin kun ennuste on ollut pienempi kuin toteuma, ja vastaavasti positiivinen silloin kun ennuste on ollut toteutunutta lukua suurempi.

Jotta virheitä voi verrata toisiinsa, ne pitää skaalata yhteismitallisiksi. Vaihtoehtoisia skaalaustapoja on monta, sen mukaan, mihin virheet suhteutetaan. Yksi luonteva tapa suhteuttaa virheitä olisi suhteuttaa niitä kunkin vuoden toteumalukuun. Tässä on se ongelma, että jos virhe aiheutuu suuresta tiputuksesta toteumassa, suhteellinen virhe nousee erittäin suureksi. Jos toteuma on tippunut nolnaan tuona vuonna, ei suhteellista virhettä voi laskea ollenkaan.

Näiden ongelmien vuoksi yhden vuoden toteuma ei välttämättä edusta tulosrivin suuruusluokkaa kovin hyvin. Havainnollisempi ja stabiilimpi tapa skaalata on suhteuttaa virheet tulosrivin toteuman keskiarvoon. Näin skaalaten otoksessa on 1873 tulosrivin keskiarvoon suhteutettua eri aikajaksoille tehtyjen ennusteiden virhettä. Kuva 7 esittää ennustevirheiden jakaumaa. Kuvassa jakauman päälle on piirretty normaalijakauma helpottamaan jakauman muodon arviointia.



Kuva 7. Histogrammi tulosrivin toteuman keskiarvoon suhteutetuista ennustevirheistä.

	Keskiarvo	Mediaani	Keskihajonta
Virheet (%)	7.23	3.41	69.51

Kuten graafisestakin tarkastelusta nähdään, toteuman keskiarvoon suhteutetut virheet eivät ole normaalijakautuneet, $W = 0.94$, $p < 0.001$, $D = 0.52$, $p < 0.001$. Jakauma on kuitenkin melko symmetrinen, ja muistuttaa normaalijakaumaa riittävästi, jotta parametrista testaamista voidaan tehdä, jos testien tulokset eivät ole tilastollisen merkitsevyyden rajoilla.

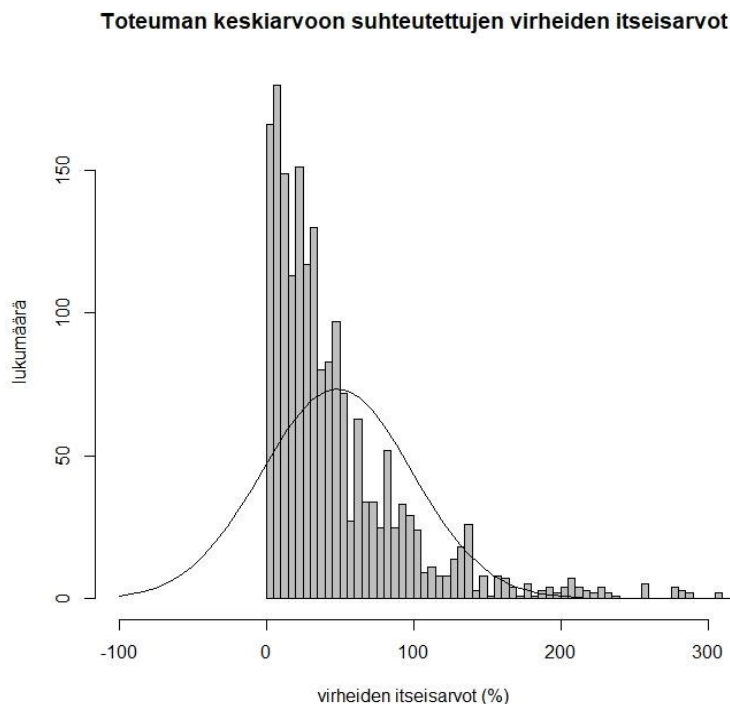
Ennusteen harhaisuus ja tarkkuus

Se, onko tarkasteltavana suureena ennustevirheen absoluuttinen arvo vai itseisarvo, riippuu siitä, tutkitaanko ennusteen harhaisuutta vai tarkkuutta. Ennusteen harhaisuus tarkoittaa sitä, että ennusteet ovat systemaattisesti toteuman yläpuolella (positiivinen vääristymä) tai alapuolella (negatiivinen vääristymä). Ennusteen tarkkuus tarkoittaa sitä, kuinka kaukana toteumasta (kumpaan tahansa suuntaan) ennusteet ovat olleet.

Kun halutaan arvioida ennusteiden harhaisuutta, huomioidaan virheiden etumerkki. Jos ennusteet eivät ole systemaattisesti toteumaa suurempia tai pienempiä, näin laskettujen virheiden keskiarvo on 0, riippumatta siitä, kuinka suurina virheet ovat olleet (kuinka kaukana ennuste on toteumasta, suuntaan tai toiseen). Mikä tahansa joukko ennusteita, joiden virheiden keskiarvo on 0, on harhaton.

Tarkasteltaessa ennusteiden tarkkuutta, käytetään tulosrivin keskiarvoon suhteutettujen virheiden itseisarvoa. Mitä pienempi virheiden itseisarvojen keskiarvo on, sitä tarkempia ennusteet ovat olleet. Luvun itseisarvo on yhtä suuri tai suurempi kuin 0, minkä vuoksi virheiden itseisarvojen jakauma on hyvin epäsymmetrinen. Kuva 8 esittää ennustevirheiden itseisarvojen jakaumaa. Kuvassa jakauman päälle on piirretty normaalijakauma helpottamaan jakauman muodon arviointia.

Kuvasta 8 nähdään, että virheiden itseisarvojen jakauma ei muistuta laisinkaan normaalijakaumaa. Normaalisuutta olettava parametrisen testaus edellyttää muuttujamuunnoksia. Suurimmassa osassa regressioita on selitettävänä muuttujana joko ennustevirhe tai ennustevirheen itseisarvo, jotka on kaikissa tapauksissa suhteutettu tulosrivin toteuman keskiarvoon.



Kuva 8. Histogrammi tulosrivin toteuman keskiarvoon suhteutettujen virheiden itseisarvoista.

Suunnitelmaluvut

Ennustevirheiden lisäksi osassa regressiomalleista käsitellään suunnitelmalukuja (ts. ennusteita, virheiden sijaan). Erikokoisten tilastorivien ennusteet eivät ole keskenään vertailukelpoisia. Regressioita varten ennusteista laskettiin, kuinka paljon yhden vuoden päähän tehtyä ennustetta oli muutettu (suhteessa aiemmin tehtyyn ennusteeseen), tämän suhteellisen muutoksen itseisarvo, sekä binäärimuuttuja (1 tai 0) sen mukaan, oliko ennustetta muutettu. Näin laskettuja lukuja voi vertailla tilastorivin suuruusluokasta riippumatta. Suunnitelmaluvun suhteellinen muutos ei ollut tilastollisesti yhteydessä tulosrivin keskiarvon suuruusluokkaan, $r = 0.03$, $p = 0.53$.

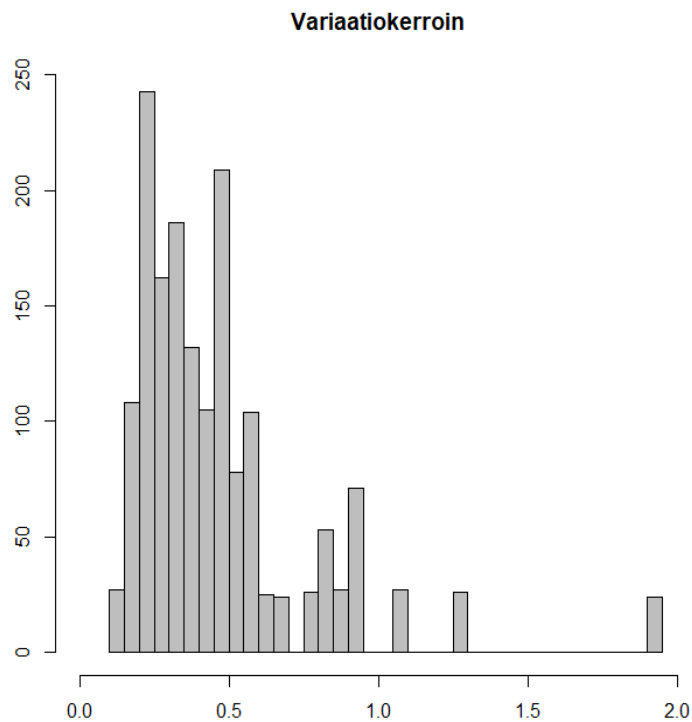
Toteumaluvut

Myös toteumaluvuista laskettiin tiettyjä tunnuslukuja regressioita varten. Keskenään vertailukelpoisia näistä ovat toteuman prosentuaalinen muutos (siinä toteumatiedossa, joka ennusteen laatijalla olisi sillä hetkellä ollut käytössään), binäärimuuttuja (1 tai 0) sen mukaan, oliko viimeisin toteuma pienempi kuin sitä edellinen (laskeva trendi), sekä variaatiokerroin.

Toteumaluvuista laskettiin regressiomalleja varten myös tulosrivin toteuman keskiarvo (jota käytettiin ennustevirheiden skaalaamisen, ks. edellä), sekä keskiarvon 10-kantainen logaritmi, jota kutsutaan tulosrivin suuruusluokaksi. Logaritmin käyttö on perusteltua, koska logaritmi on monotoninen muunnos, joka ei voi aiheuttaa korrelaatiota, jos sitä ei ole muuntamattomassa muuttujassa. Logaritmin käyttö perustuu loogiseen oletukseen siitä, että erot ennusteiden laadinnan tavoissa tai suhtautumisessa ennusteen merkitykseen voisi olla yhteydessä nimenomaan suuruusluokkaan, eikä niinkään absoluuttiseen suoritteiden lukumäärään. Aineistosta voidaan empiirisesti havaita, että ilmiöt, joilla on yhteys tilastorivin kokoon, ovat voimakkaammin yhteydessä juuri keskiarvon logaritmiin.

Huono odotettu ennustettavuus

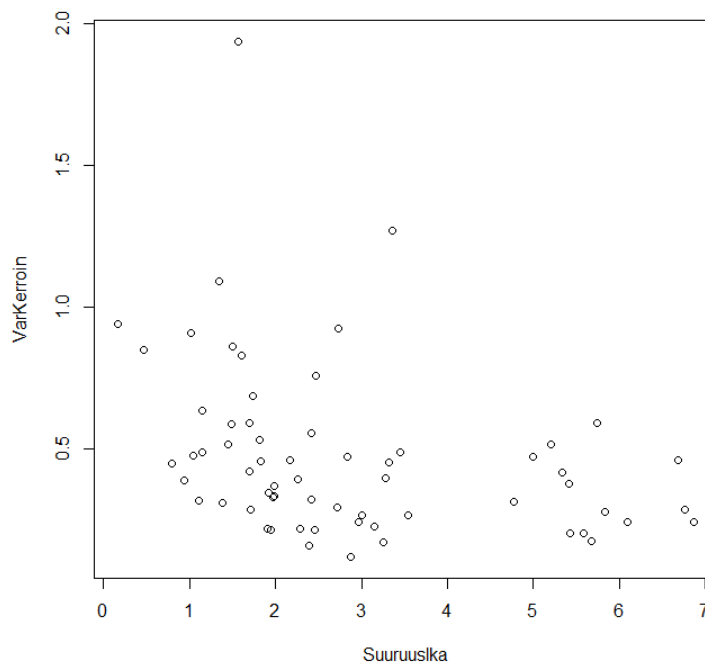
Huonoa ennustettavuutta ei voida suoraan mitata, mutta korvaavina muuttujina aineistosta voidaan käyttää tulosrivin toteuman variaatiokerrointa ja aikaisempien ennusteiden virheitä. Variaatiokerroin on dimensioton suure, joka muodostetaan jakamalla keskihajonta keskiarvolla. Tällöin eri suuruusluokkaa olevien tilastorivien hajontaluvuista tulee vertailukelpoisia keskenään. Tarkastelua varten tehdään olettaus, että mitä enemmän tilastorivin toteumaluvut heittelevät, sitä vaikeampi sitä on ennustaa. Tämän olettamuksen oikeellisuus ei välttämättä ole kiistatonta. Variaatiokertoimen etuna, huonoa ennustettavuutta kuvaavana korvaavana muuttujana, on se, että tämä tieto olisi käytettävissä myös silloin kun johonkin tilastoitavaan suureeseen liittyvää ennustetta tehdään ensimmäistä kertaa. Kuvassa 9 on esitetty histogrammilla se, millaisia lukuarvoja aineistossa olevien tilastorivien variaatiokertoimet saavat.



Kuva 9. Tilastorivien variaatiokertoimien jakauma tutkimusaineistossa.

Koska variaatiokertoimessa otoskeskihajonta jaetaan keskiarvolla, erisuuruisten tilastorivien variaatiokertoimien pitäisi olla keskenään vertailukelpoisia. Koska kuitenkin esimerkiksi miljoonan suoritteen tilastorivin vaihtelu koostuu miljoonasta yksittäisestä tapauksesta, jotka toteutuvat tai eivät toteudu jollain todennäköisyydellä, ja vastaavasti pienempi tilastorivi muodostuu pienemmästä joukosta tällaisia yksittäisiä todennäköisyyksiä, suhteellisen vaihtelun voi olettaa korreloivan negatiivisesti suuruusluokan kanssa. Näin on myös tarkasteltavassa aineistossa, $r = -0.33$, $p < 0.001$ (ks. kuva 10). Suuruusluokka selittää kuitenkin vain osan variaatiokertoimen vaihtelusta, $R^2 = 0.11$.

Suuruusluokka on sisällytetty kontrollimuuttujana regressiomalleihin siten, että voitaisiin tarkastella sitä variaatiokertoimen kuvaamaa tulosrivin toteuman suhteellista vaihtelua, joka ei selity pelkästään suuruusluokalla. Kuten edellä todettu, suuruusluokalla kontrolloitua variaatiokerrointa käytetään korvaavana muuttujana huonolle odotetulle ennustettavuudelle. Suuruusluokka on tulosrivin toteuman keskiarvon 10-kantainen logaritmi.



Kuva 10: Variaatiokertoimen ja tilastorivin suuruusluokan yhteys tutkimusaineistossa.

Aiemmat virheet

Jos tilastoitavaan suureeseen liittyviä ennusteita on tehty jo aikaisemmin (kuten tässä regressiomalleissa käytettävässä aineistossa on), olisi ennusteen laatijalla helposti käytettävissään myös tieto siitä, ovatko aiemmat ennusteet olleet oikean suuntaisia. Toisena huonoa ennustettavuutta kuvaavana korvaavana muuttujana käytetään tässä tarkastelussa aiempia vuosien toteumatietoja koskeneiden ennusteiden virheiden itseisarvojen keskiarvoja. Ennustevirheet on jälleen suhteutettu tilastorivin toteumien keskiarvoon. Tarkastelua varten tehdään olettaus, että aiempien ennusteiden virheellisyys sisältäisi jotain informaatiota siitä, miten vaikeaa tiettyä tilastoriviä on ennustaa. Tämänkin lähtöolettamuksen oikeellisuutta on syytä tarkastella myöhemmin.

Koska aineiston ensimmäisille ennusteille ei voi laskea ”aiempia ennustevirheitä”, otos on pienempi ($N = 755$) niissä regressiomalleissa, joissa aiempia virheitä on käytetty selittävänä muuttujana. Aiempien virheiden sisällyttäminen regressiomalliin toimii samalla myös paneelidatan autokorrelaation korjaavana toimenpiteenä.

Regressiivinen ennuste

Jotta voidaan vertailla tilastollisen ennusteen (liukuva keskiarvo) ja asiantuntijaennusteen (aineistosta) onnistumista, rakennetaan selitettävä muuttuja ”Regressiivisen ennusteen paremmuus”, joka on laskettu asiantuntijaennusteen virheen (itseisarvon) ja regressiivisen ennusteen virheen (itseisarvon) erotuksena. Näin, mikäli regressiivinen ennuste on osunut lähemmäksi toteumaa, muuttujan ’Regressiivisen ennusteen paremmuus’ arvo on positiivinen, ja toisin päin. Muuttujan nimi ei sisällä mitään normatiivisia kannanottoja, vaan on tarkoitettu ainoastaan merkitsemään sitä, miten päin etumerkki on tulkittava (jos muuttuja on positiivinen, tilastollisen ennusteen virhe on ollut pienempi).

Vertailukohtana käytettävän tilastollisen ennusteen on oltava yksinkertainen, jotta vertailu olisi mielekästä. Yksinkertaiseksi tilastolliseksi ennusteeksi valittiin ennustetta laadittaessa käytettävissä olleista toteumatiedoista kolmen viimeisimmän toteuman liukuva keskiarvo. Näin laskettua arvoa käytettiin ”yksinkertaisena tilastollisena ennusteena” kaikille aikaväleille.

Toteuman keskiarvon käyttäminen ennusteena olisi ollut kaikkien ennusteita laatineiden käytössä. Vertailun validiteetin kannalta on keskeistä, ettei tilastollista ennustetta soviteta dataan millään tavalla. Kolmen vuoden liukuva keskiarvo on menetelmänä niin yleinen ja yksinkertainen, että sen voi katsoa täyttävän tämän ehdon.

Aikavälit

Tutkimusaineistossa Rajavartiolaitoksen suoritteiden määriä on ennustettu viidelle eri aikavälille siten, että lyhyimmän aikavälin ennusteet koskevat samaa vuotta, jona ne on tehty; ts. kesällä on arvioitu, mikä toteuma tulee olemaan koko vuoden osalta. Samaa vuotta koskevan ennusteen lisäksi on tulosrivien toteumalle annettu ennustettu määrä neljälle seuraavalle vuodelle.

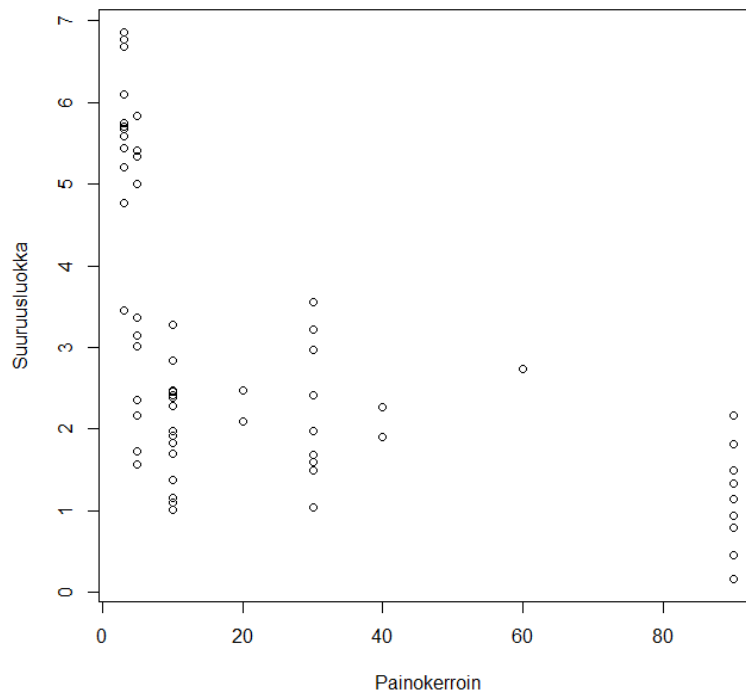
Aikaväliä on merkitty numeerisesti sen mukaan, monenko vuoden päähän ennuste on laadittu. Samaa vuotta koskevan ennusteen 'Aikaväli' on tässä merkintätavassa 0, seuraavaa vuotta koskevan ennusteen 1, ja niin edelleen. Näin saadaan diskreetti numeerinen muuttuja, jota voidaan käyttää mm. lineaarisen regressioon selittävänä muuttujana.

Painokerroin ja suuruusluokka

Osa tilastoriveistä vaikuttaa suoraan hallintoyksikön ja viraston tuloksellisuuden laskentaan, ja niille oli annettu tilastorivin tärkeyttä kuvaava painokerroin. Kaikilla otoksessa mukana olleilla tilastoriveillä ei ollut painokerrointa, vaan ne huomioitiin tuloksellisuuden laskennassa joko osana joltain summamuuttujaa tai ei ollenkaan.

Painokerrointen on tuloksellisuuden laskennassa tarkoitus kuvata tietyn toimenpiteen tuloksellista arvoa vain saman toimialan muihin suoritteisiin verrattuna. Eri tilastorivien erilaiset painokertoimet eivät ole vertailukelpoisia keskenään, koska eri toimialoilla (rajatarkastus, rikostorjunta, meripelastus) painokertoimet on asetettu eri tavoin. Tästä johtuen painokerrointa voidaan käyttää (joskin epätäydellisenä) korvaavana muuttujana toimialalle. Toisin sanoen, painokerrointa tarkastelemalla voidaan erottaa eri toimialoja toisistaan. Tilastolliset yhteydet painokertoimen ja ennustevirheiden välillä tarkoittaisivat eroja ennusteiden laadinnassa eri toimialojen välillä. Tulosten tarkempi tulkinta on diskussiossa.

Toinen tapa erottaa eri toimialojen ja erilaisten tulosrivien ennustevirheet toisistaan on tarkastella tulosrivin suuruusluokkaa. Rajatarkastuksia voi olla isossa hallintoyksikössä useita miljoonia vuodessa, kun taas meripelastuksen tai rikostorjunnan suoritteita voi olla keskimäärin esimerkiksi joitakin kymmeniä.



Kuva 11. Painokerroin ja suuruusluokka aineistossa.

Organisaation osaa kuvaavat binäärimuuttujat

Dataa on mukana kuudesta hallintoyksiköstä. Sitä, mistä hallintoyksiköstä on kyse, kuvataan viidellä binäärimuuttujalla (1 tai 0). Hallintoyksikön nro 6 osalta kaikki viisi organisaatiota osoittavaa binäärimuuttujaa saavat arvon 0. Tällöin hallintoyksikön nro 6 arvo voidaan lukea suoraan vakiotermitä. Muiden hallintoyksiköiden osalta estimaattorin oikea tulkinta saadaan laskemalla yhteen vakiotermi ja organisaatiokohtaisen binäärimuuttujan kohdalle merkitty arvo.

