

PRO GRADU -TUTKIELMA

**Janne Kaseva**

Väkivalta- ja omaisuusrikollisuus Suomessa

TAMPEREEN YLIOPISTO

Matematiikan ja tilastotieteen laitos

Tilastotiede

Toukokuu 2010

Tampereen yliopisto

Matematiikan ja tilastotieteen laitos

KASEVA, JANNE: Väkivalta- ja omaisuusrikollisuus Suomessa

Pro gradu -tutkielma, 69 s., 6 liites.

Tilastotiede

Toukokuu 2010

---

## Tiivistelmä

Tutkimuksen tavoitteena on tarkastella ja vertailla suomalaisten kuntien rikollisuuden rakennetta ja siinä tapahtuneita muutoksia. Vertailukohtana on aiempi tutkimus, joka käsitteli 2000-luvun alun rikostilannetta kunnissa. Tutkimuksessa tarkasteltiin ennen kaikkea väkivalta- ja omaisuusrikollisuutta.

Kuntien väkivalta- ja omaisuusrikollisuutta selitettiin regressiomallien avulla. Mallien antamia ennusteita verrattiin poliisin tilastoimaan rikostasoon. Tulosten perusteella kunnat luokiteltiin viiteen ryhmään ja niiden rikostasoluokitusta verrattiin aiemman tutkimuksen tuloksiin. Vertailussa huomattiin, että kahdessa-toista kunnassa rikostasoa oli pysynyt poikkeuksellisen korkeana. Kuntia, joissa sekä väkivalta- että omaisuusrikostasoa olivat molemmat poikkeuksellisen korkeita, oli myös 12.

Väkivalta- ja omaisuusrikoksia selittävien regressiomallien selitysasteet vaihtelivat välillä 0,48–0,80. Kaupunkimaisten kuntien ryhmässä saavutettiin selvästi korkeimmat selitysasteet. Malleista huomattiin, että alkoholilla oli selkeä yhteys väkivaltarikollisuuteen. Omaisuusrikollisuutta selittivät alkoholin käytön ohella myös huumausainerikokset. Muita rikollisuutta selittäviä muuttujia olivat muun muassa koulutus ja muuttoliike. Perhesuhteilla ja varallisuudella oli myös keskeinen asema useissa malleissa.

Faktorianalyysin avulla pyrittiin löytämään muuttujien taustalla olevia rakenteita. Faktorimallien selitysasteet vaihtelivat välillä 0,63–0,76. Parhaimmat selitysasteet saatiin kaupunkimaisten kuntien ryhmässä. Kaikissa ryhmissä parhaaksi faktoreiden lukumääräksi osoittautui viisi. Faktorit *koulutus ja varallisuus* ja *lapsiperheet* olivat mukana kaikissa malleissa. Myös *työttömyys*, *kaupunkimaisuus* ja *sosiaaliset ongelmat* –faktorit esiintyivät useissa malleissa. Faktoreiden suhteellisen selkeät tulokset kertovatkin osaltaan kuntien rakenteista.

**Asiasanat:** faktorianalyysi, rikollisuus kunnissa, väkivaltarikollisuus, omaisuusrikollisuus, eliminoiva regressioanalyysi

# Sisällysluettelo

<b>1 Johdanto</b>	<b>5</b>
1.1 Tutkimuksen lähtökohdat ja tavoitteet .....	6
1.2 Rikollisuuden tutkimuksen historiaa .....	6
<b>2 Tutkimuksen toteutus</b>	<b>8</b>
2.1 Aineiston kuvaus .....	8
2.2 Kuntien ryhmittely .....	9
2.3 Selittävät muuttujat .....	10
2.4 Muuttujien välisten rakenteiden tutkiminen B-Course –ohjelmalla	11
<b>3 Väkivaltarikollisuus kunnissa</b>	<b>14</b>
3.1 Kaupunkimaiset kunnat .....	14
3.2 Taajaan asutut kunnat .....	18
3.3 Maaseutumaiset kunnat .....	21
3.4 Väkivaltarikosmallien vertailu .....	24
<b>4 Omaisuusrikollisuus kunnissa</b>	<b>26</b>
4.1 Kaupunkimaiset kunnat .....	26
4.2 Taajaan asutut kunnat .....	29
4.3 Maaseutumaiset kunnat .....	32
4.4 Omaisuusrikosmallien vertailu .....	35

<b>5</b>	<b>Faktorianalyysi</b>	<b>37</b>
5.1	Faktorianalyysin tausta . . . . .	37
5.2	Faktorimallin muodostus . . . . .	37
5.3	Faktorointi . . . . .	39
5.4	Rotatointi . . . . .	40
5.4.1	Ortogonaaliset rotaatiot . . . . .	40
5.4.2	Vinokulmaiset rotaatiot . . . . .	42
5.5	Faktoripistemäärät . . . . .	43
<b>6</b>	<b>Faktorimallien sovitus aineistoon</b>	<b>45</b>
6.1	Kaikki yli tuhannen asukkaan kunnat . . . . .	46
6.2	Kaupunkimaiset ja taajaan asutut kunnat . . . . .	49
6.3	Kaupunkimaiset kunnat . . . . .	53
6.4	Taajaan asutut kunnat . . . . .	55
6.5	Maaseutumaiset kunnat . . . . .	58
<b>7.</b>	<b>Tutkimustulosten analysointi</b>	<b>61</b>
<b>8.</b>	<b>Yhteenveto</b>	<b>64</b>
	<b>Lähdeluettelo</b>	<b>68</b>
	<b>Liite A: Muuttujaluettelo</b>	<b>70</b>
	<b>Liite B: Kolmiluokkainen kuntaryhmittely</b>	<b>72</b>
	<b>Liite C: Väkivaltarikosmalleihin valitut muuttajat</b>	<b>74</b>
	<b>Liite D: Omaisuusrikosmalleihin valitut muuttajat</b>	<b>75</b>

# 1 Johdanto

Aluksi esitellään tutkimuksen lähtökohdat ja tavoitteet. Lisäksi paneudutaan rikollisuuden tutkimuksen historiaan ja tutkimustuloksiin. Erityisesti kiinnitetään huomiota alueiden rikollisuutta selittäviin tekijöihin ja jätetään yksilölliset ominaisuudet ja taipumukset rikollisuuteen vähemmälle tarkastelulle.

Toisessa luvussa esitellään kunta-aineisto ja siihen poimitut muuttujat, jotka on pyritty valitsemaan kattavasti erilaisilta kunnan rakennetta kuvaavilta osalualueilta. Kunnat on myös jaettu kolmeen ryhmään niiden rakenteen perusteella. Viimeisessä alaluvussa tutkitaan muuttujien välisiä yhteyksiä B-Course –ohjelman avulla.

Luvuissa 3 ja 4 tutkitaan kuntien väkivalta- ja omaisuusrikollisuutta. Kuntaryhmille muodostetaan molempia rikostyyppjä selittävät regressiomallit, joiden pohjalta kuntien rikostasoa voidaan selittää. Malleja verrataan poliisin tilastoihin rikollisuuteen, ja tulosten perusteella kunnat luokitellaan viiteen rikostasoluokkaan. Luokituksen perusteella nähdään missä kunnissa rikostasoa on poikkeuksellisen korkea ennustettuun tasoon nähden. Molempien lukujen viimeisissä alaluvuissa on vertailtu kuntaryhmien rikollisuutta selittäviä regressiomalleja ja yksittäisten kuntien rikosluokitusta Heiskanen tutkimuksen vastaaviin malleihin ja luokituksiin (Heiskanen, Siren & Roivainen 2003).

Luvussa 5 esitellään faktorianalyysin historiaa ja teoriaa. Faktorianalyysi on tilastollinen menetelmä, jolla pyritään etsimään muuttujien taustalla piileviä rakenteita tiivistämällä muuttujien sisältämää informaatiota. Luvussa 6 etsitäänkin kuntien rikollisuutta selittäviä rakenteita faktorianalyysin avulla. Faktorianalyysi tehdään kaikissa kolmessa kuntaryhmässä, minkä lisäksi ryhmiä myös yhdistellään. Analyysiin ei valita kaikkia muuttujia vaan muuttujat valitaan tapauskohtaisesti ryhmien rakenteiden ja ominaisuuksien perusteella. Muodostuneille faktoreille pyritään aina löytämään tulkinta, jonka perusteella ne voidaan myös nimetä. Faktoripisteiden avulla voidaan tarkastella faktoreiden ja rikostasojen välisiä korrelaatioita.

Luvussa 7 vertaillaan yksittäisten kuntien väkivalta- ja omaisuusrikostasoja rinnakkain. Toisin sanoen tarkastellaan, missä kunnissa molemmat rikostasot poikkeavat odotetusta rikostasosta. Näin voidaan arvioida yksittäisten kuntien yleistä rikostasoa 2000-luvun jälkipuoliskolla. Myös faktorianalyysin tulokset esitellään tässä luvussa. Eri kuntaryhmissä muodostetut faktorit on monesti nimetty samoin, mutta ne ovat kuitenkin sisällöltään ja painotuksiltaan erilaisia.

Luvussa 8 kerrotaan tiivistetysti tutkimuksen lähtökohdat, toteutus ja olennaisimmat tulokset. Kuntien rikostasojen vertailua, sekä keskenään että aiemman tutkimuksen pohjalta, on niin paljon, ettei sitä voi kattavasti esitellä tässä yhteydessä. Tarkempia tuloksia kannattaakin etsiä kappaleista, joissa ne alun perin esitettiin.

## 1.1 Tutkimuksen lähtökohdat ja tavoitteet

Suomalaisten kuntien rikollisuutta on tutkittu aiemminkin. Erityisesti kiinnitän huomiota vuonna 2003 ilmestyneeseen tutkimukseen *Rikollisuus kunnissa* (Heiskanen et al. 2003), jossa on selvitetty vuosituhannen alun rikollisuutta kunnissa. Käytän pitkälti samoja taustamuuttujia kuin kyseisessä tutkimuksessa on käytetty. Lisäksi olen täydentänyt aineistoa muutamilla muuttujilla, joilla saattaisi olla vaikutusta rikollisuuteen. Valinnat on pyritty tekemään aiempien tutkimustulosten perusteella. Kuntien ennustetut rikostasot on muodostettu eliminoivan regressioanalyysin avulla. Tilastollisissa analyyseissä olen käyttänyt SPSS-ohjelmaa.

Tutkimukseni noudattaa kappaleiden 3–5 osalta pitkälti Heiskasen tutkimuksen rakennetta, mikä helpottaa tutkimustulosten vertailua. Samasta syystä myös taulukot ja analyysit on pääsääntöisesti tehty vastaavalla tavalla. Yksi tutkimuksen keskeisin tarkoitus onkin verrata 2000-luvun alun tuloksia muutaman vuoden takaiseen tilanteeseen. Varsinainen vertailu esitetään kappaleissa 3.4 ja 4.4, minkä lisäksi liitteissä on tarkempaa tietoa analyysien suoritukseen ja tuloksiin liittyen.

Tutkimuksen toisessa osassa kunta-aineistosta on faktorianalyysin avulla pyritty löytämään ulottuvuuksia, joiden avulla rikollisuuden rakennetta kunnissa olisi helpompi ymmärtää. Faktorianalyysin avulla taustamuuttujista saatua informaatiota voidaan tiivistää. Tarkoituksena ei kuitenkaan ole ottaa kaikkia muuttujia mukaan analyysiin vaan valita sopivat muuttujat tilanteen mukaan. Tavoitteena on löytää faktoreita, joiden tulkinta ja liittäminen rikollisuuteen olisi mielekästä. Lopuksi faktoreiden ja rikostasojen välisiä korrelaatioita on vielä vertailtu faktoripisteiden avulla.

## 1.2 Rikollisuuden tutkimuksen historiaa

Rikollisuuteen vaikuttavia tekijöitä on tutkittu paljon viidenkymmenen viime vuoden aikana. Täysin yksiselitteisiä vastauksia ei kuitenkaan ole löydetty, vaan tutkimustulokset ovat monelta osin ristiriitaisia. Selitystä on haettu niin sosiaalisista, taloudellisista kuin yhteiskunnallisistakin tekijöistä. Alla on esitelty muutamia keskeisiä teorioita. Ensisijaisesti on keskitytty alueellisia ja rakenteellisia tekijöitä koskeviin tutkimuksiin eikä niinkään yksilön toimintaa selittäviin teorioihin.

*Sosiaalisen kontrollin teorian* mukaan ihmiset, joilla on kiinteät suhteet johonkin sosiaaliseen viiteryhmään, tekevät rikoksia harvemmin kuin ne, joilla tällaisia siteitä ei ole. Viiteryhmiä ovat muun muassa perhe, koulu ja ystävät, joihin kiintymisen johdosta yksilö sisäistää yhteisön arvot ja normit. Myös monet muut teoriat tukevat rikollisuuden lisääntymistä lähiympäristön epävirallisen kontrollin vähetessä. Tällaisilla alueilla on usein tarjolla runsaasti rikostilaisuuksia mutta vähän mahdollisuuksia ammatilliseen koulutukseen tai työelämään pääsyyn. (Tolvanen 2005, 23–25., 46.)

Usein myös kuvitellaan, että köyhillä alueilla rikollisuus on yleisempää kuin rikkailla alueilla. Jo klassisten tutkimustulosten mukaan rikkailla alueilla esiintyy kyllä vähemmän väkivaltarikoksia, mutta omaisuusrikollisuuden määrä kasvaa potentiaalisten anastuskohteiden lisääntyessä. Merkitseviä näyttivät olevan varallisuuserot saman alueen sisällä, ei alueiden välillä. Talouserojen yhteyttä rikollisuuteen hämmentävät myös lama-ajat, jolloin rikollisuuden on todettu pysyvän ennallaan tai jopa laskevan. (Tolvanen 2005, 23–25., 46.)

Myös työttömyyden yhteyttä rikollisuuteen on tutkittu paljon. Tutkimustulokset ovat ristiriitaisia ja joissakin tutkimuksissa työttömyyden on jopa todettu laskevan rikollisuutta. Syyksi tähän on epäilty vanhempien parantuneita mahdollisuuksia valvoa lapsiaan. Vaikka rikollisuuteen vaikuttavia tekijöitä, ja niiden yksittäistä vaikutusta, on vaikea arvioida tarkasti, voidaan todeta tiettyjen tekijöiden olevan tyyppisiä korkean rikostason alueilla. Näitä ovat muun muassa korkeat avioeroluvut, yksinhuoltajaperheiden suuri määrä, tiheä asutus, puutteelliset asumisolot, huonot koulutus- ja sosiaalipalvelut, muuttoliike sekä väestöryhmien kärjistyneet suhteet. (Tolvanen 2005, 24–25.)

*Rodney Stark* on tutkinut erityisesti kaupunkimaisen ympäristön rikollisuuteen vaikuttavia rakenteellisia tekijöitä. Hänen mukaansa näitä olivat erityisesti tiheä asutus, köyhyys, sekalainen maankäyttö (teollisuus, liikkeet ja asutus samoilla alueella), muuttoliike ja rakennusten rapistuminen. Nämä rakenteelliset tekijät synnyttävät asukkaissa kyynistä asennetta, lisäävät rikostilaisuuksia ja rikoksen tekemisen motivaatiota sekä heikentävät lähikontrollia. Kehitys johtaa siihen, että lain rikkomiseen taipuvaiset yksilöt muuttavat alueelle ja lakia kunnioittavat pyrkivät sieltä pois. Tämä taas johtaa rikosten lisääntymiseen. (Tolvanen 2005, 31.)

*Robert J. Sampson* (Tolvanen 2005) on esittänyt, että köyhyys yhdessä vilkkaan muuttoliikkeen ja perheiden hajoamisen kanssa lisää rikollisuutta. Sampsonin sosiaalisen disorganisaation teorian mukaan kyse on ennen kaikkea sosiaalisen pääoman puutteesta; mitä enemmän alueella on ihmisten välisiä sosiaalisia suhteita, sitä vähemmän alueella tehdään rikoksia. Sekä Starkin että Sampsonin mukaan rikollisuutta selittää ennen kaikkea anonymiteetti ihmisten välillä, josta seuraa kyvyttömyys ylläpitää järjestystä julkisilla paikoilla yhteisön toimesta. Tämä ilmenee yksilötasolla erityisesti välinpitämättömyytenä puuttua havaittuihin epäkohtiin. (Tolvanen 2005, 31.)

## 2 Tutkimuksen toteutus

### 2.1 Aineiston kuvaus

Tutkimusaineisto koostuu erilaisista Suomen kuntia koskevista tilastotiedoista. Aineisto on kerätty Internetin tietokannoista, joihin on vapaa pääsy. Käytetyt tietokannat, Altika, Sotkanet ja Statfin, ovat varsin kattavat, joten puuttuvaa tietoa on vähän. Muuttujien valinnassa käytettiin pohjana Heiskanen samaa aihetta käsittelevää tutkimusta (Heiskanen et al. 2003). Osaa muuttujista ei kuitenkaan ollut saatavilla vuodelta 2007 tai ne olisi ollut hankittava muuta kautta, joten muuttujalista täydennettiin muuttujilla, joilla saattaisi olla vaikutusta rikollisuuteen. Muuttujat selityksineen on esitelty liitteessä A.

Tutkimuksessa tarkastellaan kuntien rikostasoa kahden muuttujan avulla. Nämä ovat

- 1) väkivaltarikokset, johon sisältyvät murhat ja tapot sekä henkeen ja terveyteen kohdistuneet rikokset.
- 2) omaisuusrikokset, joihin kuuluvat varkaudet, ryöstöt, petokset ja maksuvälinepetokset.

Suurin osa tiedoista on vuodelta 2007, mutta muutamat muuttujat ovat vuodelta 2006, koska vuoden 2007 tietoja ei ollut saatavilla. Rikollisuutta ja muutto-liikettä koskevissa muuttujissa on lisäksi käytetty useamman vuoden keskiarvoja luotettavuuden varmistamiseksi. Muuttujat on useimmiten suhteutettu väestöön tai muuhun perusjoukkoon esimerkiksi *lasten osuus kunnan väestöstä, työttömät työvoimasta* tai *alkoholin myynti litroina asukasta kohden*.

Kuntien määrä pieneni vuoden 2009 alussa 415:sta 348:aan. Kaikki tiedot onkin päivitetty vuoden 2009 kuntaluokituksen mukaiseksi. Kunnista 332 sijaitsee Manner-Suomessa ja 16 Ahvenanmaalla. Kaksikielisiä kuntia on 34 ja ruotsinkielisiä 19. Kuntien keskimääräinen asukasluku on 15 231 ja mediaani 5 763. Tämä kertoo siitä, että Suomessa on huomattavasti enemmän asukasluvultaan pieniä kuin suuria kuntia. Yli 100 000 asukkaan kaupunkeja on vain seitsemän, kun taas alle 2 000 asukkaan kuntia on peräti 48.

Pienet kunnat aiheuttivatkin osaltaan ongelmia tutkimuksessa. Osa muuttujista ei saanut niissä lainkaan arvoja tai tapauksia ei ilmoiteta tietosuojasyistä. Tämä johtuu siitä, ettei Tilastokeskus julkaise tietoa, jos kunnassa ilmenee alle neljä kyseiseen muuttujaan liittyvää tapausta (Heiskanen et al. 2003). Tästä syystä kaikkia haluttuja muuttujia ei voitu ottaa mukaan tutkimukseen. Muutamissa muuttujissa puuttuvia tietoja on korvattu vuoden 2006 tai 2008 tiedoilla, jos ne ovat olleet saatavilla. Korvattujen havaintojen lukumäärä on kuitenkin vain 20.



Pienissä, muutaman sadan asukkaan kunnissa, tilastoituja rikoksia saattaa olla erittäin vähän, mikä vaikeuttaa luotettavan ennusteen tekemistä. Alle tuhannen asukkaan kunnat onkin jätetty pois mallintamisesta. Näin menetettiin 14 kuntaa, joista yhdeksän sijaitsee Ahvenanmaalla. Ahvenanmaan tapauksessa tilastot eivät myöskään ole täysin luotettavia, koska turistien aiheuttama rikollisuus on poikkeuksellisen suurta.

## 2.2 Kuntien ryhmittely

Kunnat ovat keskenään hyvin erilaisia, mikä hankaloittaa ennusteen tekoa. Tästä syystä kunnat on ryhmitelty niiden koon mukaan, jolloin ennusteet voidaan tehdä ryhmien sisällä. Ryhmittelyyn on käytetty Tilastokeskuksen kehittämää kaupunkimaisuuden kuvaajaa, joka perustuu väestölaskentojen yhteydessä kerran viidessä vuodessa tehtävään taajamarajaukseen, ja siitä saatuun tietoon taajamaväestöstä. Kunnat jaetaan taajamaväestön osuuden ja suurimman taajaman väkiluvun perusteella kaupunkimaisiin, taajaan asuttuihin ja maaseutumaisiin kuntiin. Tutkimuksessa on käytetty vuoden 2009 ryhmittelyä. Liitteessä B kunnat on ryhmitelty luokituksen mukaisesti.

**Kaupunkimaisia kuntia** ovat kunnat, joiden väestöstä vähintään 90 % asuu taajamissa tai suurimman taajaman väkiluku on vähintään 15 000. Tällaisia kuntia oli aineistossa 63. Asukasluvultaan suurin kaupunkimainen kunta on Helsinki (568 000 asukasta) ja pienin Kaskinen (1 500 asukasta). Ryhmän kuntien keskikoko on 30 000 asukasta. Luokittelun jälkeenkin kuntaerot ovat suuria, mutta tilanne on huomattavasti parempi kuin ilman luokittelua.

Toisin kuin voisi odottaa, väestön määrä korreloi *kaupunkimaisten kuntien* ryhmässä heikosti väkivallan yleisyyden kanssa ( $r = 0,19$ ). Sen sijaan omaisuusrikollisuuden kanssa sillä on selkeä positiivinen korrelaatio ( $r = 0,47$ ). Rikostyyppien keskinäinen korrelaatio ( $r = 0,69$ ) on yllättävänkin korkea. Heiskasen tutkimuksessa vuoden 2001 vastaava korrelaatio ( $r = 0,42$ ) oli huomattavasti matalampi. Vahva korrelaatio kertoo erityyppisten rikosten kasautumisesta, joten olisi sikiin mielenkiintoista selvittää, onko tilanne todella muuttunut näin paljon muutamassa vuodessa.

**Taajaan asuttuja kuntia** ovat kunnat, joiden väestöstä vähintään 60 %, mutta alle 90 %, asuu taajamissa ja suurimman taajaman väkiluku on vähintään 4 000 mutta alle 15 000. Tällaisia kuntia oli aineistossa 66. Ryhmän kuntien asukasluvut vaihtelivat välillä 6 000–29 000 ja keskikoko oli 11 000 asukasta.

Väestön määrä korreloi hieman paremmin rikollisuuden kanssa *taajaan asutuissa kunnissa*. Omaisuusrikosten kanssa korrelaatio ( $r = 0,56$ ) on jälleen korkeampi kuin väkivallan kanssa ( $r = 0,27$ ). Rikostyyppien välinen korrelaatio ( $r = 0,62$ ) sen sijaan on pienempi kuin kaupunkimaisten kuntien ryhmässä. Heiskasen tutkimuksessa rikostyyppit eivät kuitenkaan korreloineet lainkaan, joten ero on jälleen merkittävä.

**Maaseutumaisia kuntia** ovat kunnat, joiden väestöstä alle 60 % asuu taajamissa ja suurimman taajaman väkiluku on alle 15 000, sekä kunnat, joiden väestöstä vähintään 60 %, mutta alle 90 %, asuu taajamissa ja suurimman taajaman väkiluku on alle 4 000. Tällaisia kuntia oli aineistossa 219. Ryhmän kuntien asukasluvut vaihtelivat välillä 116–12 000 ja keskipöytä oli 3 500 asukasta.

*Maaseutumaisissa kunnissa* väestön määrä korreloi huonosti rikollisuuden kanssa. Omaisuusrikosten kanssa korrelaatio ( $r = 0,22$ ) on hieman korkeampi kuin väkivallan kanssa ( $r = 0,17$ ). Rikostyyppien välinen korrelaatio ( $r = 0,44$ ) sen sijaan on edelleen selkeä. Heiskanen tutkimuksessa rikostyyppien välinen korrelaatio ( $r = 0,15$ ) oli kuitenkin heikko, joten ero on jälleen merkittävä.

**Taulukko 2.1.** Poliisin tietoon tullut rikollisuus kuntaryhmittäin 10 000 asukasta kohden. Rikosmäärät ovat vuosien 2005 – 2007 keskiarvoja.

	Väkivaltarikokset	Omaisuusrikokset
Kaupunkimaiset kunnat	67	465
Taajaan asutut kunnat	53	282
Maaseutumaiset kunnat	42	206

Taulukossa 2.1 on esitetty rikollisuus kuntaryhmittäin. Taulukosta nähdään, että *kaupunkimaisissa kunnissa* rikosten määrä on suurin ja *maaseutumaisissa kunnissa* pienin. Rikollisuus näyttäisi siis lisääntyvän taajamien ja kuntakoon kasvaessa. Tämä on toki luonnollista, koska tällöin myös ihmisten väliset kontaktit ja rikosmahdollisuudet lisääntyvät. Myös sosiaalisen kontrollin on todettu heikenevän kaupungistumisen myötä. *Maaseutumaisissa kunnissa* ihmiset tuntevat toisensa paremmin, mikä oletettavasti vähentää rikollisuutta.

## 2.3 Selittävät muuttujat

Aineistossa on useita *väestörakenteeseen* liittyviä muuttujia, joista *väkiluku* ja *väestöntiheys* ovat keskeisimpiä kunnan rakennetta kuvaavia muuttujia. Rikollisuuden kannalta olennaisia ovat myös miesten, ja ennen kaikkea nuorten miesten, suhteelliset osuudet. Tutkimusten mukaan näissä ryhmissä taipumus rikollisuuteen on huomattavasti muita väestöryhmiä suurempaa (Tolvanen, 2005). Myös muita ikäryhmämuuttujia on käytetty analyysissä.

*Sosiaalisista rakennetekijöistä* mukana ovat muun muassa *avioerojen määrä naimisissa olevia kohti* ja *yksinhuoltajien osuus lapsiperheistä*. Myös koulutustasoa on mitattu kahdella ja muuttoliikettä kolmella muuttujalla. Heiskanen tutkimuksessa ulkomaalaisten määrää kuvaavaa muuttujaa ei otettu mukaan, koska sitä ei pidetty kuntatasolla merkittävänä väestörakennetta erilaistavana tekijänä. Maahanmuutto on kuitenkin kasvanut vuosi vuodelta, ja täten myös vierasta

äidinkieltä puhuvien osuus kasvaa; vuonna 2001 vierasta kieltä äidinkielenään puhui 21 ihmistä tuhannesta, kun sama luku vuonna 2007 oli jo 33. Nouseva trendi voidaan nähdä Sotkanetin tilastotietokannasta.

*Sosiaalista ympäristöä* kuvaavia muuttujia ovat ennen kaikkea *vuokra-asuntojen osuus asunnoista, äänestysaktiivisuus ja toimeentulotukimenot väestöön suhteutettuna. Alkoholin myynti, päihtyneiden säilöönnotot ja huumausainerikosten määrä* kertovat päihtyneiden osuudesta rikollisuuteen. Lukuisat tutkimukset puoltavat etenkin alkoholin osuutta väkivaltarikoksiin yksikäsitteisesti. Voidaan todeta, että ainakin kahdessa kolmesta väkivaltatilanteesta tekijä, uhri tai molemmat ovat päihtyneitä (Siren & Honkatukia 2005; Kivivuori & Salmi 2005).

*Taloudellisista* muuttujista keskeisiä ovat ainakin *kunnan yleinen pienituloisuusaste* ja asukkaiden *valtionveronalaiset tulot*. Kuntataloutta kuvaavia mittareita ovat muun muassa *sosiaali- ja päihdehuollon kustannukset. Moottorijouneuvojen ja kesämökkien suhteelliset osuudet* kertovat sekä asukkaiden taloustilanteesta että kuntarakenteen luonteesta. Toisin kuin Heiskasen tutkimuksessa, kesämökkien lukumäärä on suhteutettu väkilukuun eikä pinta-alaan.

