

HELSINKILÄISTEN KERROSTALOASUNTOJEN HINNANMUODOSTUS

Jani Hakala

Helsingin yliopisto
Matemaattis-luonnontieteellinen
tiedekunta
Tilastotiede
Pro gradu -tutkielma
Maaliskuu 2018



HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

MATEMAATTIS-LUONNONTIETEELLINEN TIEDEKUNTA
MATEMATISK-NATURVETENSKAPLIGA FAKULTETEN
FACULTY OF SCIENCE

Tiedekunta – Fakultet – Faculty		Koulutusohjelma – Utbildningsprogram – Degree programme	
Matemaattis-luonnontieteellinen		Tilastotiede	
Tekijä – Författare – Author			
Jani Hakala			
Työn nimi – Arbetets titel – Title			
Helsingiläisten kerrostaloasuntojen hinnanmuodostus			
Työn laji – Arbetets art – Level	Aika – Datum – Month and year	Sivumäärä – Sidoantal – Number of pages	
Pro gradu -tutkielma	Maaliskuu 2018	57 s. + 6 liites.	
Tiivistelmä – Referat – Abstract			
<p>Tutkielman tavoitteena on luoda lineaarisen regressioanalyysin avulla hinnoittelumalli helsinkiläisille kerrostaloasunnoille, minkä avulla pyritään selvittämään asuntojen hinnanmuodostukseen vaikuttavia tekijöitä ja niiden vaikutusten suuruutta. Tutkielman alussa kuvataan asuntomarkkinoita Suomessa keskittyen erityisesti hinnanmuodostukseen mahdollisesti vaikuttaviin tekijöihin. Teoriaosiossa käsitellään lineaarista regressioanalyysiä, jota voidaan käyttää yhden selitettävän muuttujan ja yhden tai usean selittävän muuttujan välisen lineaarisen yhteyden mallintamiseen ja selittämiseen.</p> <p>Tutkielman empiiristä osiota varten on poimittu aineisto Kiinteistövälitysalan Keskusliitto ry:n KVKL HSP -hintaseurantapalvelusta. Palvelusta löytyy suomalaisten kiinteistövälittäjien tekemät asuntokaupat sisältäen yksityiskohtaista tietoa myydyistä asunnoista. Aineistoksi rajattiin tammikuun 2016 ja elokuun 2017 välillä tapahtuneet helsinkiläiset kerrostaloasuntokaupat. Lisäksi aineistoa on rikastettu lisäämällä kohteiden sijaintikoordinaatit avoindata.fi -palvelusta, mistä löytyy Väestönrekisterikeskuksen ylläpitämät tiedot, sekä R-ohjelmiston avulla laskettu erilaisia matka-aikoja ja -pituuksia kohteista keskustaan hyödyntäen Google Maps -karttapalvelun ohjelmointirajapintaa.</p> <p>Tutkielmassa muodostetaan pienimmän neliösumman menetelmällä yhteensä kolme eri estimointimallia, joiden avulla kuvataan ja selitetään eri ominaisuustekijöiden vaikutuksia asunnon hintaan. Mallit kuvaavat niissä olevien yksittäisten selittäjien arvojen vaihtelun vaikutusta, kun muiden selittäjien vaikutus on vakioitu. Tutkielmassa onnistuttiin selittämään asuntojen hinnanmuodostusta melko tarkasti. Saadut tulokset ovat odotetun suuntaisia suhteessa taustateoriaan, ja niiden perusteella helsinkiläiset kerrostaloasunnot ovat hyvin moniulotteisia hyödykkeitä, joiden hinnat määräytyvät useiden hintatekijöiden yhteisvaikutuksen perusteella. Tutkielman tuloksissa näkyi lisäksi selvästi asuntomarkkinoiden suhdanteen nousukehitys.</p>			
Avainsanat – Nyckelord – Keywords			
Asuntojen hinnat, lineaarinen regressioanalyysi, PNS-menetelmä, asuntomarkkinat			
Säilytyspaikka – Förvaringställe – Where deposited			
Muita tietoja – Övriga uppgifter – Additional information			

Sisällysluettelo

1	JOHDANTO	1
2	ASUNTOMARKKINAT SUOMESSA	3
2.1	Asuntokanta ja asuinolot	3
2.2	Rahoitus ja verotus	5
2.3	Omassa vai vuokralla	7
2.4	Asunto-osakkeen hinnan arviointi	8
2.5	Asuntojen hintaeroihin vaikuttavat tekijät	10
2.5.1	Sijainti	10
2.5.2	Rakennusvuosi ja kunto	11
2.5.3	Asuntokohtaiset ja muut vaikuttavat tekijät	12
3	TUTKIMUSAINEISTO	14
3.1	Aineiston rajaus ja laatu	14
3.2	Aineiston kuvaus	16
3.2.1	Sijaintitekijöitä mallintavat muuttajat	17
3.2.2	Taloyhtiötä ja asuinrakennusta mallintavat muuttajat	20
3.2.3	Asunnon ominaisuuksia mallintavat muuttajat	23
3.2.4	Hintakehitys aineistossa	25
4	LINEAARINEN REGRESSIOMALLI	27
4.1	Pienimmän neliösumman menetelmä	28
4.2	Regressiokertoimien ja mallin merkitsevyys sekä mallin selitysaste	29
4.3	Muuttujien muunnokset	31
4.4	Regressiodiagnostiikka	33
4.5	Mallin valinta	36
5	AINEISTON ANALYYSI	38
5.1	Päämalli asunnon velattomalle neliöhinnalle	38
5.2	Lisämallit asunnon velattomalle neliöhinnalle	44
5.3	Estimaattien kuvailu	46
5.3.1	Sijaintitekijät	46
5.3.2	Taloyhtiötä ja asuinrakennusta kuvaavat tekijät	48
5.3.3	Asunnon ominaisuuksia kuvaavat tekijät	49
5.3.4	Ajallinen hintakehitys	50
6	JOHTOPÄÄTÖKSET	51
	LÄHTEET	56
	LIITTEET	

1 JOHDANTO

Lähtökohtaisesti jokaisen ihmisen tulee asua jossakin, minkä vuoksi omistusasunto on monille melko luonteva varallisuuden muoto. Oma asunto onkin kotitalouksien merkittävin varallisuuserä; vuonna 2013 noin 68 prosenttia suomalaisista kotitalouksista asui omassa asunnossa ja se muodosti 53 prosenttia kotitalouksien kokonaisvarallisuudesta. Omaan käyttöön tarkoitettun asunnon lisäksi suomalaisten kotitalouksien kokonaisvarallisuudesta noin 16 prosenttia tuli muusta asuntovarallisuudesta (sijoitusasunnot, vapaa-ajanasunnot ja muut asunnot), jota oli noin 24 prosentilla kotitalouksista. Kokonaisasuntovarallisuus oli siis noin 69 prosenttia kokonaisvarallisuudesta ja sitä oli 70 prosentilla kotitalouksista – eli muuta asuntovarallisuutta omaavilla kotitalouksilla oli myös hyvin todennäköisesti omaan käyttöön tarkoitettu asunto. Yleensä asunnon hankinta rahoitetaan osittain pankkilainan avulla, minkä vuoksi asuntovelka on suomalaisten kotitalouksien selvästi merkittävin velkaerä. Vuonna 2013 kotitalouksien veloista lähes 70 prosenttia oli asuntovelkaa, jota oli noin joka kolmannelle kotitaloudella; keskimäärin 92 000 euroa velallista kotitaloutta kohden. (Tilastokeskus 2015.)

Kaupungistuminen on yksi voimakkaimmista yhteiskunnallisista trendeistä ympäri maailmaa. Kaupunkien houkuttelevuus yritysten ja asukkaiden näkökulmasta perustuu lyhyihin välimatkoihin, kun palveluiden ja tavaroiden vaihdanta on helpompaa ja taloudellisempaa esimerkiksi maaseutuun verrattuna. Myös Suomessa on tapahtunut ja tapahtuu edelleen selvää kaupungistumista, mutta Suomen nähdään olevan selkeästi jäljessä kaupungistumisessa verrattaessa muihin vastaavan tuloluokan maihin. Suomen kaupungistuminen näkyy luonnollisesti Helsingin väestönkehityksessä; vuosina 1990–2015 Helsingin asukasluku kasvoi noin 27,5 prosenttia, kun samassa ajassa koko Manner-Suomen asukasluku kasvoi 9,7 prosenttia. Erityisesti Helsingin seudulla asuntotuotanto ei ole vastannut väestönkehitystä ja Teknologian tutkimuskeskus onkin arvioinut vuoden 2015 lopussa Helsingin seudun asuntovajeen olleen noin 20 000 asuntoa. Lisäksi on mahdotonta arvioida, kuinka moni on jättänyt muuttamatta Helsingin seudulle, koska ei ole löytänyt itselleen sopivaa asuntoa. (Laakso & Loikkanen 2004; Vainio 2016.) Asuntojen selvästi tarjontaa korkeampi kysyntä on kiihdyttänyt hintojen kehitystä Helsingissä ja muissa kasvukeskuksissa. Esimerkiksi vuosina 2010–2017 vanhojen kerrostalo-osakeasuntojen hinnat ovat nousseet Helsingissä noin 21,9 prosenttia, kun puolestaan pääkaupunkiseudun ulkopuolella hinnat ovat nousseet vain noin 3,7 prosenttia. Kun vielä otetaan inflaatio

huomioon, ovat Helsingissä hinnat nousseet reaalisesti noin 10,7 prosenttia ja pääkaupunkiseudun ulkopuolella puolestaan laskeneet reaalisesti noin 5,8 prosenttia. (Tilastokeskus 2017a.)