Eri toimialojen roolia painokertoimen merkityksen aiheuttamisessa voi välillisesti yrittää tutkia tarkastelemalla organisaation eri osia. Eri vartiostoissa eri toimialojen osuus toiminnassa on erilainen. Lisäksi, koska ennusteet on tarkistettu Rajavartiolaitoksen esikunnassa, ei hallintoyksiköiden välillä teoriassa pitäisi olla merkitseviä eroja.

3.6. Menetelmät

Huomautus termeistä

Monille tilastollisille termeille on useita synonyymejä, joiden käyttö riippuu lähinnä kunkin tieteenalan traditiosta. Tässä työssä puhutaan regressiomallin selitettävästä muuttujasta (jota joissain yhteyksissä kutsutaan myös mm. vastemuuttujaksi, endogeeniseksi muuttujaksi tai tulosmuuttujaksi), selittävästä muuttujasta (tarkoittaen samaa kuin mm. eksogeeninen muuttuja, regressori, ennustemuuttuja tai riippumaton muuttuja) ja residuaaleista eli jäännöstermistä (toisilta nimiltään mm. satunnaistekijä, virhetermi tai muiden tekijöiden vaikutus).

Hajontalukujen kuvaaminen ja vertailu ryhmittäin

Ennustevirheitä ja osaa muista suureista kuvataan graafisesti ja tunnusluvuin (keskiarvo, mediaani, keskihajonta, kvartiilivälit). Graafisen tarkastelun lisäksi muuttujien, otosten tai otoksen osien (kuten eri aikaväleille laadittujen ennusteiden virheiden) hajontalukuja voidaan tutkia tilastollisilla testeillä, joilla selvitetään laskennallista todennäköisyyttä sille, että varianssien ero olisi satunnaisvaihtelun aiheuttamaa. Jos testien osoittama todennäköisyys on tarpeeksi pieni, voidaan nollahypoteesi otosten varianssien samankaltaisuudesta hylätä. Varianssien samankaltaisuuden testaaminen, kuten moni muukin parametrinen testi, olettaa muuttujien normaalisuutta. Tulosrivin keskiarvoon suhteutetut virheet eivät ole normaalijakautuneita (ks. sivu 48).

F-testillä tutkitaan kahden muuttujan varianssin suhdetta. Normaalisuusehdon rikkominen saattaa aiheuttaa F-testiin virheen. Koska F-testi on hyvin herkkä normaalisuusehdon rikkomiselle, tarkastellaan samaa asiaa myös Levenen testillä, jossa analysoidaan datapisteiden etäisyyttä otoksen keskiarvosta, ja joka ei edellytä normaalisuutta. Koska Levenen testin tulokseen vaikuttavat keskiarvoa vääristävät poikkeavat havainnot, otetaan mukaan myös Brown-Forsythe -testi, joka käyttää keskiarvon sijaan vertailulukuna mediaania.

Kahden muuttujan suhde

Osana aineiston kuvaamista on tarpeen tarkastella kahden muuttujan välistä suhdetta (ks. esim. sivu 51, suuruusluokka ja variaatiokerroin). Tämäkin voidaan tehdä sekä graafisesti, että tilastollisesti. Tässä työssä tarkastellaan nk. Pearsonin korrelaatiokertoimia, sekä niihin liittyvän korrelaatiotestin tuloksia. Korrelaatiotestillä selvitetään laskennallinen todennäköisyys (p) sille, että korrelaatio olisi otoksessa satunnaisvaihtelun aiheuttamaa. Pearsonin korrelaation osalta on huomattava, että sen arvo kuvaa ainoastaan muuttujien välistä lineaarista korrelaatiota. Korrelaatiokerroin lasketaan tekemällä muuttujien välisistä kovariansseista yhteismitallisia jakamalla ne keskihajontojen tulolla, minkä vuoksi korrelaatiokerroin saa aina arvon välillä $[-1, 1]$. Korrelaatiokertoimen arvo 1 kuvaa täydellistä positiivista korrelaatiota ja -1 täydellistä negatiivista korrelaatiota.

Lineaarinen regressio

Tärkein menetelmä tässä työssä on lineaarinen regressioanalyysi. Linearisessa regressiossa tarkastellaan selitettävän muuttujan lineaarista riippuvuutta selittävistä muuttujista. Analyysin tuloksena saadaan nk. estimaattorit, eli numeroarvot, joiden perusteella selitettävän muuttujan arvoa voitaisiin estimoida, kun tiedetään selittävien muuttujien arvot.

Selitettävän muuttujan todellisen (ts. aineistossa kussakin tietueessa olevan arvon) ja estimoidun arvon välistä erotusta kutsutaan tilastotieteessä joskus virheeksi (varsinkin englanniksi käytetään usein sanaa *error*), mutta – kuten edellä todettu – työn aiheesta johtuen, sekaannusten välttämiseksi, tässä työssä todellisen ja estimoidun arvon erotusta kutsutaan residuaaliksi eli jäännökseksi.

Regressiomallien yhteydessä on tulosluvussa raportoitu residuaalien neliöiden neliöjuurten keskiarvo (*Root-mean-square error*, RMSE), joka kuvaa sitä, kuinka kaukana selitettävän muuttujan estimoitu arvo keskimäärin on todellisesta arvosta. RMSE:n osalta on huomioitava, että sen kokoluokka riippuu selitettävän muuttujan kokoluokasta, ts. että se ei ole eri regressioanalyysien välillä yhteismitallinen luku.

Sitä osaa selitettävän muuttujan arvojen vaihtelusta, joka regressiomallin mukaan selittyy selittävien muuttujien arvojen vaihtelulla, sanotaan selitysasteeksi, R^2 . Koska huonojenkin selittävien muuttujien lisääminen regressiomalliin kasvattaa selitysastetta laskennallisesti, ilmoitetaan tulosluvussa myös nk. korjattu selitysaste, $R^2_{korj.}$, jossa selittävien muuttujien lukumäärän lisäys on kompensoitu. Sosiaalitieteissä regressiomallien selitysasteet ovat usein hyvin alhaisia, koska ihmisten käytökseen ja elämään vaikuttavia tekijöitä on loputtoman paljon. Alhaiset selitysasteet kuvaavat olennaisesti tutkimusasetelmaa, jossa pyritään selvittämään, onko tietyillä asioilla ylipäätään yhteyttä ennustevirheisiin. Sen sijaan asetelma, jossa pyrittäisiin selittämään ennustevirheet, edellyttäisi korkeita selitysasteita.

Selitysasteiden ja RMSE:n lisäksi jokaisesta regressiomallista on raportoitu otoskoko (N). Osassa malleista on selittävänä muuttujana myös kahden muun selittävän muuttujan tulo (tai useampi tällainen). Näitä kutsutaan interaktiotermeiksi. Interaktiotermillä tutkitaan oletusta siitä, että yhden selittävän muuttujan vaikutus selitettävään muuttujaan on riippuvainen toisen selittävän muuttujan arvosta.

Regressioanalyysin lähtöoletukset

Jotta regressioanalyysin tuloksia voitaisiin pitää valideina, tulee regressiomallin täyttää tiettyjä oletuksia. T-testin edellyttämän normaalisuuden lisäksi lineaarisen regression validiteetin edellyttämät, nk. Gauss-Markov -oletukset ovat:

- residuaalien odotusarvo on 0
- selittävät muuttujat ja residuaalit ovat toisistaan riippumattomia
- residuaalien varianssi on vakio (ei heteroskedastisuutta)
- residuaalit eivät ole keskenään korreloituneita (ei autokorrelaatiota)

Gauss-Markov -oletusten täytyessä lineaariset estimaattorit ovat harhattomat ja parhaat lineaariset estimaattorit todellisista (tuntemattomista) parametreistä. Gauss-Markov -oletukset ovat kuitenkin hyvin ankarat, eivätkä ne usein käytännön sovellutuksissa päde. Tämä pitää paikkansa myös tämän työ osalta. Lineaarinen estimaattori on kuitenkin harhaton, vaikka jäännöstermi olisi autokorreloitunut ja heteroskedastinen, jos jäännöstermin keskiarvo on 0, ja jäännöstermi ja selittävät muuttujat ovat toisistaan riippumattomia (Verbeek 2012, s. 16).

Heteroskedastisuus ja sen testaaminen

Mikäli residuaalien (jäännöstermin) varianssi ei ole vakio, regressiomalli on heteroskedastinen. Tyypillisesti heteroskedastisuus ei tee estimaattoreista harhaisia, mutta estimaattoreiden keskivirheet ovat väärin. Tällöin regressiomallin tilastollinen merkitsevyys on epäluotettava.

Heteroskedastisuutta voi arvioida tarkastelemalla residuaaleja graafisesti tai testata mm. Breusch-Pagan -testillä tai White-testillä, joista ensin mainittu tuo esille ennen kaikkea lineaarista heteroskedastisuutta. Kun heteroskedastisuuden muodosta ei ole tarkkaa oletusta, graafinen tarkastelu voi olla paras tapa (Verbeek 2012, s. 107). Residuaalianalyysijä ei ole sisällytetty tuloslukuun. Regressiomallien yhteydessä on raportoitu joko Breusch-Pagan tai White -testin p sen mukaan, millainen heteroskedastisuuden muodon on tulkittu olevan.

Multikollinearisuus

Lineaarisia estimaattoreita ei voida laskea, jos joku selittävistä muuttujista muodostuu suoraan toisista selittävistä muuttujista (ns. multikollinearisuus). Tyypillinen esimerkki tällaisesta tapauksesta on ns. binäärimuuttuja-ansa, eli tilanne, jossa binäärimuuttujia on mallissa yksi liikaa. Tämän vuoksi regressioissa, joissa on binäärisiä luokkia (esimerkiksi 6 eri hallintoyksikköä, joista SMA-lukuja on koottu), on luokkia kuvaavia binäärimuuttujia yksi vähemmän kuin luokkia. Jäljelle jäävää luokkaa kuvaa tilanne, jossa kaikkien muiden binäärimuuttujien arvo on 0.

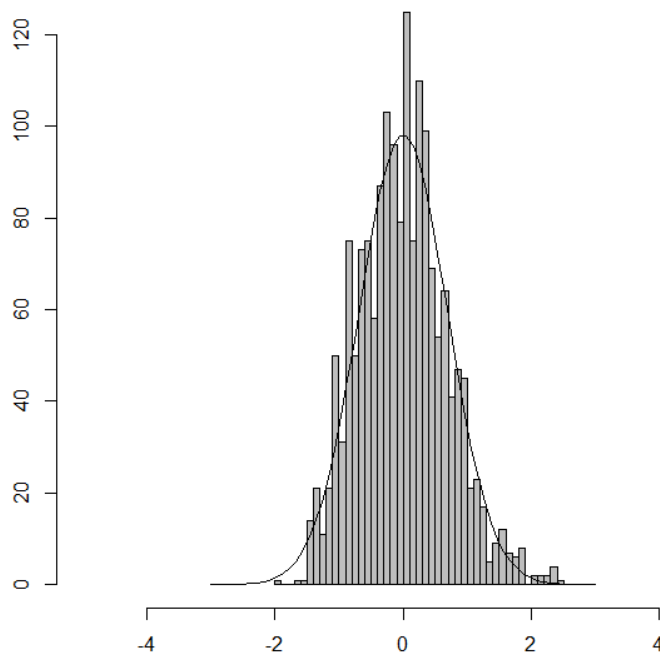
Normaalisuus ja sen testaaminen

Jos residuaalit ovat normaalijakautuneita, silloin myös estimaattori on normaalijakautunut. Jos estimaattori on myös harhaton, eli riittävät ehdot täyttyvät, niin estimaattori on normaalijakautunut todellisen (tuntemattoman) parametrin ympärillä. Tällöin t-testin edellyttämä t noudattaa Studentin t -jakaumaa, otoskoosta ja mallista riippuvilla vapausasteilla. Ja, edelleen, tällöin regressioanalyysin tärkeintä tulosta, eli t-testin tilastollista merkitsevyyttä, voidaan pitää luotettavana.

Residuaalien normalisuutta voidaan tarkastella graafisesti tai testata. Tässä työssä on raportoitu normalisuuden osalta Shapiro-Wilk -testin ja Kolmogorov-Smirnov -testin tuloksia. Näistä etenkin jälkimmäinen on erittäin herkkä hylkäämään nollihypoteesin normalisuudesta, etenkin kun otoskoko on suuri. Siksi residuaaleja on tarkasteltu graafisesti, jotta voidaan todeta jakauman muistuttavan likimääräisesti normaalijakaumaa. Jos residuaalit eivät ole normaalijakautuneita, t voi kuitenkin asymptoottisesti noudattaa t-jakaumaa, kunhan otoskoko on riittävän suuri.

Kuvassa 12 on esitetty esimerkki (taulukon 10 regressiomallista 4) residuaalien asymptoottisesta normalisuudesta, joka voidaan havaita graafisessa tarkastelussa. Testeillä voidaan todeta, ettei jakauma ole muuttujamuunnoksen jälkeenkään normaalijakautunut, $W = 0.99$, $p < 0.001$, $D = 0.52$, $p < 0.001$. Kuitenkin kuvasta voidaan nähdä, että jakauma muistuttaa likimääräisesti normaalijakaumaa. Normaalijakauma on piirretty kuvassa histogrammin päälle havainnollistamisen vuoksi.

Lineaarisia (OLS) estimaattoreita käytetään toivoen, että otoskoko on riittävän suuri siten, että estimaattoreilla on likimäärin tarvittavat ominaisuudet. Tämä on tyypillistä OLS estimaattoreita sovellettaessa käytännössä (Verbeek 2012, s. 35). Asia tulee huomioida tuloksia tarkasteltaessa. Tehtäessä päätelmiä estimaattoreiden numeerisista arvoista pitää olla hyvin varovainen.



Kuva 12. Asymptoottinen normalisuus, esimerkki.

Hyvin heikkojenkin ehtojen täyttyessä (Verbeek 2012, s. 34) otoskoon kasvaessa riittävästi, otoksen ominaisuudet konvergoituvat kohti populaation ominaisuuksia, jolloin estimaattori on konsistentti. Tähän vaaditaan vain, että jäännöstermin odotusarvo on 0, ja että se on riippumaton selittävistä muuttujista. Sitä, milloin otoskoko on riittävän suuri, ei voida mitenkään yksikäsitteisesti määrittää (Verbeek 2012, s. 36).

Autokorrelaatio paneelidatassa

Paneelidata on poikkileikkauksen ja aikasarjan yhdistelmä, jossa on tietoa useista yksilöistä, ja samoista yksilöistä useilta eri ajanjaksoilta. ”Yksilö” on tässä käsittelyssä olevassa otoksessa yhden organisaation yksi tulosrivi, esimerkiksi: ”Suomenlahden merivartioston rajatarkastukset”. Tämän vuoksi paneelidatassa ei voida perustellusti olettaa samasta yksilöistä tehtyjen perättäisten havaintojen olevan toisistaan riippumattomia. (Verbeek 2012, s. 372-373.)

Autokorrelaatio tarkoittaa tilannetta, jossa residuaalit ovat keskenään korreloituneita. Tässä työssä käytetyssä paneelidatassa residuaalit ovat autokorreloituneita, kuten paneelidatan kohdalla usein on. Kuten heteroskedastisuuden osalta, autokorrelaatiokaan ei välttämättä tee OLS estimaattorista harhaista, mutta estimaattoreiden keskivirheet lasketaan väärin (Verbeek 2012, s. 112).

Tässä työssä autokorrelaatiota ei voidakaan täysin erottaa tutkittavasta ilmiöstä. On tarkoin pohdittava, onko ennustevirheiden autokorrelaatiolle rationaalisia syitä, jotka eivät liittyisi teorialuvussa esitettyyn. Esimerkiksi tilanteessa, jossa – teorialuvussa kuvatun mukaisesti – ennusteen laatija ei riittävästi reagoisi toteumadataan tai aiempien ennusteiden virheellisyyteen, olisivat saman tulosrivin perättäisten vuosien ennustevirheet samansuuntaisia.

Autokorrelaation ja heteroskedastisuuden huomioivat keskivirheet

Kuten edellä on kerrottu, lineaarisessa regressiossa lasketaan estimaattorille todennäköisin lukuarvo (odotusarvo), ja keskivirhe. Jos oletetaan, että estimaattorin saamat mahdolliset

lukuarvot ovat normaalijakautuneet, voidaan keskivirheen avulla tehdä tilastollinen testi (t-testi) siitä, kuinka todennäköistä on, että estimaattorin todellinen (tuntematon) arvo poikkeaa nollasta. Estimaattorien tavallisten (ordinary least squares, OLS) keskivirheiden käyttö testauksessa edellyttää, että lineaarisen regression lähtöoletusten mukaisesti:

- jäännösten (residuaalien) varianssi pysyy samankaltaisena riippumatta regressiomallin muuttujien arvoista (ei heteroskedastisuutta), ja
- jäännökset eivät ole keskenään korreloituneita (ei autokorrelaatiota).

Joissain malleissa heteroskedastisuus voidaan korjata regressiomallia muuttamalla tai selitettävän muuttujan Box-Cox -muunnoksella (ks. seuraava kappale). Osassa malleista muuttujamuunnos ei kuitenkaan korjaa heteroskedastisuutta. Tällöin on yleisesti käytettyjen (OLS) keskivirheiden sijasta käytetty merkitsevyyden laskennassa autokorrelaation ja heteroskedastisuuden huomioivia keskivirheitä (*heteroskedasticity-robust standard errors*). (Verbeek 2012, s. 126)

On myös muita tapoja laskea autokorrelaation ja heteroskedastisuuden kompensoivat keskivirheet, mutta tässä työssä erot Whiten keskivirheiden, ja muiden robustien keskivirheiden välillä tulivat esiin kolmannessa merkitsevässä numerossa, eli näillä eroilla ei ole merkitystä tutkimustulosten kannalta.

Muuttujamuunnokset

Osassa regressiomalleista jäännöstermin jakauma ei muistuta normaalijakaumaa riittävästi, ja/tai lineaarisen regression edellytys jäännösten varianssin pysymisestä vakiona (homoskedastisuus) ei täyty. Esimerkiksi tarkasteltaessa ennustetarkkuutta, on selitettävänä muuttujana ennustevirheen itseisarvo, jonka jakauma on epäsymmetrinen. Joissain malleissa normaalisuuteen ja heteroskedastisuuteen liittyvät ongelmat voidaan korjata muuttujamuunnoksella. Muuttujamuunnos tarkoittaa, että muuttujan arvo korvataan toisella luvulla, joka on laskettu muuttujan arvosta. Tilastollinen testaaminen voisi olla validia ilman muunnostakin, mutta regression tulosten oikeellisuus on herkkä lähtöoletusten rikkomiselle (Box & Cox 1964, 227).

Yksi tässä työssä käytetyistä muunnoksista on 10-kantainen logaritmi. Tulosrivin toteuman keskiarvon logaritmia käytetään kuvaamaan tulosrivin toteuman suuruusluokkaa, tulosrivin

toteuman keskiarvon sijasta. Osassa regressioista selitettävä muuttuja on kerrottu sadalla, luettavuuden parantamiseksi, mikä on myös muuttujamuunnos. Osassa malleista tehdään kuitenkin näennäisesti monimutkaisempi Box-Cox -muunnos selitettävälle muuttujalle.

Regressiomallin selitettävän muuttujan muunnoksen voi tehdä, jos muuttuja ei ole ekstensiivinen. Tämä tarkoittaa, että muuttuja ei ole riippuvainen sen systeemin koosta, jonka osa se on. Esimerkiksi ennustevirheitä ei voi mielekkäästi laskea yhteen kuten esimerkiksi massaa tai tilavuutta.

Toinen edellytys muuttujamuunnokselle on, että muunnos on monotoninen. Monotoninen muunnos tarkoittaa, ettei muuttujan arvojen järjestys vaihdu. Jos $y_1 > y_2$, niin muunnoksen f jälkeen on oltava $f(y_1) > f(y_2)$. Jos näin on, kovarianssi ei voi aiheutua muunnoksesta. Toisin sanoen, selitettävän ja selittävän muuttujan välinen korrelaatio ja sen tilastollinen merkitsevyys voidaan todeta luotettavasti.

Muunnoksen jälkeen estimaattoreiden numeroarvoja ei kuitenkaan voi suoraan tulkita. Numeroarvojen suoran tulkitsemisen lisäksi ei kuitenkaan ole mitään muita syitä, joissa regressioanalyysin näkökulmasta alkuperäinen ei-ekstensiivinen muuttuja olisi parempi kuin mikään sen monotoninen muunnos (Box & Cox 1964, 214).

Box-Cox -muunnos edellyttää oikeaan osuneiden ennusteiden huomioimatta jättämistä, koska muunnettavan selitettävän muuttujan tulee olla positiivinen (eri suuri kuin 0). Otoskoko on siksi pienempi niissä regressiomalleissa, joissa muunnosta on käytetty. Muunnokseen liittyvä parametri (λ) valitaan sen mukaan, mikä arvo parhaiten tuottaa normaalijakaumaa vastaavan jäännösten jakauman.