*Työssäkäynnin* ja etenkin työttömyyden yhteys rikollisuuteen on ilmeinen, ja siksi siihen liittyviä muuttujia on mukana paljon. Näitä ovat muun muassa *työttömät työvoimasta, nuorisotyöttömät, kunnassa olevien työpaikkojen määrä, pendelöinti* ja työvoiman osuudet eri toimialoilla. Työssäkäyntiä sivuavia muuttujia ovat myös *sairastavuusindeksi* ja *demografinen huoltosuhde*, jolla tarkoitetaan ei-työikäisten osuutta työikäisiin nähden. Toisin sanoen mitä enemmän kunnassa on lapsia ja eläkeikäisiä, sitä korkeampi on huoltosuhde.

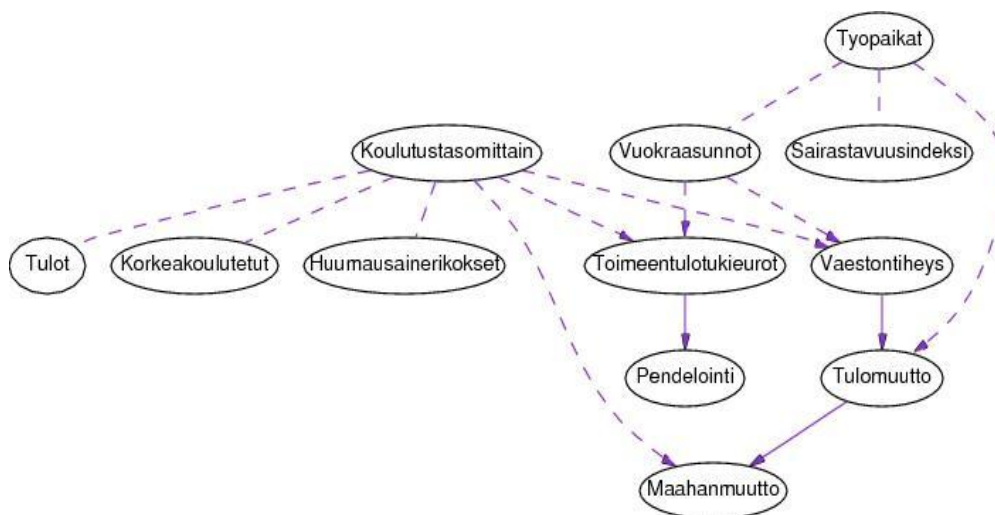
## 2.4 Muuttujien välisten rakenteiden tutkiminen B-Course –ohjelmalla

Muuttujien välisiä rakenteita ja kausaalisuhteita on pyritty mallintamaan B-Course –ohjelmalla, joka on Helsingin Yliopiston ja Teknillisen korkeakoulun yhteisen tietotekniikan tutkimuslaitoksen tutkijaryhmän (CoSCo) kehittämä web-pohjainen datan analysointityökalu. Sen toiminta perustuu bayesiläiseen mallintamiseen, ja ohjelman avulla pyritään löytämään muuttujien välisiä yhteyksiä. Ohjelmaan, ja sen käyttämiin menetelmiin, voi paremmin tutustua kehittäjäryhmän julkaiseman artikkelin avulla (Myllymäki, Silander, Tirri & Uronen 2002).

Ohjelman vahvuuksiin kuuluu sen kyky etsiä myös muita kuin lineaarisia yhteyksiä muuttujien välillä. Heikkoutena voidaan pitää 15 minuutin aikarajaa mallien etsimisessä. Laajan aineiston analysointiin aika on varsin vähäinen. Ohjelma ehtii käydä läpi noin 20 000–100 000 mallia, kun erilaisia malleja on jo seitsemän muuttujan tapauksessa yli kaksi miljoonaa. Kahdenkymmenen muuttujan tapauksessa erilaisia malleja on peräti  $1,6 * 10^{57}$ .

Tuloksia voidaan kuitenkin pitää suuntaa antavina, ja useissa malleissa samat muuttujat liittyvätkin toisiinsa. Alla on esitetty kaksi mallia, jotka ovat molemmat ei-naiiveja kausaalimalleja. Tämä tarkoittaa, että malleissa on otettu huomioon myös ulkopuoliset syyt, jotka saattavat vaikuttaa muuttujien suhteisiin. Naiiveissa malleissa sen sijaan kaikki riippuvuussuhteet on yleensä esitetty suorina kausaalisuhteina. Ohjelman rajoitusten vuoksi malleissa eivät suinkaan ole mukana kaikki muuttujat, vaan niihin on valittu 15–25 muuttujaa, joiden välisiä suhteita haluttiin selvittää.

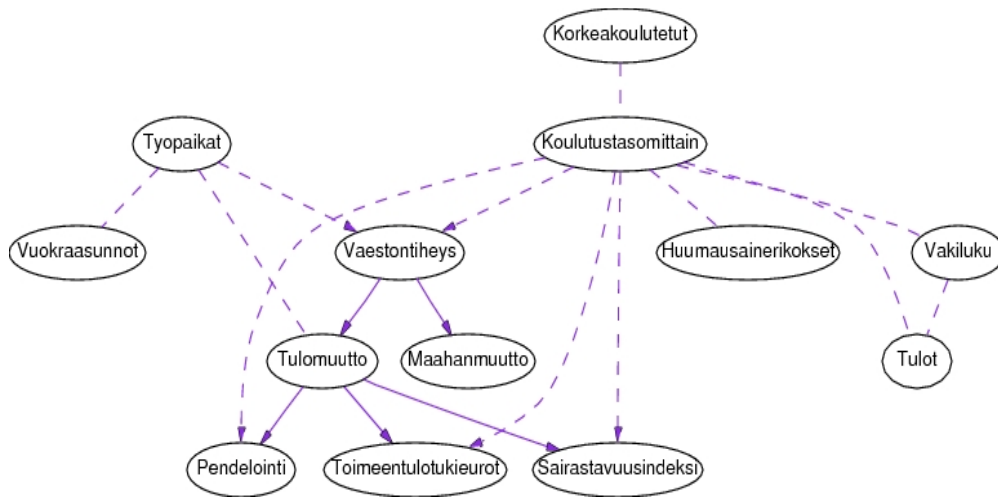
*Väkivalta-* ja *omaisuusrikokset* olivat mukana kaikissa malleissa, mutta riippuvuussuhteita muiden muuttujien kanssa ei ilmennyt. Useista malleista tuli varsin ympäröityjä, mutta esiteltävissä malleissa on mielenkiintoisia riippuvuussuhteita. Mallien paremmuuden vertaileminen ei ole yksiselitteistä, eikä olekaan mitään takeita, että esitellyt mallit olisivat parhaat mahdolliset. On jopa todennäköistä, että on olemassa lukuisia parempia malleja. Tämä ei silti tarkoita, että mallit olisivat huonoja tai yhdentekeviä. Ratkaisujen ja johtopäätösten suhteen kannattaa kuitenkin olla kriittinen.



**Kuvio 2.1.** Aineiston muuttujien ensimmäinen kausaalimalli.

Kuviosta 2.1 nähdään, että malliin valikoitui 12 muuttujaa. Eniten riippuvuussuhteita vaikuttaisi olevan koulutustasomittaimella. Muuttujien välisiä suhteita voidaan tulkita seuraavasti: suora nuoli A:sta B:hen tarkoittaa, että A:lla on suora kausaalisuhte B:hen. Näin ollen esimerkiksi *väestöntiheys* selittäisi *tulomuuttoa*. Katkoviivainen nuoli A:sta B:hen tarkoittaa, että A saattaa olla syy B:hen, mutta voi myös olla, että on olemassa jokin latenti syy, joka aiheuttaa A:n ja B:n välisen suhteen. Esimerkiksi *vuokra-asuntojen* ja *toimeentulotuen* välillä näyttäisi olevan tällainen yhteys. Pelkkä katkoviiva ilman nuolta tarkoittaa muuttujien välisen

yhteyden vaikeasti tulkittavaa luonnetta. Tällöin kumpi vain voi olla syy toiseen tai voi olla, että on olemassa jokin latentti syy, joka aiheuttaa A:n ja B:n välisen suhteen. Tällaisia suhteita mallissa on eniten, esimerkiksi *koulutustasomittaimen* ja *huumausainerikosten* välillä. Suorista nuolista nähdään, että mallin mukaan *maahanmuutto* on seurausta *tulomuutosta*, joka taas on seurausta *väestötiheydestä*. Vastaavasti *pendelöinti* on riippuvainen *toimeentulotukieuroista*. Muiden muuttujien väliset riippuvuudet eivät mallin mukaan ole yhtä selkeitä.



**Kuvio 2.2.** Aineiston muuttujien toinen kausaalimalli.

Vaikka kuvio 2.2 on hieman sekavan oloinen, ei sen tulkinta ole juurikaan hankalampaa kuin Kuvion 2.1. *Väestötiheydellä* on suorat vaikutukset *tulo-* ja *maahanmuuttoon*. *Pendelöinti*, *toimeentulotukieurot* ja *sairastavuusindeksi* taas ovat riippuvaisia *tulomuutosta*. Kuten edellisessäkin mallissa, *kunnassa olevilla työpaikkojen määrällä* on yhteys *vuokra-asuntojen osuuden* ja *tulomuuton* kanssa. *Koulutustasomittain* taas on molemmissa malleissa riippuvuussuhteessa *toimeentulotukieurojen*, *väestötiheyden*, *huumausainerikosten*, *korkeakoulutettujen* ja *tulojen* kanssa.

Yleisesti voidaankin todeta *koulutuksen* ja korkeiden *tulojen* olevan suurille kaupungeille ominaisia piirteitä. *Koulutustasomittaimen* ja *huumausainerikosten* välinen yhteys on mallin mukaan vaikeasti tulkittavissa, mikä voi johtua monesta syystä. Suurissa kaupungeissa *huumausainerikokset* ovat yleisiä, mutta toisaalta korkeasti koulutettujen keskuudessa *huumausainerikokset* ovat varmasti vähäisempiä. Vastaavasti esimerkiksi *työpaikkojen* ja *vuokra-asuntojen* välinen yhteys on vaikeasti tulkittavissa.

## 3 Väkivaltarikollisuus kunnissa

Väkivaltarikokset ovat lisääntyneet yleisesti 2000-luvulla; koko maan keskiarvo nousi 10 000 asukasta kohden 59:sta 71:een vuosina 2001–2007. Nouseva trendi ilmenee Sotkanetin tilastotietokannasta. Piilorikollisuustutkimusten perusteella tiedetään, että suuri osa varsinkin selvin päin tehdyistä väkivallanteoista jää ilmoittamatta poliisille (Kivivuori & Honkatukia 2005). On kuitenkin arveltu, että lievempien väkivaltarikosten kohdalla ilmoitus- ja ilmitulokynnys on viime vuosina madaltunut.

Vuositasolla noin 80 % henkirikoksista ja 70 % pahoinpitelyistä tehdään alkoholin vaikutuksen alaisina (Oikeuspoliittinen tutkimuslaitos 2006). Erittäin suuri osuus väkivaltarikoksista tehdään siis päihteiden vaikutusten alaisena, eikä alkoholin osuutta väkivaltarikoksiin tule aliarvioida. Henkirikosten ja ryöstöjen määrät ovat laskeneet vuoden 2001 tasosta, mutta pahoinpitelyt ovat sen sijaan lisääntyneet vuosi vuodelta.

### 3.1 Kaupunkimaiset kunnat

*Kaupunkimaisten kuntien* väkivaltarikostasoa voidaan tarkastella Taulukosta 3.1. Ääripäiden erot ovat suuria, mikä johtuu osittain Maarianhaminan ja Hangon poikkeuksellisen korkeista rikosluvuista. Mielenkiintoinen havainto on myös, että Helsinkiä ja Vaasaa lukuun ottamatta taulukon kunnat ovat väkiluvultaan suhteellisen pieniä. Keskimääräinen väkivaltarikostasoa kaupunkimaisissa kunnissa oli 70 tapausta 10 000 asukasta kohti, kun se vuonna 2001 oli 64 (Heiskanen et al. 2003).

**Taulukko 3.1.** Kymmenen tilastoidulta väkivaltatasoltaan korkeinta ja matalinta kaupunkimaista kuntaa ja kuntien väkiluku.

Korkein tilastoitu väkivaltataso			matalin tilastoitu väkivaltataso		
	Väkivalta- rikokset/ 10 000 asukasta	Väkiluku 2007		Väkivalta- rikokset/ 10 000 asukasta	Väkiluku 2007
Maarianhamina	150	10 902	Hollola	22	21 520
Hanko	129	9 708	Kauniainen	24	8 511
Pietarsaari	106	19 569	Oulunsalo	25	9 431
Helsinki	104	568 531	Pirkkala	29	15 788
Iisalmi	102	22 298	Kempele	31	15 218
Kemi	99	22 669	Kaskinen	31	1 485
Vaasa	98	57 998	Lempäälä	32	19 271
Harjavalta	90	7 646	Nurmijärvi	32	38 633
Forssa	89	17 835	Kiiminki	33	12 448
Kajaani	88	38 089	Kaarina	36	29 928

Yksittäisistä muuttujista vahvimmin väkivaltarikosten kanssa korreloivat *pendelöinti* ( $r = 0,76$ ), *kunnassa olevien työpaikkojen määrä* ( $r = 0,72$ ) ja *omistusasunto kaikista asunnoista* ( $r = -0,72$ ). Myös *vuokra-asunto kaikista asunnoista*, *päihtyneiden säilöönnotot*, *yksinhuoltajaperheiden määrä lapsiperheistä* ja *alkoholijuomien myynti asukasta kohti* korreloivat positiivisesti ( $> 0,50$ ) väkivaltarikollisuuden kanssa.

Kaupunkimaisten kuntien väkivaltarikollisuutta selittävää mallia etsittiin vaiheittaisen regressioanalyysin avulla. Menetelmänä käytettiin eliminoivaa muuttujien valintaa (backward elimination), joka on kattavasti esitetty teoksessa *"Applied regression analysis"* (Draper & Smith 1998). Eliminoiva menetelmä valikoi selittävät muuttujat alkuaan 23 muuttujan ryhmästä. Muuttujat löytyvät liitteestä C. Menetelmä poistaa yksi kerrallaan muuttujat, jotka kulloinkin vähiten heikentävät selitystasetta, kunnes muuttujan poisto ei enää paranna mallia. Muuttujan poiston perusteena käytettiin p-testisuureen arvoa 0,10.

Selittäviksi muuttujiksi valittiin ennen kaikkea muuttujia, joilla epäiltiin olevan aitoa korrelaatiota väkivallan kanssa. Hieman korkeampiakin selitystasetta olisi saatu ottamalla enemmän muuttujia mukaan analyysiin. Muuttujien rajaaminen oli kuitenkin välttämätöntä myös muuttujien multikolinearisuuden vuoksi. Maarianhamina pudotettiin analyysistä poikkeavana havaintona. Perusteluna käytettiin *Cookin etäisyyttä* (Cook's distance), minkä toimintaperiaate on esitetty teoksessa *"Detection of Influential Observations in Linear Regression"* (Cook 1977; Metsämuuronen 2003).

**Taulukko 3.2.** Kaupunkimaisten kuntien regressiomalli.

Havaintoja	Selitystaste	Korjattu selitystaste	Ennusteen keskivirhe	F-suure	P-arvo
62	0,77	0,74	11,93	25,5	0,000

	Kerroin	Beta	t-testi	p-arvo
Vakio	26,813		1,47	0,146
Alkoholijuomien myynti asukasta kohti	1,553	0,14	1,77	0,083
Korkea-asteen koulutuksen saaneet, %	-2,235	-0,58	-4,66	0,000
Kunnan yleinen pienituloisuusaste	-1,824	-0,32	-2,70	0,009
Päihtyneiden säilöönnotot / 1 000 asukasta	0,925	0,40	3,74	0,000
Maahanmuuttajat / 1 000 asukasta	2,074	0,18	1,81	0,076
Pendelöinti	0,372	0,37	3,15	0,003
Palvelualan työpaikkojen osuus, %	0,742	0,33	2,83	0,007

Taulukosta 3.2 nähdään, että seitsemän selittävän muuttujan avulla mallin selitysasteeksi saadaan 74 %, mitä voidaan pitää hyvänä. Malliin valikoituivat muuttujat *alkoholijuomien myynti asukasta kohti, korkea-asteen koulutuksen saaneet, kunnan yleinen pienituloisuusaste, päihtyneiden säilöönnotot, maahanmuuttajien osuus, pendelöinti ja palvelualan työpaikkojen osuus*.

Regressiokertoimien suuruuksiin ei suuremmin kannata keskittyä vaan olennaisia ovat kertoimien etumerkit. *Alkoholijuomien myynti, päihtyneiden säilöönnotot, maahanmuuttajien osuus, pendelöinti ja palvelualan työpaikkojen osuus* näyttäisivät lisäävän väkivaltarikollisuutta. *Alkoholin myynnin ja päihtyneiden säilöönnottojen* positiivinen kerroin kertoo, että alkoholilla on usein osuutta väkivaltarikoksiin. Tähän viittaa myös *päihtyneiden säilöönnottojen* korkea beta-kerroin. *Pendelöinnin* seurauksena kunnan alueella olevien määrä kasvaa, joten on helppo kuvitella rikostenkin lisääntyvän. Kasvava *maahanmuutto ja palvelualan työpaikkojen* suuri osuus taas ovat suurkaupunkimaisia piirteitä, ja ne kuvaavat ennen kaikkea kunnan rakenteellisia ulottuvuuksia.

*Korkea-asteen koulutus ja kunnan pienituloisuusaste* sen sijaan näyttäisivät vähentävän väkivaltarikollisuutta. Betakertoimen perusteella koulutusmuuttuja on myös mallin tärkein selittäjä. Korkea koulutusaste liitetään usein suuriin kaupunkeihin, mutta viime vuosina keskusta-asumisen kallistuessa ja väestötiheyden kasvaessa useat korkeakoulutetut perheet ovat muuttaneet myös lähikuntiin. Lisääntyvässä määrin onkin muodostumassa niin sanottuja hyväosaisten lähiöitä, joissa rikollisuus on matalampaa. *Kunnan pienituloisuusasteen* väkivaltaa vähentävä vaikutus sen sijaan on hankalammin selitettävissä.

Taulukkoon 3.3 on ryhmitelty *kaupunkimaiset kunnat* sen mukaan onko regressiomallilla estimoitu tulos korkeampi vai matalampi kuin tilastoitu väkivaltataso kunnassa. Tulokset on luokiteltu viiteen ryhmään. Ennustettua korkeamman väkivaltatason kuntiin kuuluivat ne kunnat, joissa tilastoitu väkivaltataso oli korkeampi kuin ennustettu. Jakamalla tilastoidun ja ennustetun väkivallan erotus ennustetulla väkivallan määrällä ja kertomalla osamäärä sadalla saadaan poikkeamaprosentti. Odotetun väkivaltatason kunnissa poikkeama on alle 10 %. Vastaavasti 10–25 % poikkeamat luokitellaan melko korkean ja melko matalan väkivaltatason kuntiin. Korkean ja matalan väkivaltatason kunnissa poikkeamat ovat yli 25 %. Todettakoon, että luokkarajat ovat sopimuksenvaraisia, joten osa kunnista on rajatapauksia. Luokkien sisällä eroja ei ole huomioitu vaan kunnat ovat aakkosjärjestyksessä.

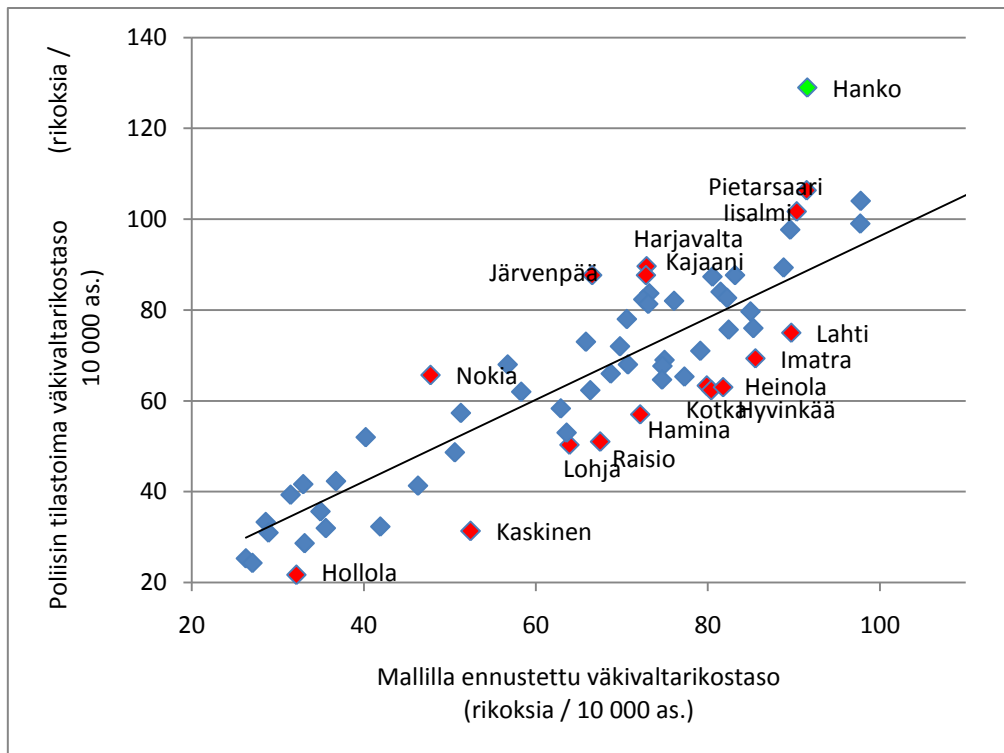


**Taulukko 3.3.** Kaupunkimaisten kuntien luokittelu väkivallan tason mukaan. Kunnat aakkosjärjestyksessä ryhmien sisällä.

Korkea yli 25 %	Melko korkea 10 – 24,9 %	Odotettu 9,9 – -9,9 %	Melko matala -10 – -24,9 %	Matala yli -25 %
Hanko	Akaa	Espoo	Hamina	Hollola
Haukipudas	Harjavalta	Forssa	Heinola	Kaskinen
Järvenpää	Iisalmi	Helsinki	Hyvinkää	
Kirkkonummi	Joensuu	Kaarina	Hämeenlinna	
Nokia	Jyväskylä	Kemi	Imatra	
	Kajaani	Kempele	Kauniainen	
	Kangasala	Kerava	Kotka	
	Kiiminki	Kouvola	Lahti	
	Kokkola	Kuopio	Lempäälä	
	Loviisa	Lappeenranta	Lohja	
	Pietarsaari	Oulu	Maarianhamina	
	Raahe	Oulunsalo	Mikkeli	
	Valkeakoski	Pori	Nurmijärvi	
	Ylöjärvi	Porvoo	Pirkkala	
		Rauma	Raisio	
		Riihimäki	Salo	
		Rovaniemi	Tampere	
		Savonlinna	Tuusula	
		Seinäjoki	Vantaa	
		Tornio		
		Turku		
		Vaasa		
		Varkaus		

Kuviossa 3.1 on verrattu *kaupunkimaisten kuntien* mallilla ennustettua ja poliisin tilastoimaa väkivaltatasoa myös graafisesti. Ennuste on laskettu yllä muodostetun mallin (Taulukko 3.2) perusteella. Punaisella merkityt kunnat poikkeavat regressiosuorasta enemmän kuin yhden ja vihreällä enemmän kuin kahden hajonnan verran. Näin ollen regressiosuoran yläpuolella olevissa kunnissa rikollisuus on ennustettua yleisempää ja alapuolella harvinaisempaa. Erityisesti Hangossa väkivaltataso näyttäisi olevan suhteellisen korkea, kun taas esimerkiksi Kaskisilla ja Hollolassa se on varsin alhainen.

Maarianhamina on niin poikkeava havainto, että se jää reilusti kuvion ulkopuolelle. Sen kohdalla ennuste on kuitenkin niin suuri, että vaikka asukaslukuun suhteutettuna siellä esiintyy eniten väkivaltarikoksia, kuuluu se kuitenkin ennusteen mukaan melko matalan rikollisuuden kuntiin. Tulos tukee hyvin päätöstä jättää Maarianhamina ennusteen ulkopuolelle.



**Kuvio 3.1.** Tilastoitu ja mallilla ennustettu väkivaltarikollisuus kaupunkimaisissa kunnissa. Rikoksia väestön 10 000 asukasta kohden.

## 3.2 Taajaan asutut kunnat

**Taulukko 3.4.** Kymmenen tilastoidulta väkivaltatasoltaan korkeinta ja matalinta taajaan asuttua kuntaa ja kuntien väkiluku.

Korkein tilastoitu väkivaltataso

	Väkivalta- rikokset/ 10 000 asukasta	Väkiluku 2007
Karkkila	106	8 996
Kuusamo	94	16 899
Pieksämäki	88	20 476
Laitila	87	8 484
Uusikaupunki	84	15 926
Kemijärvi	77	8 882
Äänekoski	77	20 404
Raasepori	74	28 731
Orivesi	70	9 555
Keuruu	69	10 919

Matalin tilastoitu väkivaltataso

	Väkivalta- rikokset/ 10 000 asukasta	Väkiluku 2007
Mustasaari	19	17 936
Ilmajoki	25	11 635
Masku	26	9 093
Keminmaa	29	8 712
Kontiolahti	30	13 326
Sipoo	31	19 470
Lieto	32	15 607
Hämeenkyrö	32	10 276
Noormarkku	32	6 158
Paimio	32	10 010

Yllä olevasta taulukosta (Taulukko 3.4) nähdään *taajaan asuttujen kuntien* väkivaltarikosten ääripäät. Karkkilassa väkivaltarikoksia tehtiin eniten, kun taas Mustasaarella vähiten. Keskimääräinen väkivaltarikoksien määrä 10 000 asukasta kohti oli 53, kun se vuonna 2001 oli vain 40 (Heiskanen et al. 2003). Tämä kertoo omalta osaltaan väkivaltarikosten lisääntymisestä.

Yksittäisistä muuttujista vahvimmin väkivaltarikosten kanssa korreloivat *vuokra-asunto kaikista asunnoista* ( $r = 0,65$ ) ja *omistusasunto kaikista asunnoista* ( $r = -0,63$ ). Myös *päihtyneiden säilönotot*, *alkoholijuomien myynti asukasta kohti*, *toimeentulotukieurot asukasta kohti*, *omassa kunnassa työssäkäyvien osuus työssäkäyvistä* ja *pendelöinti* korreloivat positiivisesti ( $> 0,50$ ) väkivaltarikollisuuden kanssa.

*Taajaan asutuille kunnille* muodostettiin vastaava regressiomalli kuin *kaupunkimaisille kunnille*. Poikkeavana havaintona aineistosta pudotettiin Karkkila. Taulukosta 3.5 nähdään, että kuuden selittävän muuttujan avulla mallin selitysas- teeksi saatiin 63 %, mitä voidaan pitää hyvänä. Malliin valikoituivat muuttujat *0–15-vuotiaiden osuus*, *alkoholijuomien myynti asukasta kohti*, *päihtyneiden säilönotot*, *vuokra-asunto kaikista asunnoista*, *maassamuuton tulomuutto* ja *kunnan nettomuutto*.

Alkoholin osuus väkivaltaan on merkittävä tekijä myös taajaan asutuissa kun- nissa; *alkoholin myynnin* ja *päihtyneiden säilönottojen* positiiviset kertoimet kertovat alkoholin merkittävästä osuudesta väkivaltarikollisuuteen. Myös korkeat betakertoimet tukevat väitettä. *Vuokra-asuntojen osuus* ja *0–15-vuotiaiden osuus* taas kuvaavat kuntien rakenteellisia tekijöitä. Mallin mukaan myös muut- toliike selittää rikollisuutta; tulomuuton kasvaessa väkivalta näyttäisi vähenevän, kun taas nettomuutto saa positiivisen kertoimen. *Nettomuutto* voi kuitenkin saa- da myös negatiivisia arvoja, joten muuttujan tulkinta ei ole yksiselitteinen.