Oma asunto on useimmille elämän suurin yksittäinen hankinta, minkä vuoksi kaupan kohteena olevan asunnon markkina-arvon arvioiminen on tärkeää. Tähän avuksi löytyy esimerkiksi Tilastokeskuksen hintatilastot ja mediassa aiheesta uutisoidaan hyvin ahkerasti, joten jonkinlaisen käsityksen saaminen lienee helppoa. Useimmiten erilaisissa tilastoissa tai hintavertailuissa tarkastelu jää hyvin pintapuoliseksi, esimerkiksi tarkasteltaessa neliohintojen kehitystä postinumeroalueittain otetaan huomioon useimmiten vain huoneiston koko, rakennustyyppi, huoneiden lukumäärä ja postinumeroalue. Aiheeseen liittyvästä uutisoinnista ei aina selviä selkeästi syy-seuraussuhteet, vaikka uutisointi olisikin tehty hyvin huomiota herättävästi, mikä saattaa olla jopa harhaanjohtavaa. Esimerkiksi Helsingin Sanomissa 7.5.2017 kirjoitettiin ”jos ostit asunnon vuonna 2006 Helsingin Kallioista 100 000 eurolla, sen arvo on nyt 159 000 euroa”. Tämän kaltaista päättelyä ei voi suoraan tehdä postinumeroalueen keskineliöhinnan kehityksen perusteella, koska luvusta ei selviä, johtuuko muutos asuinalueen arvostuksen noususta vai siitä että esimerkiksi kyseisen alueen asuntokannasta 75 prosenttia olisi linjasaneerattu vuoden 2006 jälkeen.

Tämän tutkielman tavoitteena on luoda lineaarisen regressioanalyysin avulla hinnoittelumalli helsinkiläisille kerrostaloasunnoille, minkä avulla pyritään selvittämään asuntojen hinnanmuodostukseen vaikuttavia tekijöitä ja niiden vaikutusten suuruutta. Tutkimuksen aineistoksi on kerätty tietoja vuosina 2016–2017 toteutuneista kerrostaloasuntokaupoista Helsingissä. Tutkielmassa keskitytään nimenomaan Helsingissä tehtyihin kerrostaloasuntokauppoihin, koska Helsingissä on keskimääräisesti koko Suomen korkein hintataso ja Helsinki on Suomen suurin kaupunki.

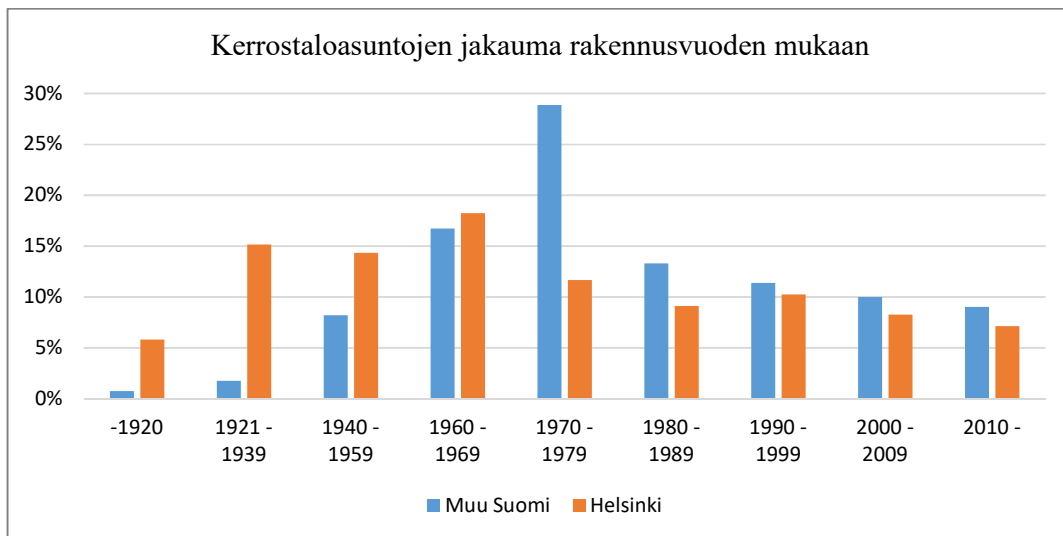
Tutkielma etenee siten, että luvussa 2 käsitellään Suomen asuntomarkkinoita keskittyen erityisesti hinnanmuodostumiseen mahdollisesti vaikuttaviin tekijöihin. Tämän jälkeen luvussa 3 esitellään ja kuvaillaan tutkielmassa käytettävää aineistoa. Luvussa 4 esitellään tutkielmassa käytettävä menetelmä. Empiirinen osuus on luvussa 5, jossa käydään läpi tutkimusmenetelmän soveltamista tutkielmassa ja sen tuottamia tuloksia. Johtopäätökset ovat lopulta luvussa 6.

2 ASUNTOMARKKINAT SUOMESSA

Tässä luvussa tarkastellaan asuntomarkkinoita Suomessa keskittyen erityisesti helsinkiläisiin asuinkerrostaloihin.

2.1 Asuntokanta ja asuinolot

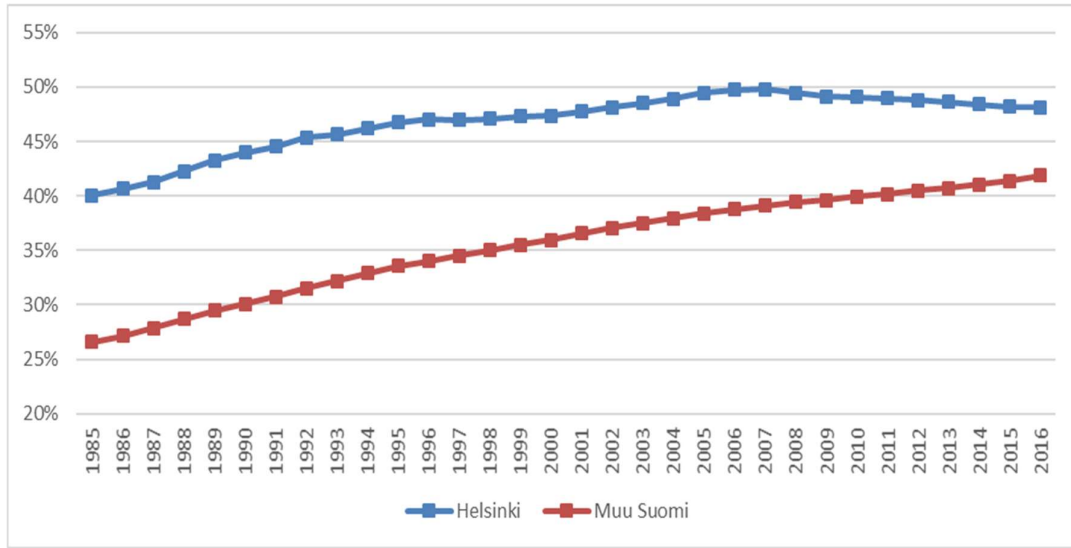
Suomessa oli vuoden 2016 päättyessä yhteensä 2 968 200 asuntoa, joista 356 975 eli noin 12 prosenttia sijaitsi Helsingissä. Helsingissä sijainneista asunnoista 305 053 oli kerrostaloasuntoja. Kuvios 2.1 kertoo, miten kerrostaloasunnot ovat jakautuneet vuonna 2016 rakennusvuoden mukaan Helsingissä ja muualla Suomessa. Kuviosta nähdään, että reilu viidennes helsinkiläisistä kerrostaloasunnoista oli rakennettu ennen vuotta 1940, kun puolestaan Helsingin ulkopuolella olevista kerrostaloasunnoista vain noin 2,5 prosenttia on rakennettu vastaavana ajankohtana. Toinen selkeästi poikkeava ero jakaumissa kohdistuu 1970-luvulla rakennettuihin asuntoihin; tuolloin valmistuneiden asuntojen osuus on muualla Suomessa ollut huomattavasti suurempi Helsinkiin verrattuna. (Tilastokeskus 2017b.)



Kuvio 2.1. Kerrostaloasuntojen rakennusvuosijakauma vuonna 2016. (Tilastokeskus 2017b)

Tilastokeskus määrittelee tilastoissaan ahtaasti asuviksi ne asuntokunnat, joissa asunnon huoneluku, kun keittiötä ei lasketa mukaan huonelukuun, on suurempi kuin asunnossa

vakinaisesti asuvien henkilöiden lukumäärä. Tilastokeskuksen tilastojen mukaan vuonna 2016 asuntoväestöstä ahtaasti asui Helsingissä noin 21 prosenttia ja muualla Suomessa noin 17 prosenttia. Asuntojen asuinala henkilöä kohden oli vuonna 2016 keskimäärin 41,1 neliometriä Helsingin ulkopuolella ja Helsingissä 34 neliometriä. Ero selittyy suurimmalta osin sillä, että Helsingissä – sekä muissa kaupungeissa – kerrostaloasuminen on huomattavasti yleisempää kuin taajamissa ja maaseudulla.



Kuvio 2.2. Yhden hengen asuntokuntien osuus kaikista asuntokunnista. (Tilastokeskus 2017b)

Yksinasumisen yleistymisen on jo pidempään ollut selkeä suuntaus Suomessa. Kuviosta 2.2 nähdään kuitenkin, että Helsingissä yhden hengen asuntokuntien osuus kaikista asuntokunnista on lähtenyt laskuun vuoden 2007 jälkeen, kun samaan aikaan muualla Suomessa yhden hengen asuntokuntien osuus on jatkanut tasaista kasvuaan. Helsingin viimeisen 10 vuoden kehitystä selittänee asuntojen hintojen ja niiden vuokrien kallistuminen, jonka seurauksena esimerkiksi yhteisasuminen eli 'kimppakämpääasuminen' on kasvattanut suosiotaan muun muassa siitä syntyvien taloudellisten säästöjen vuoksi. Taulukosta 1 selviää, että kaksio oli vuonna 2016 yleisin kerrostaloasuntotyyppi helsinkiläisten asuntokuntien keskuudessa, kun noin 40,9 prosenttia kyseisistä asuntokunnista asui kaksioissa. Vain 11,8 prosenttia asuntokunnista asui kerrostalohuoneistossa, jossa huoneita (pois lukien keittiö) oli yli kolme kappaletta.