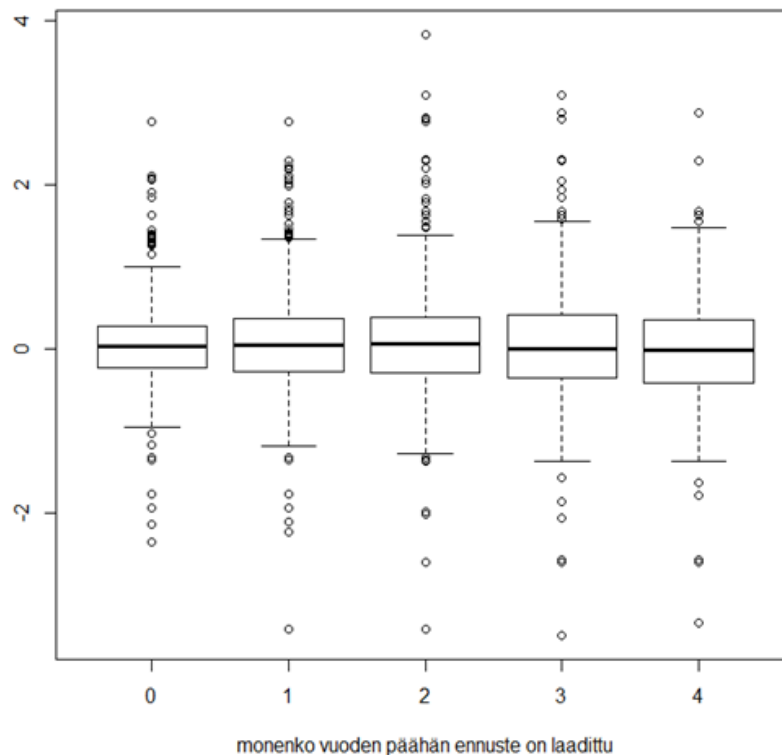
Niissä regressiomalleissa, joissa tutkitaan ennustetarkkuutta, tarkastellaan ennustevirheiden sijaan ennustevirheiden itseisarvoja. Toinen tapa merkitä samaa asiaa on virheen neliön neliöjuuri. Kumpikin johtaa siihen, että virheiden alkuperäinen etumerkki jää pois, ja tarkasteltavana on ainoastaan ennusteen ja toteuman välinen etäisyys, ts. ainoastaan virheen koko, ei sen suunta. Muuttujamuunnosten yhteydessä on huomautettava, että itseisarvon ottaminen ei ole monotoninen muunnos, vaan selitettävän muuttujan arvojen järjestys vaihtuu täysin toisenlaiseksi, ja regression aiheen oleva asia muuttuu itseisarvon ottamisen jälkeen aivan toiseksi.

4. TULOKSET

4.1. Eri aikaväleille laaditut ennusteet

Vertaillaan eri aikaväleille tehtyjen ennusteiden tulosrivin keskiarvoon suhteutettuja virheitä. Ennustetarkkuutta voidaan havainnollistaa tutkimalla ennustevirheiden jakauman muita ominaisuuksia kuin keskiarvoa. Kuvassa 13 on kuvattu laatikkodiagrammi eri aikaväleille laadittujen ennusteiden virheistä. Kuvaajassa on esitetty paksulla mustalla viivalla virheiden mediaani. Mediaanin ympärillä on laatikko, joka kuvaa ensimmäisen ja kolmannen kvartiilirajan väliin jäävää aluetta (kvartiiliväli). Toisin sanoen puolet ennustevirheistä on laatikon sisäpuolella, ja alin ja ylin neljännes laatikon ulkopuolella.

Laatikon ylä- ja alapuolella on ohuemmilla viivoilla erotettu katkoviivalla merkitty alue. Katkoviivan pituus on 1.5 kertaa kvartiiliväli. Ennustevirheet, jotka ovat tätä kauempana kvartiilivälin ylä- tai alarajasta, on esitetty yksitellen palloilla (*outliers*; poikkeavat havainnot). Mitä kauemmas tulevaisuuteen ennuste on laadittu, sitä suuremmat ovat toteumaan liittyvät epävarmuustekijät. Looginen oletus ennustetarkkuuden heikentymisestä tarkoittaisi, että virheiden hajonta kasvaa aikavälin kasvaessa.



Kuva 13. Eri aikaväleille laadittujen ennusteiden tulosrivin keskiarvoon suhteutetut virheet.

Taulukko 2.

Eri aikaväleille tehtyjen ennusteiden virheet, suhteutettuna kunkin tulosrivin toteumien keskiarvoon.

ennuste laadittu	keskiarvo	mediaani	keskihajonta	kvartiiliväli		n
				alaraja	yläraja	
samalle vuodelle	0.075	0.025	0.57	-0.23	0.28	422
seuraavalle vuodelle	0.097	0.042	0.67	-0.28	0.38	415
kahden vuoden päähän	0.12	0.059	0.76	-0.29	0.39	415
kolmen vuoden päähän	0.050	-0.0030	0.76	-0.35	0.42	345
neljän vuoden päähän	-0.0068	-0.017	0.72	-0.41	0.35	276

Graafisesti tarkasteltuna ennustevirheiden hajonta kasvaa oletuksen mukaisesti. Kuvassa 13 havaitaan positiivinen harhaisuus, joka on suurimmillaan kahden vuoden päähän tehdyissä ennusteissa. Ennustetarkkuuden heikentyminen on nähtävissä kuvaajassa jakauman levenemisenä. Aikavälin pidentyessä poikkeavia havaintoja on enemmän ja ne ovat äärimmäisempiä. Poikkeuksen muodostavat 3 ja 4 vuoden päähän laaditut ennusteet, joiden hajonta vaikuttaa graafisesti tarkastellen samankokoiselta.

Taulukko esittää kuvassa 13 kuvattujen mediaanin ja kvartiilivälin numeroarvot, ja näiden lisäksi otosten keskiarvot ja keskihajonnat. Aivan kuten graafisesta tarkastelusta havaittiin, ennustevirheiden hajonta kasvaa ja kvartiiliväli levenee, pl. kolmen ja neljän vuoden päähän tehtyjen ennusteiden välillä. Ennustevirheissä on positiivinen harha viimeistä ryhmää lukuun ottamatta. Ensimmäisellä kolmella aikavälillä ennusteet ovat keskimäärin 7.5 %, 9.7 % ja 11.8 % toteuman yläpuolella. Aineistossa on myös pieni positiivinen vinouma; keskiarvot ovat mediaaneja suurempia. Toisin sanoen, positiivisia ääritapauksia on enemmän ja positiiviset virheet ovat hieman suurempia kuin negatiiviset.

Seuraavaksi tutkitaan tilastollisilla testeillä, onko kuvassa 13 ja taulukossa 2 havaittava ennusteiden epätarkkuuden kasvu tilastollisesti merkitsevää. F-testin lisäksi tehtiin Levenen testi ja Brown-Forsythe -testi, koska tiedämme (ks. sivu 48), etteivät ennustevirheiden otokset ole normaalijakautuneita, eivätkä täysin symmetrisiä poikkeavien havaintojen osalta. Koska testien tulokset poikkeavat toisistaan, on kaikkien testien tulokset raportoitu alla.

Niiden ryhmien osalta, joiden välillä nollahypoteesi hajonnan samankaltaisuudesta jää kumoamatta, on raportoitu myös varianssianalyysi (ANOVA), joka kertoo, eroavatko ryhmien keskiarvot merkitsevästi toisistaan. On huomattava, että koska tässä ei tarkastella virheiden itseisarvoja, keskiarvon muutos ei kerro ennustetarkkuuden muutoksesta, vaan eroista ennusteiden keskimääräisessä harhaisuudessa.

Taulukko 3.
Otosten hajontalukujen vertailu.

Vertailtavat	Aikaväli numeroina	F-testi	Levene	Brown- Forsythe	ANOVA (keskiarvot)
Samalle vuodelle ja 1 vuoden päähän tehdyt ennusteet	0 ja 1	1.35 ** $p = 0.002$ (414,421)	6.16 * $p = 0.013$ (1, 835)	6.04 * $p = 0.014$ (1, 835)	–
yhden ja kahden vuoden päähän tehdyt ennusteet	1 ja 2	1.28 * $p = 0.012$ (414, 414)	2.25 $p = 0.13$ (1, 828)	2.13 $p = 0.15$ (1, 828)	0.18 $p = 0.67$ (1, 828)
kahden ja kolmen vuoden päähän tehdyt ennusteet	2 ja 3	0.998 $p = 0.98$ (414, 344)	0.17 $p = 0.68$ (1, 758)	0.23 $p = 0.63$ (1, 758)	1.54 $p = 0.22$ (1, 758)
kolmen ja neljän vuoden päähän tehdyt ennusteet	3 ja 4	1.11 $p = 0.36$ (344, 275)	0.21 $p = 0.65$ (1, 619)	0.18 $p = 0.67$ (1, 619)	0.90 $p = 0.34$ (1, 619)

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Vapausasteet sulkeissa. Keskihajonnat ja kvartiilivälit taulukossa 2.

Kaikkien testien mukaan vuoden päähän tehdyt ennusteet ovat merkitsevästi erilaisia (tarkkuus heikkenee) kuin samaa vuotta koskevat ennusteet. Tarkasteltaessa ennusteiden hajontaa graafisesti, vaikuttaa siltä, että vastaava vaikutus on nähtävissä aina kolmen vuoden päähän tehtyihin ennusteisiin asti. Tilastollisesti efekti kuitenkin vaimenee.

F-testin mukaan tarkkuus vähenee edelleen merkitsevästi siirryttäessä 1 vuoden päähän laadituista 2 vuoden päähän laadittuihin ennusteisiin. Normaalisuusehdon rikkominen saattaa aiheuttaa F-testiin virheen. Muilla testeillä nollahypoteesi homoskedastisuudesta jää täpärästi kumoamatta. Kaikilla tässä käytetyillä testeillä on tyypillistä saada virhe, jossa oikea nollahypoteesi kumotaan aiheettomasti. Näin ollen testien luonteestakin päätellen on todennäköisempää, että tältä osin nimenomaan F-testin tulos on virheellinen. Toisaalta graafinen tarkastelu tuottaa samansuuntaisen havainnon F-testin kanssa. Tulokset ovat tältä osin ristiriitaiset.

Kahden ja kolmen vuoden päähän laadittujen ennusteiden välillä vaikuttaa graafisesti tarkastellen (kuva 13) siltä, että hajonta kasvaa edelleen ajanjaksojen välillä. Tilastollisesti näin ei kuitenkaan ole. Mikään testeistä ei kykene hylkäämään nollahypoteesia varianssien samankaltaisuudesta. Toisin sanoen, ennustetarkkuus ei merkitsevästi heikkene, verrattaessa kahden vuoden päähän ja sitä kauemmas tehtyjä ennusteita.

Varianssianalyysistä (ANOVA) voidaan todeta, että eri aikaväleille tehtyjen ennusteiden harhaisuudessa ei ole tilastollisesti merkitseviä eroja, vaikka graafisesti tarkasteltuna siltä vaikuttaakin.

Poikkeavien havaintojen (ääritapausten) merkityksen arvioimiseksi tehtiin myös Levenen testi siten, että 5% suurimpia ja pienimpiä arvoja jätettiin huomioimatta. Tulokset olivat hyvin samankaltaiset kuin edellä. Tulokset eivät siis ole poikkeavien havaintojen aiheuttamia.

Huolellisuuden vuoksi vastaavat testit tehtiin myös suhteellisilla virheillä (kunkin vuoden toteumaan suhteutetut virheet). Suhteellisten virheiden jakauman vuoksi (voimakkaasti epäsymmetrinen, ei normaalijakauma, ks. sivu 46) ainoa soveltuva testi on Brown-Forsythe – testi. Merkitseviä eroja eri aikaväleille tehtyjen ennusteiden tarkkuuksissa ei kyetty suhteellisilla virheillä havaitsemaan, mikä liittyy jakaumien äärimmäisyyteen ja epäsymmetrisyyteen. Suhteellisilla virheillä jakauman muoto ja ääripisteet peittävät alleen varianssin vaihtelun.

Ennustetarkkuuden huonontumista sen mukaan, miten kauas tulevaisuuteen ennuste kohdistuu, voidaan kuvata myös lineaarisella regressiolla. Selitettäväksi muuttujaksi otetaan tulosrivin keskiarvoon suhteutettujen virheiden itseisarvo. Virheiden itseisarvojen jakauma on hyvin epäsymmetrinen, ja rikkoo lineaarisen regression edellyttämää normaalisuusehtoa räikeästi. Regressiota varten tehdään selitettävän muuttujan Box-Cox –muunnos ($\lambda = 0.37$). Muunnos edellyttää muuttujan positiivisuutta, minkä vuoksi regressiossa jätetään huomioimatta oikeaan osuneet ennusteet (joiden osalta $|\text{virhe}| = 0$). Muuttujamuunnos poisti heteroskedastisuuden, Breusch-Pagan $p = 0.37$).

Taulukko 4.
Regressioanalyysi ennustetarkkuuden huononemisesta
ennusteaikavälin mukaan.

	Malli 1
Aikaväli (vuosissa)	0.062 *** (0.012)
Vakiotermi	-0.98 *** (0.03)
R^2	0.0093
R^2 korj.	0.0088
RMSE	0.51
N	1815

Keskivirhe on sulkeissa

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Ennustevirheet ovat sitä suurempia, mitä kauemmas tulevaisuuteen ne on tehty. Yhteys on tilastollisesti merkitsevä, mutta aikavälin vaikutus selittää erittäin pienen osan virheiden varianssista. Regression tuoma lisäarvo onkin lähinnä havainnollistaa vaikutuksen suuruutta.

Graafisesta tarkastelusta on selvää, että aikavälin ja ennustevirheiden yhteys ei ole lineaarinen. Jonkinlainen ennustetarkkuuden muutos on nähtävissä lineaarisessa mallissa, mutta ryhmien pareittaiset vertailut antavat huomattavasti informatiivisemman tuloksen. Breusch-Pagan testin perusteella tiedetään muuttujamuunnoksen korjanneen heteroskedastisuuden. Taulukossa 5 on raportoitu vain F-testin tulokset. Vertailtavien

ryhmien varianssien samankaltaisuus testattiin myös Levenen testillä ja Brown-Forsythe -testillä, jotka jätetty luettavuuden vuoksi raportoimatta tässä. Vertailussa on käytetty samaa muuttujamuunnosta kuin taulukossa 4 ($\lambda = 0.37$).

Taulukko 5.
Ennustetarkkuuden vertailu ajanjaksojen välillä.

Tulosrivin keskiarvoon suhteutettujen virheiden itseisarvojen Box-Cox -muunnos ($\lambda = 0.37$)

Aikaväli (vuodet)	F-testi	ANOVA	N
0 ja 1	0.95 $p = 0.63$ (391, 414)	8.96 ** $p = 0.003$ (1, 805)	807
1 ja 2	0.89 $p = 0.24$ (414, 386)	2.59 $p = 0.11$ (1, 800)	802
2 ja 3	1.03 $p = 0.75$ (386, 344)	0.15 $p = 0.70$ (1, 730)	732
3 ja 4	1.08 $p = 0.52$ (344, 275)	0.036 $p = 0.85$ (1, 619)	621

Vapausasteet sulkeissa

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Samaa vuotta koskevat ennusteet ovat merkitsevästi tarkempia kuin seuraavaa vuotta koskevat ennusteet, ja muissa vertailuissa ei todeta merkitseviä eroja. Linearisessa mallissa (taulukko 4) todettu vaikutus aiheutuu täysin erosta samalle vuodelle ja tätä kauemmas tehtyjen ennusteiden välillä.

4.2. Painokerroin ja toteuman suuruusluokka

Tarkastellaan ennustetarkkuuden yhteyttä painokertoimeen sekä tulosrivin keskiarvoon ja suuruusluokkaan. Tarkasteltaessa ennustetarkkuutta selitettävänä muuttujana on käytettävä tulosrivin keskiarvoon suhteutettujen virheiden itseisarvoja. Käytetään selitettävän muuttujan Box-Cox -muunnosta ($\lambda = 0.30$). Malleissa 4 ja 5 Breusch-Pagan testin p on muunnoksen jälkeen 0.0065 ja 0.0026, minkä vuoksi tavalliset lineaarisen regression (OLS) keskivirheet korvattu heteroskedastisuuden huomioivilla Whiten keskivirheillä.

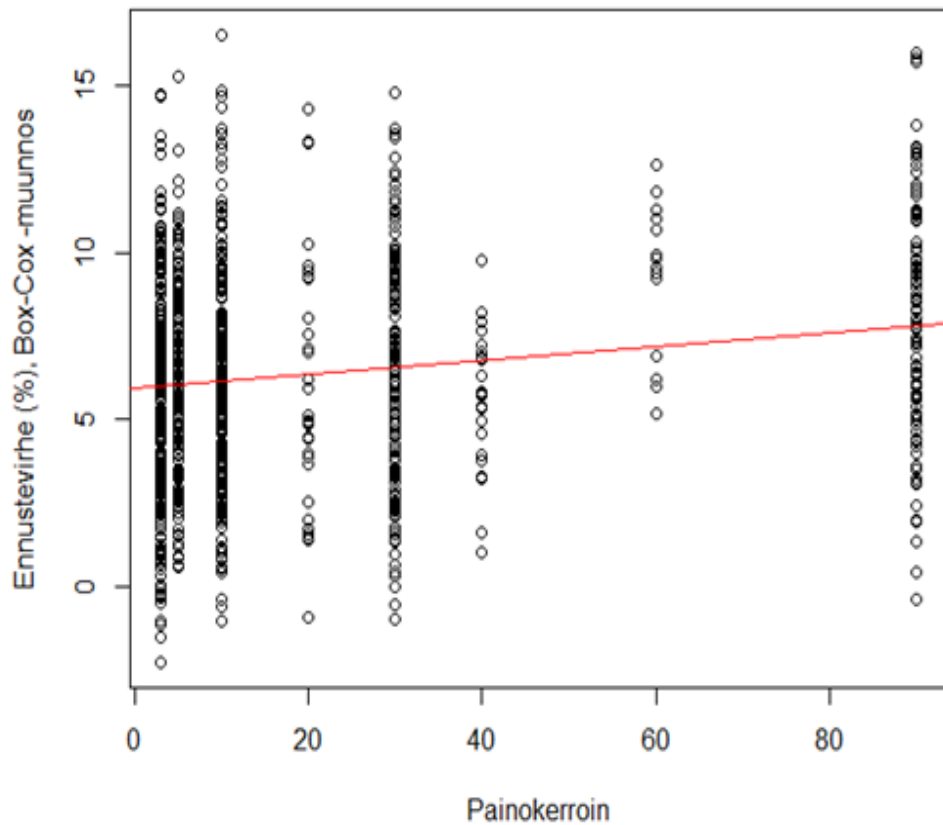
Taulukko 6.
Ennustetarkkuuden yhteys tulosrivillä tilastoitavan suoritteiden suuruusluokkaan ja painokertoimeen.

	Malli 1	Malli 2	Malli 3	Malli 4	Malli 5
Muunnos ($\lambda = 0.3$)					
Painokerroin	0.021 *** (0.0025)	0.013 *** (0.0031)	0.013 *** (0.003)	0.027 *** (0.005)	0.030 *** (0.005)
Toteuman keskiarvo (ka)		-1.1 * 10 ⁻⁷ (7.2 * 10 ⁻⁸)			-1.3 * 10 ⁻⁶ (1.3 * 10 ⁻⁶)
Painokerroin * Keskiarvo					3.95 * 10 ⁻⁷ (4.3 * 10 ⁻⁷)
Suuruusluokka (log ₁₀ (ka))		-0.20 *** (0.06)	-0.25 *** (0.05)	-0.19 *** (0.05)	-0.12 . (0.07)
Painokerroin * Suuruusluokka				-0.011 *** (0.003)	-0.013 *** (0.003)
Vakiotermi	5.92 *** (0.10)	6.72 *** (0.23)	6.86 *** (0.22)	6.82 *** (0.22)	6.65 *** (0.24)
R ²	0.039	0.054	0.053	0.0615	0.0652
R ² korj.	0.038	0.052	0.052	0.0598	0.0623
RMSE	3.10	3.08	3.08	3.063	3.059
N	1597	1597	1597	1597	1597

Keskivirhe on sulkeissa

. $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

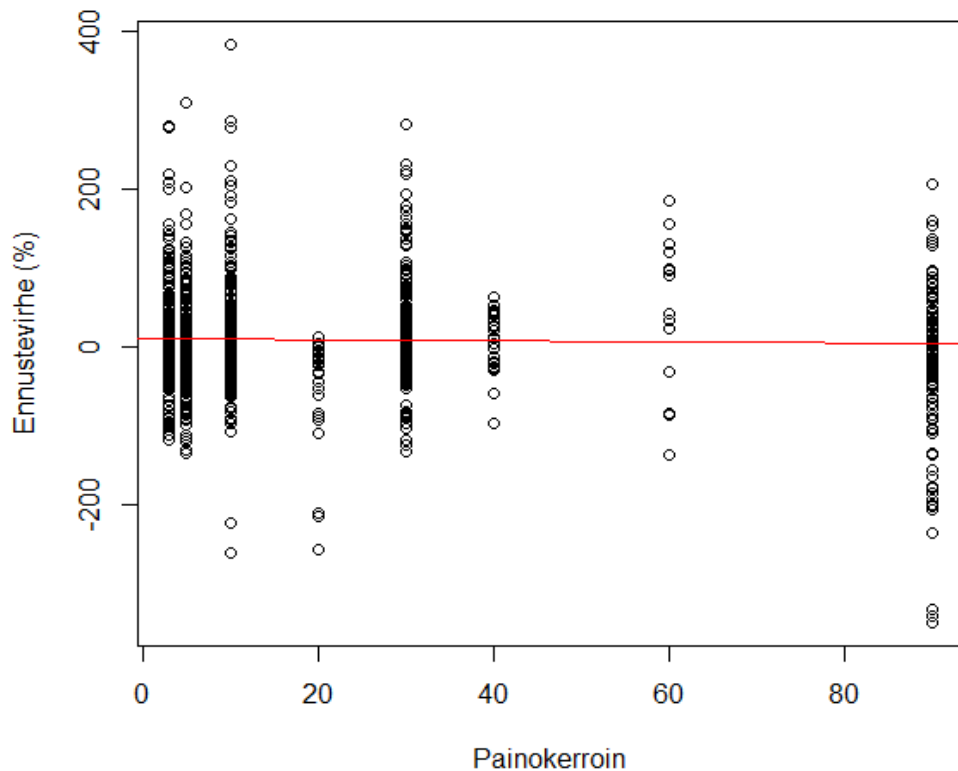
Painokertoimella, tulosrivin suuruusluokalla ja ennustetarkkuudella on yhteys. Painokerroin ennustaa positiivisesti ennustevirhettä. Suuruusluokka ennustaa virhettä negatiivisesti. Malli 1 on kuvattu alla (kuva 14).