**Taulukko 3.5.** Taajaan asuttujen kuntien regressiomalli.  
Analyysiin valittiin alun perin 24 muuttujaa.

Havainnot	Selitysas- te	Korjattu selitysas- te	Ennusteen keskivirhe	F-testi	P-arvo
65	0,63	0,58	10,93	15,07	0,000

	Kerroin	Beta	t-testi	p-arvo
Vakio	4,742		0,30	0,769
0-15 -vuotiaat, % väestöstä	1,302	0,25	2,12	0,038
Alkoholijuomien myynti asukasta kohti	1,685	0,25	2,23	0,030
Päihtyneiden säilönotot / 1000 asukasta	0,633	0,40	2,98	0,004
Vuokra-asunto kaikista asunnoista, %	1,126	0,32	2,09	0,041
Maassamuuton tulomuutto (%) / väestö	-4,701	-0,39	-3,10	0,003
Kunnan nettomuutto / 1 000 asukasta	0,439	0,24	1,83	0,073

Taulukkoon 3.6 on ryhmitelty taajaan asutut kunnat sen mukaan, onko yllä muodostetulla regressiomallilla estimoitu tulos korkeampi vai matalampi kuin tilastoitu väkivaltataso kunnassa. Tulokset on luokiteltu viiteen ryhmään kuten kaupunkimaisten kuntien kohdalla. Myös luokitteluasteikko on sama kuin kaupunkimaisten kuntien ryhmässä.

**Taulukko 3.6.** Taajaan asuttujen kuntien luokittelu väkivallan tason mukaan. Kunnat aakkosjärjestyksessä ryhmien sisällä.

Korkea yli 25 %	Melko korkea 10 – 24,9 %	Odotettu 9,9 – -9,9 %	Melko matala -10 – -24,9 %	Matala yli -25 %
Hattula	Haapavesi	Huittinen	Asikkala	Hämeenkyrö
Il	Kauhajoki	Ikaalinen	Eura	Ilmajoki
Karkkila	Muhos	Jämsä	Haapajärvi	Janakkala
Kokemäki	Muurame	Kalajoki	Kannus	Keminmaa
Kuusamo	Mynämäki	Kankaanpää	Kurikka	Länsi-Turunmaa
Laitila	Nakkila	Kauhava	Mänttä-Vilppula	Mustasaari
Loimaa	Raasepori	Kemijärvi	Noormarkku	
Naantali	Uusikaupunki	Keuruu	Nurmes	
Orimattila		Kontiolahti	Paimio	
Orivesi		Kuhmo	Parkano	
		Laihia	Siilinjärvi	
		Lapua	Sipoo	
		Laukaa	Sodankylä	
		Lieksa	Suonenjoki	
		Lieto	Ylivieska	
		Masku		
		Mäntsälä		
		Mäntyharju		
		Nastola		
		Nivala		
		Oulainen		
		Outokumpu		
		Pieksämäki		
		Sastamala		
		Ulvila		
		Vihti		
		Äänekoski		

### 3.3 Maaseutumaiset kunnat

Taulukosta 3.7 nähdään *maaseutumaisten kuntien* väkivaltarikosten ääripäät. Kuntien väliset erot tulevat taulukosta hyvin esiin, mistä kertoo se, että Vaalassa rikoksia tehtiin 15 kertaa enemmän kuin Kruunupyssä tai Luodossa. Keskimääräinen väkivaltarikoksien määrä 10 000 asukasta kohti oli 43. Jos alle tuhannen asukkaan kunnatkin otettaisiin mukaan, niin joissakin kunnissa rikosluvut olisivat vieläkin matalampia. Esimerkiksi Sottungassa ei tehty yhtään väkivaltarikosta. Useissa kunnissa rikoksia on tehty vain yksi tai muutama, mikä hankaloittaa vakaiden mallien muodostamista tällaisille kunnille. *Maaseutumaisten kuntien* ryhmässä mallin selitysaste jääkin huomattavasti pienemmäksi kuin muissa ryhmissä.

Yksittäisistä muuttujista vahvimmin väkivaltarikosten kanssa korreloivat *päihtyneiden säilöönnotot* ( $r = 0,61$ ), *vuokra-asunto kaikista asunnoista* ( $r = 0,47$ ) ja *työttömät miehet työvoimasta* ( $r = 0,45$ ). Myös *nuorisotyöttömyys* ja *alkoholin myynti asukasta kohti* korreloivat positiivisesti ( $> 0,40$ ) rikostason kanssa.

**Taulukko 3.7.** Kymmenen tilastoidulta väkivaltatasoltaan korkeinta ja matalinta maaseutumaista kuntaa ja kuntien väkiluku.

Korkein tilastoitu väkivaltataso			Matalin tilastoitu väkivaltataso		
	Väkivalta- rikokset/ 10 000 asukasta	Väkiluku 2007		Väkivalta- rikokset/ 10 000 asukasta	Väkiluku 2007
Vaala	109	3 561	Kruunupyy	7	6 716
Sotkamo	102	10 716	Luoto	7	4 572
Inari	101	6 954	Soini	8	2 526
Taivalkoski	92	4 621	Pyhäntä	11	1 763
Puumala	91	2 723	Kuhmalahti	12	1 076
Juankoski	90	5 454	Pedersören kunta	13	10 665
Rääkkylä	89	2 735	Kihniö	13	2 280
Kaavi	88	3 490	Pernaja	13	4 002
Kaustinen	85	4 298	Saltvik	14	1 731
Pyhäjärvi	85	6 124	Tervo	15	1 792

*Maaseutumaisille kunnille* muodostettiin vastaava regressiomalli kuin muillekin kuntaryhmille. Poikkeavina havaintoina aineistosta poistettiin neljä kuntaa. Tämän lisäksi 14 alle 1 000 asukkaan kuntaa jätettiin analyysin ulkopuolelle. Taulukosta 3.8 nähdään, että kahdeksan selittävän muuttujan avulla mallin selitysasteeksi saadaan 49 %, mitä voidaan pitää kohtalaisen hyvänä. Malliin valikoituivat muuttujat *0–15-vuotiaiden osuus*, *alkoholijuomien myynti asukasta kohti*, *toimeentulotukimenot*, *lapsiperheiden osuus*, *päihtyneiden säilöönnotot*, *äänestys-*

*prosentti, alkutuotannon työpaikkojen osuus ja huoltosuhde.* Näistä suuri *lapsiperheiden osuus, korkea äänestysprosentti ja huoltosuhde* näyttäisivät vähentävän rikollisuutta.

Suuri *0–15-vuotiaiden osuus* puolestaan lisää *taajaan asuttujen kuntien* väkivaltarikollisuutta. *Päihtyneiden säilönnötojen ja alkoholijuomien myynnin* positiiviset kertoimet kertovat jälleen, että alkoholilla on selvästi osuutta väkivaltarikoksiin. Suuret *toimeentulotukimenot* voidaan puolestaan tulkita huono-osaisuuden mittariksi. Muut muuttujat kertovat ennen kaikkea maaseutumaisesta kuntarakenteesta. Betakertoimien perusteella *0–15-vuotiaiden ja lapsiperheiden osuudet sekä päihtyneiden säilönnöt* ovat mallin tärkeimmät selittäjät.

**Taulukko 3.8.** Maaseutumaisien kuntien regressiomalli.  
Analyysiin valittiin alun perin 25 muuttujaa.

Havainnot	Selitysaste	Korjattu selitysaste	Ennusteen keskivirhe	F-suure	P-arvo
201	0,49	0,46	13,75	21,79	0,000

	Kerroin	Beta	t-testi	p-arvo
Vakio	101,795		3,87	0,000
0-15 -vuotiaat, % väestöstä	2,263	0,48	2,41	0,017
Alkoholijuomien myynti asukasta kohti	0,901	0,28	4,35	0,000
Toimeentulotuki, euroa / asukas	0,085	0,12	1,88	0,061
Lapsiperheet, % perheistä	-1,674	-0,49	-2,43	0,016
Päihtyneiden säilönnöt / 1 000 asukasta	1,523	0,43	6,40	0,000
Vaalit, äänestysprosentti	-0,435	-0,13	-1,99	0,048
Alkutuotannon työpaikkojen osuus	0,361	0,14	2,31	0,022
Huoltosuhde, demografinen	-0,595	-0,19	-2,85	0,005

Taulukkoon 3.9 on ryhmitelty *maaseutumaiset kunnat* sen mukaan onko yllä muodostetulla regressiomallilla estimoitu tulos korkeampi vai matalampi kuin tilastoitu väkivaltataso kunnassa. Tulokset on luokiteltu viiteen ryhmään kuten muidenkin kuntaryhmien kohdalla. Koska *maaseutumaiset kunnat* eroavat muita ryhmiä enemmän toisistaan, myös estimoitu väkivaltataso eroaa aiempaa enemmän tilastoidusta väkivaltatasosta. Taulukon luokitusta onkin muutettu vastaamaan paremmin saatuja tuloksia; odotetun väkivaltatason kunnissa poikkeama on alle 20 % ja korkean sekä matalan väkivaltatason kunnissa poikkeamat ovat yli 50 %. Todettakoon jälleen, että luokkarajat ovat sopimuksenvaraisia, joten osa kunnista on rajatapauksia. Neljästätoista alle tuhannen asukkaan kunnasta peräti kaksitoista poikkesi suuresti ennustetusta rikostasosta, mikä kertoo pienten kuntien rikostasojen ennustamisen vaikeudesta. Kyseiset kunnat on tummennettu taulukkoon.

**Taulukko 3.9.** Maaseutumaisten kuntien luokittelu väkivaltarikostason mukaan. Kunnat ovat aakkosjärjestyksessä ryhmien sisällä.

A) Korkean väkivaltarikostason maaseutumaiset kunnat

Korkea väkivaltarikostaso (yli 50 %)			
<b>Brändö</b>	Kinnula	Nummi-Pusula	Rääkkylä
<b>Geta</b>	Korsnäs	Oravainen	Sotkamo
Halsua	Kärkölä	Perho	<b>Suomenniemi</b>
Kaustinen	<b>Kökar</b>	Pornainen	Taivassalo
Keitele	Köyliö	Puumala	Tervola
Kesälahti	Miehikkälä	Pöytyä	Vaala

B) Melko korkean väkivaltarikostason maaseutumaiset kunnat

Melko korkea väkivaltarikostaso (20 – 49,9 %)				
Alajärvi	Inkoo	Myrskylä	Reisjärvi	Taivalkoski
Askola	Juankoski	Nousiainen	Ristijärvi	Tarvasjoki
<b>Eckerö</b>	Kivijärvi	Pomarkku	Rusko	Urjala
Enonkoski	Koski	Posio	Siikajoki	Vimpeli
Hartola	Kylmäkoski	Puolanka	Siikalatva	Virolahti
Ilomantsi	Kyyjärvi	Pyhäjärvi	Sonkajärvi	Ylitornio
Inari	Liminka	Pälkäne	Sulkava	Ähtäri

C) Melko matalan väkivaltarikostason maaseutumaiset kunnat

Melko matala väkivaltarikostaso (-20 – -49,9 %)				
Alavieska	Jämijärvi	<b>Lumparland</b>	Pyhäjoki	Tuusniemi
Alavus	Karjalohja	Luoto	Rautavaara	Varpaisjärvi
Eurajoki	Karvia	Maalahti	Rautjärvi	Vehmaa
<b>Föglö</b>	Kuhmalahti	Marttila	Ruokolahti	Vesilahti
<b>Hailuoto</b>	Kuhmoinen	Multia	Savukoski	Vähäkyrö
Hankasalmi	Lappajärvi	Närpiö	Sievi	Vöyri-Maksamaa
Heinävesi	Lavia	Oripää	Säkylä	Yli-ii
Hämeenkoski	Lemi	Parikkala	Taipalsaari	Ylämaa
Isojoki	Lemland	Pedersören kunta	Tammela	
Isokyrö	Liljendal	Pello	Teuva	
Jomala	Liperi	Pertunmaa	Tohmajärvi	
Juuka	<b>Luhanka</b>	Pyhtää	Toholampi	

D) Matalan väkivaltarikostason maaseutumaiset kunnat

Matala väkivaltarikostaso (alle -50 %)			
Honkajoki	Kärsämäki	Saltvik	Tervo
Kihniö	Pelkosenniemi	Simo	Toivakka
Kruunupyy	Pernaja	Soini	<b>Vårdö</b>
<b>Kumlunge</b>	Pyhäntä	<b>Sottunga</b>	

### 3.4 Väkivaltarikosmallien vertailu

Väkivaltarikollisuutta selittäviin malleihin valikoitui osittain samoja muuttujia kuin Heiskasen tutkimuksessa. Etenkin *kaupunkimaisten kuntien* väkivaltarikollisuutta selittävässä malleissa on paljon yhtäläisyyksiä. Alla on vertailtu väkivaltarikollisuutta selittäviä ennustemalleja poliisin tilastoimaan rikostasoon. Vertailtaessa kuntien luokituksia, puhutaan nimenomaan ennusteisiin perustuvista luokituksista. Korkeaan rikosluokkaan kuuluvassa kunnassa ei välttämättä tehdä määrällisesti erityisen paljon rikoksia vaan niitä tapahtuu poikkeuksellisen paljon kunnan koko ja rakenne huomioon ottaen.

**Kaupunkimaisten kuntien** väkivaltarikosmalleissa on paljon yhtäläisyyksiä; *palvelualan työpaikkojen osuus, pendelöinti, alkoholin myynti, pienituloisuusaste ja korkea-asteen koulutus* löytyvät myös Heiskasen mallista. Uuteen malliin valikoitui lisäksi muuttajat *päihtyneiden säilöönnotot ja maahanmuuttajat*, jotka eivät kumpikaan olleet mukana Heiskasen mallia rakennettaessa. Kaiken kaikkiaan mallit muistuttavat erittäin paljon toisiaan, mikä vahvistaa käsitystä *kaupunkimaisten kuntien* väkivaltarikollisuuden taustalla olevista rakenteista.

Yksittäisten kuntien rikostasoja vertailtaessa huomattiin, että Järvenpäässä väkivaltarikostasoo on pysynyt poikkeuksellisen korkeana. Heiskasen mallin korkean rikostason kunnista myös Iisalmessa ja Kokkolassa rikostasoo on pysynyt melko korkeana. Salossa ja Maarianhaminassa rikostasoo on puolestaan pudonnut korkeasta melko matalaksi. Maarianhaminan kohdalla malli ei ole erityisen luotettava, koska siellä tehdään selvästi eniten väkivaltarikoksia väestöön suhteutettuna. Tämä johtuu kuitenkin siitä, että Ruotsiin matkaavilla laivoilla tapahtuneet rikokset rekisteröidään useimmiten Maarianhaminaan. Ilman näitä rikoksia Maarianhamina saattaisi hyvinkin kuulua melko matalan rikostason kuntiin.

Uudessa mallissa Hanko, Haukipudas ja Nokia puolestaan ovat nousseet odotetun rikostason kunnista korkeaan. Kirkkonummella rikostasoo oli melko korkea jo Heiskasen mallissa. Mallien antamat tulokset ovatkin hyvin samansuuntaisia, mistä kertoo se, että melko korkean rikostason kunnista peräti puolet sijaitti samassa kategoriassa myös Heiskasen mallissa. Matalan rikostason kunnista Kasinen ja Hollola ovat säilyttäneet asemansa myös uudessa mallissa. Kotkassa, Raisiossa ja Tampereella rikostasoo on pysynyt melko matalana. Harjavalta sen sijaan on muuttunut matalan rikostason kunnasta melko korkean rikostason kunnaksi.

**Taajaan asuttujen kuntien** tapauksessa ainoa täysin sama muuttuja malleissa oli *vuokra-asuntojen osuus*. Alkoholin käyttöä kuvaava muuttuja löytyy myös molemmista malleista. *Tulo- ja nettomuutolla* näyttäisi olevan keskeinen merkitys uudessa mallissa, kun taas Heiskasen mallissa painopiste oli työssäkäynnillä ja *avioliitossa olevien osuudella*. Selitysasteet jäivät kuitenkin molemmissa malleissa alle 65 prosenttiin, mikä kertoo vakuuttavien mallien luonnin vaikeudesta *taajaan asuttujen kuntien* ryhmälle.



*Taajaan asutuista kunnista* lissä, Kokemäellä, Kuusamossa ja Laitilassa rikostaso on noussut melko korkeasta korkeaksi. Eniten nousua on tapahtunut Hattulassa, jossa rikostaso on noussut melko matalasta korkeaksi. Myös Hämeenkyrössä rikostaso on noussut matalasta melko korkeaksi. Sipoon ja Noormarkun rikostasot ovat vastaavasti pudonneet korkeasta melko matalaksi. Myös Jämsän, Nivalan ja Äänekosken kohdalla kehitys on ollut positiivista. Keminmaa ja Mustasaari puolestaan ovat säilyttäneet asemansa matalan rikostason kuntina.

**Maaseutumaisten kuntien** ryhmässä selitysasteet jäivät alle 50 prosenttiin. Mallien vertailussa ilmeni kuitenkin kolme yhteistä muuttujaa: *alkoholin myynti, äänestysprosentti ja toimeentulotukimenot*. Muista selittäjistä *0–15-vuotiaiden ja lapsiperheiden osuudet* esiintyvät uudessa mallissa erittäin vahvasti, kun taas Heiskasen mallissa painopiste oli työssäkäyntiin liittyvissä muuttujissa.

*Maaseutumaisten kuntien* ryhmä on huomattavasti alttiimpi muutoksille, joten malleissakin on enemmän eroja. Myös kohtalaisen alhaiset selitysasteet kertovat, etteivät mallit ole erityisen hyviä tälle ryhmälle. Poikkeuksellisen korkeana rikostaso on pysynyt Getassa, Keiteleellä, Kinnulassa, Kökarissa, Köyliössä, Nummi-Pusulassa, Perhossa, Pornaisissa, Puumalassa, Sotkamossa, Taivassalossa ja Vaalassa. Poikkeuksellisen matalan rikostason kuntina asemansa ovat säilyttäneet ainoastaan Kihniö ja Kruunupyö. Yllättävän moni kunta on kuitenkin muuttanut rikostasoltaan matalasta korkeaksi. Nämä kunnat ovat Kumlinge, Pernaja, Saltvik, Sottunga ja Vårdö. Sottunga onkin hyvä esimerkki mallien ongelmista pienten kuntien tapauksessa; kunnassa ei tapahtunut ainoatakaan väkivalta-rikosta 10 000 asukasta kohti vuosien 2005–2007 keskiarvolla mitattuna.

## 4 Omaisuusrikollisuus kunnissa

Omaisuusrikokset ovat laskeneet yleisesti 2000-luvulla; koko maan keskiarvo las-ki 10 000 asukasta kohden 552:sta 462:een vuosina 2001–2006. Vuodesta 2007 omaisuusrikokset ovat taas alkaneet kasvaa. Kuitenkaan varkaudet ja törkeät varkaudet eivät ole kasvaneet vaan kasvua on erityisesti lievemmissä omaisuusrikoksissa. Ennen kaikkea myymälävarkaudet ovat lisääntyneet. (Oikeuspoliittinen tutkimuslaitos 2008). Omaisuusrikoksille on myös tyypillistä, että samat henkilöt syyllistyvät lukuisiin rikoksiin. Etenkin pienillä paikkakunnilla tämä saattaa aiheuttaa suuriakin vaihteluita eri vuosien rikostilastoissa.

### 4.1 Kaupunkimaiset kunnat

Asukaslukuun suhteutettu omaisuusrikollisuus on laajinta Helsingissä. Myös Turussa ja Forssassa omaisuusrikostasoa on erittäin korkea asukasmäärään nähden. Taulukosta 4.1 huomataan, että omaisuusrikollisuutta ilmenee erityisen paljon suurkaupungeissa. Vastaavasti matalin rikostasoa on suhteellisen pienillä kaupunkikunnilla. Kiiminki erottuu joukosta selvästi matalimmalla rikostasollaan. Siellä omaisuusrikoksia tehdään asukasmäärään nähden yli viisi kertaa vähemmän kuin Helsingissä.

**Taulukko 4.1.** Kymmenen tilastoidulta omaisuusrikostasoltaan korkeinta ja matalinta kaupunkimaista kuntaa ja kuntien väkiluku.

Korkein tilastoitu omaisuusrikostasoa			Matalin tilastoitu omaisuusrikostasoa		
	Omaisuus- rikokset / 10 000 asukasta	Väkiluku 2007		Omaisuus- rikokset / 10 000 asukasta	Väkiluku 2007
Helsinki	866	568 531	Kiiminki	168	12 448
Turku	819	175 286	Oulunsalo	205	9 431
Forssa	782	17 835	Kaskinen	215	1 485
Lahti	742	99 308	Ylöjärvi	228	29 148
Maarianhamina	741	10 902	Nurmijärvi	235	38 633
Vaasa	730	57 998	Kauniainen	277	8 511
Oulu	721	135 055	Lempäälä	280	19 271
Järvenpää	713	37 989	Pirkkala	283	15 788
Vantaa	665	192 522	Haukipudas	300	18 113
Kemi	664	22 669	Nokia	310	30 485

Yksittäisistä muuttujista vahvimmin omaisuusrikosten kanssa korreloivat *yksinhuoltajaperheiden osuus lapsiperheistä* ( $r = 0,79$ ), *omistusasunto kaikista asunnoista* ( $r = -0,74$ ) ja *avioeroja 25–64-vuotiailla* ( $r = 0,65$ ). Erilaisilla perhesuhteilla vaikuttaisi siis olevan yhteyttä korkeaan omaisuusrikostasoon. Myös *vuokra-asuntojen osuus asunnoista*, *kunnassa olevien työpaikkojen määrä* ja *poliisin tietoon tulleet huumausainerikokset* korreloivat erittäin vahvasti ( $> 0,60$ ) omaisuusrikollisuuden kanssa. Vaikka korrelaatiot ovat suuria, ei kannattane tehdä suoria johtopäätöksiä muuttujien kausaliteeteista.

Omaisuusrikollisuutta mallinnettiin myös regressioanalyysin avulla. Helsinki poistettiin analyysistä poikkeavana havaintona. Malliin lähtötilanteessa valitut muuttujat löytyvät liitteestä D. Taulukosta 4.2 nähdään, että yhdeksän selittävän muuttujan avulla mallin selitysasteeksi saadaan 80 %, mitä voidaan pitää hyvänä. Malliin valikoituivat muuttujat *väestötiheys*, *0–15-vuotiaiden osuus*, *alkoholijuomien myynti asukasta kohden*, *avioeroja 25–64-vuotiailla*, *korkea-asteen koulutuksen saaneet*, *huumausainerikokset*, *maahanmuutto*, *pendelöinti* ja *huoltosuhde*. Näistä *koulutusaste* ja *huoltosuhde* vaikuttaisivat vähentävän rikollisuutta.

*Väestötiheys* ja *pendelöinti* kertovat kunnan keskuskaupunkimaisesta luonteesta. Tällaisissa kunnissa, joissa väestötiheys on suuri, rikoksetekomahdollisuudet kasvavat ja sosiaalinen kontrolli laskee. Lisäksi *alkoholi*, *huumausaineet* ja *avioerot* kertovat sosiaalisista ongelmista, jotka lisäävät omaisuusrikollisuutta. Tulokset tukevat myös yleistä käsitystä huumausaineiden omaisuusrikollisuutta lisäävästä vaikutuksesta. *Maahanmuutto* ja *0–15-vuotiaiden* suuri osuus näyttäisivät myös lisäävän rikoksia. Betakertoimista nähdään, että *koulutus* ja *pendelöinti* selittävät mallia parhaiten, kun taas *maahanmuutto* ja *huumausainerikokset* vähiten.

#### Taulukko 4.2. Kaupunkimaisten kuntien regressiomalli.

Analyyysiin valittiin alun perin 26 muuttujaa.

Havaintoja	Selitysaste	Korjattu selitysaste	Ennusteen keskivirhe	F-suure	P-arvo
62	0,80	0,76	75,40	21,95	0,000

	Kerroin	Beta	t-testi	p-arvo
Vakio	287,550		1,15	0,255
Väestötiheys	0,150	0,27	2,91	0,005
0-15 -vuotiaat, % väestöstä	11,660	0,29	2,15	0,036
Alkoholijuomien myynti asukasta kohti	19,483	0,27	3,27	0,002
Avioeroja 25–64-vuotiailla	11,546	0,25	3,00	0,004
Korkea-asteen koulutuksen saaneet	-9,346	-0,36	-3,44	0,001
Huumausainerikokset / 1 000 asukasta	19,135	0,21	2,88	0,006
Maahanmuutto (ka 05-07) / 1 000 asukasta	15,271	0,18	2,08	0,042
Pendelöinti	2,208	0,33	2,49	0,016
Huoltosuhde, demografinen	-10,163	-0,31	-4,07	0,000

Taulukkoon 4.3 on ryhmitelty kaupunkimaiset kunnat sen mukaan, onko yllä muodostetulla regressiomallilla estimoitu tulos korkeampi vai matalampi kuin poliisin tilastoima omaisuusrikostaso. Tulokset on luokiteltu viiteen ryhmään samoin rajoin kuin kaupunkimaisten kuntien väkivaltarikoksissa. Luokkien sisällä eroja ei ole huomioitu vaan kunnat ovat aakkosjärjestyksessä.

**Taulukko 4.1.3.** Kaupunkimaisten kuntien luokittelu omaisuusrikostason mukaan. Kunnat aakkosjärjestyksessä ryhmien sisällä.

Korkea yli 25 %	Melko korkea 10 – 24,9 %	Odotettu 9,9 – -9,9 %	Melko matala -10 – -24,9 %	Matala yli -25 %
Hamina	Forssa	Akaa	Espoo	Harjavalta
Hanko	Imatra	Haukipudas	Joensuu	Maarianhamina
Hollola	Kaarina	Heinola	Kaskinen	Nokia
Oulu	Kajaani	Helsinki	Kauniainen	Nurmijärvi
Oulunsalo	Kotka	Hyvinkää	Lappeenranta	Tornio
Vaasa	Kouvola	Hämeenlinna	Mikkeli	Valkeakoski
	Pietarsaari	Iisalmi	Rovaniemi	
	Pori	Jyväskylä	Salo	
	Tuusula	Järvenpää	Vantaa	
		Kangasala	Varkaus	
		Kemi		
		Kempele		
		Kerava		
		Kiiminki		
		Kirkkonummi		
		Kokkola		
		Kuopio		
		Lahti		
		Lempäälä		
		Lohja		
		Loviisa		
		Pirkkala		
		Porvoo		
		Raahe		
		Raisio		
		Rauma		
		Riihimäki		
		Savonlinna		
		Seinäjoki		
		Tampere		
		Turku		
		Ylöjärvi		

## 4.2 Taajaan asutut kunnat

Taulukosta 4.4 nähdään, että asukaslukuun suhteutettu omaisuusrikollisuus on laajinta Mäntyharjulla, Mäntsälässä ja Naantalissa. Ilmajoki taas erottuu joukosta selvästi matalimmalla omaisuusrikostasollaan. Yksittäisistä muuttujista vahvimmin omaisuusrikosten kanssa korreloi *päihdehuollon nettokustannukset* ( $r = 0,58$ ). Muut yli 0,3 korrelaatioon yltävät muuttujat ovat *vuokra-asunto kaikista asunnoista* ja *yksinhuoltajaperheiden määrä lapsiperheistä*. Omaisuusrikosten osalta korrelaatiot jäävätkin yllättävän pieniksi.

**Taulukko 4.4.** Kymmenen tilastoidulta omaisuusrikostasoltaan korkeinta ja matalinta taajaan asuttua kuntaa ja kuntien väkiluku.