Taulukko 1. Helsinkiläisten asuinkerrostaloissa asuvien asutokuntien lukumäärä huoneiden lukumäärän mukaan vuonna 2016. (Tilastokeskus 2017b)

	Asutokuntien lukumäärä	%
Yhteensä	278 136	
1 huone	68 398	24,6 %
2 huonetta	113 720	40,9 %
3 huonetta	62 288	22,4 %
4+ huonetta	32 907	11,8 %
Tuntematon	823	0,3 %

2.2 Rahoitus ja verotus

Kuten johdantoluvussa mainittiin, on asunto yleensä ihmisen elämän taloudellisesti merkittävin hankinta. Asuntojen hinnat ovat varsinkin Helsingissä yleensä moninkertaisia suhteessa ostajien nettovuosituloihin, minkä vuoksi normaalisti asunnonhankinta rahoitetaan osittain velkarahalla. Yleisin yksityishenkilöiden velkaraahoituksen muoto on henkilökohtainen pankkilaina asunnon ostoa varten, eli asuntolaina. Asuntolainan enimmäismäärää säätelee laissa määrätty enimmäisluototussuhde eli lainakatto, jota koskeva laki astui voimaan 1.7.2016. Lainakaton myötä asuntolainan määrä voi olla lainaa nostaessa enintään 90 prosenttia kyseisen asunnon vakuuksien käyvästä arvosta, eli vähintään kymmenen prosenttia asunnon hinnasta tulisi kattaa omilla säästöillä ja/tai muilla vakuuksilla. Ensiasuntoa ostettaessa lainakatto on 95 prosenttia. Asuntoa ostettaessa käyväällä arvolla tarkoitetaan asunnon hankintahintaa. (Finanssivalvonta 2016.)

Toinen hyvin yleinen asunnon rahoitusmuoto on asunto-osakeyhtiön yhtiölaina, joka voidaan jakaa jokaiselle asunto-osakkeelle. Yksittäisen asunto-osakkeen velan osuutta kutsutaan yhtiölainaosuudeksi. Yhtiölainalla katetaan yleensä kiinteistön pitkävaikutteiset ja suuret menot kuten rakentamis-, uudistamis- ja peruskorjauskustannukset. Asunto-osakkeen omistajalla on laillinen velvollisuus maksaa asunto-osakkeeseen kohdistuva yhtiövastike, joka koostuu hoito- ja rahoitusvastikkeesta. Hoitovastikkeella katetaan yhtiön tavalliset juoksevat kulut, kuten isännöinti-, vakuutus-, siivous- ja huoltokulut. Rahoitusvastikkeella puolestaan katetaan edellä mainitun yhtiölainan kulut eli lyhennykset ja korot. Vastikkeiden jyvitykset eri asunto-osakkeille määräytyy yhtiöjärjestyksessä, jossa jyvityksen perusteena on usein esimerkiksi asuntojen pinta-ala. Vuonna 2017 myynissä olleissa uudiskohteissa yhtiölainan osuus asunnon velattomasta hinnasta on yleensä ollut

enimmillään 70 prosenttia. Kauppahinta (velaton hinta – yhtiölainaosuus) katetaan esimerkiksi omilla säästöillä ja/tai asuntolainalla. Osakkailla on myös oikeus halutessaan maksaa yhtiölainaosuus kerralla pois joko asunnon oston yhteydessä tai yhtiöjärjestyksessä määrättyinä ajankohtina, jolloin yhtiölainaosuuden kokonaan maksanut osakas vapautuu kyseisen lainan perusteena olevan rahoitusvastikkeen maksamisesta. (Haulos, Nevala, Palo & Sirén 2015.)

Asuntokauppaan liittyy myös verotuksellisia seikkoja, jotka tulee ottaa huomioon. Asunto-osakkeen ostajan tulee maksaa asuntokaupasta varainsiirtovero, jonka suuruus on kaksi prosenttia asunto-osakkeen velattomasta hinnasta. Ensiasuntoa ostettaessa ei varainsiirtoveroa tarvitse maksaa. Maksu tulee suorittaa kahden kuukauden sisällä siitä, kun on allekirjoitettu kauppakirja tai muu sopimus. Vero maksetaan kaupanteon yhteydessä, jos kauppa tehdään kiinteistönvälittäjän kautta. Uudiskohteiden tapauksessa varainsiirtovero maksetaan kahden kuukauden sisällä siitä, kun omistusoikeus on siirtynyt ostajalle – riippumatta siitä, oliko kiinteistönvälittäjää käytetty.

Aiemmin mainitun asuntolainan korkojen mukaan voi tehdä korkovähennyksen pääomatulojen verotuksessa. Vuonna 2017 korkovähennyksen suuruus on 45 prosenttia, vuonna 2018 35 prosenttia ja vuodesta 2019 lähtien 25 prosenttia asuntolainan koroista. Jos korkovähennysten määrä ylittää pääomatulojen määrän, muodostaa se alijäämän, josta voi 30 prosenttia vähentää ansiotulojen verotuksessa. Ensiasunnon ostaja voi vähentää 32 prosenttia alijäämästä ansiotulojen verotuksessa. Yhtiölainan koroista ei saa korkovähennystä. Edellä mainitut ehdot pätevät, jos asunto ostetaan omaan käyttöön. Asuntosijoittajat voivat puolestaan vähentää pankkilainan korot kokonaan pääomatulojen verotuksessa ja myös rahoitusvastikkeen, jos se on tuloutettu taloyhtiön kirjanpidossa. Myös asunto-osakkeen myyjällä on mahdollisia veroseuraamuksia. Jos asunto myydään voitollisena, katsotaan myynnistä saatu voitto luovutusvoitoksi, joka lasketaan verotuksessa pääomatuloksi. Asunnon saa myytyä verovapaasti, jos on omistanut asunnon vähintään kaksi vuotta ja omistusajan aikana asunnossa on asunut itse tai joku perheestä yhtäjaksoisesti vähintään kahden vuoden ajan. Mahdollisen luovutustappion voi puolestaan vähentää muista pääomatuloista paitsi, jos kyseisen asunnon myynti voitollisena olisi ollut verovapaata. (Verohallinto 2017.)

2.3 Omassa vai vuokralla

Ihmisen elämäntilanteen tullessa siihen pisteeseen, että muutto on edessä, on yleensä tehtävä päätös siitä, että haluaako asua vuokralla vai ostaa oman. Kaikilla ei tosin välttämättä ole tätä valinnan mahdollisuutta taloudellisten syiden vuoksi. Esimerkiksi Helsingissä asuntojen hintataso on muuhun maahan verrattuna korkea, jolloin asunnon omarahoitusosuuden hankkiminen esimerkiksi säästämällä ei välttämättä ole realistista kaikille – säästämismahdollisuuksia rajoittaa myös korkea vuokrataso. Lisäksi, vaikka vaaditun omarahoitusosuuden saisi kasaan, voi riittävän asuntolainan saaminen pankista olla edelleen vaikeaa. Tähän voi olla syynä esimerkiksi luottohäiriömerkinnät, niin sanottu ”pätkätyöläisyys”, pankin näkökulmasta heikot tulevaisuuden työllistymismahdollisuudet tai riittämättömät tulot lainan kokoon nähden.

Jos oman asunnon hankinnalle ei ole selkeää estettä, voi pohtia valintaa vuokra- ja omistusasumisen välillä. Usein valintaa lähestytään taloudellisesta näkökulmasta, eli kumpi vaihtoehto olisi taloudellisesti kannattavampaa. Useimmiten omistusasuminen on taloudellisesti kannattavampaa, koska samalla kertyy asuntovarallisuutta lainaa lyhennettäessä, kun taas vuokralla asukkaalle ei kerry vuokranmaksusta mitään, vaan todennäköisesti sillä katetaan vastikkeiden lisäksi ainakin osittain vuokranantajan lainan lyhennyksiä ja korkoja. Mutta useimmiten tilanne ei ole näin yksinkertainen, vaan on otettava huomioon myös muita asioita. Vuokra-asuminen on varteenotettava vaihtoehto esimerkiksi, jos asumisratkaisu on todennäköisesti lyhytaikainen. Vuokra-asunto sitoo vähemmän, koska sen irtisanominen on huomattavasti vaivattomampaa kuin omistusasunnon myyminen. Vuokralaiselle ei myöskään muodostu asunnosta taloudellisista syistä stressiä, kun ei tarvitse olla huolissaan esimerkiksi lainanhoidosta tai lähestyvistä taloyhtiön linjasaneerauksista.