Kuva 14. Painokertoimen yhteys ennustetarkkuuteen. Taulukon 6 regressiomalli 1:n mukaisesti.

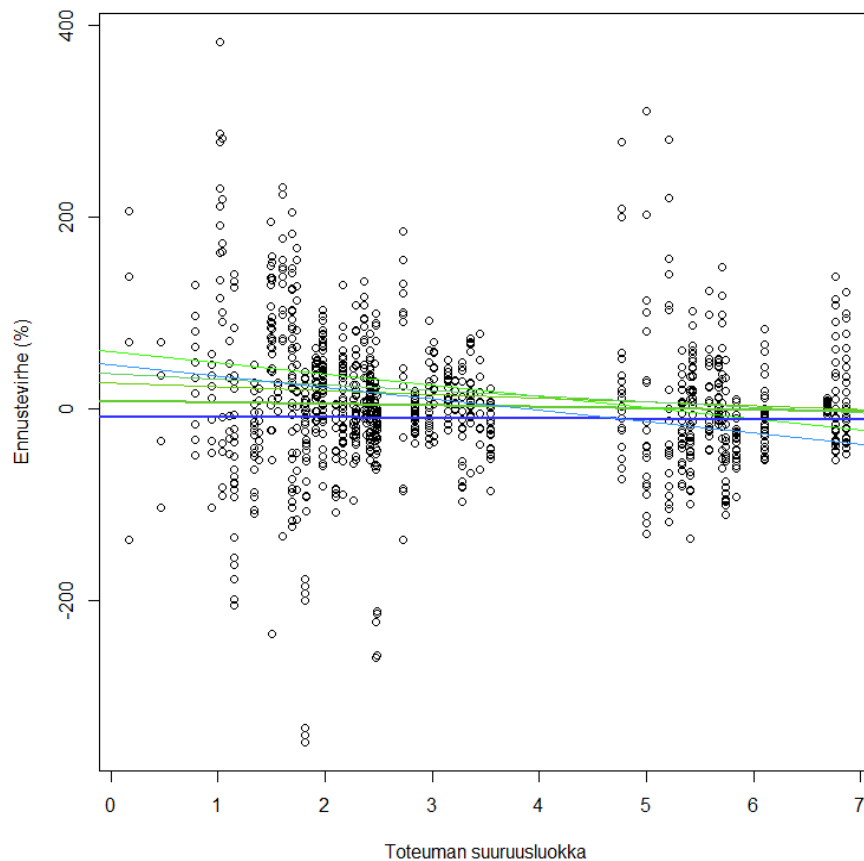
Malleissa 4 ja 5 on sisällytetty regressioon painokertoimen ja suuruusluokan interaktiotermin, joka niin ikään on merkitsevästi yhteydessä ennustevirheen kokoon, mikä selventää painokertoimen ja ennustevirheen välistä yhteyttä. Interaktiotermin merkitsevyys tarkoittaa, että painokertoimen yhteys ennustetarkkuuteen erilainen eri suuruusluokkaa olevilla tilastoriveillä. Siirryttäessä lukumääräisesti suurempiin tilastoriveihin, painokertoimen ja ennustevirheen yhteys kääntyy negatiiviseksi. Suurempi tai pienempi painokerroin ei sen sijaan muuta suuruusluokan ja ennustevirheen negatiivista korrelaatiota positiiviseksi. Suuruusluokka ennustaa siis ennustevirheen itseisarvoa negatiivisesti painokertoimesta riippumatta.

Painokerroin ei yksinään ole merkitsevästi yhteydessä ennusteen harhaisuuteen (heteroskedastisuuden huomioivilla keskivirheillä $p = 0.36$). Virheiden jakaumat eri painokertoimien suhteen, ilman muuttujamuunnosta ja itseisarvoa, on esitetty kuvassa 15.



Kuva 15. Painokertoimen yhteys ennusteen harhaisuuteen.

Kuvassa 16 on esitetty tulosrivin suuruusluokan ja ennusteen harhaisuuden yhteys. Estimaattorit ja merkitsevyys ovat eri hallintoyksiköitä koskevissa aineiston osissa erilaiset. Rajavartiostot (organisaatiot 1-4) on merkitty kuvaan virheällä, ja merivartiostot (organisaatiot 5-6) sinisellä värillä. Regression tulokset on raportoitu lyhyesti alla. Ennustevirheet on ilmaistu prosentteina, mutta muuttujalle ei ole tehty muuta muunnosta. Kuten aiemmin todettu, suuruusluokka on yhtä kuin toteuman keskiarvon 10-kantainen logaritmi. Toisin sanoen, suuruusluokka 2 tarkoittaa sataa suoritetta, ja suuruusluokka 6 tarkoittaa miljoonaa suoritetta.



Kuva 16. Ennustevirheet ja toteuman suuruusluokka.

Aiemmissä tarkasteluissa havaittu positiivinen harha on merkittävämpi niillä tulosriveillä, joiden toteumaluvut ovat pienempiä. Yhteys on tilastollisesti merkitsevä (heteroskedastisuuden huomioivilla keskivirheillä $p < 0.001$). Tämä ei kuitenkaan päde hallintoyksiköillä 1 ja 5 (Kaakkois-Suomen rajavartiosto ja Suomenlahden merivartiosto), jotka erosivat muista merkitsevästi ($p = 0.06$ ja $p = 0.04$). Jos huomioidaan otos vain näiden kahden hallintoyksikön osalta, ei yhteyttä toteuman suuruusluokan ja ennustevirheiden välillä ole ($p = 0.50$). Toinen huomioitava asia on, ettei kaikilla hallintoyksiköillä ole yhtään tulosriviä isoimmista suuruusluokissa. Hallintoyksiköiden eroja tarkastellaan seuraavassa tarkemmin.

4.3. Hallintoyksiköiden erot

Taulukossa 7 on esitetty tulokset regressioista, joilla tutkittiin (tuloksellisuuden laskennassa käytettävän) painokertoimen ja organisaation osaa (hallintoyksikköä) osoittavien binäärimuuttujien yhteyttä ennustevirheisiin. Kaikille tietokantaan otetuille tulosriveille (suoritteille) ei ole tuloksellisuuden laskennassa määritetty painokerrointa, ja tällaiset tulosrivit on jätetty näissä regressiomalleissa huomioimatta (N=1597).

Taulukko 7.
 Regressioanalyysi ennustevirheistä painokertoimen ja organisaation suhteen.

Luettavuuden vuoksi selitettävä muuttuja on kerrottu 100:lla.

	Harhaisuus		Tarkkuus	
	Malli 1	Malli 2	Malli 1	Malli 2
Painokerroin (Pk)		-1.18 *** (0.11)		0.013 * (0.006)
Org.bin.1	-17.65 ** (5.91)	-64.33 *** (6.88)	-0.69 ** (0.24)	-1.16 *** (0.31)
Org.bin.2	-1.59 (6.23)	-40.53 *** (7.70)	-1.01 *** (0.2)	-0.88 * (0.36)
Org.bin.3	-6.21 (6.08)	-39.72 *** (7.00)	-1.07 *** (0.27)	-1.36 *** (0.33)
Org.bin.4	-1.58 (8.35)	-51.48 *** (10.06)	0.71 * (0.30)	0.21 (0.38)
Org.bin.5	-29.98 *** (6.00)	-48.92 *** (7.03)	-1.67 *** (0.25)	-1.75 *** (0.33)
Pk * org.bin.1		1.77 *** (0.16)		0.015 * (0.007)
Pk * org.bin.2		1.61 *** (0.17)		-0.0036 (0.0083)
Pk * org.bin.3		1.36 *** (0.17881)		0.016 * (0.007)
Pk * org.bin.4		2.11 *** (0.27)		0.025 ** (0.008)
Pk * org.bin.5		0.74 *** (0.21)		0.0033 (0.0089)
Vakiotermi	20.69 *** (4.89)	50.52 *** (5.85)	7.09 *** (0.25)	6.77 *** (0.25)
R ²	0.025	0.116	0.058	0.104
R ² korj.	0.022	0.110	0.055	0.098
RMSE	72.52	69.21	3.07	3.00
N	1597	1597	1597	1597

*Keskivirhe on sulkeissa. Org.bin. = hallintoyksikköä kuvaava binäärimuuttuja. Pk = Painokerroin.
 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$*

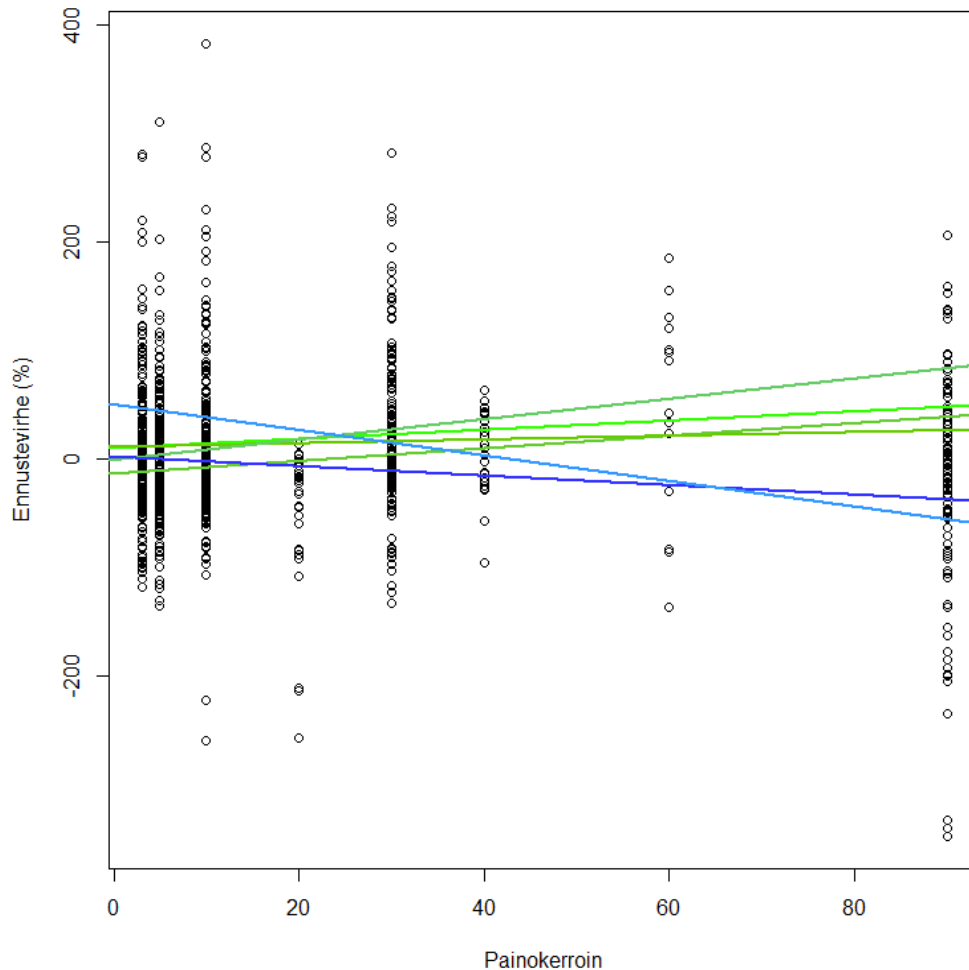
Taulukossa vasemmalla olevat mallit kuvaavat harhaisuutta, eli sitä, onko ennustevirhe systemaattisesti tietyn suuntainen. Näissä selitettävänä muuttujana on tulosrivin toteuman keskiarvoon suhteutettu ennustevirhe. Oikeanpuoleiset regressiomallit kuvaavat ennustetarkkuutta, eli ennusteen ja toteuman välistä etäisyyttä. Näissä malleissa selitettävänä muuttujana on tulosrivin keskiarvoon suhteutetun ennustevirheen itseisarvo. Selitettävä muuttuja on kaikissa malleissa kerrottu 100:lla, luettavuuden parantamiseksi.

Mallista 1 käy ilmi, että ennusteiden harhaisuuden osalta vain organisaatiot 1 ja 5 (Kaakkois-Suomen rajavartiosto ja Suomenlahden merivartiosto) poikkeavat merkitsevästi toisistaan ja muista. Organisaation 6 ennusteet ovat tarkasteluajanjaksolla keskimäärin 20.7 % toteumaa suuremmat. Organisaatioiden 2, 3 ja 4 osalta lukema on pienempi, mutta erot eivät ole tilastollisesti merkitseviä. Organisaation 1 ennusteet ovat keskimäärin 12.3 % toteuman yläpuolella, ja ero muihin on merkitsevä ($p < 0.01$). Organisaatio 5 on ainoa, jolla harhaisuus on erisuuntainen kuin muilla; sen ennusteet ovat keskimäärin 9.3 % toteutuneiden lukujen alapuolella. Tämäkin ero muihin on tilastollisesti merkitsevä ($p < 0.001$).

Painokerroin ei yksinään ollut merkitsevästi yhteydessä ennusteen harhaisuuteen (kuva 15). Tämä on jätetty luettavuuden parantamiseksi merkitsemättä erikseen taulukkoon 7. Mutta kun painokertoimen ja ennusteen harhaisuuden yhteyttä tarkastellaan siten, että huomioidaan myös hallintoyksiköt, nousee tilastollisesti merkitsevä yhteys esiin, ja kaikki hallintoyksiköt erottuvat toisistaan tilastollisesti merkitsevästi Mallissa 2. Interaktiotermin tilastollinen merkitsevyys tarkoittaa, että toinen muuttuja vaikuttaa siihen, millainen yhteys toisella muuttujalla on selitettävän muuttujan kanssa; ts. että eri toimialojen ennusteet eroavat paitsi toisistaan, myös eri toimialojen erot ovat erilaisia eri hallintoyksiköissä.

Tämä tarkoittaa, että eri toimialojen ja eri hallintoyksiköiden ennusteen laadinta ja ennustevirheet ovat erilaisia. Rajatarkastukseen liittyvillä suoritteilla on tuloksellisuuden laskennassa pienet painokertoimet, ja niitä on lukumääräisesti eniten. Suurimmat painokertoimet ovat rikostorjunnan suoritteilla, joita on vastaavasti lukumääräisesti vähän. Rajatarkastuksen ennustaminen on harhattominta, ja eri hallintoyksiköillä melko samanlaista, mutta muilla toimialoilla vartiostot eroavat selvästi toisistaan. Rikostorjunta eroaa selvästi sekä rajatarkastuksesta että myös toisten vartiostojen rikostorjunnasta. Löydöksen luotettavuutta lisännee se, että rajavartiostot (organisaatiot 1-4) ovat keskenään samankaltaisia, samoin kuin merivartiostot (organisaatiot 5-6) keskenään.

Interaktioiden estimaattoreiden merkitystä on hyvin vaikea tulkita taulukosta ilman graafista esittämistä. Kuva 17 esittää painokerroimen ja ennustevirheiden (harhaisuuden) suhdetta taulukon 7 Mallin 2 mukaisesti. Rajavartiostot on merkitty kuvaan virheällä ja merivartiostot sinisellä värillä. Hallintoyksiköiden välinen merkitsevä ero havaittiin aiemmin myös toteuman suuruusluokan ja ennustevirheen välillä (ks. s. 72).



Kuva 17. Organisaatioiden ja painokerrointen yhteys ennustevirheeseen (harhaisuus).

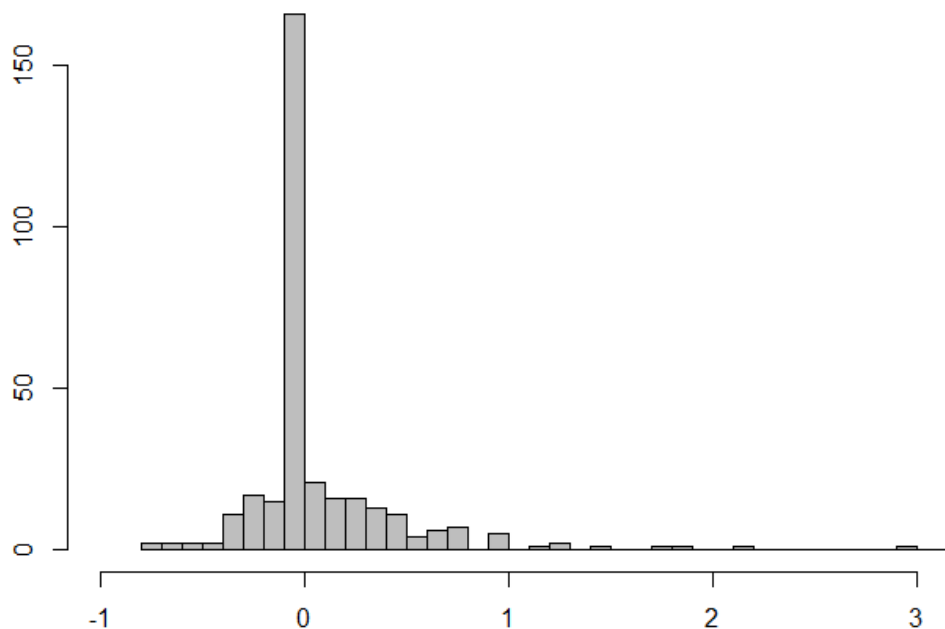
Ennustetarkkuuden osalta (Tarkkuus, Malli 1) kaikki hallintoyksiköt eroavat toisistaan merkitsevästi. Pienimmät ennustevirheet ovat organisaatiolla 5 (Suomenlahden merivartiosto), ja suurimmat organisaatiolla 4 (Lapin rajavartiosto). Keskimääräisten ennustevirheiden lukuarvoja ei voi lukea taulukosta, koska ennustetarkkuutta kuvaavissa malleissa selitettävästä muuttujasta on tehty muunnos (Box-Cox -muunnos, $\lambda = 0.30$). Lähimpänä satunnaista eroa on organisaatioiden 4 ja 6 välinen ero ($p = 0.016$).

Painokerroin ennustaa positiivisesti ennustevirhettä kaikissa hallintoyksiköissä, mutta suhde on erilainen eri organisaatioissa. Suurten painokerrointen toimialojen ennustevirheiden ero pienten painokerrointen toimialoihin on organisaatiossa 4 lähes kolminkertainen verrattuna organisaatioihin 2, 5 ja 6.

4.4. Ennusteiden muuttaminen

Yksi teorian implikaatioista on, että ennustetta laatiessaan (tai korjatessaan) ihminen ankkuroituu aiempiin lukuihin, eikä korjaa käsitystään riittävästi uuden informaation perusteella silloinkaan, kun omasta mielestään päivittää tai laatii uuden ennusteen.

Tarkastellaan aluksi sitä, kuinka paljon aiemmin tehtyjä ennusteita muutetaan. Kuvassa 18 on histogrammi seuraavaa vuotta koskevista ennusteista sen mukaan, kuinka paljon ne suhteessa eroavat edellisenä vuonna kahden vuoden päähän (eli samaa toteumavuotta koskevasta) tehdystä ennusteesta. Vaaka-akselilla on suhteellinen ero aiemmin tehtyyn ennusteeseen, pystyakselilla on havaintojen lukumäärä.



Kuva 18. Seuraavaa vuotta koskevat ennusteiden tarkentuminen.

Noin puolta (47 %) ennusteista ei ole muutettu, vaan aiempi (kahden vuoden päähän aikanaan tehty) ennuste on jätetty voimaan. Ennusteen nostaminen on aineistossa hieman yleisempää kuin ennusteen laskeminen. Jakauman keskiarvo on 0.13, mediaani 0.

Sitä, millainen yhteys ennustetarkkuudella on ennusteeseen tehdyn tarkistuksen *koolla* (kuinka suuri oli ennusteen % -muutos), ei voida aineistosta tutkia, koska jakauma on liian äärimmäinen (havaintomassa nollassa on liian suuri). Sitä, onko ennusteen muuttaminen ollut yhteydessä ennustetarkkuuden paranemiseen tai heikkenemiseen, voidaan kuitenkin tarkastella lineaarisella (binäärisellä) regressiolla. Selittävänä muuttujana on binäärimuuttuja, joka saa arvon 1, jos ennustetta oli muutettu, tai 0, jos ennustetta ei ollut muutettu. Kuten aiemmissakin ennustetarkkuutta käsittelevissä regressiomalleissa, selitettävänä muuttujana on tulosrivin keskiarvoon suhteutettujen virheiden itseisarvon Box-Cox -muunnos ($\lambda = 0.18$). Muuttujamuunnos korjaa heteroskedastisuuden (Breusch-Pagan -testin p muunnoksen jälkeen 0.622).

Taulukko 8.
Regressioanalyysi ennustetarkkuudesta sen mukaan, onko ennustetta muutettu edellisenä vuonna laaditusta.