Korkein tilastoitu väkivaltataso			Matalin tilastoitu väkivaltataso		
	Omaisuusrikokset / 10 000 asukasta	Väkiluku 2007		Omaisuusrikokset / 10 000 asukasta	Väkiluku 2007
Mäntyharju	487	6 647	Ilmajoki	149	11 635
Mäntsälä	481	18 980	Mustasaari	178	17 936
Naantali	479	18 117	Kannus	179	5 844
Uusikaupunki	440	15 926	Sodankylä	179	8 982
Laitila	396	8 484	Kurikka	186	14 928
Outokumpu	387	7 688	Hattula	188	9 444
Orimattila	381	14 705	Kuhmo	191	9 943
Oulainen	376	7 998	Kontiolahti	192	13 326
Vihti	375	27 040	Hämeenkyrö	193	10 276
Nastola	374	15 007	Kauhava	193	17 974

Eliminoivan regressioanalyysin avulla etsittiin parasta *taajaan asuttujen kuntien* omaisuusrikollisuutta selittävää mallia. Mäntyharju, Mäntsälä ja Naantali pudotettiin analyysistä poikkeavina havaintoina. Taulukosta 4.5 nähdään, että viiden selittävän muuttujan avulla mallin selitysasteeksi saadaan 50 %. Malliin valikoituivat muuttujat *avioeroja 25–64-vuotiailla, kunnan yleinen pienituloisuusaste, huumausainerikokset, vuokra-asuntojen osuus ja valtionveronalaiset tulot*.

Laajamittainen huumausainerikollisuus lisää tunnetusti omaisuusrikollisuutta, mikä näkyy myös mallista. *Kunnan yleinen pienituloisuusaste* taas kuvaa taloudellisia ongelmia, jotka johtavat omaisuusrikollisuuteen. *Vuokra-asuntojen ja eroneiden suuret osuudet* kertovat osaltaan kaupunkimaisesta elinympäristöstä, jossa omaisuusrikollisuus on laajempaa. Luultavasti tähän viittaa myös *valtionveronalaisten tulojen* kasvun rikollisuutta lisäävä vaikutus, koska muuten se vaikuttaisi erikoiselta. Betakertoimista nähdään, että *valtionveronalaiset tulot* ja *pienituloisuusaste* selittävät parhaiten mallia.

**Taulukko 4.5.** Taajaan asuttujen kuntien regressiomalli.  
Analyysiin valittiin alun perin 27 muuttujaa.

Havaintoja	Selitysaste	Korjattu selitysaste	Ennusteen keskivirhe	F-testi	P-arvo
63	0,50	0,45	49,81	10,49	0,000

	Kerroin	Beta	t-testi	p-arvo
Vakio	-485,561		-2,49	0,016
Avioeroja 25–64-vuotiailla / 1 000 avioliitossa	5,750	0,28	2,69	0,009
Kunnan yleinen pienituloisuusaste	9,861	0,62	2,41	0,020
Huumausainerikokset / 1 000 asukasta	21,063	0,29	2,79	0,007
Vuokra-asunto kaikista asunnoista, %	4,806	0,35	2,79	0,007
Valtionveronalaiset tulot, euroa/asukas 2006	0,019	0,69	2,76	0,008

Taulukkoon 4.6 on ryhmitelty taajaan asutut kunnat sen mukaan, onko regressiomallilla estimoitu tulos korkeampi vai matalampi kuin poliisin tilastoima kunnan omaisuusrikostaso. Taulukossa on käytetty samaa luokitusta kuin *kau-punkimaisten kuntien* tapauksessa.

**Taulukko 4.6.** Taajaan asuttujen kuntien luokittelu omaisuusrikostason mukaan. Kunnat ovat aakkosjärjestyksessä ryhmien sisällä.

Korkea yli 25 %	Melko korkea 10 – 24,9 %	Odotettu 9,9 – -9,9 %	Melko matala -10 – -24,9 %	Matala yli -25 %
li	Asikkala	Eura	Haapajärvi	Hämeenkyrö
Kalajoki	Huittinen	Haapavesi	Hattula	Ilmajoki
Keuruu	Kokemäki	Janakkala	Ikaalinen	Sodankylä
Laitila	Kuusamo	Jämsä	Kankaanpää	
Loimaa	Laihia	Karkkila	Kannus	
Länsi-Turunmaa	Lapua	Kauhajoki	Kemijärvi	
Masku	Nivala	Kauhava	Kontiolahti	
Mäntsälä	Paimio	Keminmaa	Nakkila	
Mäntyharju	Uusikaupunki	Kuhmo	Sipoo	
Naantali	Vihti	Kurikka		
Nastola	Ylivieska	Laukaa		
Orimattila		Liekka		
Oulainen		Lieto		
Raasepori		Muhos		
Sastamala		Mustasaari		
		Muurame		
		Mynämäki		
		Mänttä-Vilppula		
		Noormarkku		
		Nurmes		
		Orivesi		
		Outokumpu		
		Parkano		
		Pieksämäki		
		Siilinjärvi		
		Suonenjoki		
		Ulvila		
		Äänekoski		

### 4.3 Maaseutumaiset kunnat

Taulukosta 4.7 nähdään, että Töysässä, Ristiinassa ja Kaavissa omaisuusrikollisuus on laajinta väestöön suhteutettuna. Kruunupyö sen sijaan erottuu muista maaseutumaisista kunnista selvästi matalimmalla rikossuhteellaan. Yksittäisistä muuttujista vahvimmin omaisuusrikosten kanssa korreloivat *alkoholin myynti asukasta kohden* ( $r = 0,48$ ) ja *yksinhuoltajaperheiden osuus lapsiperheistä* ( $r = 0,40$ ). Myös *päihtyneiden säilöönnotot*, *toimeentulotukimenot* ja *kesämökkien lukumäärä* korreloivat positiivisesti ( $> 0,35$ ) omaisuusrikollisuuden kanssa.

**Taulukko 4.7.** Kymmenen tilastoidulta omaisuusrikostasoltaan korkeinta ja matalinta maaseutumaista kuntaa ja kuntien väkiluku.

Korkein tilastoitu väkivaltataso			Matalin tilastoitu väkivaltataso		
	Omaisuus- rikokset / 10 000 asukasta	Väkiluku 2007		Omaisuus- rikokset / 10 000 asukasta	Väkiluku 2007
Töysä	453	3 258	Kruunupyö	71	6 716
Ristiina	425	4 984	Karvia	78	2 776
Kaavi	423	3 490	Toholampi	80	3 590
Iitti	398	7 208	Pedersören kunta	81	10 665
Taivassalo	383	1 730	Reisjärvi	81	3 039
Pertunmaa	367	2 019	Soini	83	2 526
Pyhtää	357	5 140	Kihniö	95	2 280
Sotkamo	356	10 716	Teuva	99	6 158
Kittilä	353	5 967	Hammarland	100	1 409
Juva	350	7 213	Luoto	100	4 572

Omaisuusrikollisuutta mallinnettiin jälleen myös regressioanalyysin avulla. Viisi korkeimman rikollisuustason omaavaa kuntaa pudotettiin analyysistä poikkeavina havaintoina. Nämä kunnat olivat Töysä, Ristiina, Kaavi, Iitti ja Taivassalo. Mallintamisen ulkopuolelle jätettiin myös alle tuhannen asukkaan kunnat, joita oli kaikkiaan 14. Taulukosta 4.8 nähdään, että seitsemän selittävän muuttujan avulla mallin selitysasteeksi saadaan 48 %. Malliin valikoituivat muuttujat *alkoholijuomien myynti asukasta kohti*, *toimeentulotukimenot*, *yksinhuoltajaperheiden osuus lapsiperheistä*, *päihtyneiden säilöönnotot*, *maassamuuton tulomuutto*, *valti-onveronalaiset tulot asukasta kohti* ja *kesämökkien lukumäärä*.



Kaikki selittäjät saavat positiivisen kertoimen. *Päihtyneiden säilöönottojen ja alkoholijuomien myynnin* positiiviset kertoimet kertovat jälleen, että alkoholilla on usein osuutta rikollisuuteen. *Toimeentulotukimenojen ja yksinhuoltajaperheiden* suuret osuudet taas kertovat osaltaan taloudellisesta ja sosiaalisesta syrjäytymistä. *Kesämökkien lukumäärä* sen sijaan kertoo lomailijoiden suuresta määrästä, mikä lisää omaisuusrikollisuutta. Toisaalta myös mökkimurrot kasvattavat omaisuusrikostilastoja.

**Taulukko 4.8.** Maaseutumaisten kuntien regressiomalli.  
Analyyssissa oli alun perin mukana 26 muuttujaa.

Havainnot	Selitysaste	Korjattu selitysaste	Ennusteen keskivirhe	F-suure	P-arvo
200	0,48	0,46	50,62	23,99	0,000

	Kerroin	Beta	t-testi	p-arvo
Vakio	-125,308		-2,74	0,007
Alkoholijuomien myynti asukasta kohti	2,372	0,20	3,02	0,003
Toimeentulotuki, euroa / asukas	0,353	0,13	2,05	0,042
Yksinhuoltajaperheiden osuus lapsiperheistä	3,513	0,17	2,70	0,008
Päihtyneiden säilöönnotot / 1000 asukasta	3,877	0,28	4,60	0,000
Maassamuuton tulomuutto / 1000 asukasta	8,070	0,17	2,60	0,010
Valtionveronalaiset tulot, euroa / asukas	0,008	0,26	3,56	0,000
Kesämökkien lukumäärä / 1000 asukasta	0,087	0,27	4,59	0,000

Taulukkoon 4.9 on ryhmitelty *maaseutumaiset kunnat* sen mukaan, onko yllä muodostetulla regressiomallilla estimoitu tulos korkeampi vai matalampi kuin tilastoitu omaisuusrikostasotaso kunnassa. Koska *maaseutumaiset kunnat* eroavat muita ryhmiä enemmän toisistaan, myös estimoitu omaisuusrikostasotaso eroaa aiempaa enemmän tilastoidusta omaisuusrikostasosta. Taulukon luokitusta onkin muutettu vastaamaan paremmin saatuja tuloksia; odotetun omaisuusrikostasotason kunnissa poikkeama on alle 20 % ja korkean sekä matalan omaisuusrikostasotason kunnissa poikkeamat ovat yli 40 %. Myös alle tuhannen asukkaan kunnille on laskettu ennusteet, ja kyseiset kunnat on tummennettu taulukkoon. Neljästätoista alle tuhannen asukkaan kunnasta peräti yhdeksän poikkesi suuresti ennustetusta rikostasosta, mikä kertoo näiden kuntien rikostasotason ennustamisen vaikeudesta.

**Taulukko 4.9.** Maaseutumaisten kuntien luokittelu omaisuusrikostason mukaan. Kunnat ovat aakkosjärjestyksessä ryhmien sisällä.

A) Korkean omaisuusrikostason maaseutumaiset kunnat

Korkea omaisuusrikostaso (yli 40 %)			
Alajärvi	Kristiinankaupunki	Perho	Sotkamo
<b>Geta</b>	Lapinjärvi	Pyhtää	<b>Suomenniemi</b>
Humppila	Luumäki	Pälkäne	Taivassalo
Hyrnsalmi	Luvia	Rautavaara	Tervola
Iitti	Merijärvi	Ristiina	Töysä
Juva	Nousiainen	Siikajoki	Virolahti

B) Melko korkean omaisuusrikostason maaseutumaiset kunnat

Melko korkea omaisuusrikostaso (20 – 39,9 %)				
Aura	Keitele	Merikarvia	Pomarkku	Siikalatva
<b>Eckerö</b>	Kinnula	Miehikkälä	Pyhäranta	Somero
Eurajoki	Kivijärvi	Oravainen	Pöytyä	Tammela
Hämeenkoski	Kärkölä	Padasjoki	Ruotsinpyhtää	Tuusniemi
Kaavi	Kärsämäki	Pertunmaa	Rääkkylä	Ujala
Kangasniemi	Liperi	Pihtipudas	Savitaipale	Utajärvi

C) Melko matalan omaisuusrikostason maaseutumaiset kunnat

Melko matala omaisuusrikostaso (-20 – -39,9 %)				
Askola	Karstula	Lumijoki	Rautalampi	Teuva
Enonkoski	Karvia	Marttila	Rautjärvi	Toholampi
Evijärvi	Kemiönsaari	Muonio	Ruokolahti	Toivakka
Hankasalmi	Kihniö	Pielavesi	Sievi	Utsjoki
Honkajoki	Konnevesi	Punkaharju	Siikainen	Uurainen
Jämijärvi	Kruunupyy	Puolanka	Sund	Vesilahti
Kannonkoski	Lemland	Ranua	Tervo	Virrat

D) Matalan omaisuusrikostason maaseutumaiset kunnat

Matala omaisuusrikostaso (alle -40 %)			
<b>Föglö</b>	Juupajoki	Reisjärvi	Ylämaa
<b>Hailuoto</b>	<b>Kumlinge</b>	<b>Sottunga</b>	
Hammarland	<b>Lestijärvi</b>	<b>Vårdö</b>	

## 4.4 Omaisuusrikosmallien vertailu

Alla on vertailtu omaisuusrikollisuutta selittäviä ennustemalleja poliisin tilastoi-  
maan rikostasoon. Vertailtaessa kuntien luokituksia tutkimuksissa, puhutaan  
nimenomaan ennusteisiin perustuvista luokituksista. Korkeaan rikosluokkaan  
kuuluvassa kunnassa ei välttämättä tehdä määrällisesti erityisen paljon omai-  
suusrikoksia vaan niitä tapahtuu poikkeuksellisen paljon kunnan koko ja rakenne  
huomioon ottaen.

**Kaupunkimaisten kuntien** omaisuusrikollisuutta selittäviin malleihin valikoitui  
7-9 muuttujaa. *Alkoholin myynti* ja *avioerojen määrä* selittävät molempia malle-  
ja. Heiskasen mallissa avioliittoa kuvaavat muuttujat olivat tosin *avioliitossa ole-  
vien* ja *yksinhuoltajaperheiden osuudet*. *Maahanmuutto* ja *huumausainerikokset*  
eivät sisältyneet Heiskasen malliin, mutta selittävät uutta mallia. Heiskasen mal-  
lissa selittäjiksi valikoituivat ikäryhmiin ja muuttoon liittyvät muuttujat sekä  
*vähittäiskaupan liikevaihto*, joka ei ollut uudessa tutkimuksessa mukana lain-  
kaan.

*Kaupunkimaisista kunnista* ainoastaan Vaasassa omaisuusrikostaso on pysy-  
nyt poikkeuksellisen korkeana. Myös Forssassa, Kotkassa, Kouvolassa, Porissa,  
Oulussa ja Oulunsalossa rikostaso on pysynyt korkeana tai melko korkeana. Han-  
gossa rikostaso on noussut kautta linjan, mikä näkyy myös omaisuusrikosten kas-  
vussa; Hangon omaisuusrikostaso on noussut melko matalasta korkeaksi. Es-  
poossa, Joensuussa, Lappeenrannassa, Torniossa, Valkeakoskella ja Varkaudes-  
sa sen sijaan omaisuusrikostaso on pysynyt melko matalana tai matalana. Harja-  
vallassa, Maarianhaminassa, Nokiolla ja Nurmijärvellä ennustettu rikostaso on  
puolestaan vähentynyt odotetusta tasosta matalaksi.

**Taajaan asuttujen kuntien** ryhmän molemmissa malleissa esiintyy parisuh-  
teeseen ja ansiotasoon liittyviä muuttujia. Lisäksi *vuokra-asuntojen osuus* ja  
*huumausainerikokset* selittävät uutta mallia. Heiskasen mallissa muuttujia on  
enemmän ja ne liittyvät ikäryhmiin, koulutukseen, työttömyyteen ja muuttoon.  
Uudessa mallissa pääpaino on ansiotasoon liittyvillä muuttujilla, kun taas avioliit-  
tot ja työttömyys ovat parhaita selittäjiä Heiskasen mallissa. Molempien mallien  
selitysaste on noin 50 %.

*Taajaan asutuista kunnista* ainoastaan lissä omaisuusrikostaso on pysynyt  
poikkeuksellisen korkeana. Myös Huittisissa, Keuruulla, Kokemäellä, Kuusamossa,  
Laitilassa, Mäntsälässä, Paimiossa, Uudessakaupungissa ja Ylivieskassa omaisuus-  
rikostaso on pysynyt korkeana tai melko korkeana. Vastaavasti Ilmajoella, Hattu-  
lassa, Kankaanpäässä ja Kannuksessa rikostilanne on pysynyt matalana tai melko  
matalana. Nastolassa ja Orimattilassa omaisuusrikokset vaikuttaisivat kasvaneen  
melko matalasta tasosta korkeaan. Myös Mäntyharjun, Naantalın ja Oulaisen  
ennustettu rikostaso on noussut huomattavasti. Hämeenkyrön ennuste on puo-  
lestaan laskenut odotetusta tasosta matalaan.

**Maaseutumaisten kuntien** ryhmässä *alkoholin myynti ja kesämökkien lukumäärä* selittävät molempia malleja. *Kesämökkien lukumäärä* on tosin suhteutettu uudessa tutkimuksessa asukasluvun eikä pinta-alan mukaan. Myös avioeroja ja muuttoliikettä kuvaavat muuttajat löytyvät molemmista malleista. Uuteen malliin sisältyy lisäksi kaksi toimeentuloon liittyvää muuttujaa, kun taas Heiskasen mallia selittävät *vuokra-asuntojen ja alueen ulkopuolella työssäkäyvien osuudet*. Molempien mallien selitysasteet ovat noin 50 %.

Alajärvellä, Humppilassa, Lapinjärvellä, Rautavaarassa, Ristiinassa ja Taivassalossa omaisuusrikollisuus on pysynyt poikkeuksellisen korkeana. Kuntia, joissa rikostaso on molempien ennusteiden mukaan korkea tai melko korkea, ovat Eurajoki, Juva, Kärämäki, Luumäki, Merikarvia, Oravainen, Padasjoki, Perho, Pertunmaa, Pomarkku, Pyhtää, Siikajoki, Somero, Tammela ja Virolahti. *Maaseutumaisten kuntien* rikostason ailahtelusta kertovat Luvian, Merijärven ja Töysän omaisuusrikosennusteet, jotka ovat nousseet matalasta korkeaksi.

Poikkeuksellisen matalana omaisuusrikollisuus on pysynyt ainoastaan Hammarlandissa ja Juupajoella. Melko matala tai matala rikostaso on molempien ennusteiden mukaan Enonkoskella, Enontekiössä, Hankasalmella, Honkajoella, Karviassa, Kihniössä, Lemillä, Lemlandissa, Lumijoella, Muoniossa, Puolangassa, Pyhäjoella, Rautjärvellä, Reisjärvellä, Sundissa, Tervossa, Toivakassa, Uuraisissa, Vehmaassa, Vesilahdessa ja Vetelissä. Mielenkiintoista on, ettei yhdessäkään kunnassa omaisuusrikostaso ole laskenut korkeasta matalaksi. Melko korkeasta matalaksi sen sijaan ovat laskeneet Lestijärven ja Ylämaan ennusteet.

# 5 Faktorianalyysi

## 5.1 Faktorianalyysin tausta

Faktorianalyysissa etsitään havaintoyksiköiden ominaisuuksia kuvaavasta muuttujajoukosta piileviä yhdenmukaisuuksia eli faktoreita. Taustalla on ajatus, ettei tiettyjä havaintoyksikköjen ominaisuuksia pystytä havainnoimaan suoraan, vaan niistä saadaan ainoastaan epäsuoraa tietoa. Faktori voidaan ymmärtää teoreettisena käsitteenä, jonka olemassaolo päätellään konkreettisista havainnoista. Tällaisia käsitteitä ovat esimerkiksi *sosiaalisuus* ja *älykkyys*.

Faktorianalyysi sai alkunsa 1900-luvun alussa, kun englantilainen psykologi Charles Spearman julkaisi älykkyystestejä koskevat tutkimuksensa (Spearman 1904). Francis Galton tosin tutki älykkyyttä ja esitti latentin muuttujan idean jo muutama vuosikymmen aiemmin. Spearman vei analyysin teorian kuitenkin uudelle tasolle kehittämällä yhden yleisen faktorin ratkaisun. Tämän jälkeen faktorianalyysin teoria kehittyi erityisesti R. Cattelin, D. N. Lawleyn ja L. L. Thurstonen toimesta. Myöhemmin päärooliin ovat tietokoneiden kehittymisen myötä nousseet laskentaintensiiviset tekniikat.

## 5.2 Faktorimallin muodostus

Oletetaan, että havaittavat satunnaismuuttujat  $x_1, x_2, \dots, x_p$  voidaan esittää ei-havaittavien satunnaismuuttujien  $f_1, f_2, \dots, f_r$  ja  $u_1, u_2, \dots, u_r$  avulla seuraavasti:

$$x_i = \mu_i + a_{i1}f_1 + a_{i2}f_2 + \dots + a_{ir}f_r + u_i, \quad i = 1, 2, \dots, p, \quad r \leq p.$$

Kun merkitään  $\mathbf{x} = (x_1, x_2, \dots, x_p)'$ ,  $\mathbf{f} = (f_1, f_2, \dots, f_r)'$  ja  $\mathbf{u} = (u_1, u_2, \dots, u_r)'$  voidaan faktorimalli esittää muodossa

$$(5.1) \quad \mathbf{x} = \boldsymbol{\mu} + \mathbf{A}\mathbf{f} + \mathbf{u},$$

missä  $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_p)'$  on havaintojen odotusarvovektori ja

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pr} \end{bmatrix}$$

on  $r \times p$  –faktorimatriisi, jossa  $r \leq p$ . Tässä vektorit ovat pystyvektoreita ja  $\mathbf{x}'$  tarkoittaa vektorin  $\mathbf{x}$  transpoosia. Koska faktorianalyysin avulla pyritään tiivistämään informaatiota muuttujien määrää vähentämällä, toivotaan että  $r$  olisi paljon pienempi kuin  $p$ . Edellä esitetyn faktorimallin osista käytetään nimityksiä

yhteisfaktori ( $\mathbf{f}$ ), ominaisfaktori ( $\mathbf{u}$ ) ja  $i$ . muuttujan lataus  $j$ . faktorille ( $a_{ij}$ ). Mallissa lataukset  $a_{ij}$  ovat estimoitavia parametreja. (Mellin 2008, 3–4.).

Faktorimallia koskevat seuraavat oletukset (1) – (6):

(1)  $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ , missä  $\boldsymbol{\Sigma}$  on positiivisesti definiitti matriisi ( $\boldsymbol{\Sigma} > 0$ ),

(2)  $\mathbf{f} \sim N(\mathbf{0}, \boldsymbol{\Phi})$ ,

(3)  $\mathbf{u} \sim N(\mathbf{0}, \boldsymbol{\Psi})$ , missä  $\boldsymbol{\Psi} = \text{diag}(\psi_1^2, \psi_2^2, \dots, \psi_p^2)$ ,

(4) ominaisfaktorit  $\mathbf{u}$  ovat riippumattomia yhteisfaktoreista  $\mathbf{f}$ ,

(5)  $r(\mathbf{A}) = r \leq p$

(6)  $\text{cov}(\mathbf{f}) = \mathbf{I}$  eli faktorit ovat ortogonaaliset.

Oletukset antavat kovarianssimatriisille  $\boldsymbol{\Sigma}$  tietynlaisen rakenteen. Helposti nähdään, että ehtojen (1) – (5) pätiessä

$$\begin{aligned}\boldsymbol{\Sigma} &= E(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})' = E(\boldsymbol{\mu} + \mathbf{A}\mathbf{f} + \mathbf{u} - \boldsymbol{\mu})(\boldsymbol{\mu} + \mathbf{A}\mathbf{f} + \mathbf{u} - \boldsymbol{\mu})' \\ &= E(\mathbf{A}\mathbf{f} + \mathbf{u})(\mathbf{A}\mathbf{f} + \mathbf{u})' = E(\mathbf{A}\mathbf{f}\mathbf{f}'\mathbf{A}') + E(\mathbf{u}\mathbf{u}') = \text{cov}(\mathbf{A}\mathbf{f}) + \text{cov}(\mathbf{u}) \\ &= \mathbf{A}\text{cov}(\mathbf{f})\mathbf{A}' + \text{cov}(\mathbf{u}) = \mathbf{A}\boldsymbol{\Phi}\mathbf{A}' + \boldsymbol{\Psi}.\end{aligned}$$

Tästä saadaan faktorianalyysin perusyhtälö

$$(5.2) \quad \text{cov}(\mathbf{x}) = \boldsymbol{\Sigma} = \mathbf{A}\boldsymbol{\Phi}\mathbf{A}' + \boldsymbol{\Psi},$$

josta faktoreiden ortogonaalisuusehdon (6) ja riippumattomuusehdon (4) perusteella saadaan

$$(5.3) \quad \text{cov}(\mathbf{x}) = \boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}' + \boldsymbol{\Psi} \text{ ja } \text{cov}(\mathbf{x}, \mathbf{f}) = \mathbf{A}. \quad (\text{Mustonen 1995, 75–76.})$$

Muuttujan  $\mathbf{x}$  *systemaattiseksi osaksi* kutsutaan satunnaismuuttujaa  $\mathbf{y} = \mathbf{x} - \boldsymbol{\mu} - \mathbf{u} = \mathbf{A}\mathbf{f}$ , joka voidaan esittää muodossa

$$y_i = a_{i1}f_1 + a_{i2}f_2 + \dots + a_{ir}f_r, \quad i = 1, 2, \dots, p, \quad \mathbf{y} \sim N(\mathbf{0}, \mathbf{A}\mathbf{A}').$$

Komponentit  $y_i$  voidaan esittää  $r$  –ulotteisen faktoriavaruuden pisteinä

$$\mathbf{a}_i = (a_{i1}, a_{i2}, \dots, a_{ir}) \quad , i = 1, 2, \dots, p.$$

Vektorin  $\mathbf{a}_i$  pituus saadaan yhtälöstä

$$\|\mathbf{a}_i\|^2 = \mathbf{a}_i' \mathbf{a}_i = h_i^2.$$

Siten

$$\|\mathbf{a}_i\| = h_i$$

ja vektoreiden  $\mathbf{a}_i$  ja  $\mathbf{a}_j$  välisen kulman  $\varphi_{ij}$  kosini on

$$\cos(\varphi_{ij}) = \cos(\mathbf{a}_i, \mathbf{a}_j) = \text{cor}(\mathbf{y}_i, \mathbf{y}_j) = \frac{\mathbf{a}_i' \mathbf{a}_j}{h_i h_j} = \frac{\rho_{ij}}{h_i h_j},$$

jossa  $\rho_{ij} = \text{cor}(\mathbf{x}_i, \mathbf{x}_j)$ .

Havaittavan muuttujan  $\mathbf{x}$  varianssi taas voidaan esittää muodossa

$$D^2(\mathbf{x}_i) = \sigma_{ii} = a_{i1}^2 + a_{i2}^2 + \dots + a_{ir}^2 + \psi_i^2 = h_i^2 + \psi_i^2,$$

jossa  $h_i^2$  on muuttujan  $i$ . kommunaliteetti ja  $\psi_i^2$  muuttujan  $i$ . ominaisvarienssi. Kommunaliteetti kuvaa muuttujan systemaattista, ei-havaittavilla satunnaismuuttujilla  $f_i$  selitettävää osaa. Jos yksittäisen muuttujan kommunaliteetti on pieni, kannattaa harkita, onko muuttujaa ylipäänsä syytä sisällyttää analyysiin. Kun faktorointi tehdään korrelaatiomatriisista  $\mathbf{R}$ , on muuttujan  $i$ . kommunaliteetilla  $h_i^2$  ominaisuudet

$$h_i^2 = 1 - \psi_i^2 < 1 \quad \text{ja} \quad \text{cor}(\mathbf{x}_i, \mathbf{y}_i) = h_i. \quad (\text{Mellin 2008, 7–8.})$$

### 5.3 Faktorointi

Jos tutkimusongelmaa koskevan tiedon perusteella ei voida etukäteen kiinnittää faktoreiden lukumäärää, sitä voidaan arvioida tilastollisesti analyysiin valittujen  $x$ -muuttujien otoskorrelaatiomatriisin  $\mathbf{R}$  ominaisarvojen avulla. Valitaan esimerkiksi niin monta faktoria kuin on yli ykkösen olevia ominaisarvoja. Tarkasteluun voidaan lisätä näkökulma, jossa lukumääräraja näkyy selvänä katkoksenä ominaisarvojen suuruudessa. Jos muuttujien oletetaan noudattavan moniulotteista normaalijakaumaa ja faktoroinnissa käytetään suurimman uskottavuuden menetelmää, voidaan faktoreiden lukumäärää arvioida myös  $\chi^2$ -testin avulla. (Nummenmaa 1997, 243–244.)