Lyhytaikainen omistusasunto on varsinkin verotuksellisista syistä usein hyvin kallis ratkaisu. Esimerkiksi edellisessä luvussa mainittu varainsiirtovero nostaa huomattavasti asumisen kustannuksia, jos omistusasuminen jää lyhyeksi. Lisäksi omassa käytössä olleen omistusasunnon myyntivoiton verovapaus kahden vuoden asumisen jälkeen kannustaa pidempään asumiseen. Asuntojen hinnat ovat historiassa nousseet pitkällä aikavälillä, mutta lyhyessä omistusasumisessa on hintariski mahdollisten laskusuhdanteiden vuoksi. Omistusasumisen taloudellinen hyöty korostuu esimerkiksi silloin, kun on päässyt siihen

pisteeseen, että asunto on velaton, jolloin asumiskustannus tulee ainoastaan hoitovastikkeesta. Muita selkeitä positiivisia puolia omistusasumisessa ovat erilaiset vaikuttamismahdollisuudet. Esimerkiksi halutessaan asunnon seinät voi maalata haluamansa väriksi tai vaihtaa kaapistot mieleisekseen. Asunto-osakkeen omistajana on myös mahdollisuus vaikuttaa asunto-osakeyhtiön asioihin yhtiökokouksessa (Haulos et al. 2015).

Jussila (2017) huomauttaa omistusasumiseen liittyvän myös kulttuurillisia ja psykologisia syitä. Hän mainitsee, että länsimaissa omistusasunto on statussymboli ja sen ostaminen koetaan osana aikuistumista ja itsenäistymistä, kun taas vuokralla asumiseen on liitetty päinvastaisia merkityksiä. Asunnon omistamisen on koettu luoneen monilla turvallisuuden tunteen – varsinkin perheellisillä. Omistusasuminen saattaa myös täyttää inhimillisiä perustarpeita paremmin kuin vuokra-asuminen, kun asunnon voi selkeämmin kokea omaksi kodikseen, jonka suhteen itsellä on päätösvalta.

2.4 Asunto-osakkeen hinnan arviointi

Kasso (2016) mainitsee kolme keskeistä menetelmää kiinteistön arvon määrittelemiseksi: kustannusarvo-, tuottoarvo- ja kauppa-arvomenetelmä. Kustannusarvomenetelmässä arvioidaan kohderakennusta vastaavan rakennuksen teoreettiset rakentamiskustannukset. Kustannusarviossa otetaan huomioon rakennuksen ikä ja kunto, koska vanhemman rakennuksen arvo ei olisi sama kuin uuden, vaikka molempien uudelleen rakentaminen maksaisi yhtä paljon. Kun rakentamiskustannuksista on vähennetty rakennuksen iästä ja kunnosta johtuva arvonalennus, jäljellä jäävää arvoa kutsutaan rakennuksen tekniseksi nykyarvoksi (Haulos et al. 2015). Tontin arvo ei sisälly kustannusarvoon, vaan se on arvioitava erikseen. Kustannusarvomenetelmä sopii lähinnä kokonaisten asuinrakennusten arvioimiseen – varsinkin uudisrakentamisen ollessa todellinen vaihtoehto. (Kasso 2016.)

Tuottoarvomenetelmää käytetään varsinkin toimitilojen arvioinnissa, mutta myös tukevana menetelmänä asuntosijoituskohteissa. Nimensä mukaisesti tuottoarvomenetelmässä kohteen arvo määräytyy siitä saatavan tuoton ja tuotto-odotuksen mukaan. Kohteen nettotuoton laskennassa otetaan huomioon muun muassa vuokratuotto, kohteen odotettavissa oleva hintakehitys, rahoituskulut, hoitokustannukset ja keskimääräiset remonttikustannukset. Nettotuoton ja tuotto-odotuksen avulla saadaan laskettua arvo, jonka sijoittaja on valmis kohteesta maksamaan saadakseen haluamansa tuoton. Esimerkiksi, jos kohteen arvioitu vuotuinen nettotuotto on 9 000 euroa ja tuotto-odotus sijoitetulle pääomalle on

vuosittain 6 prosenttia, niin tuottoarvoksi saadaan 150 000 euroa (9000/0,06). Kasso (2016) huomauttaa, että markkinat vaikuttavat selvästi tuottovaatimukseen, mihin liittyy riskit ja arvioitu arvonnousu. Esimerkiksi, jos asunto sijaitsee suositulla alueella, missä riski kuukausista ilman vuokralaista on pieni ja odotettu arvonnousu hyvä, on tuotto-odotus normaalisti maltillisempi ja kohteen arvo korkeampi. (Haulos et al. 2015.)

Haulos et al. (2015) mukaan asunto-osakkeiden arvioimiseen käytetään pääasiassa kauppaa-arvomenetelmää. Kauppaa-arvomenetelmässä kohdetta verrataan sitä mahdollisimman samankaltaisten kohteiden kauppahintoihin. Menetelmän etuna on sen yksinkertaisuus ja vertailun tapahtuminen todellisiin toteutuneisiin kauppahintoihin. Haasteena voi olla riittävän samankaltaisten vertailukauppojen löytyminen ja niiden vähäisyys. Lisäksi vertailukohteet on käytännössä aina tehty menneisyydessä, joten ne eivät välttämättä enää kuvaa riittävästi kyseisen hetken arvoa.

Asunto-osake on ylipäätään arvioinnin kohteena haastava, koska jokainen kerrostalo-osake on yksilöllinen, jolle ei löydy identtistä vastinetta. Esimerkiksi, vaikka samassa rakennuksessa olisi myyty aineiston perusteella vastaava asunto, on vähintään huoneiston sijainti rakennuksessa aina eri, millä voi olla arvoon vaikuttavia tekijöitä. Esimerkiksi, jos huoneiston pohjaratkaisu on parempi tai huoneisto on sisäpihan puolella meluisan kadun sijaan. Lisäksi vertailtavuutta heikentää poikkeukselliset kaupat, esimerkiksi lähipiirissä tapahtuvat.

Edellä mainittujen arvioiden tai niiden yhdistelmien tueksi on hyvä tehdä markkina-analyysi, eli arvioida vallitseva markkinatilanne ja sen vaikutukset kohteen arvoon kysynnän ja tarjonnan kautta. Kasso (2016) jaottelee kysyntään ja tarjontaan vaikuttavat markkinatekijät yleisiin markkinatekijöihin, alueellisiin markkinatekijöihin, kohteen markkinatilanteeseen ja tulevaisuuteen liittyviin tekijöihin. Yleisiä markkinatekijöitä ovat esimerkiksi korkotasoa ja talouden suhdannevaihe. Alueellisia markkinatekijöitä ovat puolestaan esimerkiksi alueen työllisyystilanne, vuotuinen muuttovoitto tai -tappio ja oppilaitosten lukumäärä. Kohteen markkinatekijöitä ovat kysynnän ja tarjonnan suhde sekä kohdetta vastaava muu tarjonta. Tulevaisuuteen liittyvät tekijät saattavat usein olla enemmän tai vähemmän epävarmoja, jolloin ne ovat osittain spekulatiivisia. Tulevaisuuteen liittyviä tekijöitä ovat esimerkiksi infrastruktuuriin tai korkotasoon liittyvät muutokset, kuten päätös metrolinjan rakentamisesta tietylle alueelle. Markkina-analyysin avulla saadaan hinta-

arviota tarkennettua huomattavasti. Haulos et al. (2015) mainitsevat, että hyvissä olosuhteissa arvioinnin tarkkuus on yleensä noin +/- 10 prosenttia.

2.5 Asuntojen hintaeroihin vaikuttavat tekijät

Kerrostalo-osakeasunto on moniulotteinen hyödyke, jonka hinta määräytyy useiden eri tekijöiden perusteella. Vaikuttavien tekijöiden ja niiden vaikutusten syiden sekä suuruuden ymmärtäminen on hyvin oleellinen asia, jotta asunnon hinta-arviosta tulee riittävän tarkka. Seuraavissa alaluvuissa käydään läpi näitä tekijöitä ja niiden vaikutusta hintoihin.

2.5.1 Sijainti

Asunnon sijainti nähdään useimmiten tärkeimmäksi asunnon hintaan vaikuttavaksi tekijäksi. Hyvä sijainti ei kuitenkaan ole aina yksiselitteisesti määriteltävissä, vaan myös siihen vaikuttavat monet eri tekijät. Todennäköisesti helpoiten mitattavissa oleva sijaintitekijä on keskustaetaisyys. Normaalisti kaupungeissa keskustassa ja sen lähialueilla ovat taloudelliset toiminnot, palvelut ja ihmiset pakkautuneet lähelle toisiaan, mikä lisää kaupungin ja sen asukkaiden taloudellista tehokkuutta, kun tavaroiden ja palveluiden vaihdanta on helpompaa ja halvempaa (Laakso & Loikkanen 2004). Hyvät yksityisen ja julkisen liikenteen yhteydet lisäävät muun muassa palveluiden saavutettavuutta, mikä on selkeä arvotekijä asunnolle. Esimerkiksi Tuominen (2014) ja Laakso (1997) ovat arvioineet, että metron rakentaminen on selvästi nostanut metroasemien vaikutusalueella olevien asuntojen arvoa parantuneen saavutettavuuden takia.

Lähiasuinalueeseen liittyy monia asioita, jotka voivat vaikuttaa asunnon arvoon; muun muassa luonnon ja vesistöjen läheisyys, palvelut, alueen maine, koulut, harrastusmahdollisuudet, työpaikkojen määrä, asukkaiden ikäjakauma ja niin edelleen. Muun muassa Harjunen, Kortelainen ja Saarimaa (2014) arvioivat, että koulujen väliset erot näkyvät myös asuntojen hinnoissa. Tosin he huomauttavat, että on viitteitä siitä, että koulun oppimisvaikutusten sijaan asunnonostajille on tärkeää oppilaiden sosioekonominen tausta. Myös samalla asuinalueella toisiansa hyvin lähellä olevilla asunnoilla voi olla suuriakin sijainnin aiheuttamia hintaeroja. Esimerkiksi kaksi asuntoa, jotka molemmat ovat alle 100 metrin päässä merestä, mutta vain toisessa asunnossa on merinäköala.