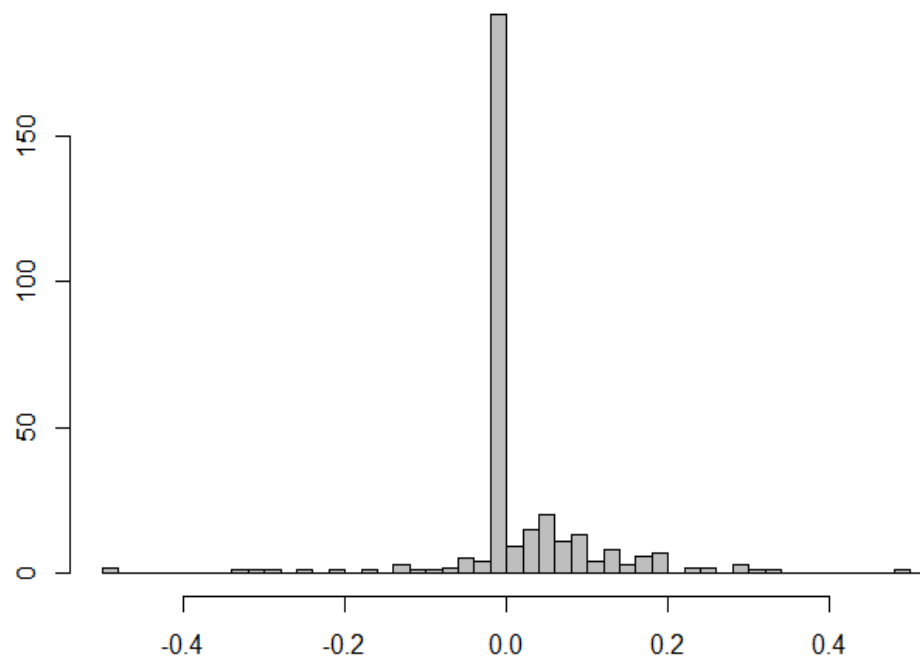
	Malli 1
Ennustetta muutettu (bin.)	-0.01 (0.22)
Vakiotermi	4.73 *** (0.16)
R^2	< 0.001
R^2 korj.	-0.003
RMSE	2.004
N	326

Keskivirhe on sulkeissa

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Yksinkertaisella lineaarisella regressiolla ei havaittu minkäänlaista yhteyttä ennustetarkkuuteen sillä, oliko ennustetta muutettu vai ei. On selvää, että samana pidetyt ennusteet sisältävät myös paljon lukuja, joita on tarkasteltu huolellisesti ja perustellusti päätetty pitää ennallaan.

Aineistosta ei voi luotettavasti testata sitä, korreloivatko ennusteet voimakkaammin aiemmin tehtyjen, samaa toteumavuotta koskevien ennusteiden vai samana vuonna tehtyjen, muita vuosia koskevien ennusteiden kanssa. Otokoko on tähän liian pieni, jakaumat liian äärimmäisiä, ja paneelidatassa on liian paljon autokorrelaatiota. Yhteyttä samaa vuotta ja seuraavaa vuotta koskevan ennusteen välillä voidaan tarkastella graafisesti histogrammilla (kuva 19). Ero samaa vuotta koskevan ja seuraavaa vuotta koskevan ennusteen välillä on samalla ennuste seuraavan vuoden aikana tapahtuvasta muutoksesta.



Kuva 19. Ennustettu muutos, eli seuraavaa vuotta koskevan ennusteen suhteellinen ero samaa vuotta koskevaan ennusteeseen.

Taulukossa 9 on malleissa 1 ja 2 selitettävänä muuttujana vuoden päähän tehtyjen ennusteiden prosentuaalinen muutos (suhteessa aiemmin tehtyihin). Malleissa 3 ja 4 selitettävänä muuttujana on saman muutoksen itseisarvo. Malleissa 5 ja 6 selitettävä muuttuja on binäärinen, sen mukaan, muutettiin ennustetta aiemmasta. Selittävinä muuttujina on toteuman muutos, joka on ennusteen laatijan käytössä olleen toteumatiedon kahden viimeisen luvun suhteellinen ero (prosentuaalinen muutos toteumassa, l. trendi). Seuraava selittävä muuttuja on binäärimuuttuja sen mukaan, onko toteumassa ollut laskeva trendi (1, jos laskeva trendi, 0, jos nouseva trendi tai täsmälleen sama toteuma kuin edeltävänä vuonna). Näiden lisäksi malleihin 2, 4 ja 6 on sisällytetty kahden ensimmäisen interaktiomuuttuja.

Jos toteumassa on ollut laskeva trendi, ennustetta on muutettu harvemmin (malli 5) ja vähemmän (mallit 3 ja 4) kuin tilanteissa, joissa toteumassa on ollut nouseva trendi. Tämä on hyvin mielenkiintoinen löydös, jonka mahdollisista selityksistä on syytä keskustella diskussiossa huolellisesti.

Taulukko 9.

Regressioanalyysi vuoden päähän tehtyjen ennusteiden muuttamisesta toteuman muutoksen ja trendin suhteen.

	Malli 1 (% muutos)	Malli 2 (% muutos)	Malli 3 (muutoksen itseisarvo)	Malli 4 (muutoksen itseisarvo)	Malli 5 (binääri- muuttuja)	Malli 6 (binääri- muuttuja)
Toteuman muutos	0.005 (0.024)	0.004 (0.024)	0.006 (0.023)	0.010 (0.023)	-0.020 (0.025)	-0.022 (0.026)
Toteumassa laskeva trendi (binäärimuuttuja)	-0.25 *** (0.06)	-0.24 ** (0.08)	-0.13 * (0.06)	-0.19 ** (0.07)	-0.12 . (0.06)	-0.09 (0.08)
Toteuman muutos * toteumassa laskeva trendi		0.051 (0.202)		-0.25 (0.19)		0.11 (0.21)
Vakiotermi	0.23 *** (0.06)	0.23 *** (0.04)	0.28 *** (0.04)	0.27 *** (0.04)	0.61 *** (0.04)	0.61 *** (0.04)
R ²	0.068	0.068	0.023	0.028	0.012	0.013
R ² korj.	0.061	0.058	0.016	0.018	0.005	0.003
RMSE	0.47	0.47	0.45	0.44	0.50	0.50
N	299	299	299	299	299	299

Keskivirhe on sulkeissa.

. $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Mallien 1 ja 2 kuvaama yhteys tarkoittaa ainoastaan sitä, että ennustetta on päivitetty todennäköisemmin alaspäin kuin ylöspäin silloin kun toteumassa on ollut laskeva trendi, mitä voisi etukäteenkin pitää itsestään selvänä.

Myös aiempien ennusteiden virheiden yhteys ennusteen muuttamiseen on heikko, $t(202) = .185, p = .85$.

4.5. Heikko etukäteinen (odotettu) ennustettavuus

Taulukossa 10 on esitetty regressioanalyysi ennustetarkkuudesta tulosrivin suuruusluokan ja variaatiokertoimen, ja aiempien ennustevirheiden suhteen. Ennustetarkkuutta on kuvattu tulosrivin toteumien keskiarvoon suhteutettujen virheiden itseisarvoilla. Tulosrivin suuruusluokkaa kuvataan tulosrivin toteumien keskiarvon kymmenkantaisella logaritmillä. Variaatiokerroin on kunkin tulosrivin toteumien otoskeskihajonta jaettuna toteumien keskiarvolla, jolloin eri suuruusluokkaa olevien toteumien hajontaluvut ovat keskenään yhteismitallisia.

Variaatiokertoimen ja suuruusluokan välillä on yhteys, koska keskiarvo on variaatiokertoimessa nimittäjänä; toisin sanoen pienempi keskiarvo aiheuttaa suuremman variaatiokertoimen, mikäli otoskeskihajonta pysyy samana. Suuruusluokka ei kuitenkaan selitä kuin osan variaatiokertoimen vaihtelusta (ks. s. 51 metodiluvussa).

Mallien 1-4 osalta $N=1695$. Aiemmat virheet -termi on muodostettu niin, että on laskettu tulosrivin keskiarvoon suhteutettujen virheiden itseisarvojen keskiarvo regressiossa terminä olevaa virhettä aiemmilta vuosilta. Tämä termi on mahdollista muodostaa vain tarkasteluajanjakson viimeisiltä kolmelta vuodelta (2012–2014 tehtyjen ennusteiden osalta), minkä vuoksi malleissa 5-8 on pienempi otoskoko, $N=755$.

Jäännökset eivät olleet normaalijakautuneita ja niiden varianssi ei ole vakio (heteroskedastisuus). Selitettävä muuttuja on korvattu muuttujamuunnoksella (Box-Cox -muunnos). Testien tulokset ja muuttujamuunnoksen λ -termi on raportoitu kunkin mallin osalta. Muuttujamuunnos korjasi heteroskedastisuuden malleissa 2 ja 5-8. Malleissa 1, 3 ja 4 on käytetty OLS-keskivirheiden sijasta Whiten keskivirheitä.

Variaatiokerroin ennustaa positiivisesti ennustevirhettä, kun aiemmat virheet eivät ole mukana mallissa selittävänä muuttujana. Toisin sanoen, mitä suurempaa on toteuman suhteellinen vaihtelu vuosien välillä, sitä suurempia ennustevirheet ovat ($p < 0.001$). Suuruusluokka on niin ikään merkitsevästi yhteydessä ennustevirheeseen siten, että mitä suurempi on tilastoitavan suureen toteumien keskiarvo, sitä pienempiä ennustevirheet keskimäärin ovat.

Taulukko 10.

Regressioanalyysi ennustetarkkuudesta tulosrivin suuruusluokan, variaatiokerroimen ja aiempien ennustevirheiden suhteen.

Tulosrivin toteumien keskiarvoon suhteutettujen virheiden itseisarvojen Box-Cox -muunnos, λ raportoitu alla.

	Malli 1	Malli 2	Malli 3	Malli 4	Malli 5	Malli 6	Malli 7	Malli 8
$\lambda =$	0.39	0.37	0.37	0.37	0.30	0.30	0.30	0.30
Variaatiokerroin	0.59 *** (0.08)		0.49 *** (0.08)	0.75 *** (0.15)			0.0038 (0.0966)	0.14 (0.14)
Suuruusluokka ($\log_{10}(\text{ka})$)		-0.054 *** (0.007)	-0.05 *** (0.01)	-0.0061 (0.0221)		-0.0064 (0.0166)		
Variaatiokerroin * suuruusluokka				-0.12 * (0.05)				
Aiemmat virheet					1.06 *** (0.07)	1.05 *** (0.08)	1.06 *** (0.08)	1.47 *** (0.30)
Variaatiokerroin * aiemmat virheet								-0.51 (0.36)
Vakiotermi	-1.12 *** (0.04)	0.64 *** (0.02)	-0.94 *** (0.05)	-1.05 *** (0.08)	-1.33 *** (0.04)	-1.31 *** (0.07)	-1.33 *** (0.05)	-1.45 *** (0.09)
R^2	0.064	0.040	0.078	0.082	0.219	0.220	0.219	0.222
R^2 korj.	0.063	0.039	0.077	0.080	0.218	0.218	0.217	0.219
RMSE	0.69	0.71	0.70	0.70	0.74	0.74	0.74	0.74
N	1695	1695	1695	1695	755	755	755	755

Keskivirhe on sulkeissa

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Malli 4 tuottaa mielenkiintoisen tuloksen. Selittäviin muuttujiin on tässä otettu myös variaatiokerroimen ja suuruusluokan interaktiotermin. Interaktiotermin sisällyttäminen malliin poistaa merkitsevyyden suuruusluokalta, mutta interaktiotermin sijaan on merkitsevästi yhteydessä ennustevirheen kokoon ($p=0.0225$). Tämä tarkoittaa, että suuruusluokka ei olisi itsessään tilastollisessa yhteydessä ennustevirheen kokoon, vaan ainoastaan sen kautta, miten voimakas on variaatiokerroimen ja ennustevirheen korrelaatio. Mitä suurempi on tilastoitavan suureen keskiarvo, sitä heikompi on variaatiokerroimen ja ennustevirheen yhteys. Variaatiokerroin ennustaisi siis suurempia ennustevirheitä erityisesti pienten tulosrivien osalta.

Malleista 5-8 nähdään, että aiempien virheiden huomioiminen regressiomallissa poistaa kuitenkin merkitsevyyden sekä variaatiokertoimelta että suuruusluokalta, eikä aiemmilla virheillä myöskään ole merkitseviä interaktioita näiden heikomman ennustetarkkuuden proxy-muuttujien kanssa.

Autokorrelaatio voi aiheuttaa osan edellä kuvatusta variaatiokertoimen ja suuruusluokan yhteydestä ennustevirheeseen. Autokorrelaation korjaamiseksi testi toistettiin käyttämällä yhtä poikkileikkaavaa otosta. Selitettäväksi muuttujaksi otettiin paneelidatasta kunkin paneelin ennustevirheiden itseisarvojen keskiarvo. Tämä menettely poistaa autokorrelaatioon liittyvän validiteettiongelman, mutta otos jää pieneksi (N=70). Regressiossa variaatiokertoimella ja suuruusluokalla vaikutti aluksi olevan merkitsevä yhteys ennustevirheeseen. Kumpikaan regressiomalli ei kuitenkaan ollut homoskedastinen eivätkä jäännöstermit olleet normaalijakautuneita, minkä vuoksi selitettävä muuttuja korvattiin muuttujamuunnoksella ($\lambda = 0.22$), ja tavalliset lineaarisen regression (OLS) keskivirheet korvattiin Whiten keskivirheillä, normaalisuusehdon rikkomisen ja heteroskedastisuuden kompensoimiseksi. Kompensoivien toimien jälkeen variaatiokertoimelle ja suuruusluokalle ei jäänyt tilastollisesti merkitsevää yhteyttä paneelikohtaiseen ennustevirheen keskiarvoon.

4.6. Vertailu tilastolliseen ennusteeseen

Ankkuroitumisen lisäksi toinen teoriaosuudessa käsiteltyyn ennusteiden laatimisen psykologiaan liittyvä ilmiö on niin sanottu edustavuusvääristymä. Teorian mukaan ennusteiden laatijat eivät kykene reagoimaan huonoon odotettuun ennustetarkkuuteen. Rationaalinen vastaus huonoon ennustetarkkuuteen olisi siirtyä kohti regressiivisempää ennustetta.

Jälkiä tästä ilmiöstä tutkitaan aineistosta vertaamalla ennusteiden tarkkuutta yksinkertaiseen tilastolliseen ennusteeseen. Vertailukohtana käytettävän tilastollisen ennusteen on oltava yksinkertainen, jotta vertailu olisi mielekäästä. Yksinkertaiseksi tilastolliseksi ennusteeksi valittiin ennustetta laadittaessa käytettävissä olleista toteumatiedoista kolmen viimeisimmän toteuman liukuva keskiarvo. Samaa lukua käytettiin tilastollisena ennusteena kaikille aikaväleille.

Toteuman keskiarvon käyttäminen ennusteena olisi ollut kaikkien ennusteita laatineiden käytössä. Vertailun validiteetin kannalta on keskeistä, ettei tilastollista ennustetta soviteta dataan millään tavalla. Kolmen vuoden liukuva keskiarvo on menetelmänä niin yleinen ja yksinkertainen, että sen voi katsoa täyttävän tämän ehdon.

Variaatiokerroin ja suuruusluokka ennustavat positiivisesti ennustevirheitä (eli heikkenevää ennustetarkkuutta). Käytetyillä korvaavilla muuttujilla vaikuttaisi siis olevan jonkinlaista tilastollista voimaa kuvata huonoa ennustettavuutta.

Jos edustavuushypoteesin mukaisesti ennusteiden laatijat eivät osaa reagoida huonoon ennustetarkkuuteen siirtymällä kohti regressiivisempää ennustetta, pitäisi riittävän suuressa otoksessa asiantuntijaennusteiden pärjätä vertailussa sitä huonommin, mitä vaikeammin ennustettava tilastorivi on kyseessä.

Kaikki taulukossa 11 esitetyt regressiomallit ovat heteroskedastisia (Breusch-Pagan -testin $p < 0.001$), ja lineaarisen regression tavallisten (OLS) keskivirheiden sijaan on käytetty Whiten keskivirheitä. Selitettävän muuttujan Box-Cox -muunnos ei ole mahdollinen, koska regressiivisen ennusteen paremmuus saa sekä negatiivisia että positiivisia arvoja. Lineaarisen regression lähtöoletusten rikkominen on huomioitava tulosten luotettavuutta arvioitaessa.

Taulukko 11.

Regressioanalyysi regressiivisen ennusteen paremmuudesta tulosrivin suuruusluokan, variaatiokerroimen ja aiempien ennustevirheiden suhteen.

	Malli 1	Malli 2	Malli 3	Malli 4	Malli 5
Variaatiokerroin	-0.13 (0.08)		-0.20 * (0.08)	0.014 (0.16)	
Suuruusluokka ($\log_{10}(ka)$)		-0.024 *** (0.006943)	-0.036 *** (0.007)	0.0043 (0.0182)	
Aiemmat virheet					0.39 *** (0.09)
Variaatiokerroin * suuruusluokka				-0.103 * (0.048)	
Aiemmat virheet * variaatiokerroin					
Aiemmat virheet * suuruusluokka					
Aiemmat virheet * variaatiokerroin * suuruusluokka					
Vakiotermi	0.11 *** (0.03)	0.12 *** (0.03)	0.24 *** (0.04)	0.15 * (0.07)	-0.033 (0.033)
R^2	0.0073	0.0080	0.023	0.029	0.089
R^2 korj.	0.0067	0.0074	0.022	0.027	0.088
RMSE	0.46	0.46	0.46	0.46	0.46
N	1695	1695	1695	1695	755

Keskivirhe on sulkeissa

. $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Malleissa 1-4 variaatiokerroin ennustaa regressiivisen ennusteen paremmuutta negatiivisesti. Mallissa 3 yhteys on tilastollisesti merkitsevä, ja mallissa 4 interaktiotermi saa tilastollisen merkitsevyyden. Mallin 2 mukaan regressiivinen ennuste (liukuva keskiarvo) on parempi alle 100 000 suoritteen tilastoriveillä. Mallin 4 mukaan tämä kynnys on sitä matalammalla, mitä suurempi on tilastoitavan suureen varianssi.

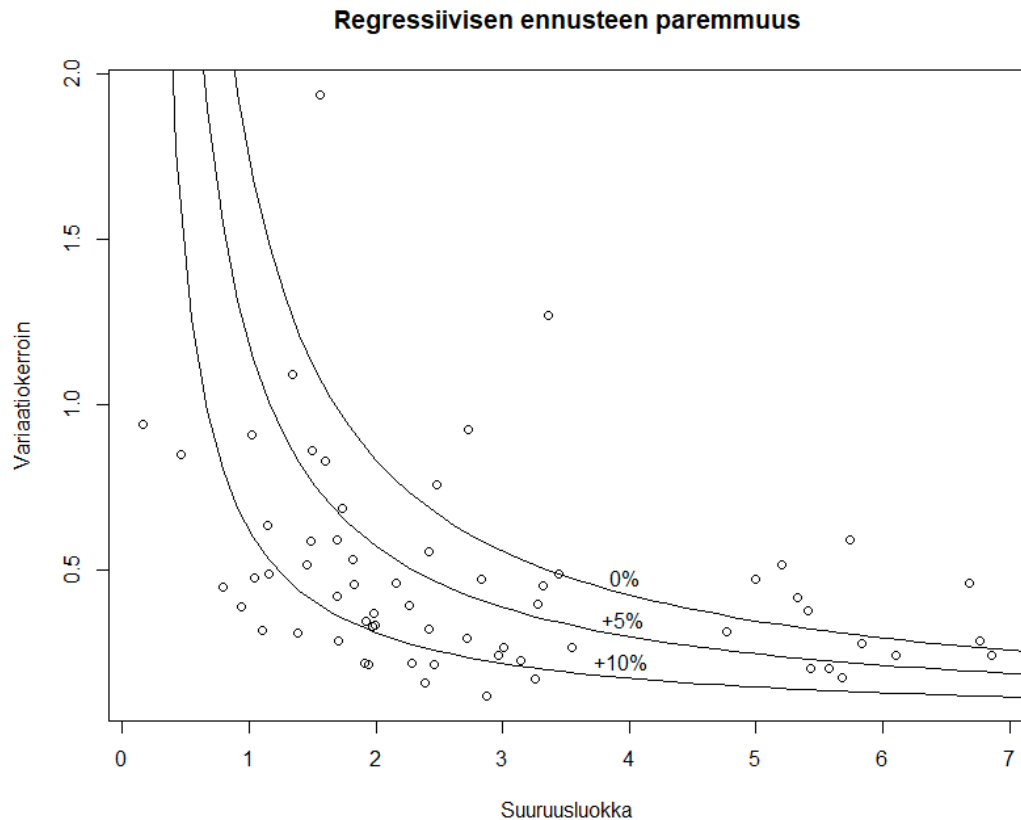
Taulukko 11. (jatkoa edelliseltä sivulta)

Malli 6	Malli 7	Malli 8	Malli 9	Malli 10	Malli 11
		-0.16 *** (0.04)	0.071 (0.104)	-0.13 ** (0.04)	0.32 . (0.18)
0.027 * (0.011)	-0.027 (0.020)			0.023 . (0.012)	0.070 (0.050)
0.44 *** (0.10)	0.20 (0.15)	0.45 *** (0.10)	1.12 *** (0.25)	0.47 *** (0.10)	0.95 * (0.40)
					-0.21 * (0.10)
			-0.84 * (0.37)		-0.97 . (0.57)
	0.18 * (0.07)				-0.060 (0.184)
					0.41 (0.31)
-0.13 ** (0.05)	-0.05 (0.06)	0.018 (0.034)	-0.17 ** (0.06)	-0.071 (0.050)	-0.26 * (0.11)
0.097	0.12	0.097	0.11	0.10	0.14
0.095	0.12	0.094	0.11	0.10	0.13
0.46	0.46	0.46	0.46	0.46	0.45
755	755	755	755	755	755

Keskivirhe on sulkeissa

. $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Interaktiotermin sisältävä regressiomalli on vaikea tulkita ilman visualisointia. Kuvassa 20 on esitetty regressiivisen ennusteen paremmuus mallin 4 tuottamien arvojen mukaan, kolmiulotteisessa kuvaajassa, tasoina suuruusluokan ja variaatiokertoimen päällä. Kuten kuvasta voidaan havaita, vaikka variaatiokertoimen estimaattori on mallissa 4 positiivinen, regressiivisen ennusteen paremmuus kuitenkin laskennallisesti heikkenee variaatiokertoimen kasvaessa, kun interaktiotermin vaikutus huomioidaan.



Kuva 20. Taulukon 11 regressiomalli 4.