Käytetyimmät faktorointimenetelmät ovat pääakseliratkaisu (principal axis solution) ja suurimman uskottavuuden menetelmä (maximum likelihood solution). Faktoroinnin yhteydessä lasketaan kommunaliteetit, jotka kertovat, kuinka suuren osan havaittujen muuttujien vaihtelusta pystytään selittämään faktoreiden avulla. Mitä suurempi on muuttujan  $x_i$  kommunaliteetti, sitä paremmin faktorit selittävät sen vaihtelua.

*Pääakselifaktoroinnissa* tarkastellaan systemaattisen osan  $\mathbf{y}$  kovarianssimatriisia

$$\text{cov}(\mathbf{y}) = E(\mathbf{A}\mathbf{f}\mathbf{f}'\mathbf{A}') = \mathbf{A}\mathbf{A}' = \mathbf{R} - \mathbf{\Psi}.$$

Jos kommunaliteetit tunnettaisiin, saataisiin

$$\mathbf{R} - \mathbf{\Psi} = \begin{bmatrix} h_1^2 & r_{12} & \cdots & r_{1p} \\ r_{21} & h_2^2 & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \cdots & h_p^2 \end{bmatrix}$$

ja faktorimatriisi  $\mathbf{A}$  voitaisiin määrätä minimoimalla lauseke

$$\| \mathbf{R} - \mathbf{\Psi} - \mathbf{A}\mathbf{A}' \|^2.$$

Koska käytännössä kommunaliteetteja ei kuitenkaan tunneta, ne joudutaan korvaamaan sopivilla estimaateilla. Tavallisimmat estimaatit ovat

- 1)  $h_i^2 = \max_{i \neq j} |r_{ij}|,$
- 2)  $h_i^2 = R_{i, 1, 2, \dots, i-1, i+1, \dots, p}^2,$
- 3)  $h_i^2 = 1.$

Kun  $p$  on parinkymmenen luokkaa tai suurempi, on melko samantekevää, miten kommunaliteetit arvioidaan. Lopulliset estimaatit saadaan pääkseliratkaisun jälkeen laskemalla ne estimoidusta faktorimatriisista  $\mathbf{A}$ . Pienillä muuttujamäärillä kommunaliteetit on paras määrätä iteroimalla eli syöttämällä estimoidusta matriisista  $\mathbf{A}$  lasketut kommunaliteetit takaisin korrelaatiomatriisiin lävistäjälle ja toistamalla tätä menettelyä, kunnes kommunaliteetit tulevat riittävän vakaiiksi. Käytännössä muutama iteraatio riittää. (Mustonen 1995, 78–79.)

## 5.4 Rotatointi

Graafiset rotaatiot olivat pitkään käytössä, mutta iterointimenetelmien kehittymisen myötä ne ovat jääneet vähäiselle huomiolle. Analyttiset menetelmät jakautuvat ortogonaalisiin ja vinokulmisiin rotaatioihin. Ortogonaalisissa rotaatioissa faktorit jäävät keskenään korreloimattomiksi, kun taas vinoissa rotaatioissa faktoreiden sallitaan korreloida keskenään. Alla on esitelty muutamia tunnettuja ja analyyseissa käytettyjä rotaatiomenetelmiä.

### 5.4.1 Ortogonaaliset rotaatiot

Faktorit  $\mathbf{f}$  eivät määräydy oletuksista (1) – (5) ja perusyhtälöstä yksikäsitteisesti, vaan jäljelle jää rotaatiomahdollisuus. Olkoon  $\mathbf{f}^* = \mathbf{T}'\mathbf{f}$ , missä  $\mathbf{T}$  on säännöllinen  $r \times r$ -matriisi, jolloin  $\mathbf{f}^* \sim N(\mathbf{0}, \mathbf{T}'\mathbf{\Phi}\mathbf{T})$ . Jos merkitään  $\mathbf{\Phi}^* = \mathbf{T}'\mathbf{\Phi}\mathbf{T}$  ja  $\mathbf{A}^* = \mathbf{A}(\mathbf{T}^{-1})'$ , saadaan muuttujalle  $\mathbf{x}$  esitys

$$\mathbf{x} = \boldsymbol{\mu} + \mathbf{A}^*\mathbf{f}^* + \mathbf{u} \text{ ja } \boldsymbol{\Sigma} = \mathbf{A}^*\mathbf{\Phi}^*(\mathbf{A}^*)' + \mathbf{\Psi},$$

josta seuraa, että jokainen  $\mathbf{f}^*$  ja sitä vastaava  $\mathbf{A}^*$  on myös mahdollinen ratkaisu.



Oletuksen (6) perusteella rotaatiomatriiseiksi  $\mathbf{T}$  kelpaavat vain ortogonaaliset matriisit eli  $\mathbf{T}\mathbf{T}' = \mathbf{T}'\mathbf{T} = \mathbf{I}$ . Havainnollisuuden kannalta on hyvä muistaa, että ortogonaalisella matriisilla kertominen vastaa rotaatiota eli kiertoa, jolloin vektorien pituudet ja niiden väliset kulmat säilyvät. Nyt faktorimalliyhtälö (4.1) voidaan esittää muodossa

$$\mathbf{x} - \boldsymbol{\mu} = \mathbf{A}\mathbf{f} + \mathbf{u} = \mathbf{A}\mathbf{T}\mathbf{T}'\mathbf{f} + \mathbf{u} = \mathbf{A}^*\mathbf{f}^* + \mathbf{u} = \mathbf{A}^*\mathbf{f}^* + \mathbf{u},$$

missä  $\mathbf{A}^* = \mathbf{A}\mathbf{T}$  ja  $\mathbf{f}^* = \mathbf{T}'\mathbf{f}$ . Helposti nähdään, että  $E(\mathbf{f}^*) = \mathbf{0}$  ja  $\text{cov}(\mathbf{f}^*) = \mathbf{I}$ . Käytännössä tämä tarkoittaa, että faktoreilla  $\mathbf{f}$  ja  $\mathbf{f}^*$  on samat ominaisuudet ja latauksilla  $\mathbf{A}$  ja  $\mathbf{A}^*$  sama kovarianssimatriisi  $\boldsymbol{\Sigma}$ . Nähdään myös, että korrelaatiomatriisi  $\mathbf{R}$  on muotoa

$$\text{cor}(\mathbf{x}) = \mathbf{R} = \mathbf{A}^*(\mathbf{A}^*)' + \boldsymbol{\Psi}.$$

Faktorimatriisia  $\mathbf{A}$  pyritään rotatoimaan matriisiin  $\mathbf{T}$  avulla niin, että faktoreiden tulkinta helpottuu. Yleensä pyritään sellaiseen tilanteeseen, että latausmatriisi  $\mathbf{A}$  sisältää kullakin rivillä vain muutaman itseisarvoltaan suuren latauksen ja mahdollisimman monta nollaa lähellä olevaa latausta. (Mustonen 1995, 76.)

Tunnetuimmat ortogonaaliset rotaatiomenetelmät ovat Quartimax ja Varimax. Ne perustuvat faktorilatausten neljänsien potenssien maksimoinnille. Quartimax-rotaatiossa maksimoidaan latausten neljänsien potenssien summa

$$Q = \sum_{j=1}^m \sum_{l=1}^p a_{jl}^4, \text{ missä}$$

$Q$  = maksimoitava arvo,

$a_{jl}$  = muuttujan  $x_j$  faktorilataus faktorilla  $S_l$  (Rummel 1970, 391.).

Varimax-menetelmässä puolestaan tarkastellaan faktorilatausten variansseja sarakkeittain. Varianssi kasvaa sen mukaan, mitä enemmän yksittäisellä faktorilla on sekä suuria että pieniä latauksia. Tällöin myös faktorilatausten neliöiden varianssi kasvaa. Suurin varianssi saavutetaan kun lataukset ovat nollija tai ykkösiä. Näin ollen ortogonaalinen Varimax-rotatio saadaan maksimoimalla latausten neliöiden faktoreittain laskettujen varianssien summa eli

$$V = m \sum_{i=1}^p \sum_{j=1}^m \left( \frac{a_{jl}}{h_j} \right)^4 - \sum_{i=1}^p \left( \sum_{j=1}^m \frac{a_{jl}^2}{h_j^2} \right)^2, \text{ missä}$$

$V$  = normalisoitujen faktoreiden varianssi,

$a_{jl}$  = muuttujan  $x_j$  faktorilataus faktorilla  $S_l$ ,

$h_j^2$  = muuttujan  $x_j$  kommunaliteetti (Rummel 1970, 391–392.).

## 5.4.2 Vinokulmaiset rotaatiot

*Vinoissa rotaatioissa* faktoreiden sallitaan korreloida keskenään, mikä kuvastaa monissa tilanteissa luonnollisemmin todellisuutta. Jos alkuperäinen faktorirakenne on ortogonaalinen, rotaatiossa  $\mathbf{A}^* = \mathbf{A}(\mathbf{T}^{-1})'$ , perusyhtälö (6.2) saa muodon

$$\mathbf{\Sigma} = \mathbf{A}^* \mathbf{T}' \mathbf{T} (\mathbf{A}^*)' + \mathbf{\Psi}.$$

Nyt korrelaatiomatriisin  $\mathbf{T}'\mathbf{T} = \mathbf{\Phi}$  lävistäjäalkiot ovat ykkösiä ja matriisin  $\mathbf{T}$  sarakkeet ovat yksikkövektorin mittaisia. Nämä sarakkeet koostuvat faktoriavaruuden uusien vinojen akselien suuntakosineista alkuperäisten ortogonaalisten akselien suhteen. Tällöin matriisi

$$\mathbf{S} = \mathbf{A}\mathbf{T}$$

antaa muuttujien systemaattisten osien projektiot uusilla vinokulmaisilla koordinaattiakseleilla. Varsinainen rotatoitu faktorimatriisi  $\mathbf{A}^*$ , joka antaa vinokulmaiset koordinaatit, on tällöin

$$\mathbf{A}^* = \mathbf{A}(\mathbf{T}^{-1})' = \mathbf{S} \mathbf{T}^{-1} (\mathbf{T}^{-1})' = \mathbf{S} \mathbf{\Phi}^{-1}. \quad (\text{Mustonen 1995, 83–84.})$$

Käytetyin vinorotaatiomenetelmä on Oblimin, jossa minimoidaan lauseke

$$\sum_{j=1}^r \sum_{k=1}^{j-1} \left( \sum_{i=1}^p a_{ik}^{*2} a_{ij}^{*2} - \frac{\delta}{p} \sum_{i=1}^p a_{ik}^{*2} \sum_{i=1}^p a_{ij}^{*2} \right)$$

rotaatiomatriisin  $\mathbf{T}$  suhteen ehdolla  $\text{diag}(\mathbf{T}'\mathbf{T}) = \mathbf{I}$ . Minimoitava lauseke riippuu parametrasta  $\delta$ , jonka arvoksi valitaan 0 tai negatiivinen luku. Mitä negatiivisempi  $\delta$  on, sitä ortogonaalisemmaksi ratkaisu muodostuu. (Mustonen 1995, 84.)

Toinen paljon käytetty vinorotaatiomenetelmä on Promax, joka löytää ratkaisun Oblimin-menetelmää nopeammin ja on käsitteellisesti yksinkertaisempi. Sen lähtökohdista on ortogonaalinen Varimax-rotaatio, jonka jälkeen mallin yksinkertaistamiseksi luovutaan faktoreiden ortogonaalisuudesta. Normalisoimalla ortogonaalinen matriisi riveittäin ja sarakkeittain, ja korottamalla lataukset  $k$ . potenssiin päästään tilanteeseen, jossa faktoreiden vahvat lataukset vahvistuvat ja heikot lataukset heikentyvät entisestään. Lopuksi näin saadulle matriisille etsitään pienimmän neliösumman sovite ortogonaalisen matriisin pohjalta. (Rummel 1970, 419.)

Normalisointivaiheessa latauksien painotuksissa käytetään potenssina parametria  $k$  (*kappa*), jolle annetaan useimmiten arvoja väliltä 2 – 4. Tällä pyritään pakottamaan latausten rakenne bipolaariseksi. SPSS käyttää oletusarvoa  $k = 4$ , kun taas esimerkiksi SAS käyttää arvoa  $k = 3$ . Promax-menetelmän kehittäjät Hendrickson ja White ovat todenneet, että aineistoilla, joissa on suhteellisen yksinkertainen rakenne, saavutetaan parempia tuloksia pienemmillä parametrin  $k$  arvoilla. (Hendrickson & White, 1964)

## 5.5 Faktoripistemäärät

Faktoripisteillä tarkoitetaan faktoreiden  $\mathbf{f}$  esittämistä muuttujan  $\mathbf{x}$  avulla. Muuttujan  $\mathbf{x}$  arvot ovat havaittuja ja keskiarvovektori  $\boldsymbol{\mu}$  on estimoitu faktoroinnin yhteydessä. Alla on esitetty niin sanottu regressiomenetelmä, jota olen käyttänyt myös tutkimuksessani, faktoripisteiden laskemiselle. Merkitään

$$(5.4) \quad \mathbf{f} = \mathbf{K}(\mathbf{x} - \boldsymbol{\mu}),$$

jossa  $\mathbf{K}$  on  $r \times p$  -matriisi. Pyritään löytämään sellainen matriisi  $\mathbf{K}$ , joka toteuttaisi yhtälön mahdollisimman hyvin. Valitaan  $\mathbf{K}$  siten, että erotuksen

$$\mathbf{f} - \mathbf{K}(\mathbf{x} - \boldsymbol{\mu})$$

normin odotusarvo minimoituu. Tutkitaan siksi neliösummaa

$$\begin{aligned} \varphi(\mathbf{K}) &= E\|\mathbf{f} - \mathbf{K}(\mathbf{x} - \boldsymbol{\mu})\|^2 \\ &= E[(\mathbf{f} - \mathbf{K}(\mathbf{x} - \boldsymbol{\mu}))' (\mathbf{f} - \mathbf{K}(\mathbf{x} - \boldsymbol{\mu}))]. \end{aligned}$$

Voidaan osoittaa, että matriisi  $\mathbf{K}$ , joka minimoi neliösumman  $\varphi(\mathbf{K})$  on muotoa

$$(5.5) \quad \mathbf{K} = \mathbf{A}' \boldsymbol{\Sigma}^{-1},$$

josta seuraa yhtälön (5.4) mukaan

$$\mathbf{f} = \mathbf{A}' \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}).$$

Faktorianalyysin perusyhtälön (5.3) avulla voidaan muodostaa yhtälö

$$\begin{aligned} \mathbf{A}' \boldsymbol{\Psi}^{-1} \boldsymbol{\Sigma} &= \mathbf{A}' \boldsymbol{\Psi}^{-1} (\mathbf{A} \mathbf{A}' + \boldsymbol{\Psi}) \\ &= \mathbf{A}' \boldsymbol{\Psi}^{-1} \mathbf{A} \mathbf{A}' + \mathbf{A}' \boldsymbol{\Psi}^{-1} \boldsymbol{\Psi} \\ &= \mathbf{A}' \boldsymbol{\Psi}^{-1} \mathbf{A} \mathbf{A}' + \mathbf{A}' \\ &= (\mathbf{A}' \boldsymbol{\Psi}^{-1} \mathbf{A} + \mathbf{I}) \mathbf{A}'. \end{aligned}$$

Kun edellä mainittu yhtälö kerrotaan oikealta matriisilla  $\boldsymbol{\Sigma}^{-1}$ , saadaan

$$\mathbf{A}' \boldsymbol{\Psi}^{-1} = (\mathbf{A}' \boldsymbol{\Psi}^{-1} \mathbf{A} + \mathbf{I}) \mathbf{A}' \boldsymbol{\Sigma}^{-1},$$

josta yhtälön (5.5) mukaan seuraa

$$\mathbf{A}' \boldsymbol{\Psi}^{-1} = (\mathbf{A}' \boldsymbol{\Psi}^{-1} \mathbf{A} + \mathbf{I}) \mathbf{K}.$$

Matriisi  $\mathbf{K}$  voidaan tällöin esittää muodossa

$$\mathbf{K} = (\mathbf{A}' \boldsymbol{\Psi}^{-1} \mathbf{A} + \mathbf{I})^{-1} \mathbf{A}' \boldsymbol{\Psi}^{-1},$$

jolloin faktoripisteiden regressioestimaattien  $\mathbf{f}$  yhtälöksi saadaan

$$\mathbf{f} = (\mathbf{A}'\Psi^{-1}\mathbf{A} + \mathbf{I})^{-1}\mathbf{A}'\Psi^{-1}(\mathbf{x} - \boldsymbol{\mu}) .$$

Korvaamalla saaduissa lausekkeissa matriisit  $\mathbf{A}$ ,  $\Psi$  ja  $\boldsymbol{\mu}$  estimaateillaan, saadaan faktoripisteiden regressioestimaatit:

$$\hat{\mathbf{f}} = \hat{\mathbf{A}}' \hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{x} - \bar{\mathbf{x}}) = (\hat{\mathbf{A}}' \hat{\boldsymbol{\Psi}}^{-1}\hat{\mathbf{A}} + \mathbf{I})^{-1}\hat{\mathbf{A}}' \hat{\boldsymbol{\Psi}}^{-1}(\mathbf{x} - \bar{\mathbf{x}}) .$$

Näin vältetään matriisin  $\boldsymbol{\Sigma}$  kääntäminen, mikä saattaa olla laskennallisesti raskasta, jos muuttujien  $p$  lukumäärä on suuri. Menetelmän esitteli ensimmäisenä W. Ledermann vuonna 1938. (Mellin 2008, 17–19.; Ledermann, 1938)

Edellä muodostetussa yhtälössä on oletettu faktorit korreloimattomiksi. Jos faktorit kuitenkin korreloivat eli  $\text{cov}(\mathbf{f}) = \boldsymbol{\Phi}$ , niin neliösumma  $\varphi(\mathbf{K})$  minimoituu, kun  $\mathbf{K} = \boldsymbol{\Phi}\mathbf{A}'\boldsymbol{\Sigma}^{-1}$ . Tällöin Ledermannin lyhennetty muoto on

$$\mathbf{K} = (\mathbf{A}'\Psi^{-1}\mathbf{A} + \boldsymbol{\Phi}^{-1})^{-1}\mathbf{A}'\Psi^{-1} \quad (\text{Mustonen, 1995}).$$

## 6 Faktorimallien sovitus aineistoon

Kunta-aineistosta on faktorianalyysin avulla pyritty löytämään ulottuvuuksia, joiden avulla rikollisuuden rakennetta kunnissa olisi helpompi ymmärtää. Faktorianalyysin yhteydessä havainnoille lasketaan myös faktoripistemäärät, joiden avulla faktoreiden ja rikostasojen välisiä korrelaatioita voidaan tarkastella.

Analyysi on suoritettu useassa kuntaryhmässä, koska kunnat ovat rakenteeltaan hyvin erilaisia. Faktorianalyysin kannalta on suositeltavaa, että havaintojen määrä olisi mahdollisimman suuri. Kunta-aineistossa kuntien määrä on kuitenkin rajallinen. Aluksi faktorianalyysi tehdään *kaikille yli tuhannen asukkaan kunnille*, jolloin havaintojen lukumäärä on 333. Seuraavaksi *kaupunkimaiset ja taajaan asutut kunnat* analysoidaan yhdessä, jolloin havaintojen lukumäärä tippuu 128:aan. Tämän jälkeen tehdään analyysit *kaupunkimaisille ja taajaan asutuille kunnille* erikseen, jolloin havaintojen määrät ovat enää vain 62 ja 66. Faktorianalyysin kannalta nämä eivät enää ole suositeltavan kokoisia otoksia, etenkin kun analyysihin valittavien muuttujien lukumäärät ovat suuria, molemmissa tapauksissa lähes 30. Lopuksi analysoidaan *maaseutumaisten kuntien* ryhmä, johon kuuluu 205 kuntaa.

Usein suositellaan, että havaintojen lukumäärä olisi vähintään sata ja ainakin viisinkertainen muuttujien määrään nähden. Minimioletukset riippuvat kuitenkin myös aineistosta; jos kommunaliteetit ovat korkeita, voi pienempikin aineisto tulla kysymykseen. Etenkin psykologian alalla minimirajana pidetään yleensä havaintojen kaksinkertaista määrää muuttujiin nähden. Koska erityisesti *kaupunkimaisten ja taajaan asuttujen kuntien* ryhmien analyyseissa aineisto on ongelmallinen, on näiden tulkinnessa noudatettu erityistä varovaisuutta.

Aineistojen sopivuutta faktorianalyysiin on testattu Kaiser-Meyer-Olkinin KMO-testillä ja Bartlettin testillä (Bartlett's Test of Sphericity). Molemmat testit perustuvat korrelaatiomatriisin rakenteen tutkimiseen. Ensimmäisessä verrataan osittaiskorrelaatioiden summaa tavallisten korrelaatiokertoimien summaan. Ihannetilanteessa osittaiskorrelaatiot ovat pieniä, jolloin KMO-arvo on lähellä ykköstä. Yleisenä ohjenuorana pidetään, että testisuureen arvon tulisi olla suurempi kuin 0,6 (Metsämuuronen 2003, 524.).

Aineistosta lasketut arvot vaihtelivat välillä 0,743–0,829 eli testien mukaan aineistot sopivat faktorianalyysiin. Bartlettin testissä taas testataan hypoteesia, jossa kaikkien alkuperäisten muuttujien keskinäiset korrelaatiokertoimet ovat nolliä. Kaikissa tapauksissa p-arvot olivat alle 0,000 eli myös Bartlettin testien mukaan aineistot soveltuvat faktorianalyysiin.

## 6.1 Kaikki yli tuhannen asukkaan kunnat

Aluksi faktorianalyysi tehdään *kaikille yli tuhannen asukkaan kunnille*. Helsinki on kuitenkin poistettu analyysistä poikkeavuutensa takia. Analyysi suoritettiin pääakselimenetelmällä, ja siihen valikoitui lopulta 27 muuttujaa. Alun perin analyysiin valittiin 32 muuttujaa, joista viisi poistettiin pienten kommunaliteettien takia. Nämä muuttujat olivat *avioerot, huumausainerikokset, maahanmuutto, pitkäaikaistyöttömät ja väestötiheys*. Analyysi tuotti lopulta viiden faktorin mallin, jonka selitysasteeksi saatiin 69,4 %. Löydetty faktorit ovat *koulutus ja varallisuus, työttömyys, kaupunkimaisuus, lapsiperheet sekä alkoholi ja monikulttuurisuus*. Taulukossa 6.1 on esitetty latausmatriisi, jonka perusteella faktorit on tulkittu. Itseisarvoltaan yli 0,5 vahvuiset lataukset on tummennettu selkeyden vuoksi.

**Taulukko 6.1.** Yli tuhannen asukkaan kuntien rotaatiomatriisi.

27 muuttujaa	Promax-rotatoitu faktorimalli pääakselimenetelmällä				
	Koulutus ja varallisuus	Työttömyys	Kaupunkimaisuus	Lapsiperheet	Alkoholi ja monikulttuurisuus
Valtionveronalaiset tulot	<b>,935</b>	-,091	-,071	-,009	-,030
Korkea-asteen koulutetut	<b>,918</b>	,111	-,011	,066	,011
Kunnan yleinen pienituloisuusaste	<b>-,856</b>	,236	,042	-,104	,070
Koulutustasomittain	<b>,824</b>	,175	,050	,239	,033
Sairastavuusindeksi	<b>-,774</b>	<b>,523</b>	-,027	,101	-,074
Maassamuuton tulomuutto	<b>,711</b>	,173	-,368	,059	,098
Kunnan nettomuutto	<b>,663</b>	-,012	-,250	-,039	-,103
Huoltosuhte, demografinen	<b>-,603</b>	-,151	-,212	,052	-,095
Alkutuotannon työpaikkojen osuus	<b>-,593</b>	-,285	-,203	-,080	-,089
Päihdehuollon nettokustannukset	,430	,162	,323	-,215	,019
Väkiluku	,416	,232	,395	-,051	-,074
Nuorisotyöttömät	-,402	<b>,846</b>	-,113	,081	,012
Vaalit, äänestysprosentti	-,178	<b>-,791</b>	,001	,086	,356
Yritystoimipaikkojen lukumäärä	-,257	<b>-,778</b>	,096	-,261	,036
Työttömät	<b>-,563</b>	<b>,753</b>	-,199	-,063	,121
Päihtyneiden säilönnöt	,006	<b>,660</b>	,332	-,050	-,126
Toimeentulotukimenot	-,003	<b>,551</b>	,162	-,143	,066
Palvelualan työpaikkojen osuus	,270	,425	-,164	,049	,312
Kunnassa olevien työpaikkojen lkm	-,034	-,209	<b>,975</b>	-,056	,031
Pendelöinti	-,442	,023	<b>,919</b>	,000	,036
16–24-vuotiaat miehet	-,179	-,024	<b>,832</b>	<b>,528</b>	,008
Vuokra-asunto kaikista asunnoista	,005	,482	<b>,559</b>	-,079	,030
0–15-vuotiaat	,242	,023	,069	<b>,901</b>	,001
Lapsiperheet	,469	,043	,085	<b>,761</b>	,031
Yksinhuoltajaperheet	,184	,353	,088	-,479	,132
Alkoholijuomien myynti	-,096	,034	,052	-,039	<b>,800</b>
Vieras äidinkieli	,078	-,303	,042	,027	<b>,793</b>

**Koulutus ja varallisuus** –faktorilla vahvimmin latautuivat *korkeakoulutuksen saaneiden osuus, valtionveronalaiset tulot ja kunnan yleinen pienituloisuusaste*. Myös *koulutustasomittain, sairastavuusindeksi ja tulomuutto* latautuivat vahvasti faktorilla. Faktori kuvaa selkeästi kuntia, joissa koulutus- ja varallisuusaste ovat korkeita. Positiivinen netto- ja tulomuutto sekä korkea väkiluku taas kertovat suuresta ja edelleen kasvavasta kunnasta. Tällaisia kuntia ovat ennen kaikkea suuret keskuskunnat, erityisesti yliopistokaupungit. Matala huoltosuhte, alhainen työttömyysaste ja alkutuotannon työpaikkojen vähyys ovat myös ominaisia piirteitä suurille kaupungeille.

**Työttömyys** –faktorilla latautuivat vahvimmin *nuorisotyöttömät, yritystoimipaikkojen lukumäärä ja äänestysprosentti*. Myös *työttömät, päihtyneiden säilönotot ja toimeentulotukimenot* latautuivat vahvasti tällä faktorilla. *Yritystoimipaikkojen lukumäärä ja äänestysprosentti* saivat negatiivisen latauksen, mikä sopii hyvin faktorin luonteeseen. Onkin luonnollista, että työttömyys, yritystoimipaikkojen vähyys, korkea pienituloisuusaste ja toimeentulotukimenot ovat selvästi yhteydessä toisiinsa. Matala äänestysprosentti puolestaan kertoo osittain asukkaiden tyytymättömyydestä omiin vaikuttamismahdollisuuksiinsa. Heiskasen tutkimuksessa äänestysprosentin katsottiin myös mittaavan ohuesti yhteisön sosiaalista kiinteyttä. Tällöin alhainen äänestysaktiivisuus kertoisi kiinteyden puutteesta, millä on rikollisuutta lisäävä vaikutus.