2.5.2 Rakennusvuosi ja kunto

Sijaintitekijöiden jälkeen merkittävimiksi tekijöiksi asunnon hinnanmuodostuksessa nousee talon ja asunnon rakennusvuosi sekä kunto, jotka ovat myös selvästi yhteydessä keskenään. Luonnollisesti uusien asuntojen hinnat ovat korkeimmat, ja normaalin kulumisen vuoksi talon ja huoneistojen korjaus- ja huoltotarpeet kasvavat, mikä laskee vanhempien asuntojen arvoa. Niissä taloyhtiöissä, joissa huoltoon panostetaan ja korjaustarpeet suoritetaan hyvissä ajoin – eikä vasta pakon edessä – asunnot säilyttävät arvonsa paremmin. Tehdyt remontit parantavat rakennuksen ja asuntojen kuntoa, mikä nostaa niiden arvoa. Tämän vuoksi esimerkiksi vasta remontoitu 50 vuotta vanha asunto voi olla arvokkaampi kuin vastaavanlainen 20 vuotta nuorempi asunto.

Korjausremonteista johtuvat kustannukset ovat asunnon omistajille hyvin huomattava kuluera; Hietala et al. (2015) ovat arvioineet, että suomalaisten asuinrakennusten vuotuinen tekninen korjaustarve on keskimäärin noin 3,5 miljardia euroa vuosina 2016–2025. He tosin huomauttavat, että taloudellisesti perusteltu korjaustarve on keskimäärin noin 3,2 miljardia euroa, koska esimerkiksi maaseudulla sijaitsee asuinrakennuksia, joiden korjaaminen ei ole välttämättä taloudellisesti kannattavaa, koska niihin ei kohdistu kysyntää.

Korjausremonteista kallein on linjasaneeraus eli putkiremontti, joka vaikuttaa selvästi asunnon arvoon. Putkistojen tekninen elinkaari vaihtelee normaalisti kolmen- ja viidenkymmenen vuoden välillä. Nikola (2011) on arvioinut, että asunnon arvo alkaa laskea selvästi noin kuusi vuotta ennen lähestyvää putkiremonttia, mikä jatkuu remontin valmistumiseen asti. Valmistunut putkiremontti puolestaan nostaa huomattavasti asunnon arvoa. Nikola (2011) ja Venäläinen (2013) ovat kuitenkin arvioineet, että putkiremontin kustannukset eivät täysin välity asuntojen arvoihin, eli remonttikustannukset ovat keskimäärin suuremmat kuin asunnon arvonnousu remontin myötä. Tästä he ovat päätelleet, että keskimäärin asunto on kannattavampaa myydä ennen remonttia kuin sen jälkeen.

län aiheuttamien vaikutusten lisäksi rakennusvuosi on kunnan indikaattori, koska eri vuosikymmenten välillä on eroja käytettyjen rakennustekniikoiden ja -materiaalien suhteen, jotka vaikuttavat rakennusten kuntoon. Esimerkiksi vuosina 1969–1975 oli Suomessa erittäin voimakas muuttoaalto maalta sodan jälkeen syntyneiden suurten ikäluokkien siirtäessä työmarkkinoille (Laakso & Loikkanen 2004). Tämä asetti haasteita, koska piti rakentaa nopeasti ja edullisesti uusia asuntoja. Tähän yhtenä ratkaisuna oli rakentaa beto-

nielementtitaloja, joiden ongelmana on ollut huono ilmanvaihto, mikä on aiheuttanut sisäilmaongelmia, jotka ovat saattaneet tehdä joistakin rakennuksista jopa asumiskelvottomia.

2.5.3 Asuntokohtaiset ja muut vaikuttavat tekijät

Edellä mainittujen sijaintia ja asuinrakennusta koskevien tekijöiden lisäksi asunnon hintaan vaikuttaa luonnollisesti myös asuntokohtaiset tekijät. Asunnon pinta-alaa pidetään melko itsestään selvänä vaikuttimena asunnon hintaan, minkä vuoksi useimmiten asuntojen hintoja tarkastellaan ja verrataan niiden neliöhintojen mukaan. Pinta-alan merkitys riippuu myös kokonaispinta-alasta, koska esimerkiksi viiden lisäneliön merkitys on huomattavasti suurempi 30 neliön asunnon kohdalla kuin 70 neliön asunnon. Hyvä pohjaratkaisu todetaan useasti vetovoimatekijäksi. Esimerkiksi heikosti hyödynnettävissä olevat neliöt asunnossa laskevat asunnon arvoa verrattaessa vastaavan suuruisiin asuntoihin, joissa niin sanottuja ”hukkaneliöitä” ei ole. Asuntojen eri pohjaratkaisujen vertailun ongelmana on niiden heikko mitattavuus. Samankokoisten asuntojen kohdalla huonelukumäärällä on usein hintaa nostava vaikutus, jos huoneet ovat käyttökelpoisia kokonsa ja pohjansa puolesta. Kerrostaloissa useimmiten ensimmäisen kerroksen asunnot ovat selvästi edullisempia ylempien kerrosten asuntoihin verrattuna. Ensimmäisestä kerroksesta on normaalisti vaatimattomammat näkymät ikkunasta, suuremmat meluhaitat (esimerkiksi liikenteestä) ja heikompi yksityisyyden tunne, jos esimerkiksi kadulta tai pihasta näkee verhojen ollessa auki melko esteettömästi sisälle. Ylimmät kerrokset eivät välttämättä ole puolestaan aina alimpia arvokkaampia, jos rakennuksesta puuttuu hissi. Saunan ja parvekkeen olemassaoloa pidetään myös usein positiivisena asiana asunnon kannalta.

Hoitovastike ja yhtiölainaosuus ovat puolestaan tekijöitä, jotka vaikuttavat suoraan asumiskustannuksiin ja sitä kautta myös sen arvoon. Hoitovastikkeella on useimmiten asunnon arvoa laskeva vaikutus. Hoitovastikkeen suuruuteen vaikuttaa merkittävästi se, että onko asunto-osakeyhtiön tontti oma vai vuokralla. Tontin omistaminen kasvattaa asunto-osakkeen arvoa, kun puolestaan vuokratontilla tonttivuokra peritään osana hoitovastiketta. Asunnot, joilla on yhtiölainaa, voivat olla houkuttelevampia, jos esimerkiksi yhtiölainan korko on matalampi kuin mitä henkilökohtaiseen pankkilainaan saisi. Asuntosijoittajalle voi olla suurikin merkitys sillä, että tuloutetaanko vai rahastoidaanko mahdol-

lisesta yhtiölainasta perityt rahoitusvastikkeet taloyhtiön kirjanpidossa. Jos rahoitusvastike tuloutetaan eli näytetään taloyhtiön tuloslaskelmassa, ovat ne vähennyskelpoisia vuokratulon verotuksessa (Verohallinto 2017).

Kun tarkastellaan tilastojen tai tilastollisten mallien avulla eri tekijöiden vaikutuksia asunnon arvoon, on hyvä muistaa, että on kyse keskimääräisistä vaikutuksista. Asunto on moniulotteiden hyödyke, jonka ostamiseen ihmisillä liittyy rationaalisten tekijöiden lisäksi myös paljon tunneperäisiä tekijöitä. Asumispreferenssit voivat vaihdella paljonkin eri ihmisten välillä, jolloin joidenkin tekijöiden arvioiminen yleisellä tasolla voi olla jopa lopulta haastavaa. Esimerkiksi sauna voi olla sellainen ominaisuus, jota toiset pitävät mielekkäänä arjen ylellisyytenä kodissaan, mutta toisille se puolestaan voi olla täysin turha asia, joka vie turhaan paljon tilaa asunnosta. Tämän vuoksi eri hintatekijöiden vaikutusten syiden analysointi voi tosinaan olla hyvinkin monimutkaista. Lisäksi asunnon toteutuvaan kauppahintaan voi vaikuttaa suuresti valittava kiinteistönvälittäjä ja hänen ammattitaitonsa.

3 TUTKIMUSAINEISTO

Tutkielman empiirisessä osiossa tarkastelun kohteena ovat helsinkiläisten kerrostaloasuntojen hinnat. Tätä varten on poimittu aineisto Kiinteistönvälitysalan Keskusliitto ry:n KVKL HSP -hintaseurantapalvelusta. Hintaseurantapalvelu on tarkoitettu ensisijaisesti kiinteistönvälitys- ja rakennusalalla toimivien yritysten tilasto- ja hakupalveluksi, joka toimii työkaluna muun muassa markkina-arvon määrittelemisessä sekä arvioiden ja neuvotteluiden tueksi. Hintaseurantapalvelusta löytyy suomalaisten kiinteistönvälittäjien tekemät asuntokaupat sisältäen yksityiskohtaista tietoa myydyistä asunnoista. Tähän tutkielmaan on poimittu tammikuun 2016 ja elokuun 2017 välillä hintaseurantapalveluun lisätyt helsinkiläiset kerrostaloasuntokaupat. Aineistosta löytyy kohteiden katuosoitteet, joiden avulla haettiin kyseisten asuinrakennusten sijaintikoordinaatit avoindata.fi-palvelusta, mistä löytyy Väestönrekisterikeskuksen ylläpitämät tiedot. Gmapsdistance-paketissa olevalla samannimisellä funktiolla on R-ohjelmistolla haettu lisätietoja hyödyntäen koordinaattitietoja. Funktio käyttää Google Maps -karttapalvelun ohjelmointirajapintaa laskeakseen matka-ajan ja -pituuden eri sijaintien välillä (Melo & Zarruk 2016). Tätä tutkielmaa varten kohteille haettiin funktion avulla tieto kävelymatkasta keskustaan sekä matka-ajasta keskustaan julkisilla ja autolla.