Kun aiemmat virheet otetaan regressiomalliin mukaan (mallit 5 – 11), variaatiokertoimen yhteys regressiivisen ennusteen paremmuuteen pysyy negatiivisena. Kuten mallissa 4, malleissa 9 ja 11 negatiivinen yhteys sisältyy interaktiitermien estimaattoreiden arvoihin (ks. edellinen kappale). Näin ollen variaatiokertoimen tilastollinen yhteys regressiivisen ennusteen paremmuuteen on kaikissa malleissa erisuuntainen kuin hypoteesin perusteella oletettiin. Tilastollisesta merkitsevyydestä huolimatta regression tulokset eivät siis tue hypoteesia niiltä osin kuin variaatiokertoimen ajatellaan edustavan huonoa ennustettavuutta.

Taulukossa 12 esitetään regressiivisen ennusteen paremmuutta verrattuna eri aikaväleille laadittuihin ennusteisiin. Aiemmat virheet on sisällytetty regressiomalliin 2 kontrolloimaan autokorrelaatiota. Mallissa 2 on heteroskedastisuuden (Breusch-Pagan -testin $p < 0.001$) vuoksi käytetty Whiten keskivirheitä.

Koko aineistossa yksinkertainen tilastollinen ennuste (kolmen vuoden liukuva keskiarvo) on 4.9 prosenttiyksikköä parempi kuin asiantuntijaennuste (aineistosta), ts. regressiivisen ennusteen virhe on keskimäärin pienempi kuin asiantuntijaennusteen virhe. Äärihavaintojen merkitys keskiarvossa on kuitenkin suuri, eikä tästä luvusta todennäköisesti voida päätellä mitään. Tarkastelu ajanjaksoittain antaa tilanteesta tarkemman kuvan.

Taulukon 12 regressiomalleissa vakiotermin kuvaama tilannetta samalle vuodelle laadittujen ennusteiden osalta. On huomionarvoista, että asiantuntijaennusteen laatijalla olisi ollut käytössään enemmän dataa (alkuvuoden aikana kertynyt osittainen toteumatieto) kuin mitä tilastollinen ennuste huomioi. Tästä huolimatta tilastollisen ennusteen virhe on mallin 1 mukaan ollut parempi kaikilla muilla aikaväleillä paitsi neljän vuoden päähän laadituissa ennusteissa.

Taulukko 12.

Regressioanalyysi regressiivisen ennusteen paremmuudesta verrattuna eri aikaväleille laadittuihin ennusteisiin

	Malli 1	Malli 2
Seuraavalle vuodelle, binäärimuuttuja	0.0082 (0.0338)	0.030 (0.037)
Kahden vuoden päähän, bin.	0.036 (0.034)	0.11 * (0.05)
Kolmen vuoden päähän, bin.	0.019 (0.035)	0.13 * (0.06)
Neljän vuoden päähän, bin.	-0.048 (0.038)	0.025 (0.070)
Aiemmat virheet		0.39 *** (0.05)
Vakiotermi	0.043 . (0.024)	-0.090 * (0.043)
R ²	0.0031	0.099
R ² korj.	0.0007	0.093
RMSE	0.46	0.46
N	1695	755

Keskivirhe on sulkeissa

. $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Kun aiempien ennusteiden virheet otetaan mukaan kontrollimuuttujaksi mallissa 2, tilanne muuttuu siten, että samalle vuodelle laadittujen ja seuraavaa vuotta koskevien ennusteiden osalta asiantuntijaennuste on parempi. Tilastollinen ennuste on myös mallissa 2 asiantuntijaennustetta parempi kahden ja kolmen vuoden päähän laadittujen ennusteiden osalta.

Lopuksi tarkastellaan vertailtua tilastollisen ennusteen ja asiantuntijaennusteen keskinäisestä paremmuudesta regressoituna painokertoimen ja suuruusluokan suhteen. Aiemmat virheet on jälleen sisällytetty malliin 3 korjaamaan autokorrelaatiota.

Taulukko 13.

Regressioanalyysi tilastollisen (regressiivisen) ennusteen paremmuudesta painokertoimen, suuruusluokan ja aiempien ennusteiden virheiden suhteen.

	Malli 1	Malli 2	Malli 3
Painokerroin	-0.0011 * (0.0005)	0.0005 (0.0010)	-0.0044 ** (0.0016)
Suuruusluokka ($\log_{10}(\text{ka})$)	-0.029 ** (0.009)	-0.023 * (0.009)	0.0010 (0.0131)
Painokerroin * suuruusluokka		-0.0012 * (0.0006)	0.0004 (0.0009)
Aiemmat virheet			0.487 *** (0.091)
Vakiotermi	0.17 *** (0.04)	0.17 *** (0.04)	0.0021 (0.0543)
R^2	0.008	0.012	0.132
R^2 korj.	0.007	0.010	0.127
RMSE	0.49	0.48	0.47
N	1637	1637	722

Keskivirhe on sulkeissa

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Koska selitettävänä muuttujana on regressiivisen ennusteen paremmuus, negatiiviset luvut tarkoittavat asiantuntijaennusteen paremmuutta (pienempää virhettä). Malleista 1 ja 2 nähdään, että asiantuntijaennusteet ovat suhteessa tilastolliseen ennusteeseen sitä parempia, mitä suurempi on ollut tulosrivin toteuman suuruusluokka ja painokerroin tuloksellisuuden laskennassa. Koska painokerroin ei voi loogisesti olla missään itsenäisessä yhteydessä ennustettavuuteen, taulukon 13 tulokset eivät liity huonoon ennustettavuuteen.

Kuten taulukon 11 malleissa, tässäkin yhteydessä muuttujamuunnos ei ole mahdollinen, koska selitettävä muuttuja saa myös negatiivisia arvoja. Heteroskedastisuuden kompensoimiseksi on käytetty Whiten keskivirheitä.

5. DISKUSSIO

5.1. Samojen virheiden toistaminen

Kenties kaikkein perustavanlaatuisin ja eniten pohdintaa aiheuttava löydös on samojen ennustevirheiden toistaminen. Noin puolia ennusteista ei ollut muutettu verrattuna samasta asiasta aiemmin laadittuun ennusteeseen. Keskustelussa on erotettava toisistaan samojen ennusteiden toistaminen ja samojen ennustevirheiden toistaminen. Samana pidettyihin ennusteisiin kuuluu varmasti myös lukuja, jotka on kriittisen tarkastelun jälkeen tietoisesti päätetty pitää ennallaan. Saman luvun käyttäminen voi olla täysin rationaalista ja perusteltua, vaikka aiemmat ennusteet olisivat virheellisiä. Asiantuntijaennusteiden laatijoilla on ollut käytössään huomattavasti enemmän tietoa kuin vain aiempien vuosien toteumaluvut: raportteja, uutisia, yhteistyökumppaneilta saatua tietoa, kapasiteetin muutoksia, investointeja, liikenneyhteyksien ja säädösten muutoksia ja niin edelleen. Aiemman luvun käyttämiseen liittyvät mahdolliset rationaaliset ja loogiset syyt selittävät kuitenkin vain saman ennusteen toistamista, eivät saman virheen toistamista.

Virheet aiempien vuosien ennusteissa ennustivat seuraavien vuosien ennustevirheitä lähes yhden suhde yhteen. Toisin sanoen, jos ennuste oli virheellinen, seuraavan vuoden ennusteessa samasta asiasta oli odotusarvoisesti samankokoinen ja -suuntainen virhe. Virheitä aiheutuu monista eri syistä, mutta selitysasteet olivat verraten korkeita niissä regressiomalleissa, joissa aiemmat virheet olivat mukana. Aiempien virheiden toistaminen selitti yli 20 prosenttia ennustevirheiden vaihtelusta.

Virheiden toistumisen syyksi voitaisiin tarjota sitä, että toimintaympäristö on erilainen eri paikoissa ja että erilaisuus on yllättänyt ennusteiden laatijat. Tämä ei kuitenkaan tunnu uskottavalta, koska se tarkoittaisi sitä, että toimintaympäristö on yllättänyt samalla tavalla vuodesta toiseen. Samankaltaisena toistuvia virheitä voisivat aiheuttaa myös saman tyyppisenä jatkuvat kehityskulut toimintaympäristön muutoksessa. Tästä ei kuitenkaan vaikuta olevan kyse ottaen huomioon sen, miten suuri osa ennusteista oli samoja aiemmin tehtyihin verrattuina, vaan kyse oli nimenomaan virheiden toistamisesta.

On huomionarvoista, että aiempien ennusteiden virheellisyyden yhteys ennusteen muuttamiseen oli heikko. Tämä jo sinänsä osoittaa, että joko aiempien ennusteiden virheisiin ei ollut reagoitu lainkaan tai että siihen oli reagoitu liian vähän. Tämä toteamus on hieman tautologinen eikä vielä kerro meille mitään reagoimattomuuden syistä.

Edustavuusheuristiikan yhteydessä kuvattiin, etteivät tutkimusten koehenkilöt osaa reagoida huonoksi arvioimaansa ennustettavuuteen siirtymällä kohti regressiivisempää ennustetta. Mitä epämääräisempää kaikki muu informaatio (kuin toteumatieto) on ja mitä huonommaksi ennustettavuus arvioidaan, sitä suurempi painoarvo pitäisi antaa regressiiviselle ennusteelle, kuten keskiarvolle. Ihminen ei kuitenkaan intuitiivisesti osaa toimia näin. Tämä teoriassa kuvattu mekanismi voisi aiheuttaa saman suuntaisen vaikutuksen, kuin mitä aineistosta havaittiin. Muitakin mahdollisia selityksiä kuitenkin on.

Edustavuusheuristiikka voisi vaikuttaa ennusteiden laadintaan muillakin tavoilla, jotka voisivat aiheuttaa samojen virheiden toistamista. Ensinnäkin olisi mahdollista, että asiantuntija kuvaisi tiedostamattaan jotain muutakin kuin vain taustatodennäköisyyksiä suoritteiden määrästä: esimerkiksi suoritteiden merkitystä tai tärkeyttä, koettua työpainetta tai jotain muuta sellaista, jota korkeammat ennusteet kuvaisivat eli edustaisivat hänen mielestään paremmin. Tällaisesta ei kuitenkaan voida saada viitteitä tutkimusasetelmassa, jossa ei päästä haastattelemaan ennusteiden laatijoita. Ilman jatkotutkimusta tällainen mekanismi jää pelkäksi hypoteesiksi.

Edustavuusheuristiikan yhteydessä kuvattiin myös, että ihmisillä on sellaisia intuitiivisia tapoja reagoida heikkoon ennustettavuuteen, jotka eivät ole tarkasti rationaalisia. Tällainen olisi esimerkiksi tapaus, jossa ennustetta ei muuteta siksi, että ennusteen laatija ei koe omaavansa riittävästi informaatiota ja kehitys on liian vaikeasti ennustettava, jotta ennustetta olisi perusteltua päivittää johonkin suuntaan. Tällainen hypoteesi on sinänsä ihan looginen ja uskottava, vaikkei tässä työssä voidakaan selvittää, onko tässä aineistossa mukana tällainen vaikutus. Voidaan todeta, että jos samojen ennusteiden käyttäminen olisi edes osittain intuitiivinen reaktio heikkoon ennustettavuuteen, tällöin ennusteen laatijat eivät olisi osanneet siirtyä kohti regressiivisempää ennustetta, aivan kuten kirjallisuudessa kuvattiin. Kuvatun kaltainen reaktio ennustamisen vaikeuteen on inhimillisesti ymmärrettävä, mutta ei tarkalleen ottaen oikea. Tällainen taustatodennäköisyyksien liian vähäinen huomiointi on osa edustavuusheuristiikkaa ja niitä vääristymiä, joita heuristinen ajattelu aiheuttaa.

Jos asiaa tarkastellaan sen perusteella, mitä teoriassa kuvattiin ankkuroitumisesta, aineistossa voisi olla käsillä sellainen tilanne, että aiemmat ennusteet ovat aiheuttaneet tilastolliseen ennusteeseen verratun asiantuntijaennusteen virheen ankkuroimalla ennusteen laatijan aiempiin lukuihin. Olisi mahdollista, että reagoimattomuus virheisiin johtuisi osittain siitä, että aiemmat luvut vaikuttavat ennusteisiin enemmän kuin ennusteiden laatijat itse

huomaavat. Ankkuroituminen, kognitiivisena vääristymänä, sellaisena kuin se teorialuvussa kuvattiin, on melko tarkkarajainen termi. Kaikki samojen lukujen käyttäminen ei ole ankkuroitumista. Ankkuroitumisesta voisi olla kyse siinä tapauksessa, että ennusteen laatijat olisivat omasta mielestään reagoineet ennusteiden virheellisyyteen tarkistamalla ennusteita, mutta olisivat tiedostamattaan reagoineet liian vähän siten, että efekti jää suuressa otoksessa havaittavaksi.

Tulosluvun alussa kuvattiin eri aikaväleille tehtyjen ennusteiden tarkkuutta ja harhaisuutta. Tietyin varauksin voidaan arvioida, että aineistossa vaikuttaa olevan eräänlainen ennustehorisontti eli aikaväli, jonka taakse ennustetarkkuus ei merkittävästi heikkene. Tämä horisontti on aineistossa yhdestä kahteen vuoden päässä. Graafisesti tarkastellen ennustetarkkuus vaikutti heikkenevän melko tasaisesti sen mukaan, kuinka pitkälle tulevaisuuteen ennuste oli laadittu. Kuitenkin tilastollisesti testattuna vain samaa vuotta koskevat ennusteet olivat merkittävästi tarkempia kuin sitä pitemmälle aikavälille laaditut. Yhdestä neljään vuoden päähän tehtyjen ennusteiden välillä ei ollut merkittäviä eroja ennustetarkkuudessa. Jos ennustehorisontti todella on aineistossa olemassa, mahdollinen johtopäätös tiedostamattomasta ankkuroitumisesta saa lisää merkitystä. Jos ennusteita on laadittu horisontin taakse ja sen jälkeen tiedostamatta ankkuroidutaan aiempiin lukuihin, tullaan turhaan heikentäneeksi kykyä toimintaympäristön muutosten ennakkointiin.

Voidaan todeta, että heuristisen ajattelun ja kognitiivisten vääristymien teoriassa kuvatut mekanismit voisivat aiheuttaa saman tyyppistä virheiden systematiikkaa, kuin mitä aineistosta havaittiin. Tulosten perusteella on siis mahdollista, että jokin pieni osa ennustevirheiden toistumisesta ja ennusteiden pitämisestä ennallaan liittyy edustavuusheuristiikkaan, intuitiiviseen reaktioon huonoon ennustettavuuteen tai tiedostamattomaan ankkuroitumiseen aiempiin lukuihin. Kausaliteetin osoittamiseen tarvittava tieto kuitenkin puuttuu, koska aineistossa ei ole ennusteiden laatijoille esitettyjä taustakysymyksiä.

Koska kognitiiviset vääristymät ovat teorian mukaan kaikille ihmisille yhteisiä, ei hallintoyksiköiden välillä voi suuressa otoksessa olla keskimääräisiä eroja siinä osassa aiempien virheiden toistamista, joka on heurististen ajattelun aiheuttamaa. Kun taulukon 12 regressiot toistettiin niin, että hallintoyksiköitä kuvaavat binäärimuuttujat otettiin mukaan, havaittiin hallintoyksiköiden välillä tilastollisesti merkittäviä eroja. Tämä liudentaa heuristisen ajattelun ja kognitiivisten vääristymien teorian selitysvoimaa näiden tulosten

tulkinnassa. Toisin sanoen kyse täytyy olla muustakin kuin heuristisesta ajattelusta ja tiedostamattomasta kiinnittymisestä aiempiin lukuihin.

Kausaalisten suhteiden todistamiseen liittyvistä puutteista huolimatta voidaan edetä keskustelussa siihen, miten prosessia voitaisiin tältä osin kehittää. Jos samojen virheiden toistamisessa on edes osittain kyse edellä ehdotetuista heuristisen ajattelun aiheuttamista kognitiivisista vääristymistä, prosessia on melko yksinkertaista tältä osin parantaa. Ennusteet tulisi laatia siten, etteivät aiemmat ennusteet ole näkyvillä. Johdonmukaisuutta ja eroja aiemmin arvioituihin lukuihin tulisi tarkastella vasta, kun uusi ennuste on laadittu. Jos ennustetta laadittaessa halutaan käyttää jotain lähtöarvoa, vanhoja ennusteita parempi olisi toteuman keskiarvo joltain tietyltä ajanjaksolta. Tilastorivien suhteita voisi myös merkitä näkyviksi. Esimerkiksi "rajatarkastuksista" oletettavasti syntyy jollain prosentuaalisella osuudella "havaittuja liikennejuopumuksia", ja tämän suhteen historiallinen vaihteluväli on laskettavissa aineistosta. Mahdollisena lähtöarvona etenkin pienten tilastorivien ennusteille voisi käyttää laskennallisia osuuksia muista tilastoriveistä.

Yksinkertaisen tilastollisen ennusteen (kolmen edellisen vuoden toteumien keskiarvo) virhe olisi ollut koko aineistossa keskimäärin 4.9 prosenttiyksikköä pienempi kuin tutkimusaineistona olleiden asiantuntijoiden tekemien ennusteiden. On mielenkiintoista huomata, että kun aiemmat virheet olivat mukana kontrollimuuttujana (taulukko 13, Malli 3), eivät tilastolliset ennusteet olleet keskimäärin merkitsevästi parempia kuin asiantuntijaennusteet. Toisin sanoen, jos ennusteiden laatijat olisivat riittävästi reagoineet siihen, että aiempien vuosien ennusteet ovat olleet virheellisiä, asiantuntijaennusteet olisivat olleet yhtä tarkkoja kuin tilastolliset ennusteet.

Aiempien virheiden käyttäminen regressiomallissa kontrollimuuttujana on teknisessä mielessä keino poistaa regression validiteettia uhkaavaa jäännöstermien autokorrelaatiota. Tulosten ymmärtämisen näkökulmasta sillä on myös muu kuin tekninen merkitys. Voidaan ajatella, että regressiomallit, joissa aiemmat virheet on kontrolloitu, kuvaavat sitä ennustetarkkuuden tasoa, johon ennusteiden laatijat olisivat päässeet, jos he olisivat riittävästi huomioineet aiempien ennusteiden virheellisyyden.

Perättäisten havaintojen välinen korrelaatio (autokorrelaatio) on lineaarisen regression kannalta ongelma, mutta se on myös löydös. Tietynlaisessa paneelidatassa autokorrelaatiota voidaan odottaa. Esimerkiksi, jos tutkitaan pituuden yhteyttä painoon ja käytetään perättäisiä

havaintoja samoista yksilöistä, voidaan datassa odottaa autokorrelaatiota, joka laskee tulosten luotettavuutta. Ei kuitenkaan ole samalla tavalla ilmeistä, että perättäisinä vuosina tehtyjen ennusteiden virheet olisivat väistämättä autokorreloituneita. Ennustevirheiden yhteydessä voidaan perustellusti kysyä, onko ennustevirheiden autokorrelaatiolle eri vuosien välillä mitään loogista syytä.

5.2. Usko suoritemäärien kasvuun

Verrattaessa samaa vuotta koskevia ennusteita samaan aikaan laadittuihin, seuraavia vuosia koskeviin ennusteisiin havaittiin, että aineistossa oli selvästi yleisempää ennustaa toteuman kasvua kuin laskua. Aineistossa havaittiin ennustevirheiden systemaattinen positiivinen virhe (positiivinen harhaisuus). Samaa vuotta koskevat ennusteet olivat keskimäärin 7.5 prosenttia toteumaa suuremmat, yhden vuoden päähän tehdyt 9.7 prosenttia toteumaa suuremmat ja kahden vuoden päähän tehdyt ennusteet keskimäärin 11.8 prosenttia toteumaa suuremmat. Vaikuttaa siis siltä, että ennusteen laatijoiden usko suoritemäärien kasvuun on toteumatiedon valossa ollut osin perusteeton.

Ennusteita oli muutettu useammin ylös- kuin alaspäin. Rajanylitysliikenteen odotettu kasvu tai muut vastaavat loogiset syyt eivät selitä tätä tulosta, koska ennusteen muuttamisessa on kyse odotetun muutoksen muutoksesta. Lisäksi suuria positiivisia ennustevirheitä oli tutkimusaineistossa enemmän, ja ne olivat äärimmäisempiä (itseisarvoltaan suurempia) kuin suuret negatiiviset ennustevirheet.

Toteuman muutosten ja ennusteisiin tehtyjen muutosten yhteys oli heikko. Aineistossa havaittiin viitteitä siitä, että toteuman laskuun oli reagoitu eri tavalla kuin toteuman kasvuun. Ennusteita oli muutettu vähemmän todennäköisesti silloin, jos toteumassa oli ollut muutos alaspäin. Toteuman kasvuun reagoiminen muuttamalla ennustetta ylöspäin oli merkittävästi todennäköisempää kuin toteuman laskuun reagoiminen muuttamalla ennustetta alaspäin. Ennusteisiin tehdyt muutokset olivat myös kooltaan pienempiä silloin, jos toteutuneet suoritemäärät olivat laskeneet. Voidaan siis todeta, että toteuman laskuun oli reagoitu ennustetta muuttamalla sekä harvemmin että vähemmän kuin toteuman kasvuun.

Tutkimusaineistossa havaittu ennusteiden systemaattinen positiivinen harhaisuus sekä toisistaan eroava suhtautuminen laskeviin ja nouseviin trendeihin antavat viitteitä siitä, että

ennusteiden pitämiseen ennallaan saattaa liittyä epärationaalinen, heuristinen komponentti. Tutkimustulokset antavat vähintään viitteellistä näyttöä siitä, että informaation sisältö vaikuttaa siihen, huomioidaanko sitä vai ei. Jos näin on, tämäkin olisi yhtä pitävä sen kanssa, miten edustavuusheuristiikan vaikutuksia on teorialuvussa kuvattu.