**Kaupunkimaisuus** –faktorilla latautuivat vahvimmin *kunnassa olevien työpaikkojen lukumäärä, pendelöinti ja nuorten miesten suuri osuus*. Myös *vuokra-asuntojen osuus ja väkiluku* saivat vahvan latauksen. Vaikka työssäkäyntimuuttujat latautuivat vahvimmin faktorilla, kertovat muut lataukset ennen kaikkea kaupunkimaisuudesta. Suurissa kaupungeissa työpaikkoja on paljon ja töihin tullaan kauempaakin. Korkea väkiluku, nuorten miesten ja vuokra-asuntojen osuudet sekä päihdeongelmat ovat myös suurille kaupungeille tunnusomaisia piirteitä. Muuttoa koskevat muuttujat sen sijaan latautuvat negatiivisesti, mikä puolestaan ei ole tyyppillistä suurille kaupungeille. Faktori ei olekaan tulkinnan kannalta täysin yksiselitteinen.

**Lapsiperheet** –faktorilla vahvimmin latautuivat *0–15-vuotiaiden ja lapsiperheiden osuudet*. Myös *nuorten miesten korkea ja yksinhuoltajien vähäinen osuus* latautuivat vahvasti. Koska muut muuttujat latautuivat heikosti faktorille, on sitä vaikea tulkita muuten kuin lapsiperheiden mukaan. Heikommista latauksista negatiivisesti latautuneet *päihdehuollon kustannukset ja yritystoimipaikkojen määrä* kertovat omalta osaltaan faktorin luonteesta. Faktorin kuvaamia kuntia voisivat olla esimerkiksi keskusten ympäristökunnat, joihin lapsiperheet muuttavat rauhallisemman elinympäristön perässä.

**Alkoholi ja monikulttuurisuus** –faktorilla latautuivat vahvimmin *alkoholijuomien myynti ja vieras äidinkieli*. Myös *äänestysprosentti ja palvelualan työpaikkojen osuus* latautuivat positiivisesti faktorilla. Muuttujien väliset yhteydet eivät ole selkeästi nähtävissä, joten faktori on nimetty sen vahvimpien latausten mukaan.

**Taulukko 6.2.** Rikostasojen ja faktoreiden korrelaatiomatriisi faktoripisteiden avulla laskettuna. Yli tuhannen asukkaan kunnat.

	Väki- valta- rikok- set	Omai- suus- rikok- set	Koulutus ja varal- lisuus	Työttö- myys	Kau- punkki- mai- suus	Lapsi- perheet	Alkoholi ja moni- kulttuu- risuus
Väkivaltarikokset	1	,615	,054	,621	,540	-,359	,427
Omaisuusrikokset	,615	1	,503	,593	,612	-,336	,476
Koulutus ja varallisuus	,054	,503	1	,162	,241	,238	,214
Työttömyys	,621	,593	,162	1	,437	-,280	,505
Kaupunkimaisuus	,540	,612	,241	,437	1	-,450	,413
Lapsiperheet	-,359	-,336	,238	-,280	-,450	1	-,420
Alkoholi ja monikulttuurisuus	,427	,476	,214	,505	,413	-,420	1

Rotaatiomenetelmänä analyysissä käytettiin vinokulmaista Promax-rotaatiota. Tämä tarkoittaa, että faktoreiden annettiin korreloida keskenään. Yleisimmin faktorianalyysin yhteydessä käytetään suorakulmaisia rotaatioita, mutta sosiologisessa tutkimuksessa on luonnollista ajatella, että faktoreiden välillä on olemassa aitoja korrelaatioita.

Taulukosta 6.2 nähdään, että faktorit *työttömyys* ja *kaupunkimaisuus* korreloivat vahvimmin väkivalta- ja omaisuusrikosten kanssa. *Työttömyys* korreloi hieman voimakkaammin väkivaltarikosten kanssa, kun taas *kaupunkimaisuus* on selkeämmin yhteydessä omaisuusrikollisuuteen. Viides faktori *Alkoholi ja monikulttuurisuus* korreloi myös selkeästi molempien rikostasojen ja faktoreiden kanssa. Nämä kolme faktoria ovatkin selkeimmin yhteydessä korkeisiin rikostasoihin.

*Lapsiperheet* on ainoa faktori, jolla vaikuttaisi olevan rikollisuutta vähentävä vaikutus, koska se korreloi positiivisesti ainoastaan *koulutus ja varallisuus* -faktorin kanssa. Hieman yllättäen *koulutus ja varallisuus* ei korreloi lainkaan väkivaltarikosten kanssa, mutta omaisuusrikosten kanssa korrelaatio on vahva. Faktorin luonne ei siis olekaan rikollisuuden kannalta niin positiivinen kuin olisi saattanut odottaa.



**Taulukko 6.3.** Kuntaryhmien faktoripisteiden keskiarvot.

	Koulutus ja varallisuus	Työttömyys	Kaupunkimaisuus	Lapsiperheet	Alkoholi ja monikulttuurisuus
Kaupunkimaiset kunnat	1,27	0,96	1,12	-0,22	0,59
Taajaan asutut kunnat	0,30	0,20	0,07	0,22	0,01
Maaseutumaiset kunnat	-0,49	-0,36	-0,37	0,00	-0,19

Taulukossa 6.3 on vielä esitetty kuntaryhmien faktoripisteiden keskiarvot, joiden perusteella voidaan vertailla kuntaryhmien sijoittumista faktoreiden suhteen. *Kaupunkimaiset kunnat* saivat neljällä faktorilla selvästi suurimmat arvot. Koulutus ja varallisuus sekä kaupunkimaisuus ovat tälle ryhmälle erittäin tunnusomaisia piirteitä. Vastaavasti *maaseutumaisissa kunnissa* nämä faktorit latautuivat heikosti. *Taajaan asutuille kunnille* koulutus ja varallisuus ovat tunnusomaisia piirteitä, mutta kaupunkimaisuus ei ilmene niissä erityisen vahvasti. Lapsiperheet sen sijaan ovat tyypillisimpiä *taajaan asutuissa kunnissa*, ja toisaalta harvinaisimpia *kaupunkimaisissa kunnissa*.

Myös luonteeltaan negatiiviset faktorit *työttömyys* ja *alkoholi ja monikulttuurisuus* ovat tyypillisimpiä *kaupunkimaisten kuntien* ryhmässä. Erityisesti työttömyys latautuu vahvasti näissä kunnissa. Myös *taajaan asutuissa kunnissa* työttömyys –faktori latautuu positiivisesti, kun taas *maaseutumaisissa kunnissa* työttömyys ei ole lainkaan tyypillistä. Myöskään faktori *alkoholi ja monikulttuurisuus* ei lataudu vahvasti *maaseutumaisten kuntien* ryhmässä.

Kuten yllä huomattiin, kuntien väliset suuret rakenne-erot hankaloittavat selvästi faktoreiden tulkintaa. Seuraavaksi tehdäänkin faktorianalyysi *kaupunkimaisten ja taajaan asuttujen kuntien* ryhmässä, jolloin erot pienenevät huomattavasti. Oletettavasti tämä parantaa mallia ja yksinkertaistaa faktoreiden tulkintaa.

## 6.2 Kaupunkimaiset ja taajaan asutut kunnat

Taulukossa 6.4 on tehty faktorianalyysi kaupunkimaisten ja taajaan asuttujen kuntien ryhmässä. Helsinki on jälleen poistettu analyysistä poikkeavuutensa takia. Analyysi suoritettiin pääakselimenetelmällä, ja siihen valikoitui lopulta 28 muuttujaa. Alun perin analyysiin valittiin 31 muuttujaa, joista kolme kuitenkin poistettiin pienten kommunaliteettien takia. Nämä muuttujat olivat *alkutuotannon työpaikkojen osuus*, *huumausainerikokset* ja *pitkäaikaistyöttömät*. Analyysi tuotti viiden faktorin mallin, jonka selitysasteeksi saatiin 75,4 %. Selitysaste siis nousi kun maaseutumaiset kunnat jätettiin analyysin ulkopuolelle. Löydetyt faktorit ovat *lapsiperheet*, *työttömyys*, *koulutus ja varallisuus*, *kaupunkimaisuus* ja *sosiaaliset ongelmat*.

**Taulukko 6.4.** Kaupunkimaisten ja taajaan asuttujen kuntien rotaatiomatriisi.

28 muuttujaa	Promax-rotatoitu faktorimalli pääakselimenetelmällä				
	Lapsiperheet	Työttömyys	Koulutus ja varallisuus	Kaupunkimaisuus	Sosiaaliset ongelmat
Lapsiperheet	<b>-,908</b>	,014	,077	,176	-,205
Pendelöinti	<b>,847</b>	,208	-,056	,193	-,110
0–15-vuotiaat	<b>-,831</b>	,090	,049	,048	-,362
Kunnassa olevien työpaikkojen lkm	<b>,810</b>	-,046	-,112	,421	-,165
Maassamuuton tulomuutto	<b>-,657</b>	-,156	,192	,156	,114
Kunnan nettomuutto	<b>-,495</b>	<b>-,492</b>	-,049	,079	,286
Vuokra-asunnot	<b>,480</b>	,270	,233	,330	,292
Nuorisotyöttömät	,038	<b>,959</b>	-,061	-,224	-,076
Työttömät	,182	<b>,835</b>	-,081	-,144	,111
Sairastavuusindeksi	,036	<b>,599</b>	-,399	-,122	,127
Yritystoimipaikkojen lukumäärä	<b>,455</b>	<b>-,585</b>	-,104	-,102	,168
Kunnan yleinen pienituloisuusaste	,380	<b>,571</b>	-,138	,026	,216
16–24-vuotiaat miehet	,125	<b>,560</b>	,365	,390	,056
Koulutustasomittain	-,397	,177	<b>,850</b>	,048	-,067
Valtionveronalaiset tulot	-,082	-,268	<b>,843</b>	-,263	,003
Korkea-asteen koulutetut	-,347	,000	<b>,829</b>	,038	-,019
Maahanmuutto	,373	-,026	<b>,638</b>	,206	-,161
Väestötiheys	-,082	-,331	<b>,546</b>	-,002	<b>,454</b>
Palvelualan työpaikkojen osuus	-,276	,179	<b>,498</b>	,174	,412
Väkiluku	-,020	-,042	,017	<b>,911</b>	-,049
Huoltosuhde	,170	,104	,120	<b>-,867</b>	-,089
Vieras äidinkieli	,283	-,401	,213	<b>,656</b>	,005
Päihdehuollon nettokustannukset	<b>,469</b>	-,416	-,081	<b>,564</b>	,029
Vaalit, äänestysprosentti	,449	-,247	<b>,484</b>	<b>-,487</b>	-,384
Avioeroja 25–64-vuotiailla	-,056	-,186	-,013	,143	<b>,785</b>
Alkoholijuomien myynti	,434	-,037	,064	-,371	<b>,672</b>
Päihtyneiden säilönnöt	,185	,300	-,020	,069	<b>,552</b>
Toimeentulotukimenot	,103	,220	,040	,288	<b>,473</b>

Faktorien nimistä nähdään, että ne ovat viimeistä lukuun ottamatta samat kuin *kaikkien yli tuhannen asukkaan kuntien* tapauksessa. *Alkoholi ja monikulttuurisuus* –faktori on korvautunut faktorilla *sosiaaliset ongelmat*. Alkoholi latautui edelleen vahvasti faktorilla, mutta *vieras äidinkieli* ei enää lainkaan. Muutos johtuu siitä, että kaupungit ovat huomattavasti monikulttuurisempia kuin pois rajatut pienet kunnat. Vaikka faktorien tulkinnot ovat pysyneet pitkälti samoina, niiden järjestys ja lataukset ovat muuttuneet. *Lapsiperheet* –faktori onkin nousut neljänneltä sijalta ensimmäiseksi. Tämä on oleellista sikäli, että ensimmäisten faktoreiden selitysosuudet ovat aina korkeampia kuin viimeisten.

**Lapsiperheet** –faktorilla vahvimmin latautuivat *lapsiperheiden osuus perheistä, pendelöinti ja 0–15-vuotiaiden osuus*. Myös *kunnassa olevien työpaikkojen määrä, maassamuuton tulomuutto ja kunnan nettomuutto* latautuivat vahvasti. Tulkinnan helpottamiseksi käännetään latauksien etumerkit, ja tulkitaan faktori seuraavasti; kun *lapsiperheiden osuus* on suuri, niin *pendelöinti ja kunnassa olevien työpaikkojen määrä* ovat vähäisiä. Tällöin on luonnollista, että myös *yrittäjätoimipaikkojen lukumäärä* on pieni. Myös *vuokra-asuntojen osuus ja päihdehuollon kustannukset* ovat vähäisiä. *Tulomuutto ja nettomuutto* sen sijaan latautuvat vahvasti faktorilla. Tulkinnan kannalta olennainen huomio on, ettei *väkiluku* lataudu lainkaan *lapsiperheet* –faktorilla.

Näyttäisi siltä, että tähän kategoriaan kuuluvat ennen kaikkea suurten kaupunkien ympäristökunnat, joihin lapsiperheet muuttavat asumaan. Viime vuosille on ollut ominaista, että näiden kuntien väestömäärä on ollut kasvussa. Usein tällaisten kuntien asukkaat käyvät kuitenkin töissä oman kuntansa ulkopuolella, lähinnä suuressa keskuskunnassa. Tampereen seudulla esimerkiksi Pirkkala ja Ylöjärvi ovat tällaisia kuntia.

**Työttömyys** –faktorilla latautuivat vahvimmin *nuorisotyöttömät, työttömien osuus ja sairastavuusindeksi*. Myös *pienituloisuusaste, yritystoimipaikkojen vähäisyys ja nuorten miesten osuus* latautuivat vahvasti. Tämä kertoo yleisemmin korkean työttömyysasteen kuntien tilanteesta; työttömyys, työpaikkojen vähäisyys, pienituloisuusaste ja sairastavuus ovat selkeästi yhteydessä toisiinsa. Kunnissa ilmenee myös muuttotappiota, kun ihmiset siirtyvät työpaikan perässä muualle. Näiden moniongelmaisten kuntien on toisaalta myös hankala houkutelua uusia asukkaita.

**Koulutus ja varallisuus** –faktorilla vahvimmin latautuivat *koulutustasomittain, korkeakoulutuksen saaneiden osuus ja valtionveronalaiset tulot*. Myös *maahanmuutto, väestötiheys ja palvelualan työpaikkojen osuus* latautuivat vahvasti. Faktori kuvaa selkeästi kuntia, joissa koulutus- ja varallisuusaste ovat korkeita. Yleensä tällaisia kuntia ovat suuret keskuskunnat, ennen kaikkea yliopistokaupungit. On selvää, että suuri väestötiheys, laaja maahanmuutto ja palvelualan työpaikkojen suuri osuus ovat myös ominaisia piirteitä tällaisille kunnille. Korkea äänestysprosentti kertoo myös osaltaan yhteisön sosiaalisesta kiinteydestä.

Neljännän faktorin tulkinta on hieman hankalampaa. Faktorilla latautuivat vahvimmin *väkiluku, huoltosuhde ja vieras äidinkieli*. Myös *päihdehuollon nettokustannukset, äänestysprosentti ja kunnassa olevien työpaikkojen määrä* saivat vahvan latauksen. Korkea väkiluku ja kunnassa olevien työpaikkojen suuri määrä kertovat kaupunkimaisuudesta. Suurissa kaupungeissa on myös paljon maahanmuuttajia ja suhteellisen alhainen huoltosuhde. Näin ollen faktori on nimetty **kaupunkimaisuudeksi**. Päihdehuollon suuret nettokustannukset ja matala äänestysprosentti kertovat omalta osaltaan kaupunkimaisuudesta sosiaalisesta ympäristöstä.

**Sosiaaliset ongelmat** –faktorilla latautuivat vahvimmin *avioerojen määrä* ja *alkoholijuomien myynti*. Myös *päihtyneiden säilöönnotot*, *toimeentulotukimenot* ja *väestöntiheys* latautuivat vahvasti. Latautuneet muuttujat, etenkin *alkoholin myynti*, *avioerojen osuus* ja *toimeentulotukimenot*, kertovat sosiaalisista ongelmista. Faktorin voidaankin todeta kuvaavan, *työttömyys* –faktorin ohella, huono-osaisuutta. *Väestöntiheyden* ja *palvelualan työpaikkojen osuuden* vahvat lataukset viittaavat sosiaalisten ongelmien olevan tunnusomaisempia suurille kaupungeille.

**Taulukko 6.5.** Rikostasojen ja faktoreiden korrelaatiomatriisi faktoripisteiden avulla laskettuna. Kaupunkimaiset ja taajaan asutut kunnat.

	Väki- valta- rikok- set	Omai- suus- rikok- set	Lapsi- perheet	Työttö- myys	Koulu- tus ja varalli- suus	Kau- punk- mai- suus	Sosiaa- liset ongel- mat
Väkivaltarikokset	1	,632	-,646	,430	-,314	,306	,538
Omaisuusrikokset	,632	1	-,506	,230	-,018	,586	,652
Lapsiperheet	-,646	-,506	1	-,221	,226	-,246	-,291
Työttömyys	,430	,230	-,221	1	-,272	,227	,455
Koulutus ja varallisuus	-,314	-,018	,226	-,272	1	,375	-,006
Kaupunkimaisuus	,306	,586	-,246	,227	,375	1	,526
Sosiaaliset ongelmat	,538	,652	-,291	,455	-,006	,526	1

Rotaatiomenetelmänä analyysissä käytettiin jälleen vinokulmaista Promax-rotatiota. Taulukosta 6.5 nähdään, että faktoreiden ja rikostasojen väliset korrelaatiot ovat muuttuneet huomattavasti *kaikkien yli tuhannen asukkaan kuntien* analyysistä. *Työttömyys* ja *kaupunkimaisuus* korreloivat edelleen positiivisesti ja *lapsiperheet* negatiivisesti väkivaltarikosten kanssa. *Sosiaaliset ongelmat*, joka on uusi faktori, korreloi erittäin voimakkaasti molempien rikostasojen kanssa. *Koulutus ja varallisuus* –faktori puolestaan korreloi negatiivisesti väkivaltarikosten kanssa, muttei lainkaan omaisuusrikosten kanssa. Omaisuusrikosten suhteen korrelaatioiden suunnat ovat samat, mutta niiden vahvuuksissa on eroavaisuuksia. *Sosiaaliset ongelmat* ja etenkin *kaupunkimaisuus* korreloivat selvästi voimakkaammin omaisuusrikosten kanssa, kun taas *työttömyyden* korrelaatio heikkenee.

Yleisluonteeltaan negatiiviset muuttujat korreloivat myös keskenään, minkä lisäksi *kaupunkimaisuus* korreloi selkeästi *koulutuksen ja varallisuuden* kanssa. Tiivistettynä voidaan todeta, että kun *maaseutumaiset kunnat* jätettiin pois ana-

lyysistä, niin faktoreiden yhteys rikollisuuteen selkiintyi. Erityisesti faktorit *koulutus ja varallisuus* ja *lapsiperheet* ovat selkeämmin yhteydessä alhaisiin rikostasoihin. Ennen kuin tehdään vastaava faktorianalyysi *maaseutumaisten kuntien* ryhmälle, niin tarkastellaan vielä erikseen *kaupunkimaisten ja taajaan asuttujen kuntien* faktorimalleja.

### 6.3 Kaupunkimaiset kunnat

Taulukossa 6.6 on tehty faktorianalyysi *kaupunkimaisten kuntien* ryhmässä, johon kuuluu 63 kuntaa. Helsinki on jälleen poistettu analyysistä poikkeavuutensa takia. Analyysi suoritettiin pääakselimenetelmällä, ja siihen valikoitui lopulta 29 muuttujaa. Analyysiin valittiin 31 muuttujaa, mutta *huumausainerikokset* ja *pitkäaikaistyöttömät* poistettiin pienten kommunaliteettien takia. Analyysi tuotti viiden faktorin mallin, jonka selitysasteeksi saatiin 76,1 %. Löydetyt faktorit ovat *lapsiperheet, koulutus ja varallisuus, työttömyys, sosiaaliset ongelmat ja kaupunkimaisuus*.

Tulokset ovat hämmästyttävän samanlaisia kuin *kaupunkimaisten ja taajaan asuttujen kuntien* tapauksessa. Latausten painotuksetkin ovat niin lähellä toisiinsa, ettei faktoreita ole mielekästä tulkita kohta kohdalta uudelleen. Ainoa olennainen muutos edelliseen kohtaan on faktoreiden järjestys; toinen ja kolmas sekä neljäs ja viides faktori ovat vaihtaneet paikkojaan. *Koulutus ja varallisuus* selittävät siis enemmän *kaupunkimaisten kuntien* varianssista kuin *työttömyys*. Tästä voidaan päätellä koulutuksen ja varallisuuden olevan työttömyyttä tunnusomaisempi piirre *kaupunkimaisille kunnille*.

**Taulukko 6.6.** Kaupunkimaisten kuntien rotaatiomatriisi.

29 muuttujaa	Promax-rotatoitu faktorimalli pääakselimenetelmällä				
	Lapsiperheet	Koulutus ja varallisuus	Työttömyys	Sosiaaliset ongelmat	Kaupunkimaisuus
Pendelöinti	<b>,874</b>	-,061	,192	-,103	,192
Lapsiperheet	<b>-,867</b>	,078	,008	-,226	,181
Kunnassa olevien työpaikkojen lkm	<b>,844</b>	-,119	-,064	-,170	,425
0–15-vuotiaat	<b>-,771</b>	,048	,074	-,377	,051
Maassamuuton tulomuutto	<b>-,662</b>	,196	-,132	,066	,178
Kunnan nettomuutto	<b>-,543</b>	-,040	<b>-,460</b>	,261	,090
Vuokra-asunnot	<b>,464</b>	,237	,277	,291	,330
Koulutustasomittain	-,367	<b>,854</b>	,183	-,097	,054
Valtionveronalaiset tulot	-,095	<b>,848</b>	-,263	,012	-,275
Korkea-asteen koulutetut	-,331	<b>,834</b>	,007	-,043	,042
Maahanmuutto	,397	<b>,639</b>	-,054	-,119	,176
Väestötiheys	-,152	<b>,558</b>	-,304	<b>,464</b>	-,012
Palvelualan työpaikkojen osuus	-,315	<b>,511</b>	,195	,423	,162
Nuorisotyöttömät	,074	-,066	<b>,944</b>	-,065	-,224
Työttömät	,188	-,081	<b>,823</b>	,146	-,157
Sairastavuusindeksi	,033	-,401	<b>,600</b>	,132	-,117
Yritystoimipaikkojen lukumäärä	,403	-,101	<b>-,573</b>	,185	-,109
Kunnan yleinen pienituloisuusaste	,369	-,136	<b>,569</b>	,237	,020
16–24-vuotiaat miehet	,163	,366	<b>,569</b>	,007	,413
Yksinhuoltajaperheet	,201	-,014	-,091	<b>,896</b>	,043
Avioeroja 25–64-vuotiailla	-,166	,004	-,150	<b>,807</b>	,133
Alkoholijuomien myynti	,323	,081	-,016	<b>,747</b>	-,410
Päihtyneiden säilöönnotot	,130	-,014	,335	<b>,522</b>	,087
Toimeentulotukimenot	,054	,050	,229	<b>,500</b>	,275
Väkiluku	,015	,023	-,063	-,040	<b>,895</b>
Huoltosuhte	,150	,120	,096	-,036	<b>-,894</b>
Vieras äidinkieli	,283	,220	-,424	,046	<b>,624</b>
Päihdehuollon nettokustannukset	<b>,461</b>	-,077	-,432	,060	<b>,540</b>
Vaalit, äänestysprosentti	<b>,483</b>	<b>,479</b>	-,251	-,401	<b>-,485</b>

Rotaatiomenetelmänä analyysissa on jälleen käytetty vinokulmaista Promax-rotatiota. Taulukosta 6.7 nähdään, että faktoreiden tulkinat ja korrelaatiot rikostasojen kanssa ovat hämmästyttävän yhtäläiset kuin *kaupunkimaisten ja taajaan asuttujen kuntien* analyysissa. *Sosiaaliset ongelmat* nousi neljänneksi faktoriksi *kaupunkimaisuuden* ohi, ja sen yhteys korkeisiin rikostasoihin onkin noussut hieman. Muut muutokset ovat niin vähäisiä, ettei niitä ole tarpeen esitellä uudelleen.

**Taulukko 6.7.** Rikostasojen ja faktoreiden korrelaatiomatriisi faktoripisteiden avulla laskettuna. Kaupunkimaiset kunnat.

	Väki- valta- rikok- set	Omai- suus- rikok- set	Lapsi- perheet	Koulu- tus ja varalli- suus	Työttö- myys	Sosiaaliset ongelmat	Kau- pun- kimai- suus
Väkivaltarikokset	1	,672	-,644	-,317	,406	,572	,299
Omaisuusrikokset	,672	1	-,498	-,014	,214	,709	,583
Lapsiperheet	-,644	-,498	1	,230	-,185	-,380	-,221
Koulutus ja varallisuus	-,317	-,014	,230	1	-,261	-,028	,381
Työttömyys	,406	,214	-,185	-,261	1	,420	,231
Sosiaaliset ongelmat	,572	,709	-,380	-,028	,420	1	,532
Kaupunkimaisuus	,299	,583	-,221	,381	,231	,532	1

## 6.4 Taajaan asutut kunnat

Taulukossa 6.8 on tehty faktorianalyysi *taajaan asuttujen kuntien* ryhmässä, johon kuuluu 66 kuntaa. Analyysi suoritettiin pääakselimenetelmällä ja siihen valikoitui lopulta 27 muuttujaa. Analyysiin valittiin alun perin 31 muuttujaa, joista neljä poistettiin pienten kommunaliteettien takia. Nämä muuttujat olivat *avioerot, huumausainerikokset, palvelualan työpaikkojen osuus ja väkiluku*. Analyysi tuotti viiden faktorin mallin, jonka selitysasteeksi saatiin 73,0 %. Löydetyt faktorit ovat *lapsiperheet, koulutus ja varallisuus, työttömyys, työssäkäynti ja monikulttuurisuus*

*Taajaan asuttujen kuntien* analyysissä käytettiin rotaatiomenetelmänä suorakulmaista Varimax-rotaatiota. Faktoreiden ei siis annettu korreloida keskenään. Faktorit ovat osittain samoja kuin *kaupunkimaisten ja taajaan asuttujen kuntien* tapauksessa; kolme ensimmäistä faktoria saavat saman tulkinnan, vaikkakin latausten painotukset ovat muuttuneet jonkin verran. *Koulutus ja varallisuus* –faktori on myös kohonnut *työttömyys* –faktoriin edelle. Kaksi viimeistä faktoria sen sijaan ovat muuttuneet huomattavasti. Näiden kahden lisäksi tarkastellaan myös muita faktoreita, koska niiden tulkinta on osittain muuttunut.