3.1 Aineiston rajausta ja laatu

Kohteiden katuosoitteet syötetään HSP-palveluun käsin, minkä vuoksi aineistossa oli runsaasti kirjoitusvirheitä. Esimerkiksi katuosoitteen kohdalla oli osalla kohteista kadun nimi ja talon numero, mutta osassa kohteita oli lisäksi tieto rapusta ja asunnon numerosta. Katuosoitteet muokattiin ja korjattiin yhteneväiseen muotoon, jotta aineistoon saatiin rikastettua kohteiden koordinaattitiedot. Pienessä osassa kohteita ei ollut ollenkaan talon numeroa, minkä vuoksi ne on jätetty otoksesta pois, koska niiden tarkan sijainnin määrittäminen ei ollut mahdollista.

Aineistoa on rajattu kohteiden rakennusvuoden mukaan siten, että pois on jätetty kohteet, joiden rakennusvuosi on 2018 tai 2019, eli kohteet ovat varmuudella vasta rakenteilla. Poimitussa aineistossa seuraavien ominaisuuksien olemassaoloa kuvataan arvolla 'TRUE': uudiskohde, sauna, hissi, vuokrattu, sauna ja parveke. Muissa tapauksissa näiden muuttujien arvo oli tyhjä. Aineistosta ei pysty päättelemään, onko arvo tyhjä sen

vuoksi, että kyseistä ominaisuutta ei ole vai tiedon puuttumisen takia. Esimerkiksi aineistosta löytyy kaksi kohdetta, jotka sijaitsevat 2015 valmistuneen rakennuksen seitsemännessä kerroksessa, mutta hissimuuttujan arvo on tyhjä, eikä ole uskottavaa, että kyse olisi hissittömistä rakennuksista. Lisäksi internetistä on helposti saatavilla tietoja asuinrakennuksista, minkä avulla pystyi tekemään pistotarkistuksia, mistä selvisi, että tyhjä arvot johtuvat usein tiedon eikä ominaisuuden puuttumisesta. Tämän vuoksi nämä muuttujat on jätetty pois, vaikka näillä ominaisuuksilla on todennäköisesti vaikutusta asunnon arvoon. Myös puuttuneisuutta esiintyi aineistossa, minkä vuoksi osa kohteista jätettiin pois. Puuttuneisuuden oletetaan olevan satunnaista. Lisäksi aineistoista jätettiin pois selvästi virheellisiä kohteita, esimerkiksi autopaikkoja ja kymmenen neliöinen kolmio. Lopulliseksi aineiston kooksi saatiin 9901 havaintoa. Taulukosta 2 nähdään muuttujat, jotka lopulta sisällytettiin tutkielman analyysiin mukaan.

Taulukko 2. Tutkimusaineiston muuttujat.

Muuttuja	Yksikkö
Postinumero	
Asuinala	m ²
Rakennusvuosi	
Huoneita	
Kerrosnumero	
Velaton hinta	€
Neliöhinta	Velaton hinta per neliö (€/m ²)
Kiinteistön kunto	Huono/Tyydyttävä/Hyvä/Erinomainen/Uusi
Kauppan pvm	
Tontin omistajuus	Oma/Vuokrattu
Hoitovastike	€
Hoitovastike per neliö	€/m ²
Kävelymatka keskustaan	km
Matka-aika keskustaan autolla	min
Matka-aika keskustaan julkisilla	min

Myös taulukossa 2 olevien muuttujien kohdalla saattaa olla yksittäisiä virheellisiä tietoja, mutta pääsääntöisesti aineiston laatu on hyvä osittain sen suuren koon ansiosta.

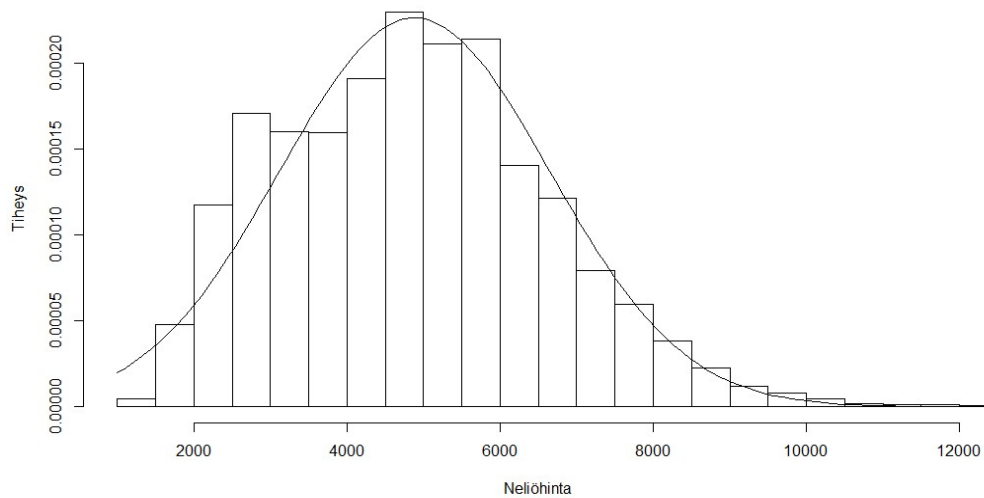
3.2 Aineiston kuvaus

Tässä luvussa kuvataan tutkielmassa käytettäviä muuttujia. Tarkastelussa muuttujia peilataan paljon neliöhintamuuttujaan, koska empiirisessä osiossa pyritään sen arvoja estimoimaan muiden muuttujien avulla. Kohteille on rakennusvuoden mukaan laskettu ikämuuttuja, jolla pyritään kuvaamaan asunnon ja rakennuksen kulumisen ja ikääntymisen vaikutusta. Taulukossa 3 on nähtävissä aineiston numeeristen muuttujien tunnusluvut.

Taulukko 3. Tunnuslukutaulukko.

	Keskiarvo	Keskihintajonta	Min	q1	Median	q3	Max
Asuinala	58,95	26,83	11	41	54	72	276
Ikä	55,62	29,25	0	37	56	79	167
Huoneita	2,28	1,00	1	2	2	3	9
Kerrosnumero	3,06	1,73	1	2	3	4	23
Velaton hinta	279718,60	179680,40	64666	172000	228500	326000	2650000
Neliöhinta	4883,91	1761,79	1026,44	3500,00	4853,66	6000,00	14583,33
Hoitovastike	252,09	115,55	3,40	174,00	234,60	304,00	1370,20
Hoitovastike per neliö	4,38	0,99	0,07	3,80	4,30	4,90	17,38
Kävellen,_km	6,20	4,40	0,23	2,30	5,19	9,49	15,64
Autolla,_min	18,17	6,50	2,23	12,32	17,97	24,07	32,85
Julkisilla, min	19,64	8,37	1,52	12,53	18,98	25,20	44,97

Nähdään, että neliöhinnan keskiarvo on hyvin lähellä neliöhinnan mediaania, mikä viittaa muuttujan jakauman symmetrisyyteen. Velattoman hinnan keskiarvo ja mediaani ovat puolestaan melko kaukana toisistaan. Kuvion 3.1 histogrammi vahvistaa neliöhinnan jakauman symmetrisyyden. Kuviosta nähdään myös, että neliöhinta on melko normaalijakautunut. Jakauman oikea häntä on vasenta pidempi, koska aineisto sisältää neliöhinnan suhteen poikkeavan suuria havaintoja.



Kuvio 3.1. Neliöhinnan jakauma.

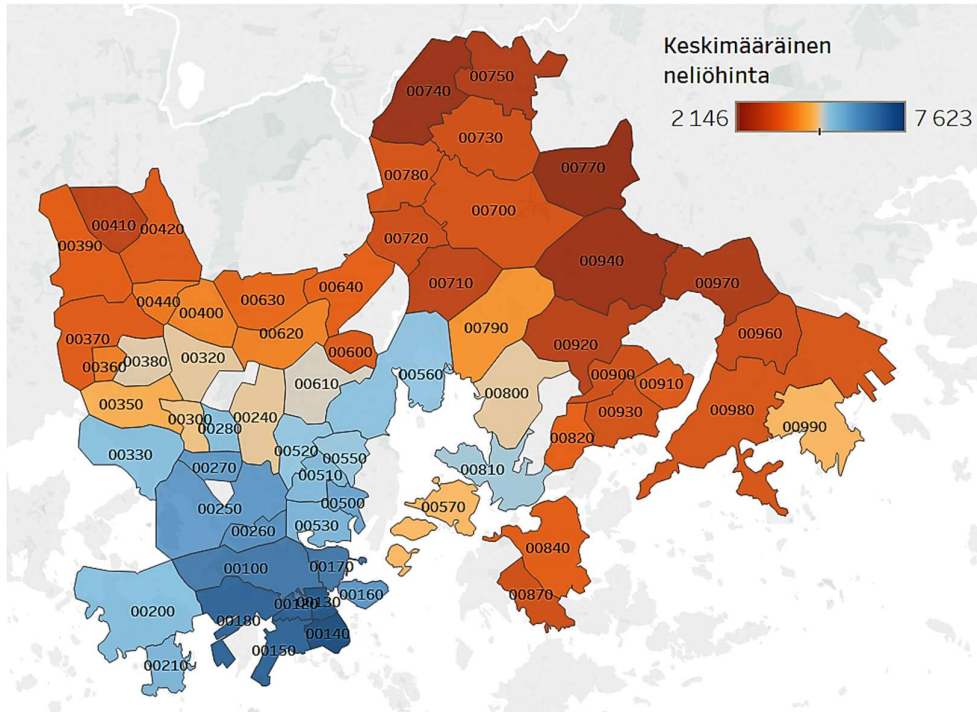
Seuraavissa alaluvuissa kuvaillaan muuttujia, joita hyödynnetään tutkielman empiirisessä osiossa selittävinä muuttujina.