Heuristiseen ajatteluun ja kognitiivisiin vääristymiin liittyy myös odotusten vahvistaminen. Ihmisillä on taipumus hyväksyä helpommin informaatiota, joka sopii yhteen heidän toiveidensa tai ennakko-odotustensa kanssa. Aineistosta voidaan todeta ennusteiden laatijoiden odottaneen kasvua, koska kasvavan trendin ennustaminen oli aineistossa selvästi yleisempää kuin laskevan trendin.

5.3. Ennustettavan tulosrivin merkitys ja tärkeys

Tulosluvussa on kerrottu tulosrivin toteuman variaatiokertoimen, suuruusluokan ja ennustettavalle tilastoriville tuloksellisuuden laskennassa annetun painokertoimen yhteyttä ennusteiden tarkkuuteen ja harhaisuuteen. Alun perin tavoitteena oli selvittää näiden käytettävyyttä ennustettavuutta kuvaavina korvaavina muuttujina. Jos esimerkiksi suuruusluokalla ja variaatiokertoimella olisi ollut selkeä yhteys ennustetarkkuuteen, toteumatietojen vaihtelusta olisi voitu saada viitteitä siitä, kuinka suuri painoarvo tulisi ennusteita laatiessa antaa regressiivisille ennusteille (kuten keskiarvolle).

Tulosrivin toteuman suurempi variaatiokerroin vaikuttaisi olleen yhteydessä erityisesti pienten tulosrivien huonoon ennustettavuuteen. Vaikuttaa kuitenkin siltä, että variaatiokerroin ei kuvaa ennustettavuutta kovin luotettavasti. Tulosrivin suuruusluokalla oli yhteys ennusteiden osuvuuteen siten, että suurilla tulosriveillä ennustetarkkuus oli suhteessa parempi.

Myös painokertoimella oli merkitseviä korrelaatioita ennustetarkkuuden kanssa. Korrelaation suunta vaikutti olevan riippuvainen tulosrivin suuruusluokasta, koska painokertoimen ja suuruusluokan interaktioterminä oli negatiivinen merkitsevä yhteys ennustevirheiden kokoon. Suuruusluokan kasvaessa painokertoimen ja ennustevirheen itseisarvon korrelaatio kääntyi negatiiviseksi. Pienimmät ennustevirheet ovat näin ollen tulosriveillä, joilla on iso suuruusluokka ja painokerroin. Asiantuntijoiden tekemät ennusteet pärjäsivät myös vertailussa tilastolliseen ennusteeseen paremmin niillä tilastoriveillä, joilla oli suurempi toteuma ja isompi painokerroin.

Heuristinen ajattelu ei teoreettisena mallina tarjoa mitään selityksiä painokertoimen ja ennustetarkkuuden yhteydelle. Havaitun yhteyden on liityttävä muihin virheisiin vaikuttaviin ilmiöihin. Suurempi painokerroin ja suoritteiden suuremmat määrät saavat laskennallisesti aikaan sen, että tilastorivillä on enemmän merkitystä tuloksellisuuden laskennassa. Voitaneen tehdä oletus, että tällaisilla tilastoriveillä on enemmän merkitystä myös ennusteiden laatijoille, eli että ennusteiden laatijat ovat pitäneet tällaisia rivejä tärkeämpinä – tiedostaen tai tiedostamattaan. Tällöin etenkin suuruusluokan ja painokertoimen interaktiotermi saattaa olla yhteydessä siihen, miten tärkeästä tulosrivistä on kyse. Jos näin on, on mahdollista, että ennusteen laadintaan on tällöin käytetty enemmän aikaa ja voimavaroja.

Asiantuntijaennusteet olivat (suhteessa tilastolliseen ennusteeseen) sitä parempia, mitä suurempi oli kyseisen tulosrivin suoritteiden lukumäärä ja mitä suurempi oli kyseiselle suoritteelle tuloksellisuuden laskennassa annettu painokerroin. Tämä voisi tukea ajatusta siitä, että asiantuntijaennusteet ovat parempia silloin, kun ennusteen laatimiseen on kiinnitetty enemmän huomiota. Tällainen tulkinta tuntuu sinänsä loogiselta. Erot siinä, kuinka paljon huolellisemmin ennuste laaditaan tulosriville, jonka merkitys on suurempi, voivat olla yksittäisen ennusteen osalta niin pieniä, etteivät ennusteen laatijat välttämättä itse edes tiedosta eroa. Suuressa otoksessa korrelaatio tulee kuitenkin kiistattomasti esiin, ja jokin selitys asialle on oltava. Se, sattuuko kullakin yksittäisellä kerralla regressiivinen ennuste vai asiantuntijan laatima ennuste lähemmäksi toteumaa, on luonnollisesti hyvin satunnaista. Tämä näkyy selvästi myös regressioiden alhaisissa selityksasteissa. Ison otoksen keskiarvot tuntuvat kuitenkin tavoittavan tutkimusaiheen kannalta relevanttia systematiikkaa ennustevirheissä.

Mielenkiintoista on myös, että painokertoimen negatiivinen yhteys regressiivisen ennusteen paremmuuteen on tilastollisesti merkitsevä silloinkin, kun aiemmat virheet ovat mukana regressiomallissa. Myös tämä voidaan tulkita siten, että asiantuntijat ovat onnistuneet sitä paremmin, mitä tärkeämmästä tulosrivistä on ollut kyse. Edelleen tulkinta perustuu oletukseen siitä, että tulosrivin suuruusluokka, painokerroin ja näiden interaktiotermi onnistuvat jollain tasolla kuvaamaan kunkin tilastorivin merkitystä viraston toiminnalle ja ennusteiden laatijoille. Regressioiden tulosten valossa tällainen tulkinta vaikuttaa ainakin jossain määrin perustellulta.

Tällaiset erot ennustetarkkuudessa eivät liity heuristiseen ajatteluun ja kognitiivisiin vääristymiin muutoin kuin välillisesti. Teorian mukaan ihminen kiinnittyy aiempiin lukuihin

enemmän kuin itse tiedostaa, ja lukuja korjataan vähemmän kuin pitäisi. Tällainenkin ilmiö vaikuttaa olevan aineistossa olemassa, eivätkä nämä kaksi selitysmallia sulje toisiaan pois. Osittain voi olla, että aiempia lukuja ei ole korjattu, koska osaa luvuista ei ole pidetty riittävän tärkeinä. Samaan aikaan voi olla, että osaa luvuista ei ole korjattu, koska tarvetta korjaamiseen ei ole kyetty intuitiivisesti tunnistamaan. Teorian kuvaamalla tavalla heuristisen ajattelun ongelmat korostuisivat juuri sellaisten ennusteiden osalta, joiden laatiminen tehdään nopeasti.

Tämän tutkimuksen menetelmällä ja aineistolla ei ennustetarkkuuden vaihtelun eri syitä voida todistaa eikä erottaa toisistaan. Jos lukumäärältään suuret suoritteet ovat viraston toiminnan ja resursoinnin kannalta keskeisempiä, on hyvin mahdollista, että muut psykologiset tekijät ovat selitysvoimaltaan merkityksellisempiä tässä yhteydessä kuin heuristinen ajattelu ja kognitiiviset vääristymät. Teoreettista pohjaa muille syille pitäisi etsiä motivaatiota ja kannustimia käsittelevistä teorioista, tai kenties organisaatioteorian alalta tai sosiaalipsykologiasta.

5.4. Mitä ennusteilla ilmaistaan?

Ennustevirheet eivät olleet samanlaisia tulosrivien kesken. Toteuman suurempi vaihtelu, jota kuvattiin variaatiokertoimella, ennusti positiivisesti ennustevirhettä. Toteuman suuruusluokka ennusti negatiivisesti ennustevirhettä. Ennustetarkkuus oli huonompi pienillä tulosriveillä, tosin tämä ei koskenut kaikkia hallintoyksiköitä.

Eri toimialojen ja eri hallintoyksiköiden välillä oli ennustevirheissä vuodesta toiseen samankaltaisena toistuvia eroja. Toimialoja ei voitu aineistossa erottaa toisistaan, mutta niiden eroavaisuudet voitiin havaita välillisesti painokertoimen kautta, koska eri toimialoilla on hyvin erilaisilla painokertoimilla olevia suoritteita. Painokertoimen yhteys ennustevirheisiin oli erilainen eri hallintoyksiköissä. Toisin sanoen eri toimialojen ennusteet erosivat toisistaan, ja toimialojen erot olivat erilaisia eri hallintoyksiköissä.

Koko aineistossa suurempi painokerroin oli yhteydessä suurempaan ennustevirheeseen (huonompaan ennustetarkkuuteen). Painokerroin ei yksinään merkitsevästi korreloinut ennusteen harhaisuuden kanssa. Aineistossa pienimmät painokertoimet liittyivät pääosin rajatarkastuksen ja suurimmat painokertoimet rikostorjunnan suoritteisiin. Rajatarkastusten

ennustaminen vaikuttaa olleen harhattominta ja lisäksi melko samanlaista eri hallintoyksiköiden kesken.

Hallintoyksiköiden ja toimialojen välillä havaittavia eroja ei kannata ajatella ennusteiden paremmuuden vertailuna. Voidaan pohtia, miksi aineistoon on jäänyt hallintoyksiköiden väliset sitkeät erot, vaikka Rajavartiolaitoksen esikunta on tarkastanut luvut ennen niiden hyväksymistä. Vaikuttaa siltä, että lukuja ei ole tarkastamisen yhteydessä muutettu niin paljon, että hallintoyksiköiden ja toimialojen erot olisivat kokonaan hävinneet. Ankkuroituminen tarjoaisi yhden mahdollisen selityksen tälle. Teorian mukaan ihminen kiinnittyy näkemiinsä lukuihin tiedostamaansa enemmän. Tämä tarjoaisi yhden mahdollisen osaselityksen sille, miksi samassa paikassa tarkastettuihin lukuihin on jäänyt havaittavaksi vuodesta toiseen säilyvät erot hallintoyksiköiden välillä.

Aineistossa saattaa olla käsillä tilanne, jossa ennusteiden laatijoiden välillä on eroja siinä, miten piste-ennusteet ymmärretään. Osa ennusteista (suunnitelmaluvuista) on selvemmin laadittu kuvaamaan tilastoitavan suoritteiden todennäköisintä lukumäärää. Tämä näkyy siinä, että osassa hallintoyksiköitä ja toimialoja ei ennustevirheissä ole suurta systemaattista virhettä, vaan virheet sijoittuvat enemmän satunnaisesti nollan ympärille. Osa ennusteiden laatijoista on kuitenkin saattanut ilmaista suunnitelmaluvulla jotain raja-arvoa tai kapasiteettia, johon tulee varautua. Tällainen mekanismi voisi selittää ennusteiden positiivista harhaisuutta ja sitä, miksi positiiviset ennustevirheet ovat aineistossa yleisempiä kuin negatiiviset. Jos ennusteet ovat tässä mielessä tarkoituksellisesti toteumaa suuremmat, eivät heuristinen ajattelu ja kognitiiviset vääristymät ole oikea teoriapohja havainnon tulkitsemiseen. Tällainen argumentti ei kuitenkaan päde koko aineistossa, koska osalla hallintoyksiköistä ennusteet ovat systemaattisesti toteumaa suuremmat ja osalla eivät ole. Tästä voidaan kenties tehdä varovainen johtopäätös siitä, että hallintoyksiköiden ja toimialojen välillä ei vallitse yhteisymmärrys siitä, mitä suunnitelmaluvulla on tarkoitus ilmaista. Tällä asialla on merkitystä tämän tutkimustyön validiteetin kannalta, koska teorian valinnassa ja tutkimustulosten tulkinnassa on lähdetty siitä oletuksesta, että suunnitelmalukujen tulkitaan olevan ennusteita suoritelmäärästä. Tämä pitää implisiittisesti sisällään oletuksen siitä, että suunnitelmaluvulla on pyritty ilmaisemaan suoritteiden todennäköisintä lukumäärää eli odotusarvoa.

Prosessin kehittämisen näkökulmasta hallintoyksiköiden väliset erot ennusteiden tarkkuudessa ja harhaisuudessa ovat viraston suunnittelun kannalta huono asia. Jos

ennusteiden laadinnassa on mukana tiedostamatonta ankkuroitumista aiempiin lukuihin, kuten tulosten valossa vaikuttaa, myös sillä on merkitystä, kuka tekee ennusteen ensin (hallintoyksikkö vai Rajavartiolaitoksen esikunta). Tämän työn aineisto koski vuosia 2009–2014, jona aikana hallintoyksikkö teki omat ennusteensa ensin ja toimitti ne tämän jälkeen Rajavartiolaitoksen esikunnalle tarkastettavaksi.

Vuoden 2014 jälkeen suunnitteluprosessia muutettiin niin, että Rajavartiolaitoksen esikunta teki omat alustavat lukunsa ensin, ja lähetti ne lausuttavaksi hallintoyksiköihin. Jos ennustevirheissä on mukana tiedostamatonta ankkuroitumista, ja jos Rajavartiolaitoksen esikunta ei tehnyt vuoden 2014 jälkeen omia ennusteitaan päivittämällä vanhoja ennusteita, hallintoyksiköiden välisten erojen olisi pitänyt periaatteessa tuolloin hävitä. Nyt käsillä olevan aineiston perusteella ei voida todeta, tapahtuiko näin. Voidaan kuitenkin todeta, että teorian perusteella tämän pitäisi olla juuri se järjestys, missä ennusteet kannattaa laatia.

Argumentteja voitaisiin esittää myös sen puolesta, että suunnitelmaluvuilla ilmaistaisiin kapasiteettia, johon suunnittelussa pitää varautua, eli käytettäisiin lukuja, jotka ovat tarkoituksellisesti suuremmat kuin suoritteiden odotusarvot. Peruste voisi olla esimerkiksi se, että ylikapasiteetilla toimiminen on yleisesti ajatellen kallista, ja jos esimerkiksi henkilöstö- ja tilahankkeet mitoitetaan tarkasti odotusarvoon, joudutaan usein toimimaan yli suunnitellun kapasiteetin. Johtamisen ja suunnittelun kannalta on kuitenkin niin, että suunnitteluvaraa ei kannata piilottaa ennusteisiin vaan ilmaista eksplisiittisesti erikseen. Jos jonkinlainen varakapasiteetti sisällytetään ennusteeseen, se jää tarkastelulta piiloon. Tämä voi olla valmisteltavan aineiston perusteella päätöksittäin tekevien virkamiesten kannalta harhaanjohtavaa. Luottamusvälien arvioiminen voisi selventää asian käsittelyä. Tämä tarkoittaa järjestelyä, jossa ennuste ilmaistaisiin esimerkiksi muodossa: "odotusarvo 970 000, 90 % luottamusväli (720 000, 1 220 000)".

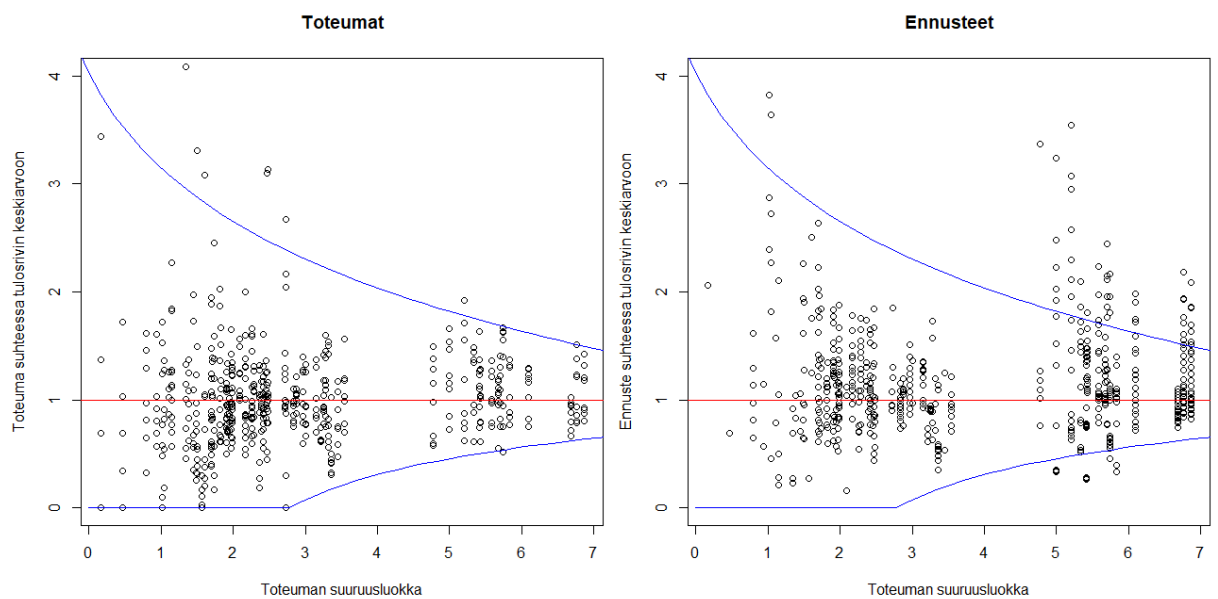
5.5. Odotukset, joita ennusteisiin kohdistetaan

Myös suhtautumista ennusteisiin, ja etenkin yhdellä luvulla ilmaistuihin piste-ennusteisiin tulisi tarkastella. Tämän mahdollistamiseksi ennusteen laatimisen yhteydessä voisi olla hyödyllistä arvioida myös luottamusvälejä ja ennustettavuutta. Aiempien ennusteiden virheitä olisi hyvä tarkastella ja raportoida tietyin väliajoin. Virheiden arvioinnin ei tarvitse olla monimutkainen tai raskas prosessi, enemmän kyse olisi siitä, että ennusteiden laatimisen

yhteydessä arvioitaisiin sitä, ovatko aiemmin tehdyt arviot osuneet oikeaan. Tällöin vältyttäisiin tilanteilta, joissa aiemmat ennusteet ovat pahasti virheellisiä, mutta niitä silti toistetaan pitkälle tulevaan.

Ennusteiden osuvuudelle voisi myös asettaa jonkinlaiset tavoitteet. Tavoitteiden ei välttämättä tarvitsisi olla numeroin ilmaistavia tavoitteita. Myös muunlainen virheisiin liittyvä ohjaus olisi mahdollista. Esimerkiksi voitaisiin antaa ohjeita siitä, onko toivottavampaa, että ennuste on suurempi vai pienempi kuin toteuma. Tavoite voidaan ymmärtää laajana käsitteenä ja tavoitteet asettaa sen mukaan, mikä viraston johdon käyttötarve ennusteille on.

Aiempien vuosien toteumatiedoista voidaan havainnollistaa sitä, kuinka suuri toteuman vaihteluväli on historiallisesti ollut (kuva 21). Tämä voi olla avuksi sekä suunnittelussa että pyrittäessä arvioimaan ennustettavuutta. Siniset viivat kuvaavat aluetta, jolla suurin osa toteuman vaihtelusta on aineistossa tapahtunut. Viivoja ei ole sijoitettu millään tarkoilta laskennallisilla perusteilla, vaan ne ovat ainoastaan graafisia apuvälineitä kuvien vertailun helpottamiseksi.



Kuva 21. Toteumien vaihteluvälin arviointia.

Kuvan 21 lukuarvoja ei kannata tulkita liian tarkasti. Luvut on suhteutettu 11 vuoden (2006–2016) keskiarvoon, jota ääritoteumat muuttavat. Toteumat olisivat siis saattaneet olla äärimmäisempiä verrattuna siihen aiempien lukujen keskiarvoon, joka olisi ollut ennusteen laatijan laskettavissa. Vastaavasti ennusteet ovat saattaneet olla lähempänä *laadintahetken* tulosrivin keskiarvoa, joka on myöhemmin aikasarjassa voinut muuttua. Se, miten lukujen

suhteutustapa valitaan, vaikuttaa lukuarvoihin. Lisäksi vaihteluvälit eivät luotettavasti päde tarkastellun otoksen ulkopuolella. Toisin sanoen kuvan 21 esittämät jakaumat ja vaihteluvälit eivät välttämättä toistu samanlaisina toisessa aineistossa.

Tarkkoja lukuarvoja olennaisempaa onkin miettiä, onko tällaisesta esitystavasta apua toteuman historiallisen vaihtelun havainnollistamisessa. Osasta tilastorivejä olisi melko vertailukelpoista aineistoa jo n. 20 vuoden ajalta. 2001–2021 aikavälille mahtuu myös useita rajanylitysliikenteeseen melko merkittävästi vaikuttaneita kriisejä, finanssikriisistä Krimin miehityksen kautta koronapandemiaan (jonka vaikutus rajanylitysliikenteeseen oli kaikkein dramaattisin mainituista). Havainnollistamalla pitkä aikasarja voitaisiin ennusteiden laatijoille ja lukijoille luoda kuvaa siitä, millainen toteumalukujen vaihtelu on aiemmin ollut. Tällöin suunnittelukautta koskeville ennusteille olisi jonkinlainen pohja arvioida esimerkiksi 90% luottamusväliä.

Jos huomioidaan, mitä kirjallisuuden perusteella voidaan olettaa piste-ennusteisiin suhtautumisesta, edellä mainitulla havainnollistamisella on vielä suurempi merkitys. Teorialuvussa kuvatus mukaisesti asiantuntijatkin aliarvioivat vaihteluvälejä, ja sijoittavat vaihteluvälin tyypillisesti yksittäisen (esim. viimeisimmän) toteutuneen luvun ympärille, rationaalisempien vaihtoehtojen (kuten keskiarvon) sijaan. Tässä työssä esitetyn näytön perusteella ei voida sanoa, pitääkö tämä paikkansa. Mutta jos pitää, se voisi tarkoittaa, että sellaiset toteumatiedot, jotka ovat itse asiassa melko odotettavia historiallisen vaihtelun perusteella, tuntuvat yllättävämiltä kuin ne ovat. Esimerkiksi, jos jokin kriisi laskee yllättäen jonkin suoritteiden määriä ko. tilastorivin keskihajonnan verran, olisi perusteetonta kutsua tapahtunutta täydelliseksi yllätykseksi - yhden keskihajonnan suuruinen heilahdus suuntaan tai toiseen on joutsenena korkeintaan vaalean harmaa.