**Taulukko 6.8.** Rotatoitu faktorimatriisi taajaan asuttujen kuntien ryhmässä.

27 muuttujaa	Varimax-rotatoitu faktorimalli pääakselimenetelmällä				
	Lapsiperheet	Koulutus ja varallisuus	Työttömyys	Työssä-käynti	Monikulttuurisuus
0–15-vuotiaat	<b>-,925</b>	,294	-,074	,037	-,028
Lapsiperheet	<b>-,851</b>	,445	-,098	-,028	,048
Yksinhuoltajaperheet	<b>,758</b>	,002	,335	-,014	,069
Vuokra-asunnot	<b>,580</b>	,075	<b>,561</b>	,439	,018
Alkoholijuomien myynti	<b>,542</b>	-,109	,185	,268	,076
Kunnan nettomuutto	-,497	,199	-,234	-,422	,312
Alkutuotannon työpaikkojen osuus	-,058	<b>-,791</b>	-,194	,105	-,212
Korkea-asteen koulutetut	-,378	<b>,742</b>	-,192	-,377	-,065
Koulutustasomittain	<b>-,540</b>	<b>,716</b>	-,080	-,262	-,097
Valtionveronalaiset tulot	-,343	<b>,712</b>	-,297	-,402	,148
Huoltosuhte	,041	<b>-,665</b>	-,039	,017	-,155
Kunnan yleinen pienituloisuusaste	,477	<b>-,622</b>	,378	,352	-,019
Väestötiheys	-,288	<b>,609</b>	-,196	-,161	,209
Maassamuuton tulomuutto	-,460	<b>,510</b>	-,024	-,417	,161
Nuorisotyöttömät	,329	-,190	<b>,825</b>	,003	-,146
Työttömät miehet	,379	-,271	<b>,818</b>	-,027	-,098
Yritystoimipaikkojen lukumäärä	,222	-,317	<b>-,792</b>	,279	,057
Toimeentulotukimenot	,404	-,037	<b>,640</b>	,200	,172
Päihtyneiden säilönotot	,439	-,106	<b>,629</b>	,260	,010
Vaalit, äänestysprosentti	,035	,000	<b>-,588</b>	-,039	-,300
Sairastavuusindeksi	,111	<b>-,549</b>	<b>,584</b>	,493	-,045
Kunnassa olevien työpaikkojen lkm	,435	-,032	-,161	<b>,786</b>	,124
16–24-vuotiaat miehet	-,057	-,216	,064	<b>,751</b>	-,235
Pendelöinti	,492	-,306	,147	<b>,725</b>	,040
Pitkäaikaistyöttömät	,145	,314	-,027	<b>-,524</b>	,069
Vieras äidinkieli	,140	,168	,079	-,072	<b>,903</b>
Maahanmuutto	-,030	,133	,017	-,054	<b>,552</b>

**Lapsiperheet** –faktorilla vahvimmin latautuivat *0–15-vuotiaiden osuus, lapsiperheiden osuus ja yksinhuoltajaperheiden osuus lapsiperheistä*. Myös *vuokra-asuntojen osuus, alkoholijuomien myynti ja koulutustasomittain* latautuivat vahvasti. Tulkinnan helpottamiseksi käännetään latauksien etumerkit ja tulkitaan faktori seuraavasti: lapsiperheiden osuus on suuri, yksinhuoltajien osuus pieni ja alkoholin myynti vähäistä. Myös vuokra-asuntoja on vähän, koulutustaso alhainen ja kunnan pienituloisuusaste korkea. Kunnan asukasmäärä on kuitenkin kasvussa, koska muuttoaiheiset muuttajat latautuvat sillä positiivisesti. Kunnassa on myös vähän työpaikkoja, joten pendelöintiä on vähän. Suuri osa asukkaista käy siis töissä oman kunnan ulkopuolella.



**Koulutus ja varallisuus** –faktorilla vahvimmin latautuivat vähäinen *alkutuo-*  
*tannon työpaikkojen osuus, koulutustasomittain, korkeakoulutettujen osuus ja*  
*valtionveronalaiset tulot.* Myös *väestötiheys ja tulomuutto* latautuivat positiivi-  
sesti. Sen sijaan *huoltosuhde, sairastavuusindeksi ja pienituloisuusaste* latautui-  
vat negatiivisesti faktorilla. Toisin sanoen kaikki muuttajat kertovat hyväosaisuus-  
desta. Faktori näyttäisi kuvaavan ennen kaikkea keskeisiä kasvukuntia, joissa  
asukkaiden varallisuus- ja hyvinvointiaste ovat korkeita. Tällaisia kuntia ovat  
etenkin keskisuuret keskuskunnat, joissa koulutusmahdollisuudet ovat hyvät.

**Työttömyys** –faktorilla latautuivat vahvimmin *nuorisotyöttömät, työttömien*  
*osuus ja yritystoimipaikkojen vähäisyys.* Myös *toimeentulotukimenot, sairasta-*  
*vuusindeksi ja päihtyneiden säilöönnotot* latautuivat positiivisesti faktorilla. *Äänes-*  
*tysprosentti* sen sijaan latautui negatiivisesti. Latautuneet muuttajat kertovat  
yleisemmin korkean työttömyysasteen kuntien tilanteesta; työttömyys, työpaik-  
kojen vähyys, toimeentulo- ja terveysongelmat ovat selkeästi yhteydessä toisiin-  
sa. Matala äänestysprosentti taas kertoo osaltaan yhteisön sosiaalisen kiinteyden  
puutteesta. Faktori onkin luonteeltaan selvästi negatiivinen ja kuvastaa kuntien  
ongelmia ja huono-osaisuutta.

Neljäs faktori on tulkittu **työssäkäynniksi**, koska sillä latautuivat vahvimmin  
*kunnassa olevien työpaikkojen määrä ja pendelöinti.* Myös suuri *nuorten miesten*  
*osuus ja vähäinen pitkäaikaistyöttömien määrä* latautuivat vahvasti faktorilla.  
Muista latauksista nähdään, että talous- ja terveysongelmat ovat tunnusomaisia  
*työssäkäynti* –faktorille. Myös muuttotappio ja vuokra-asuntojen suuri osuus  
kuvaavat faktorin luonnetta. Faktorin yksiselitteinen tulkinta onkin hankalaa.  
Useat muuttajat puhuvat kaupunkimaisuuden puolesta, mutta osa selvästi sitä  
vastaan. Tästä syystä faktori nimettiin suoraan sen vahvimpien latausten mukai-  
sesti.

**Monikulttuurisuus** –faktorin tulkinta on suhteellisen yksinkertainen, koska sil-  
lä latautuivat vahvasti ainoastaan muuttajat *vieras äidinkieli ja maahanmuutto.*  
*Taajaan asuttujen kuntien* ryhmässä monikulttuurisuus näyttäisi siis erottuvan  
omaksi kokonaisuudekseen, kun taas muissa ryhmissä monikulttuurisuutta  
kuvaavat muuttajat sekoittuivat eri faktoreihin. *Kaikkien yli tuhannen asukkaan*  
*kuntien* analyysissä monikulttuurisuus oli vahvasti esillä, mutta tällöin viidennen  
faktorin tulkinta oli *alkoholi ja monikulttuurisuus.* Myös *nettomuutto* ja matala  
*äänestysprosentti* latautuivat faktorilla, mikä tukee valittua tulkintaa. Matala  
*äänestysprosentti* kertoo osaltaan yhteisön sosiaalisen kiinteyden puutteesta,  
jota monikulttuurisuus luonnollisesti lisää.

**Taulukko 6.9.** Rikostasojen ja faktoreiden korrelaatiomatriisi faktoripisteiden avulla laskettuna. Taajaan asutut kunnat.

	Väkivalta- rikokset	Omaisuus- rikokset
Lapsiperheet	-,355	-,280
Koulutus ja varallisuus	-,078	,313
Työttömyys	,320	,103
Työssäkäynti	,456	,077
Monikulttuurisuus	,195	,207

Taulukosta 6.9 nähdään, että faktoreiden ja rikostasojen väliset korrelaatiot ovat heikompia kuin edellisissä analyyseissa, mikä johtunee osittain suorakulmaisesta rotatoinnista. Faktorit eivät siis korreloi keskenään lainkaan. *Työssäkäynti* ja *työttömyys* korreloivat positiivisesti ja *lapsiperheet* negatiivisesti väkivaltarikosten kanssa. Omaisuusrikosten kanssa vahvimmin korreloi *koulutus ja varallisuus*. *Lapsiperheet* –faktorilla on jälleen negatiivinen yhteys korkeisiin rikostasoihin. Muiden faktoreiden ja rikostasojen väliset korrelaatiot eivät ole merkityksellisiä.

## 6.5 Maaseutumaiset kunnat

*Maaseutumaisten kuntien* ryhmässä analyysin tekeminen osoittautui huomattavasti hankalammaksi kuin muissa ryhmissä. Vaikka analyysistä jätettiin pois alle tuhannen asukkaan kunnat, on ryhmä varsin hajanainen. Tästä kertoo myös alkuperäisestä mallista poistettujen muuttujien suuri määrä. Erittäin pienissä kunnissa havainnot saattavatkin vaihdella huomattavasti vuosittain tai jäädä yksinkertaisesti olemattomaksi. Näin ollen vakaiden analyysien ja mallien tekeminen on huomattavasti hankalampaa kuin suurempien kuntien tapauksissa.

Analyysi suoritettiin pääakselimenetelmällä ja siihen valikoitui lopulta 23 muuttujaa (Taulukko 6.10). Alun perin analyysiin valittiin 30 muuttujaa, joista kuitenkin peräti seitsemän poistettiin pienten kommunaliteettien tai muiden ongelmien takia. Nämä muuttujat olivat *avioerot*, *huoltosuhde*, *huumausainerikokset*, *moottoriajoneuvot*, *nettomuutto*, *päihdehuollon kustannukset* ja *väkiluku*. Analyysi tuotti lopulta viiden faktorin mallin, jonka selitysasteeksi saatiin 67,7 %. Löydetyt faktorit ovat *koulutus ja varallisuus*, *työttömyys ja sosiaaliset ongelmat*, *työssäkäynti*, *lapsiperheet* sekä *työttömyys ja matkailu*.

**Taulukko 6.10.** Rotaatiomatriisi maaseutumaisten kuntien ryhmässä.

23 muuttujaa	Varimax-rotatoitu faktorimalli pääakselimenetelmällä				
	Koulutus ja varallisuus	Työttömyys ja sosiaaliset ongelmat	Työssä-käynti	Lapsiperheet	Työttömyys ja matkailu
Valtionveronalaiset tulot	,903	-,215	-,154	,062	-,176
Korkea-asteen koulutetut	,879	-,190	-,136	,049	,148
Koulustustasomittain	,862	-,091	-,122	,258	,196
Kunnan yleinen pienituloisuusaste	-,756	,347	,084	-,199	,079
Väestöntiheys	,624	-,041	-,132	,141	-,372
Maassamuuton tulomuutto	,605	,107	-,280	-,074	,053
Alkutuotannon työpaikkojen osuus	-,586	-,071	-,269	,014	-,370
Vaalit, äänestysprosentti	-,044	-,816	,013	,120	,154
Nuorisotyöttömät	-,284	,615	-,020	-,121	,411
Sairastavuusindeksi	-,549	,599	,023	-,012	,145
Päihtyneiden säilöönnot	-,187	,583	,200	-,089	,198
Työttömät	-,441	,537	,009	-,231	,511
Yritystoimipaikkojen lukumäärä	-,146	-,534	,209	-,169	-,466
Kunnassa olevien työpaikkojen lkm	-,058	-,112	,957	-,025	-,155
Pendelöinti	-,385	,160	,880	-,012	,010
Vuokra-asunnot	-,374	,431	,491	-,144	,296
0–15-vuotiaat	,459	-,005	-,249	,739	-,050
Lapsiperheet	,597	-,025	-,196	,669	-,060
16–24-vuotiaat miehet	-,185	-,235	,306	,652	,009
Yksinhuoltajaperheet	-,054	,258	,284	-,549	,137
Kesämökkien lukumäärä	-,112	-,087	-,093	-,526	,316
Palvelualan työpaikkojen osuus	,128	,149	-,085	-,180	,616
Alkoholijuomien myynti	,024	,096	,411	-,346	,472

*Maaseutumaisten kuntien* analyysissä käytettiin rotaatiomenetelmänä suorakulmaista Varimax-rotatiota. Faktoreiden ei siis annettu korreloida keskenään. Voidaan huomata, että useat faktorit on nimetty samoin kuin *kaupunkimaisten kuntien ja taajaan asuttujen kuntien* tapauksessa. Kaikille yhteiset faktorit ovat *lapsiperheet* sekä *koulutus ja varallisuus*, eikä näiden faktoreiden sisältökään ole olennaisesti muuttunut. Työttömyys sen sijaan jakautui kahdelle faktorille. Toinen faktori, *työttömyys ja sosiaaliset ongelmat*, vastaa paremmin edellisten analyysien faktoria *työttömyys; pienituloisuusaste, sairastavuusindeksi ja päihtyneiden säilöönnot* kertovat sosiaalisista ongelmista ja alhainen *äänestysprosentti* sosiaalisen kiinteyden puutteesta.

*Työttömyys ja matkailu* –faktorin tulkinnessa on sen sijaan käytetty mielikuvi- tusta. Faktorilla latautuvien *palvelualan työpaikkojen osuuden, alkoholin myyn- nin, kesämökkien lukumäärän* ja alhaisen *väestöntiheyden* on oletettu viittaavan matkailun suureen määrään kunnan alueella. *Maaseutumaisista kunnista* useat ovatkin niin sanottuja mökkikuntia, joissa lomailijoiden määrä on suhteellisen suuri.

Kolmas faktori, *työssäkäynti*, taas vastaa lähes täysin *taajaan asuttujen kuntien* ryhmän vastaavaa faktoria. Sillä latautuivat vahvasti *kunnassa olevien työpaikkojen määrä, pendelöinti ja vuokra-asuntojen osuus*, joten sen tulkinta on suhteellisen yksiselitteinen.

Taulukosta 6.11 nähdään, että *työttömyys ja sosiaaliset ongelmat* korreloi vahvimmin väkivaltarikosten kanssa. Myös faktorit *työssäkäynti* ja *työttömyys ja matkailu* korreloivat positiivisesti väkivaltarikosten kanssa. *Lapsiperheet* ja *koulutus ja varallisuus* –faktorit sen sijaan korreloivat negatiivisesti väkivaltarikosten kanssa, mutta korrelaatiot ovat heikkoja. Omaisuusrikosten kanssa vahvimmin, ja jälleen negatiivisesti, korreloi faktori *lapsiperheet*. *Koulutus ja varallisuus, työttömyys ja sosiaaliset ongelmat* sekä *työttömyys ja matkailu* korreloivat positiivisesti omaisuusrikosten kanssa.

**Taulukko 6.11.** Rikostasojen ja faktoreiden korrelaatiomatriisi faktoripisteiden avulla laskettuna. Maaseutumaiset kunnat.

	Väkivalta- rikokset	Omaisuus- rikokset
Koulutus ja varallisuus	-,135	,215
Työttömyys ja sosiaaliset ongelmat	,473	,273
Työssäkäynti	,233	,117
Lapsiperheet	-,161	-,386
Työttömyys ja matkailu	,240	,226

## 7 Tutkimustulosten analysointi

Alla on vertailtu yksittäisten kuntien väkivalta- ja omaisuusrikostasoja. Yllättävän harvassa kunnassa molemmat rikostasot olivat poikkeuksellisen korkeita tai matalia. Kuntia, joissa molemmat rikostasot olivat vähintään melko korkeita tai melko matalia, on hieman enemmän. *Maaseutumaisten kuntien* kohdalla tällaisia kuntia oli suhteellisestikin mitattuna eniten. On kuitenkin syytä muistaa, että niiden kohdalla mallit olivat selitysteeltaan huomattavasti matalampia.

*Kaupunkimaisista kunnista* ainoastaan Hangossa sekä väkivalta- että omaisuusrikostaso oli poikkeuksellisen korkea. *Taajaan asutuista kunnista* Ii, Laitila, Loimaa, Naantali ja Orimattila puolestaan erottuivat korkeilla rikostasoiltaan. Vastaavasti *maaseutumaisista kunnista* Getassa, Perhossa, Sotkamossa, Suomensiemellä, Taivassalossa ja Tervolassa rikostasot osoittautuivat poikkeuksellisen korkeiksi molemmissa ennustemalleissa.

Useissa kunnissa rikostasot olivat vähintään melko korkeita sekä väkivalta- että omaisuusrikoksissa. *Kaupunkimaisista kunnista* ainoastaan Kajaani ja Pietarsaari sekä *taajaan asutuista kunnista* Kokemäki, Kuusamo, Raasepori ja Uusikauunki kuuluivat tähän joukkoon. *Maaseutumaisista kunnista* tällaisia olivat Alajärvi, Eckerö, Keitele, Kinnula, Kärkölä, Miehikkälä, Nousiainen, Oravainen, Pomarkku, Pälkäne, Pöytyä, Rääkkylä, Siikajoki, Siikalatva, Urjala ja Virolahti.

Vastaavasti *kaupunkimaisia kuntia*, joissa rikostaso oli matala tai melko matala molemmissa rikostyypeissä, olivat Kaskinen, Kauniainen, Maarianhamina, Mikeli, Nurmijärvi, Salo ja Vantaa. *Taajaan asuttujen kuntien* kohdalla vaihtelu oli suurempaa, koska tällaisia kuntia oli vain neljä: Haapajärvi, Kannus, Sipoo ja Sodankylä. Vähintään melko matalien rikostasojen *maaseutumaisia kuntia* puolestaan on peräti 19: Föglö, Haijoki, Hankasalmi, Honkajoki, Jämijärvi, Karvia, Kihniö, Kruunupyy, Lemland, Marttila, Rautjärvi, Ruokolahti, Sievi, Tervo, Teuva, Toholampi, Toivakka, Vesilahti ja Ylämaa.

Matalan väkivalta- ja omaisuusrikostason kuntia olivat Hämeenkyrö ja Ilmajoki *taajaan asutuista kunnista* sekä Kumlinge, Sottunga ja Vårdö *maaseutumaisista kunnista*. *Kaupunkimaisista kunnista* yksikään ei saanut matalan rikostason luokitusta molemmissa rikostyypeissä. Osasyynä tähän on se, että väkivaltarikoksissa ainoastaan kaksi kuntaa sai matalan luokituksen.

Regressiomallien lisäksi kuntien rikollisuuden rakenteita pyrittiin selvittämään myös **faktorianalyysin** avulla. Faktorianalyysi tehtiin viidessä eri ryhmässä: *kaikki yli tuhannen asukkaan kunnat, kaupunkimaiset ja taajaan asutut kunnat, kaupunkimaiset kunnat, taajaan asutut kunnat ja maaseutumaiset kunnat*. Tarkoituksena oli löytää kuntaryhmien taustalla piileviä rakenteita, jotka voisivat selittää rikollisuutta kunnissa. Faktorianalyysin avulla löydettiin yllättävänkin selkeitä rakenteita muuttujien välillä. Myös faktoreiden tulkinta osoittautui suhteellisen yksinkertaiseksi.

Faktorianalyysit tehtiin pääkselimenetelmällä ja rotaatioissa käytettiin Varimax- ja Promax-menetelmiä. *Taajaan asuttujen ja maaseutumaisten kuntien* ryhmissä käytettiin ortogonaalista Varimax-rotatiota, jossa faktoreiden ei anneta korreloida keskenään. Muissa tapauksissa faktoreiden sallittiin korreloida keskenään. Kaikissa ryhmissä parhaaksi faktoreiden lukumääräksi osoittautui viisi.

Faktorianalyysi tehtiin ensin *kaikille yli tuhannen asukkaan kunnille*. Analyysi tuotti viiden faktorin mallin, jonka selitysasteeksi saatiin 69 %. Löydetyt faktorit olivat *koulutus ja varallisuus, työttömyys, kaupunkimaisuus, lapsiperheet sekä alkoholi ja monikulttuurisuus*. Kun *maaseutumaiset kunnat* poistettiin analyysistä, selitysaste nousi 75 prosenttiin. Faktorit pysyivät viidennettä lukuun ottamatta samoina; *alkoholi ja monikulttuurisuus* –faktori korvautui faktorilla *sosiaaliset ongelmat*. *Alkoholin myynti* latautui edelleen vahvasti tällä faktorilla, mutta *vieras äidinkieli* latautui vahvasti *kaupunkimaisuus* –faktorilla. Vaikka faktorien tulkinnot pysyivät pitkälti samoina, niiden järjestys ja lataukset muuttuivat. *Lapsiperheet* –faktori nousikin neljänneltä sijalta ensimmäiseksi. Tämä on oleellista siksi, että ensimmäisten faktoreiden selitysosuudet ovat aina korkeampia kuin viimeisten.

*Kaupunkimaisten kuntien* ryhmässä faktorianalyysin selitysaste nousi 76 prosenttiin. Tulokset olivat hämmästyttävän samanlaisia kuin *kaupunkimaisten ja taajaan asuttujen kuntien* tapauksessa. Ainoa olennainen muutos edelliseen kohtaan oli faktoreiden järjestys; toinen ja kolmas sekä neljäs ja viides faktori vaihtoivat paikkojaan. *Koulutus ja varallisuus* selittävät siis enemmän *kaupunkimaisten kuntien* varianssista kuin *työttömyys*. Tästä voidaan päätellä koulutuksen ja varallisuuden olevan työttömyyttä tunnusomaisempi piirre *kaupunkimaisille kunnille*.

*Taajaan asuttujen kuntien* ryhmässä faktorianalyysi tuotti viiden faktorin mallin, jonka selitysasteeksi saatiin 73 %. Löydetyt faktorit olivat *lapsiperheet, koulutus ja varallisuus, työttömyys, työssäkäynti ja monikulttuurisuus*. Faktorit olivat siis osittain samanlaisia kuin *kaupunkimaisten ja taajaan asuttujen kuntien* tapauksessa. Kolme ensimmäistä faktoria saivat saman tulkinnan, vaikkakin latausten painotukset muuttuivat jonkin verran. *Koulutus ja varallisuus* –faktori kohosi myös *työttömyys* –faktorin edelle. Kaksi viimeistä faktoria sen sijaan muuttuivat huomattavasti.

*Maaseutumaisten kuntien* ryhmässä analyysin tekeminen osoittautui huomattavasti hankalammaksi kuin muissa ryhmissä. Malliin valikoituikin vähemmän muuttujia ja selitysaste tippui 68 prosenttiin. Löydetyt faktorit olivat *koulutus ja varallisuus, työttömyys ja sosiaaliset ongelmat, työssäkäynti, lapsiperheet sekä työttömyys ja matkailu*. Faktoreiden nimistä nähdään, että tulkintakin oli hankalampaa kuin muissa ryhmissä. Esimerkiksi työttömyys jakautui kahdelle faktorille. Toinen faktori, *työttömyys ja sosiaaliset ongelmat*, vastasi paremmin edellisten analyysien faktoria *työttömyys*. Kolmas faktori, *työssäkäynti*, taas vastasi lähes täysin *taajaan asuttujen kuntien* ryhmän vastaavaa faktoria. *Lapsiperheet* sekä *koulutus ja varallisuus* –faktorit sen sijaan löytyivät kaikista malleista.

Faktorit *lapsiperheet* ja *koulutus ja varallisuus* löytyivät kaikista malleista. Myös *työttömyys* esiintyi kaikissa malleissa, tosin viimeisessä se jakautui kahdelle faktorille. *Lapsiperheet* oli ainoa faktori, joka korreloi kaikissa malleissa negatiivisesti rikostasojen kanssa. *Koulutus ja varallisuus* –faktorin luonne sen sijaan vaihteli; väkivaltarikosten kanssa se korreloi useimmiten negatiivisesti, mutta omaisuusrikoksiin sillä vaikutti olevan positiivinen yhteys. Tämä tukee teoriaa, jonka mukaan omaisuusrikollisuus kasvaa vaurilla alueilla potentiaalisten anastuskohteiden lisääntyessä.

Yleisluonteeltaan negatiiviset faktorit *työttömyys*, *kaupunkimaisuus* ja *sosiaaliset ongelmat* korreloivat positiivisesti korkeiden rikostasojen kanssa. Erityisesti huomataan, että *työttömyys* korreloi voimakkaammin väkivaltarikosten ja *kaupunkimaisuus* ja *sosiaaliset ongelmat* omaisuusrikosten kanssa. Muut faktorit korreloivat pääsääntöisesti lievästi rikostasojen kanssa. Poikkeuksena faktorilla *työssäkäynti* vaikuttaisi olevan selkeä positiivinen yhteys korkean väkivaltarikostason kanssa.

## 8 Yhteenveto

Tutkimuksen tavoitteena oli selvittää millainen kuntien rikostasoa on, kun erilaiset rakenteelliset tekijät on otettu huomioon, ja miten se on muuttunut 2000-luvun alusta. Väkivalta- ja omaisuusrikollisuutta mallinnettiin regressioanalyysin avulla ja ennustettua tasoa verrattiin poliisin tilastoimaan rikostasoon. Erojen perusteella kunnat jaettiin viiteen ryhmään. Selittäjinä toimivat erilaiset kuntien väestörakennetta, työssäkäyntiä, taloudellista ja sosiaalista rakennetta kuvaavat muuttujat. Suuri osa muuttujista oli samoja kuin Heiskasen tutkimuksessa (Heiskanen et al. 2003), mutta mukaan otettiin myös joitakin muita sopivaksi katsottuja muuttujia.

Kunnat jaettiin kolmeen ryhmään tilastokeskuksen laatiman luokituksen mukaisesti. Jokaiselle ryhmälle sovitettiin oma regressiomalli, koska *kaupunkimaaiset, taajaan asutut ja maaseutumaiset kunnat* ovat rakenteeltaan hyvin erilaisia. Mallien avulla kuntien rikostasoa voitiin arvioida ja vertailla omissa ryhmissään. Malleihin valikoituvista selittäjistä voitiin myös tehdä varovaisia johtopäätöksiä rikollisuuden suhteen. Kausaalisuhteiden esittämisessä tulee kuitenkin olla varovainen, koska malleihin liittyy aina rajoituksia. Selittäjiä eliminoiva menetelmä saattaa tuottaa hyvinkin erilaisia malleja analyysiin valittavista muuttujista riippuen. Malleissa ei olekaan ensisijaisesti haettu korkeinta mahdollista selitystasoa, vaan niihin valittiin ennen kaikkea kriminologisesti perusteltuja muuttujia.

Mallien selitystasot vaihtelivat välillä 0,48–0,80. *Kaupunkimaisten kuntien* malleissa selitystasot olivat huomattavasti korkeampia kuin muissa ryhmissä. Vastaavasti matalimmat selitystasot saatiin *maaseutumisten kuntien* malleissa. *Taajaan asuttujen kuntien* ryhmässä väkivaltarikosten malli osoittautui selkeästi paremmaksi kuin omaisuusrikosten malli. Selitystasojen huomattavat erot johtuvat ennen kaikkea suurista rakente-eroista kuntaryhmien sisällä. Etenkin *maaseutumaiset kunnat* ovat keskenään hyvin erilaisia.

Väkivalta- ja omaisuusrikollisuutta selittäviin malleihin valikoitui osittain samoja muuttujia kuin Heiskasen tutkimuksessa. Etenkin *kaupunkimaisten kuntien* väkivaltarikollisuutta selittävässä malleissa on paljon yhtäläisyyksiä, mikä vahvistaa käsitystä *kaupunkimaisten kuntien* väkivaltarikollisuuden taustalla olevista rakenteista. Yhteisiä muuttujia olivat *palvelualan työpaikkojen osuus, pendelöinti, alkoholin myynti, pienituloisuusaste ja korkea-asteen koulutus*.

Kaikkia uusia malleja yhdistää alkoholin käyttöön liittyvä muuttuja. Etenkin väkivaltarikoksissa alkoholilla on keskeinen asema suurimmassa osassa tapauksista. *Päihtyneiden säilönotot* oli mukana kaikissa väkivaltarikosmalleissa ja latautui niissä erittäin vahvasti. Alkoholin keskeinen asema rikollisuuteen ilmeni myös Heiskasen tutkimuksesta. Myös muuttoliikkeellä vaikuttaisi olevan yhteys väkivaltarikostasoihin; tulomuuttoa koskeva muuttuja löytyy *taajaan asuttujen kuntien* mallista, ja *maahanmuutto* selittää osittain *kaupunkimaisten kuntien*



rikollisuutta. *Taajaan asuttujen kuntien* mallin mukaan *vuokra-asuntojen* korkea osuus on yhteydessä väkivaltarikollisuuteen. *Maaseutumaisten kuntien* väkivalta-tasoa taas selittävät toimeentulovaikeudet ja lapsiperheiden osuus.

Omaisuusrikollisuutta selittäviin malleihin valikoitui osittain samoja muuttujia kuin väkivaltarikosmalleihin. Alkoholien lisäksi myös huumausaineilla vaikuttaisi olevan yhteys omaisuusrikostasoon *kaupunkimaisissa ja taajaan asutuissa kun-nissa*. Myös perhesuhteet olivat jälleen keskeisessä asemassa; *avioerot* selittivät kaikkia malleja. *Kaupunkimaisissa kunnissa* myös *alle 15-vuotiaiden osuus* selitti omaisuusrikostasoa. Koulutustasolla oli selkeästi sekä väkivalta- että omaisuus-rikostasoa alentava vaikutus *kaupunkimaisissa kunnissa*.