3.2.1 Sijaintitekijöitä mallintavat muuttujat

Helsinki jakautuu maantieteellisesti 84:än eri postinumeroalueeseen. Aineistossa myyjiä kohteita on yhteensä 78:lta eri postinumeroalueelta. Kallein postinumeroalue keskimääräisen neliöhinnan perusteella on Etelä-Helsingissä sijaitseva Kaivopuisto – Ullanlinna (00140), missä keskimääräinen neliöhinta oli noin 7 633 euroa. Halvin postinumeroalue puolestaan aineiston perusteella on Jakomäki – Alppikylä (00770), joka sijaitsee Pohjois-Helsingissä Vantaan rajan tuntumassa. Siellä keskimääräinen neliöhinta oli noin 2 146 euroa, eli yli viisi tuhatta euroa vähemmän kuin Helsingin kalleimmalla alueella. Helsingin ydinkeskusta sekä kantakaupunki sijaitsevat Etelä-Helsingissä.

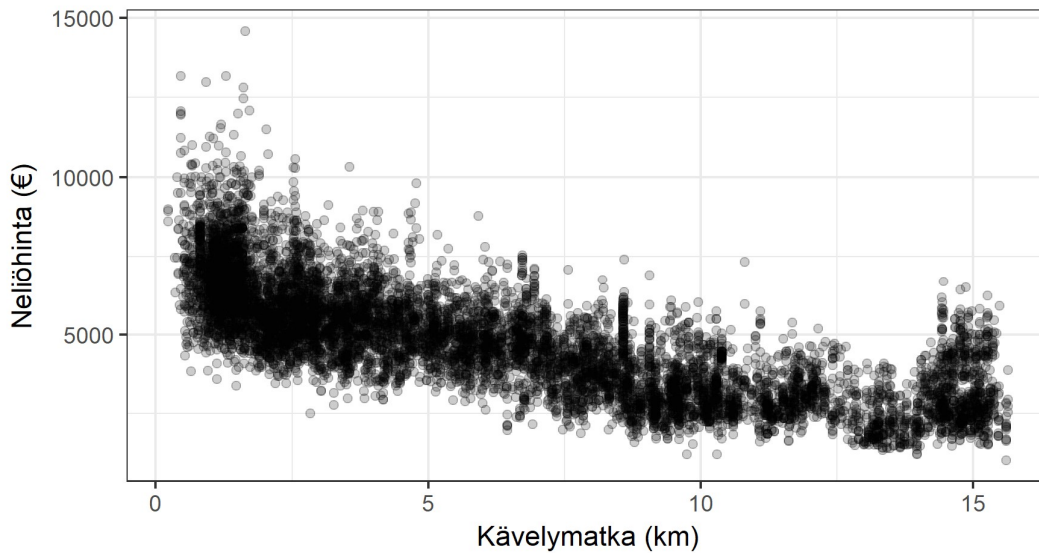
Kuviosta 3.2 nähdään Helsingin keskimääräiset neliöhinnat postinumeroalueittain. Kuviosta on jätetty pois postinumeroalueet, joiden kohdalla havaintoja on aineistossa alle kolmekymmentä kappaletta, jotta alueet olisivat riittävän edustettuja. Kuviosta pois jätettyjen alueiden havaintojen osuus kaikista havainnoista on noin 1,8 prosenttia. Kuviosta nähdään, että keskineliöhinnan perusteella kalleimmat alueet sijaitsevat Etelä-Helsingissä ja siirtyessä keskusta-alueelta pois päin keskineliöhinnat selvästi laskevat. Edullisimmat

postinumeroalueet sijaitsevat kuvion perusteella Pohjois-Helsingissä. Myös osalla vierekkäisistä postinumeroalueista saattaa olla suuriakin hintaeroja keskenään. Esimerkiksi Itä-Helsingissä sijaitsevassa Aurinkolahdessa (00990) on kuvion perusteella selvästi korkeammat neliöhinnat kuin läheisimmillä alueilla. Aurinkolahdessa sijaitsevien myytyjen asuntojen keskimääräinen ikä oli noin 11 vuotta, joka on yli neljäkymmentä vuotta vähemmän kuin koko aineiston asuntojen keskimääräinen ikä, mikä selittänee hintaeroa.



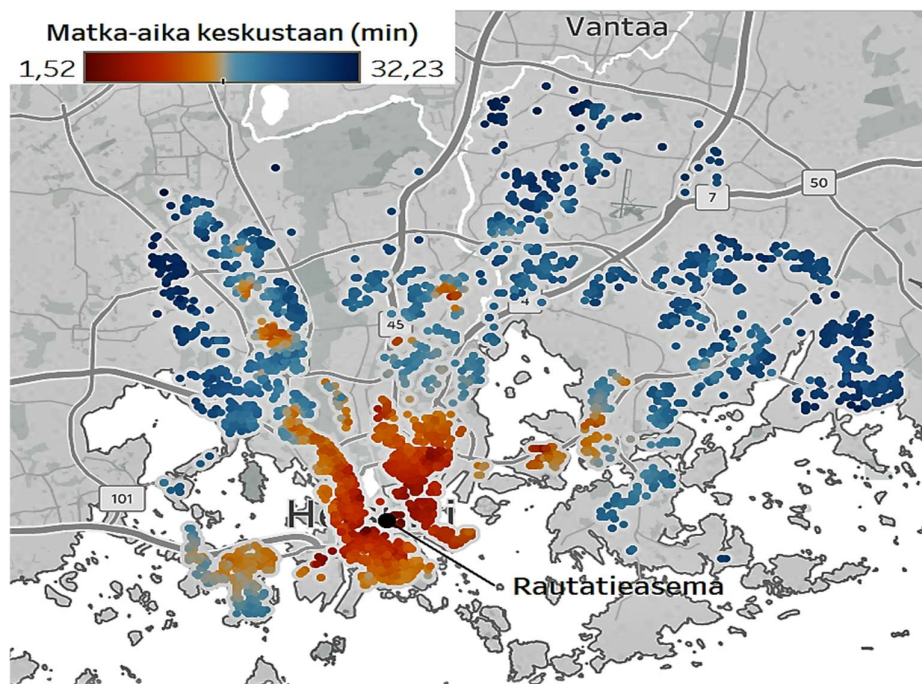
Kuvio 3.2. Keskimääräinen neliöhinta postinumeroalueittain.

Myytyjen kohteiden keskustaetäisyyttä mallinnetaan kävelymatkamuuttujalla, joka kertoo Google Maps -karttapalvelun laskeman kävelymatkan kohteesta keskustaan kilometreinä. Keskimäärin aineiston asunnoista oli kävelymatkaa keskustaan 6,2 kilometriä. Kuvion 3.2 perusteella hinnat keskimäärin laskevat siirtyessä kauemmaksi keskustasta, mikä näkyy varsin selvästi kuviossa 3.3. Yli 14 kilometrin kävelymatkan päässä sijaitsevilla kohteilla on enemmän trendistä poikkeavia neliöhintoja, mikä voi selittyä jollakin taustamuuttujalla, kuten edellä nähtiin Aurinkolahden kohdalla. Keskustan lähialueilla hintahajonta näyttäisi olevan kuvion perusteella suurinta.



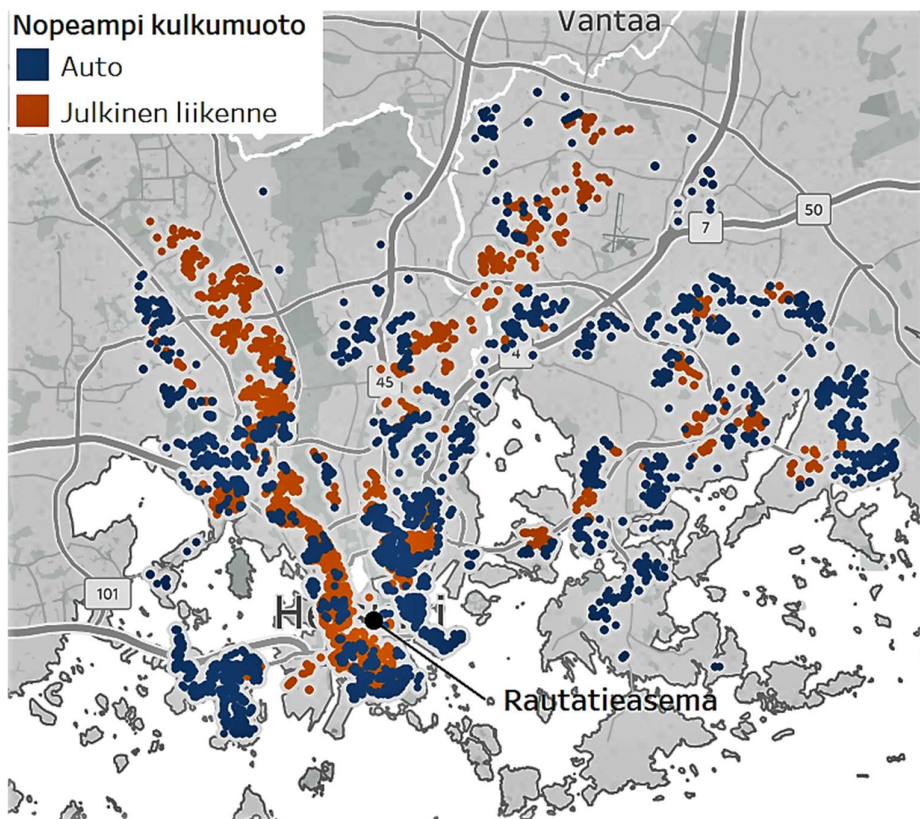
Kuvio 3.3. Neliöhinnan ja kävelymatkan välinen hajontakuvio.