Esimerkkinä voidaan pohtia seuraavaa: Vuonna 2015 Kaakkois-Suomen rajavartioston rajatarkastukset laskivat 24 % edellisestä vuodesta, mitä voi pitää melko merkittävänä häiriönä rajanylitysliikenteessä. Toisaalta voidaan sanoa, että lasku oli vain vähän yli yhden keskihajonnan (2006–2014) suuruinen (−1.1 sd), tai että vuoden 2015 rajatarkastukset olivat vain 10.9 % keskiarvon (2006–2014) alapuolella, tai että 2015 rajatarkastukset olivat vain 0.42 keskihajontaa keskiarvon alapuolella.

Luvut, vertailukohdat ja otokset voi sinänsä valita mielivaltaisesti, eikä etukäteen voi tietää, mikä on oikea referenssipiste. Lisäksi kymmenen vuotta on niin pitkä aika, että liikenteen ns.

normaalitaso (jonka ympärillä vaihtelu tapahtuu) ehtii muuttua. Tärkeämpää onkin se, miten suhtaudumme yhdellä luvulla ilmaistuun piste-ennusteeseen. Miten tahansa vertailuluvut asetetaan, joka tapauksessa piste-ennusteen perusteella olisi pitänyt etukäteen ymmärtää, että 2015 koettu muutos on hyvin mahdollinen, ja selvästi ”sinisten viivojen sisällä” (viitaten kuvaan 21). Jos tällaiset muutokset koetaan yllättävinä, pitää kenties tarkastella sitä, millaisia odotuksia piste-ennusteeseen kohdistetaan.

5.6. Tutkimuksen rajoitteet ja luotettavuuden arviointi

Tämän työn perustaksi valittiin teoria heuristisesta ajattelusta ja kognitiivisista vääristymistä, ja ennustevirheitä on analysoitu ja tutkimustuloksia tulkittu tästä näkökulmasta. Ennustevirheitä aiheuttavat monet eri tekijät, ja valittu teoria ei parhaimmillaankaan voi selittää kuin pienen osan ennustevirheistä. Ei ole mitään perusteita olettaa, että kognitiiviset vääristymät olisivat ennustevirheiden keskeisin syy.

Käytettäessä aineistoa, joka ei ole kerätty tutkimus- tai testiolosuhteissa, saavutetaan luotettavuuden kannalta positiivisena vaikutuksena se, ettei tutkija ole voinut vaikuttaa aineiston syntyyn. Tutkimusaineistona olevat luvut ovat peräisin ihmisten tosiasiallisesta toiminnasta, eikä niitä ole laadittu tutkimusta varten. Tätä voi pitää työn validiteetin kannalta hyvänä asiana. Tällaisella aineistolla on kuitenkin myös väistämättömiä heikkouksia, jotka tuovat mukanaan rajoitteita tutkimukselle: mitään lukuihin vaikuttavaa tekijää voida sulkea pois, kontrollimuuttujia tai taustatietoa ennusteiden laatijoista ei ole käytettävissä, eikä näin ollen havaittujen korrelaatioiden syitä voida mitenkään todistaa.

Tässä yhteydessä ei ole mitään tapaa kvantifioida heurististen ajattelun mahdollisten vaikutusten suuruutta. Koska tietoa ei ole, joudutaan tutkimuksen rajoitteet huomioimaan olettamalla vaikutuksen olevan hyvin pieni verrattuna kaikkien asiaan vaikuttavien seikkojen laajaan joukkoon. Etsittäessä laajasta otoksesta heikkoja korrelaatioita, joudutaan tyytymään etsimään näyttöä siitä, että jokin merkitsevä korrelaatio *on olemassa*. Tällainen lähestymistapa tutkimustuloksiin ei sinänsä välttämättä vähennä tulosten pätevyyttä. Vaikka oletettaisiin vaikutuksen olevan hyvin pieni, kaikki nolasta eroavat vaikutukset antaisivat vastauksia tutkimuskysymyksiin ja perusteita jatkotutkimukselle. Tutkimustuloksista saatava tieto on kuitenkin hyvin suppeaa, koska parhaimmillaankaan voitaisiin päätyä siihen synteisiin, että heuristisella ajattelulla *on merkitystä* ennusteiden laadinnassa. Tällainen lause voi viitata mihin tahansa kooltaan tuntemattomaan, mutta nolasta eroavaan merkitykseen.

Vaikka tutkimustulokset tulkittaisiin niin, että heuristisen ajattelun ja kognitiivisten vääristymien vaikutuksesta ennusteiden laadintaan havaittiin viitteitä, ei tämä työ tuota mitään tietoa muista ennustevirheisiin vaikuttavista mekanismeista. Ainoa poikkeus tälle on johtopäätös siitä, että ennusteet vaikuttavat olleen sitä osuvampia, mitä tärkeämpiä ne ennusteiden laatijoiden kannalta ovat. Tahattomia virheitä syntyy monista eri syistä, ja ennusteita laadittaessa tapahtuu myös tahallisia virheitä. Lisäksi jako tahallisten ja tahattomien virheiden välillä on dikotominen ainoastaan teoriassa. Käytännössä ihminen voi reagoida kannustimiin ja manipuloida lukuja tiedostamattaan, tai turvautua esimerkiksi heuristiikkaan täysin tietoisesti. Virhelähteistä ja -mekanismeista voitaisiin tehdä tutkimusta hyvin monista näkökulmista, ja tämä tutkimustyö ei selitä suurinta osaa ennustevirheiden syistä millään tavalla.

Voidaan todeta, että tässä työssä löydettiin ennusteiden virheistä tiettyä systematiikkaa ja korrelaatioita, jotka ovat samansuuntaisia kuin teorian perusteella olisi voinut olettaa. Yksilötason intuitiiviset epätarkkuudet voivat pieneltä osin selittää aineistosta tehtyjä havaintoja. Vähintään voidaan sanoa, että löydökset eivät ole ristiriidassa heuristisen ajattelun ja kognitiivisten vääristymien teorian kanssa.

Liian pitkälle vietyjen tulkintojen kanssa pitää olla varovainen, koska kausaalisuus on empiirisesti vaikea kysymys. Tilastollisella kuvaamisella ei voida sanoa minkään yksittäisen ennusteen osalta, mistä virhe johtui. Tulkintoja voidaan yrittää tehdä vain koko aineistossa systemaattisesti toistuvista vaikutuksista, jotka tulevat suuressa otoksessa esiin tilastollisesti merkitsevinä korrelaatioina. Näidenkään osalta korrelaatioista ei voi päätellä kausaalisia suhteita. Aineistosta tehtyjen havaintojen syyt ja selitykset ovat täysin keskustelun ja tulkinnan varassa.

Keskeisimmät ongelmat tutkimuksen luotettavuuden kannalta ovat datassa havaittu jäännöstermien autokorrelaatio ja otoskoko. Jaoteltaessa aineistoa hallintoyksiköittäin ja kompensoitaessa autokorrelaatiota, jää N niin pieneksi, että merkitsevien yhteyksien tulkinta muuttuu epäluotettavaksi.

Aiempien ennustevirheiden tilastollinen yhteys niitä seuraaviin ennustevirheisiin on autokorrelaatiota. Tässä tutkimustyössä se on itsessäänkin relevantti tarkasteltava asia siltä osin kuin puhutaan psykologisesta kiinnittymisestä aiempiin ennusteisiin. Autokorrelaatio

kuitenkin heikentää regressioiden tulosten luotettavuutta. Esimerkiksi se, että aiempien virheiden huomioiminen taulukon 10 malleissa 5-8 poistaa merkitsevyyden variaatiokertoimelta ja suuruusluokalta saattaa tarkoittaa, että mallien 1-4 osoittamat merkitsevät yhteydet ovat autokorrelaation aiheuttama vääristymä. Huonoa ennustettavuutta kuvaamaan valittuihin korvaaviin muuttujiin, erityisesti tulosrivin toteuman variaatiokertoimeen, liittyy edellä kuvatun kaltaisia tilastollisen validiteetin ongelmia.

Autokorrelaatiota lukuun ottamatta lineaarisen regression edellyttämät lähtöoletukset täyttyvät pääosin. Muuttujamuunnosten avulla päästiin likimääräiseen normaalisuuteen ja Whiten keskivirheet huomioivat jäännöstermien heteroskedastisuuden. Voitaneen varovasti arvioida, että t-testit eivät ole täysin epävalideja otoskoon ollessa riittävän suuri ($N > 1000$).

Tutkimustulosten ja niiden tulkinnan oikeellisuus riippuu tehtyjen oletusten oikeellisuudesta. Koko tutkimusasetelma perustuu oletukseen, että Rajavartiolaitoksen toiminnan ja talouden suunnittelussa laaditut niin sanotut suunnitelmaluvut ovat ennusteita, ja niitä voidaan tarkastella sen perusteella, mitä kirjallisuudessa sanotaan ennusteista. Kuten edellä, luvussa 5.4. on keskusteltu, ei ole täysin riidatonta, pitääkö tämä lähtöoletus paikkansa. Jos näin ei ole, on tutkimusasetelma ja tutkimustulosten tulkinnassa valittu näkökulma tältä osin virheellinen.

Tutkimustulosten tulkinta perustui osiltaan oletukseen siitä, että tulosrivin toteuman suuruusluokka ja suoritteille tuloksellisuuden laskennassa annetut painokertoimet kuvaavat jotain siitä, miten suuri merkitys kullakin tilastorivillä ennusteiden laatijoille on. Jos tämä oletus ei pidä paikkaansa, ovat tehdyt tulkinnat tältä osin vääriä.

5.7. Jatkotutkimuksen tarve

Kuten Johdannossa todettiin, suoritteiden määriä koskevat suunnitelmaluvut valittiin tämän tutkimuksen aineistoksi ennen kaikkea siksi, että ne tarjoavat helposti hyväksyttävän lähtöpisteen. Teorian implikaatiot ennusteisiin ovat verraten selviä, aineisto on numeerista ja siinä mielessä riidatonta, ennustevirheet ovat mitattavissa, ja suoritteiden toteumista on jälkikäteen muodostunut yksikäsitteinen, helposti osoitettava oikea vastaus.

Pyrittäessä hahmottamaan heuristisen ajattelun ja kognitiivisten vääristymien vaikutusta johtamiseen ja päätöksentekoon, suunnitelmaluvut eivät ole kaikkein olennaisin aineisto. Esimerkiksi siihen, mikä vaikutus mahdollisilla kognitiivisilla vääristymillä on riskianalyysiin tai operatiivisen toiminnan johtamiseen, liittyy merkitykseltään suurempia kysymyksiä.

Yksittäisen tutkimuksen perusteella ei ole syytä tehdä liian pitkälle meneviä johtopäätöksiä, ja tutkimustulosten yleistäminen käsitellyn aineiston ulkopuolelle on aina jossain määrin epäluotettavaa. Kuitenkin, jos löydökset pitävät paikkansa, ja jos varovainen tulkinta kognitiivisten vääristymien läsnäolosta pitää paikkansa, niin tällä olisi relevanssia muidenkin kuin numeeristen ennusteiden kannalta.

Jos esimerkiksi on niin, että ennusteiden laatijat toistavat vanhoja ennusteita edes marginaalisesti enemmän kuin itse tiedostavat, on vaikea keksiä perusteluja sille, miksi ilmiö ei koskisi myös laadullisia ja sanallisia ennusteita ja arvioita. Erityisen ongelmallista tämä olisi sellaisissa yhteyksissä, joissa aiempien asiakirjojen tekstiä hyödynnetään suoraan kopioimalla sitä uuden asiakirjan pohjaksi.

Riskianalyseja tai esimerkiksi todellisissa tilanteissa toimivan kenttäjohtajan tai meripelastusjohtajan tekemiä päätöksiä ei voida tutkia vastaavalla menetelmällä kuin tässä tutkimuksessa on käytetty. Tämä rajoite aiheutuu siitä, ettei tehtyjä päätöksiä tutkittaessa voida jälkikäteenkään mitenkään absoluuttisesti määrittää, mikä olisi ollut "oikea vastaus", eikä näin ollen voida myöskään kvantifioida virheitä millään tavoin. Tämä on kuitenkin ajateltava vain metodologisenä haasteena, eikä asetelman vaikeus saisi muodostua jatkotutkimuksen esteeksi.

Turvallisuusviranomaisten operatiiviseen johtamiseen kuuluu tyypillisesti seikkoja, jotka teorialuvussa esitetyn mukaan korostavat heurististen ajattelun merkitystä. Tällaisia ovat muun muassa informaation puutteellisuus, tilanteen epävarmuus ja päätöksentekoon liittyvä aikapaine. Tämäkin seikka korostaa tarvetta siirtyä seuraavissa tutkimuksissa riidattomasta ja selkeästä aineistosta kohti epäselvempiä, vaikeammin koottavia ja tulkittavia, mutta merkitykseltään tärkeämpiä aineistoja.

Eräs tapa ratkaista edellä kuvattuja metodologisia ongelmia voisi olla hyväksyä – tarvittavin varauksin – tämän työn tutkimustulos siitä, että heuristisella ajattelulla ja kognitiivisilla vääristymillä on vaikutusta päätöksentekoon todellisissa tilanteissa. Tämän tutkimustyön

rajoitteet voitaisiin huomioida esimerkiksi siten, että ei yritettäisikään kvantifioida sitä, miten suuri tämä vaikutus on. Johtopäätös voitaisiin hyväksyä jatkotutkimuksen lähtöolettamukseksi sillä tasolla, että vaikutus on suuruudeltaan tuntematon, mutta erisuuri kuin nolla. Tällöin voitaisiin seuraavana askelena tutkia esimerkiksi erilaisissa johtamistehtävissä toimivan henkilöstön niitä ominaisuuksia, jotka kirjallisuuden mukaan altistavat yksilöä intuitiivisen ajattelun aiheuttamille virheille.

Myös moni numeerisiin ennusteisiin läheisesti liittyvä kysymys jää jatkotutkimuksen varaan: Miten jäsentyneesti ennusteita hyödyntävät virkamiehet osaavat suhtautua piste-ennusteisiin? Kuinka hyviä ennusteiden laatijat itse arvioivat olevansa? Vaikuttavatko ennusteen laatijan käsitykset esimerkiksi jonkin asian tärkeydestä tai kollegoiden työn laadusta tai työpaineesta siihen, mitä he – tiedostaen tai tiedostamattaan – ilmaisevat suunnitelmaluvuilla? Nyt käytössä olleella aineistolla ja valitulla tutkimusmenetelmällä ei voitu vastata tämänkaltaisiin kysymyksiin.

LÄHTEET

- Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 26(2), 211-243.
- Böhm, G., & Brun, W. 2008. Intuition and affect in risk perception and decision making. *Judgment and Decision Making*, 3(1), 1–4.
- Camerer, C. F., & Hogarth, R. M. 1999. The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of risk and uncertainty*, 19(1-3), 7-42.
- Ehrbeck, T. & Waldmann, R. 1996. Why are professional forecasters biased? Agency versus behavioral explanations. *The Quarterly Journal of Economics*, 111(1), 21-40.
- Einhorn, H. J. 1982. Learning from experience and suboptimal rules in decision making. Teoksessa Kahneman, D., Slovic, P. & Tversky, A. (toim.) *Judgment under uncertainty: Heuristics and biases*. 28. painos. Cambridge: Cambridge University Press. 268-286.
- Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. 2000. The affect heuristic in judgments of risks and benefits. *Journal of behavioral decision making*, 13(1), 1-17.
- Fischhoff, B. 1982. Debiasing. Teoksessa Kahneman, D., Slovic, P. & Tversky, A. (toim.) *Judgment under uncertainty: Heuristics and biases*. 28. painos. Cambridge: Cambridge University Press. 422-444.
- Gilovich, T., Griffin, D. 2002. Introduction - Heuristics and biases: Then and now. Teoksessa Gilovich, T., Griffin, D.W. & Kahneman, D. (toim.) *Heuristics and biases: The psychology of intuitive judgment*. 14. painos. New York: Cambridge University Press. 1-18.
- Gigerenzer, G. 1996. On Narrow Norms and Vague Heuristics: A Reply to Kahneman and Tversky. *Psychological Review*, 103(3), 592-596.
- Johnson, E. J., Hershey, J., Meszaros, J., & Kunreuther, H. 1993. Framing, probability distortions, and insurance decisions. *Journal of risk and uncertainty*, 7(1), 35-51.

- Kahneman, D. 2003. A perspective on judgment and choice: mapping bounded rationality. *American psychologist*, 58(9), 697.
- Kahneman, D., & Frederick, S. 2002. Representativeness revisited: Attribute substitution in intuitive judgment. Teoksessa Gilovich, T., Griffin, D.W. & Kahneman, D. (toim.) *Heuristics and biases: The psychology of intuitive judgment*. 14. painos. New York: Cambridge University Press. 49-81.
- Kahneman, D. & Tversky, A. 1982. On the psychology of prediction. Teoksessa Kahneman, D., Slovic, P. & Tversky, A. (toim.) *Judgment under uncertainty: Heuristics and biases*. 28. painos. Cambridge: Cambridge University Press.
- Langer, E. 1982. The illusion of Control. Teoksessa Kahneman, D., Slovic, P. & Tversky, A. (toim.) *Judgment under uncertainty: Heuristics and biases*. 28. painos. Cambridge: Cambridge University Press.
- LeBoeuf, R. A., & Shafir, E. 2003. Deep thoughts and shallow frames: On the susceptibility to framing effects. *Journal of Behavioral Decision Making*, 16(2), 77-92.
- Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. 2001. Risk as feelings. *Psychological bulletin*, 127(2), 267.
- Martelius, Juha. Luennot Maanpuolustuskorkeakoulussa 2.9.2019.
- Mellers, B., Ungar, L., Baron, J., Ramos, J., Gurcay, B., Fincher, K., ... & Murray, T. 2014. Psychological strategies for winning a geopolitical forecasting tournament. *Psychological science*, 25(5), 1106-1115.
- Mälkki, J. 2015. Akateeminen sotapäällikkö? Sotataito, sodankäynti ja upseeriston kyvykkyys. Teoksessa Nokkala, A., Hanska J. & Häyry, M. (toim.) *Akateemisuus ja upseerius - Tieteen, tutkimuksen ja johtamisen ristivetoa puolustushallinnossa*. Tampere: Maanpuolustuskorkeakoulu.
- Newby, E. & Orjasniemi, S. 2011. Suomen Pankin ennustevirheet vuosina 2004-2010. *Euro & talous. Talouden näkymät* (3)2011, 65-72. Luettu osoitteesta <http://urn.fi/URN:NBN:fi:bof-20140812219>.
- Nisbett, R. E., Krantz, D.H., Jepson, C. & Kunda, Z. 2002. The use of statistical heuristics in everyday inductive reasoning. Teoksessa Gilovich, T., Griffin, D.W.

- & Kahneman, D. (toim.) *Heuristics and biases: The psychology of intuitive judgment*. 14. painos. New York: Cambridge University Press. 510-533.
- Popper, Karl 1963: *Science: Conjections and Refutations*. Teoksessa Balashkov, Yuri & Rosenberg, Alex (toim.) (2002): *Philosophy of Science, Contemporary Readings*. Bury St Edmunds: Routledge. 294 - 301.
- Schwarz, N. & Vaughn, L. A. 2002. The availability heuristic revisited: Ease of recall and content of recall as distinct sources of information. Teoksessa Gilovich, T., Griffin, D.W. & Kahneman, D. (toim.) *Heuristics and biases: The psychology of intuitive judgment*. 14. painos. New York: Cambridge University Press. 103-119.
- Shafir, E. 1993. Choosing versus rejecting: Why some options are both better and worse than others. *Memory & cognition*, 21(4), 546-556.
- Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. 2007. The affect heuristic. *European journal of operational research*, 177(3), 1333-1352.
- Stanovich, K. 2011. *Rationality and the reflective mind*. Oxford: Oxford University Press.
- Stanovich, K. E., & West, R. F. 2000. Individual differences in reasoning: Implications for the rationality debate? *Behavioral and brain sciences*, 23(5), 645-665.
- Stanovich, K. E., & West, R. F. 2002. Individual differences in reasoning: Implications for the rationality debate? Teoksessa Gilovich, T., Griffin, D.W. & Kahneman, D. (toim.) *Heuristics and biases: The psychology of intuitive judgment*. 14. painos. New York: Cambridge University Press. 421-440.
- Stanovich, K. & West, R. F. 2011. Intelligence as a Predictor of Performance on Heuristics and Biases Tasks. Teoksessa Stanovich, K. *Rationality and the reflective mind*. Oxford: Oxford University Press.
- Tversky, A. & Kahneman, D. 1981. The framing of decisions and the psychology of choice. *Science*, 211, 453-458.
- Tversky, A. & Kahneman, D. 1982a. Judgment under uncertainty: Heuristics and biases. Teoksessa Kahneman, D., Slovic, P. & Tversky, A. (toim.) *Judgment under uncertainty: Heuristics and biases*. 28. painos. Cambridge: Cambridge University Press. 3-20.

- Tversky, A. & Kahneman, D. 1982b. Judgements of and by representativeness. Teoksessa Kahneman, D., Slovic, P. & Tversky, A. (toim.) Judgment under uncertainty: Heuristics and biases. 28. painos. Cambridge: Cambridge University Press. 84-98.
- Yudkowsky, E. 2008. Cognitive biases potentially affecting judgment of global risks. Teoksessa Bostrom, N. & Cirković, M.M. (toim.) Global catastrophic risks. New York: Oxford University Press. 91-119.
- Verbeek, M. 2012. A guide to modern econometrics. 4th Edition. Padstow: John Wiley & Sons.
- White, H. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: journal of the Econometric Society*, 817-838.
- Wilson, T. D., Centerbar, D. B. & Brekke, N. 2002. Mental contamination and the debiasing problem. Teoksessa Gilovich, T., Griffin, D.W. & Kahneman, D. (toim.) Heuristics and biases: The psychology of intuitive judgment. 14. painos. New York: Cambridge University Press.