*Taajaan asuttujen kuntien ja maaseutumaisten kuntien* malleihin valikoitui lisäksi kaksi asukkaiden ansiotasoon liittyvää muuttujaa. *Taajaan asutuissa kun-nissa vuokra-asuntojen* korkea osuus kasvatti omaisuusrikostasoa. *Maaseutu-maisten kuntien* ryhmässä taas *kesämökkien lukumäärällä* vaikuttaisi olevan yhteys omaisuusrikostasoon. Kesämökit joutuvatkin usein omaisuusrikosten koh-teeksi, minkä lisäksi suuri mökkien määrä kertoo myös ulkopaikkakuntalaisten korkeasta määrästä.

Regressiomallilla ennustettua rikostasoa verrattiin poliisin tilastoimaan rikos-tasoon, minkä jälkeen kunnat jaettiin viiteen ryhmään. Jos ennuste poikkesi tilas-toidusta tasosta alle kymmenen prosenttia, pidettiin rikostasoa odotettuna. Yli 25 prosentin poikkeamia pidettiin korkean ja matalan rikostason rajoina. Jos poikkeama oli välillä 10–25 %, tulkittiin rikostaso melko korkeaksi tai melko matalaksi. *Maaseutumaisten kuntien* kohdalla rajat asetettiin korkeammiksi, kos-ka kuntien suurista eroista johtuen myös poikkeamat olivat suurempia. Useissa kunnissa rikosten määrä oli myös niin vähäinen, että rikostason mallintaminen oli ongelmallista.

Pääpiirteissään ennusteet vaikuttaisivat noudattavan tilastoituja rikostasoja siten, että korkean rikostason kunnissa myös tilastoitu rikostaso on korkea ja päinvastoin. Poikkeuksia on kuitenkin paljon, joista äärimmäinen tapaus on Maarianhamina. Se on tilastoidulta rikostasoltaan ylivoimaisesti korkein, mutta sijoit-tuu kuitenkin melko matalan rikostason ryhmään. Maarianhamina on kuitenkin poikkeustapaus laivaliikenteessä tapahtuvan rikollisuuden johdosta, ja se poistet-tiinkin regressiomalleista poikkeavana havaintona.

Kunnan koolla ei juurikaan vaikuttanut olevan merkitystä väkivaltaluokituk-seen. Suurimmista kaupungeista suurin osa kuului odotetun väkivaltatason ryh-mään, mutta Tampereella ja Vantaalla väkivaltataso oli melko matala. Heiskasen tutkimuksessa Tampereen rikostaso oli vieläkin alhaisempi eli matala. Myös omaisuusrikostasollaan suurimmat kaupungit sijoittuivat odotetun rikostason ryhmään. Poikkeukset olivat Oulu korkealla ja Vantaa melko matalalla rikostasol-laan. Oulun omaisuusrikostaso oli melko korkea myös Heiskasen tutkimuksessa.

*Kaupunkimaisista kunnista* ainoastaan Hangossa sekä väkivalta- että omaisuusrikostasotaso oli poikkeuksellisen korkea. *Taajaan asutuista kunnista* Ii, Laitila, Loimaa, Naantali ja Orimattila puolestaan erottuivat korkeilla rikostasoillaan. Vastaavasti *maaseutumaisista kunnista* Getassa, Perhossa, Sotkamossa, Suommenniemellä, Taivassalossa ja Tervolassa rikostasot osoittautuivat poikkeuksellisen korkeiksi molemmissa ennustemalleissa. Matalan väkivalta- ja omaisuusrikostason kuntia olivat Hämeenkyrö ja Ilmajoki *taajaan asutuista kunnista* sekä Kumlinge, Sottunga, Toivakka ja Vårdö *maaseutumaisista kunnista*. *Kaupunkimaisista kunnista* yksikään ei saanut matalan rikostason luokitusta molemmissa rikostyypeissä.

Kuntien rikollisuuden taustalla olevia rakenteita pyrittiin selvittämään myös faktorianalyysin avulla. Faktorianalyysi tehtiin viidelle eri ryhmälle: *kaikki yli tuhannen asukkaan kunnat, kaupunkimaiset ja taajaan asutut kunnat, kaupunkimaiset kunnat, taajaan asutut kunnat ja maaseutumaiset kunnat*. Analyysiin valittiin alun perin noin 30 valinnaista muuttujaa, mutta lopullisiin malleihin valikoitui 23–29 muuttujaa. Kaikissa ryhmissä parhaaksi faktoreiden lukumääräksi osoittautui viisi. Mallien selitysasteet vaihtelivat välillä 63–76 prosenttia. Parhaimmat selitysasteet saatiin *kaupunkimaisten kuntien* ryhmässä, kuten regressiomallienkin tapauksessa. *Maaseutumaisten kuntien* ryhmän suuret vaihtelut tuottivat hankaluuksia myös faktorimallin muodostamisessa, ja selitysaste tipuikin 68 prosenttiin.

Eri ryhmien faktoreilla oli paljon yhtäläisyyksiä keskenään. Tulee kuitenkin huomioida, että samoin nimetyt faktorit saattavat olla latauksiltaan erilaisia. Kaksi samoin nimettyä faktoria löytyi kaikista malleista: *koulutus ja varallisuus* sekä *lapsiperheet*. Myös *työttömyys, kaupunkimaisuus* ja *sosiaaliset ongelmat* –faktorit esiintyivät useissa malleissa. Faktoreiden suhteellisen selkeät tulkinnot kertovat omalta osaltaan kuntien rakenteista.

Faktorianalyysi tehtiin ensin *kaikille yli tuhannen asukkaan kunnille*. Löydetyt faktorit olivat *koulutus ja varallisuus, työttömyys, kaupunkimaisuus, lapsiperheet* sekä *alkoholi ja monikulttuurisuus*. Viimeisen faktorin luonne on ongelmallinen, koska sillä latautuivat vahvasti ainoastaan *alkoholin myynti* ja *vieras äidinkieli*. Muissa malleissa nämä muuttujat jakoutuivat eri faktoreille. *Kaupunkimaisten kuntien* faktorit olivat lähes samanlaiset kuin *kaupunkimaisten ja taajaan asuttujen kuntien* ryhmässä. Näistä malleista löytyvät kaikki edellä mainitut faktorit.

*Taajaan asuttujen kuntien* ryhmässä samoin nimettyjä faktoreita olivat *lapsiperheet, koulutus ja varallisuus* sekä *työttömyys*. Muut faktorit olivat *työssäkäynti* ja *monikulttuurisuus*. Samoista tulkinnoista huolimatta latausten painotukset muuttuivat ja *koulutus ja varallisuus* –faktori kohosi *työttömyys* –faktorin edelle. *Maaseutumaisten kuntien* mallissa löydetyt faktorit olivat *koulutus ja varallisuus, työttömyys ja sosiaaliset ongelmat, työssäkäynti, lapsiperheet* sekä *työttömyys ja matkailu*. Faktoreiden nimistä nähdään, että tulkinta oli hankalampaa kuin muissa ryhmissä, ja esimerkiksi *työttömyys* jakautui kahdelle faktorille.

Lopuksi pohdittiin faktoreiden välisiä korrelaatioita ja yhteyksiä rikollisuuteen. Kolmen ensimmäisen faktorimallin muodostamisessa käytettiin vinoja rotaatioita, jotka sallivat faktoreiden korreloida keskenään. *Taajaan asuttujen ja maaseutumaisien kuntien* malleissa selkeämpi tulkinta löydettiin ortogonaalisella rotaatiolla. Faktoreiden välisiä korrelaatioita tarkastelemalla voitiin tutkia faktoreiden luonnetta, ja vertailla niiden välisiä yhteyksiä. Faktoripisteiden avulla pystyttiin myös tarkastelemaan faktoreiden ja rikostasojen välisiä korrelaatioita.

*Lapsiperheet* oli ainoa faktori, joka korreloi kaikissa malleissa negatiivisesti molempien rikostasojen kanssa. *Koulutus ja varallisuus* –faktorin luonne sen sijaan vaihteli; väkivaltarikosten kanssa se korreloi useimmiten negatiivisesti, mutta omaisuusrikoksiin sillä vaikutti olevan positiivinen yhteys. Yleisluonteeltaan negatiiviset faktorit *työttömyys, kaupunkimaisuus ja sosiaaliset ongelmat* korreloivat positiivisesti korkeiden rikostasojen kanssa; *työttömyys* voimakkaammin väkivaltarikosten ja *kaupunkimaisuus* sekä *sosiaaliset ongelmat* omaisuusrikosten kanssa. Muut faktorit korreloivat huomattavasti lievemmin rikostasojen kanssa. Poikkeus oli *työssäkäynti* –faktori, jolla vaikuttaisi olevan selkeä positiivinen yhteys korkeaan väkivaltarikostasoon.

# Lähdeluettelo

Cook, R. D. (1977). "Detection of Influential Observations in Linear Regression". *Technometrics*, Vol. 19, No. 1, pp. 15-18.

Draper, N. R. & Smith, H. (1998), "Applied regression analysis", 3rd Edition, New York: Wiley.

Heiskanen, M., Siren, R. & Roivainen, O. (2003), "Rikollisuus kunnissa: Tilastollinen tutkimus Suomen kuntien väkivalta- ja omaisuusrikollisuuteen vaikuttavista tekijöistä", Edita Prima Oy, Helsinki.

Hendrickson, A. E. & White, P. O. (1964), "Promax: A quick method for rotation to oblique simple structure", *The British Journal of Statistical Psychology*, 17, 65–70.

Kivivuori, J. & Salmi, V. (2005), "Nuorten rikoskäyttäytyminen 1995–2004", Helsinki: Oikeuspoliittisen tutkimuslaitoksen julkaisuja 214.

Ledermann, W. (1938), "Shortened method of estimation of mental factors by regression", *Nature*, 141, 650.

Mellin, I. (2008), "Monimuuttujamenetelmät: faktorianalyysi.", WWW-dokumentti. Saatavilla Internetistä: <https://noppa.tkk.fi/noppa/kurssi/mat-2.3112/luennot/Mat 2 3112 faktorianalyysi.pdf>. (luettu 12.4.2010)

Metsämuuronen, J. (2003), "Tutkimuksen tekemisen perusteet ihmistieteissä", Helsinki: International Methelp Ky.

Mustonen, S. (1995), "Tilastolliset monimuuttujamenetelmät", Helsinki: Survo Systems. Saatavilla Internetistä: <http://www.survo.fi/mustonen/monim.pdf>. (luettu 12.4.2010)

Myllymäki, P., Silander, T., Tirri, H. & Uronen, P. (2002), "A Web-Based Tool for Bayesian and Causal Data Analysis. *International Journal on Artificial Intelligence Tools*, Vol.11, No. 3, 369-387. Saatavilla Internetistä: <http://cosco.hiit.fi/Articles/ijait02.pdf>. (luettu 27.4.2010)

Nummenmaa, T., Konttinen, R., Kuusinen, J. & Leskinen, E. (1997), "Tutkimusaineiston analyysi", WSOY kirjapainoyksikkö, Porvoo.

Oikeuspoliittinen tutkimuslaitos 2006, Siren, R. & Honkatukia, P., "Rikollisuustilanne 2005: väkivaltarikokset ja alkoholi", Helsinki: Oikeuspoliittisen tutkimuslaitoksen julkaisuja 220.

Oikeuspoliittinen tutkimuslaitos 2008, Siren, R. & Hinkkanen, V., "Rikollisuustilanne 2007: varkausrikokset ja luvattomat käytöt", Helsinki: Oikeuspoliittisen tutkimuslaitoksen julkaisuja 238.

Rummel, R. J. (1970), "Applied Factor Analysis", Evanston: Northwestern University Press.

Siren, R. & Honkatukia, P. (2005), "Suomalaiset väkivallan uhreina", Helsinki: Oikeuspoliittisen tutkimuslaitoksen julkaisuja 216.

Spearman, C. (1904), "General Intelligence, objectively determined and measured", American Journal of Psychology 15, 201–293

Tolvanen, M. (2005), "Johdatus kriminaalipolitiikan tutkimukseen", Joensuu: Joensuun yliopiston oikeustieteellisiä julkaisuja / n:o 14.

## Liite A: Muuttujaluettelo

Nimi	Kuvaus ja lähde
Väkivaltarikokset	Väkivaltarikokset: murhat ja tapot sekä henkeen ja terveyteen kohdistuneet rikokset. Vuosien 2005–2007 keskiarvo / keskiväkiluku 2005–2007. / 10 000 asukasta. Sotkanet.
Omaisuusrikokset	Omaisuusrikokset: varkaudet, ryöstöt, petokset ja maksuvälinepetokset. Vuosien 2005–2007 keskiarvo / keskiväkiluku 2005–2007. / 10 000 asukasta. Sotkanet.
Kuntatyyppi	Kolmiluokkainen kuntaryhmittely 2009. Tilastokeskuksen luokittelu.
Pinta-ala	Pinta-ala 2007. Sotkanet.
Väestötiheys	Väestötiheys 2007. Sotkanet.
Väestö	Väestö 2007. Sotkanet.
Väestö, M	Miehiä väestöstä (%) 2007. Sotkanet.
0–15-vuotiaat	0–15-vuotiaat väestöstä (%) 2007. Sotkanet.
16–24-vuotiaat	16–24-vuotiaat väestöstä (%) 2007. Sotkanet.
25–64-vuotiaat	25–64-vuotiaat väestöstä (%) 2007. Sotkanet.
Yli 65-vuotiaat	Yli 65-vuotiaat väestöstä (%) 2007. Sotkanet.
0–15-vuotiaat, M	0–15-vuotiaat miehet väestöstä (%) 2007. Sotkanet.
16–24-vuotiaat, M	16–24-vuotiaat miehet väestöstä (%) 2007. Sotkanet.
Alkoholin myynti	Alkoholijuomien myynti asukasta kohti 100 %:n alkoholina, litraa. Sotkanet.
Päihdehuolto	Päihdehuollon nettokustannukset, € / asukas 2007. Sotkanet.
Perusterveysmenot	Perusterveydenhuollon nettokustannukset, € / asukas 2007. Sotkanet.
Sosiaalimenot	Sosiaalitoimen nettokustannukset, € / asukas 2007. Sotkanet.
Kasvatusneuvonta	Kasvatus- ja perheneuvonnan asiakkaat vuoden aikana / 1 000 alle 18-vuotiasta 2007. Sotkanet.
Toimeentulotuki	Toimeentulotuki, € / asukas 2007. Sotkanet.
Toimeentulotuki, LP	Toimeentulotukea saaneet lapsiperheet (%) lapsiperheistä 2007. Sotkanet.
Nuorisotyöttömät	Nuorisotyöttömät (%) 15–24-vuotiaasta työvoimasta 2007. Sotkanet.
Pitkäaikaistyöttömät	Pitkäaikaistyöttömät (%) työttömistä 2007. Sotkanet.
Työttömät	Työttömät (%) työvoimasta 2007. Sotkanet.
Työttömät, M	Työttömät miehet (%) työvoimasta 2007. Sotkanet.
Lapsiperheet	Lapsiperheet (%) perheistä 2007. Sotkanet.

Nimi	Kuvaus ja lähde
Yksinhuoltajaperheet	Yksinhuoltajaperheet (%) lapsiperheistä 2007. Sotkanet.
Pienituloisuusaste, LP	Lapsiperheiden pienituloisuusaste 2007. Sotkanet.
Avioerot	Avioeroja 25–64-vuotiailla / 1 000 vastaavanikäistä naimisissa olevaa 2007. Sotkanet.
Koulutus	Korkea-asteen koulutuksen saaneet (%) 15 vuotta täyttäneistä 2007. Sotkanet.
Pienituloisuusaste	Kunnan yleinen pienituloisuusaste 2007. Sotkanet.
Päihtyneiden säilönnöt	Päihtyneiden säilönnöt / 1 000 asukasta 2007. Sotkanet.
Huumausainerikokset	Poliisin tietoon tulleet kaikki huumausainerikokset / 1 000 asukasta 2007. Sotkanet.
Sairastavuusindeksi	Sairastavuusindeksi, ikävakioitu 2007. Sotkanet.
Omistusasunnot	Omistusasunto kaikista asunnoista (%) 2007. Altika.
Vuokra-asunnot	Vuokra-asunto kaikista asunnoista (%) 2007. Altika.
Moottoriajoneuvot	Moottoriajoneuvot (henkilö- ja pakettiautot ja moottoripyörät) / väkiluku 2007. Altika.
Vaalit	Vaalit, äänestysprosentti, kaikki 2007. Altika.
Vaalit, M	Vaalit, äänestysprosentti, miehet 2007. Altika.
Maahanmuutto	Maahanmuutto, lkm (ka 05–07) / 1 000 asukasta (ka väkiluvusta) 2007. Altika.
Tulomuutto	Maassamuuton tulomuutto / väestö*100 (%) 2007. Altika.
Nettomuutto	Kunnan nettomuutto / 1 000 asukasta 2007. Sotkanet.
Ruotsinkielisten osuus	Ruotsinkielisten osuus väestöstä (%) 2007. Sotkanet.
Vieras äidinkieli	Vieras äidinkieli / 1 000 asukasta 2007. Sotkanet.
Pendelöinti	Pendelöinti (alueella työssäkäyvät / työvoima %) 2006. Sotkanet.
Yritystoimipaikat	Yritystoimipaikkojen lukumäärä / 1 000 asukasta 2007. Statfin.
Tulot	Valtionveronalaiset tulot, € / asukas 2006. Statfin.
Kesämökkit	Kesämökkien lukumäärä / 1 000 asukasta 2007. Statfin.
Työpaikat	Kunnassa olevien työpaikkojen lukumäärä / 1 000 asukasta 2006. Statfin.
Työssäkäynti, oma	Omassa kunnassa työssäkäyvien osuus työssäkäyvistä (%) 2006. Statfin.
Työ, alkutuotanto	Alkutuotannon työpaikkojen osuus (%) 2006. Statfin.
Työ, jalostus	Jalostuksen työpaikkojen osuus (%) 2006. Statfin.
Työ, palvelu	Palvelualan työpaikkojen osuus (%) 2006. Statfin.
Huoltosuhde	Huoltosuhde, demografinen 2007. Sotkanet.
Koulutustasomittain	Koulutustasomittain 2007. Sotkanet.

## Liite B: Kolmiluokkainen kuntaryhmittely

Kaupunkimaiset kunnat – 63 kuntaa			
Akaa	Järvenpää	Lappeenranta	Rauma
Espoo	Kaarina	Lempäälä	Riihimäki
Forssa	Kajaani	Lohja	Rovaniemi
Hamina	Kangasala	Loviisa	Salo
Hanko	Kaskinen	Maarianhamina	Savonlinna
Harjavalta	Kauniainen	Mikkeli	Seinäjoki
Haukipudas	Kemi	Nokia	Tampere
Heinola	Kempele	Nurmijärvi	Tornio
Helsinki	Kerava	Oulu	Turku
Hollola	Kiiminki	Oulunsalo	Tuusula
Hyvinkää	Kirkkonummi	Pietarsaari	Vaasa
Hämeenlinna	Kokkola	Pirkkala	Valkeakoski
Iisalmi	Kotka	Pori	Vantaa
Imatra	Kouvola	Porvoo	Varkaus
Joensuu	Kuopio	Raahe	Ylöjärvi
Jyväskylä	Lahti	Raisio	

Taajaan asutut kunnat – 66 kuntaa			
Asikkala	Kauhava	Masku	Outokumpu
Eura	Kemijärvi	Muhos	Paimio
Haapajärvi	Keminmaa	Mustasaari	Parkano
Haapavesi	Keuruu	Muurame	Pieksämäki
Hattula	Kokemäki	Mynämäki	Raasepori
Huittinen	Kontiolahti	Mäntsälä	Sastamala
Hämeenkyrö	Kuhmo	Mänttä-Vilppula	Siilinjärvi
Ii	Kurikka	Mäntyharju	Sipoo
Ikaalinen	Kuusamo	Naantali	Sodankylä
Ilmajoki	Laihia	Nakkila	Suonenjoki
Janakkala	Laitila	Nastola	Ulvila
Jämsä	Lapua	Nivala	Uusikaupunki
Kalajoki	Laukaa	Noormarkku	Vihti
Kankaanpää	Lieksa	Nurmes	Ylivieska
Kannus	Lieto	Orimattila	Äänekoski
Karkkila	Loimaa	Orivesi	
Kauhajoki	Länsi-Turunmaa	Oulainen	



Maaseutumaiset kunnat – 219 kuntaa				
Alajärvi	Kaavi	Lestijärvi	Punkaharju	Suomussalmi
Alavieska	Kangasniemi	Liljendal	Punkalaidun	Sysmä
Alavus	Kannonkoski	Liminka	Puolanka	Säkylä
Artjärvi	Karjajoki	Liperi	Puumala	Taipalsaari
Askola	Karjalohja	Loppi	Pyhtää	Taivalkoski
Aura	Karstula	Luhanka	Pyhäjoki	Taivassalo
Brändö	Karttula	Lumijoki	Pyhäjärvi	Tammela
Eckerö	Karvia	Lumparland	Pyhäntä	Tarvasjoki
Enonkoski	Kaustinen	Luoto	Pyhäranta	Tervo
Enontekiö	Keitele	Luumäki	Pälkäne	Tervola
Eurajoki	Kemiönsaari	Luvia	Pöytyä	Teuva
Evijärvi	Kerimäki	Maalahti	Rantasalmi	Tohmajärvi
Finström	Kesälahti	Maaninka	Ranua	Toholampi
Föglö	Kihniö	Marttila	Rautalampi	Toivakka
Geta	Kiikoinen	Merijärvi	Rautavaara	Tuusniemi
Hailuoto	Kinnula	Merikarvia	Rautjärvi	Tyrnävä
Halsua	Kitee	Miehikkälä	Reisjärvi	Töysä
Hammarland	Kittilä	Multia	Ristiina	Ujala
Hankasalmi	Kiuruvesi	Muonio	Ristijärvi	Utajärvi
Hartola	Kivijärvi	Myrskylä	Ruokolahti	Utsjoki
Hausjärvi	Kolari	Nilsjä	Ruotsinpyhtää	Uurainen
Heinävesi	Konnevesi	Nousiainen	Ruovesi	Uusikaarlepyy
Himanka	Korsnäs	Nummi-Pusula	Rusko	Vaala
Hirvensalmi	Koski	Närpiö	Rääkkylä	Valtimo
Honkajoki	Kristiinankaupunki	Oravainen	Saarijärvi	Varpaisjärvi
Humppila	Kruunupyy	Oripää	Salla	Vehmaa
Hyrnsalmi	Kuhmalahti	Padasjoki	Saltvik	Vesanto
Hämeenkoski	Kuhmoinen	Paltamo	Sauvo	Vesilahti
Iitti	Kumlinge	Parikkala	Savitaipale	Veteli
Iloantsi	Kuortane	Pedersören kunta	Savukoski	Vieremä
Inari	Kustavi	Pelkosenniemi	Sievi	Vihanti
Inkoo	Kylmäkoski	Pello	Siikainen	Viitasaari
Isojoki	Kyyjärvi	Perho	Siikajoki	Vimpeli
Isokyrö	Kärkölä	Pernaja	Siikalatva	Virolahti
Jalasjärvi	Kärsämäki	Pertunmaa	Simo	Virrat
Jokioinen	Kökar	Petäjävesi	Siuntio	Vårdö
Jomala	Köyliö	Pielavesi	Soini	Vähäkyrö
Joroinen	Lapinjärvi	Pihtipudas	Somero	Vöyri-Maksamaa
Joutsa	Lapinlahti	Polvijärvi	Sonkajärvi	Yli-ii
Juankoski	Lappajärvi	Pomarkku	Sotkamo	Ylitornio
Juuka	Lavia	Pornainen	Sottunga	Ylämaa
Juupajoki	Lemi	Posio	Sulkava	Ypäjä
Juva	Lemland	Pudasjärvi	Sund	Ähtäri
Jämijärvi	Leppävirta	Pukkila	Suomenniemi	

## Liite C: Väkivaltarikosmalleihin valitut muuttajat

Kaupunkimaiset kunnat	Taajaan asutut kunnat	Maaseutumaiset kunnat
Väestötiheys	Väestötiheys	Väestö, M
Väestö, M	Väestö, M	0–15-vuotiaat
0–15-vuotiaat	0–15-vuotiaat	16–24-vuotiaat
16–24-vuotiaat	16–24-vuotiaat	Alkoholin myynti
Alkoholin myynti	Alkoholin myynti	Toimeentulotuki
Toimeentulotuki	Toimeentulotuki	Nuorisotyöttömät
Nuorisotyöttömät	Pitkäaikaistyöttömät	Työttömät miehet
Pitkäaikaistyöttömät	Työttömät miehet	Lapsiperheet
Työttömät miehet	Yksinhuoltajaperheet	Yksinhuoltajaperheet
Yksinhuoltajaperheet	Avioerot	Avioerot
Avioerot	Koulutus	Koulutus
Koulutus	Pienituloisuusaste	Pienituloisuusaste
Pienituloisuusaste	Päihtyneiden säilöönnotot	Päihtyneiden säilöönnotot
Päihtyneiden säilöönnotot	Huumausainerikokset	Huumausainerikokset
Huumausainerikokset	Vuokra-asunnot	Sairastavuusindeksi
Sairastavuusindeksi	Vaalit	Vuokra-asunnot
Vuokra-asunnot	Maahanmuutto	Työpaikat
Vaalit	Tulomuutto	Vaalit
Maahanmuutto	Nettomuutto	Maahanmuutto
Tulomuutto	Pendelöinti	Nettomuutto
Pendelöinti	Tulot	Pendelöinti
Tulot	Työ, alkutuotanto	Tulot
Työpaikat, palvelut	Työ, palvelut	Työ, alkutuotanto
Huoltosuhde	Huoltosuhde	Työ, palvelut
		Huoltosuhde

## Liite D: Omaisuusrikosmalleihin valitut muuttajat

Kaupunkimaiset kunnat	Taajaan asutut kunnat	Maaseutumaiset kunnat
Väestötiheys	Väestötiheys	Väestötiheys
Väestö, M	Väestö, M	0–15-vuotiaat
0–15-vuotiaat	0–15-vuotiaat	16–24-vuotiaat, M
16–24-vuotiaat	16–24-vuotiaat	Alkoholin myynti
Alkoholin myynti	Alkoholin myynti	Toimeentulotuki
Toimeentulotuki	Toimeentulotuki	Nuorisotyöttömät
Nuorisotyöttömät	Nuorisotyöttömät	Työttömät
Työttömät miehet	Työttömät miehet	Lapsiperheet
Lapsiperheet	Lapsiperheet	Yksinhuoltajaperheet
Yksinhuoltajaperheet	Yksinhuoltajaperheet	Avioerot
Avioerot	Avioerot	Koulutus
Koulutus	Koulutus	Pienituloisuusaste
Pienituloisuusaste	Pienituloisuusaste	Päihtyneiden säilöönnotot
Päihtyneiden säilöönnotot	Päihtyneiden säilöönnotot	Huumausainerikokset
Huumausainerikokset	Huumausainerikokset	Sairastavuusindeksi
Sairastavuusindeksi	Vuokra-asunnot	Vuokra-asunnot
Vuokra-asunnot	Moottoriajoneuvot	Moottoriajoneuvot
Moottoriajoneuvot	Vaalit	Vaalit
Vaalit	Maahanmuutto	Tulomuutto
Maahanmuutto	Tulomuutto	Nettomuutto
Nettomuutto	Nettomuutto	Pendelöinti
Pendelöinti	Pendelöinti	Yritykset
Yritykset	Yritykset	Tulot
Tulot	Tulot	Kesämökit
Työpaikat, palvelut	Työ, alkutuotanto	Työ, palvelut
Huoltosuhde	Työ, palvelut	Huoltosuhde
	Huoltosuhde	