Kohteiden saavutettavuutta puolestaan mallinnetaan matka-ajalla keskustaan autolla ja julkisella liikenteellä liikkuen. Kuviossa 3.4 matka-aikana on käytetty kohteittain nopeamman kulkumuodon (auto tai julkinen liikenne) matka-aikaa. Nähdään, että matka-aika keskustaan kasvaa hyvin tasaisesti mentäessä keskustasta poispäin, mutta myös poikkeavia kohteita löytyy. Nämä kohteet ovat useimmiten suurien liikenneväylien tai juna-/metroasemien läheisyydessä.



Kuvio 3.4. Matka-aika keskustaan nopeimmalla kulkumuodolla.

Kuviossa 3.5 on havainnollistettu, miten kohteet jakautuvat kartalla sen mukaan, onko nopeampi kulkumuoto keskustaan kulkiessa auto vai julkinen liikenne. Keskimäärin auto on noin puolitoista minuuttia nopeampi kulkumuoto kuljettaessa kohteista keskustaan. Noin 39 prosentilla kohteista nopein kulkumuoto on julkinen liikenne ja puolestaan noin 61 prosentilla auto. Helsingin keskustasta lähtee rautatieasemalta lähijunalinjat Luoteis- ja Pohjois-Helsinkiin. Metrolinja puolestaan kulkee rautatieasemalta länteen sekä pidemmälle Itä-Helsinkiin. Kuviossa 3.5 pystyy päättelemään näiden linjojen reitit, koska niiden läheisyydestä löytyy runsaasti kohteita, joissa julkinen liikenne on autoa nopeampi vaihtoehto keskustaan mentäessä. Eli nämä julkisen liikenteen ratkaisut parantavat saavutettavuutta.

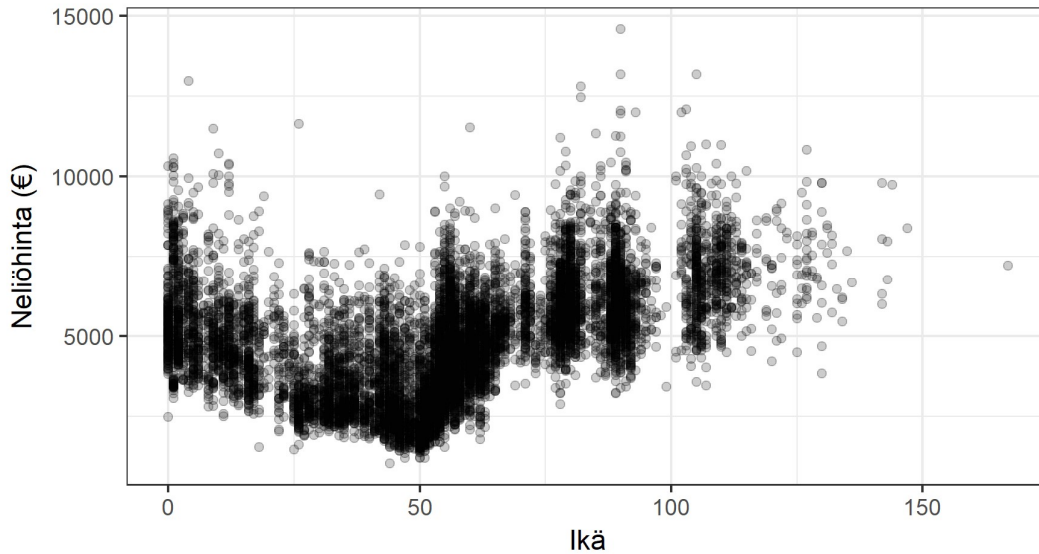


Kuvio 3.5. Kohteiden jakautuminen Helsingissä nopeamman kulkumuodon mukaan.

3.2.2 Taloyhtiötä ja asuinrakennusta mallintavat muuttujat

Aineiston kohteiden ikä on keskimäärin noin 56 vuotta, mutta alueiden välistä hajontaa on runsaasti. Eteläisimmän Helsingin kohteet ovat keskimäärin selvästi vanhimpia. Esimerkiksi Kaartinkaupungissa (00130) sijaitsevien kohteiden keskimääräinen ikä on noin

107 vuotta. Kuvio 3.6 nähdään neliöhinnan ja iän välinen yhteys. Kuvio tukee oletusta, että asuntojen arvot laskevat niiden valmistumisen jälkeen ikääntymisen myötä, mikä jatkuu linjasaneeraukseen asti, eli normaalisti noin 30 – 50 vuotta. Linjasaneerauksen valmistumisella puolestaan on asuntojen arvoa voimakkaasti nostava vaikutus, minkä jälkeen iällä on oletettavasti jälleen arvoa alentava vaikutus. Kuvion 3.6 pitkä nousevan trendin osuus selittyy edellä mainitulla seikalla, että arvokkaiden keskusta-alueiden asunnot ovat keskimäärin vanhempia muuhun aineistoon verrattuna.



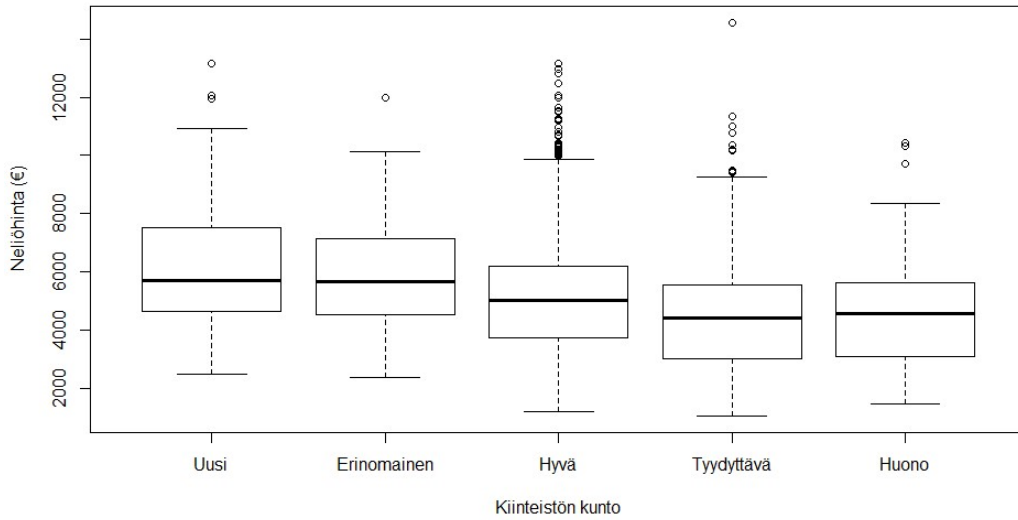
Kuvio 3.6. Neliöhinnan ja iän välinen hajontakuvi.

Kohdeasuinkiinteistöjen kunnat on luokiteltu viiteen eri luokkaan: uusi, erinomainen, hyvä, tyydyttävä ja huono. Taulukosta 4 nähdään, miten aineiston asunnot jakautuvat kiinteistön kunnan mukaan. Suurin luokka on 'hyvä', johon kuuluvat yli puolet kaikista kohteista.

Taulukko 4. Aineiston jakautuminen kiinteistöjen kunnan mukaan.

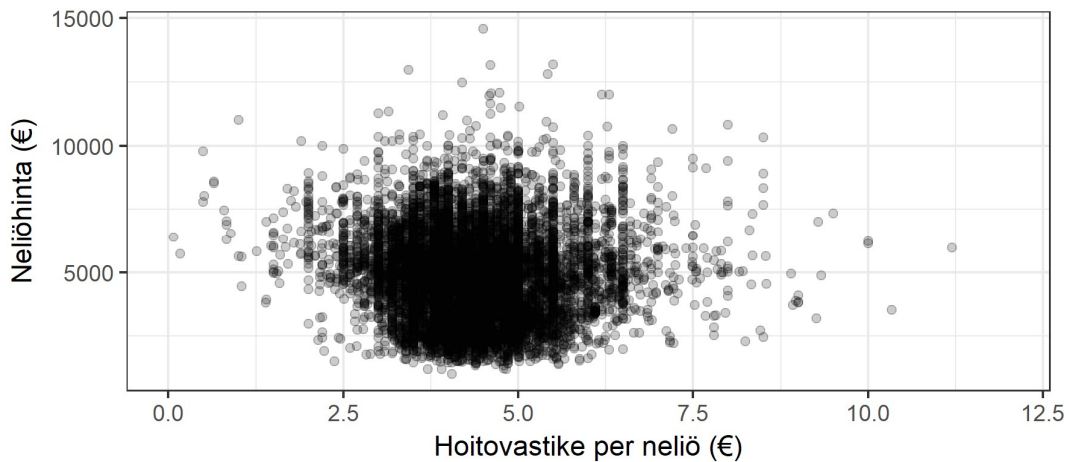
Kiinteistön kunto	N	%
Uusi	495	5,00 %
Erinomainen	181	1,83 %
Hyvä	5 325	53,78 %
Tyydyttävä	3 467	35,02 %
Huono	433	4,37 %

Uusien asuntojen mediaanineliöhinta oli luokista korkein, noin 5 706 euroa ja matalin puolestaan tyydyttäväkuntoisilla asunnoilla, noin 4 409 euroa. Kuviosta 3.7 nähdään, että uusien ja erinomaisessa kunnossa olevien asuntojen jakauma on melko lähellä toisiaan sekä myös huono- ja tyydyttäväkuntoisten asuntojen neliöhintajakauma on lähellä toisiinsa. Aineistossa on havaittavissa poikkeavia havaintoja – varsinkin hyväkuntoisten asuntojen kohdalla.



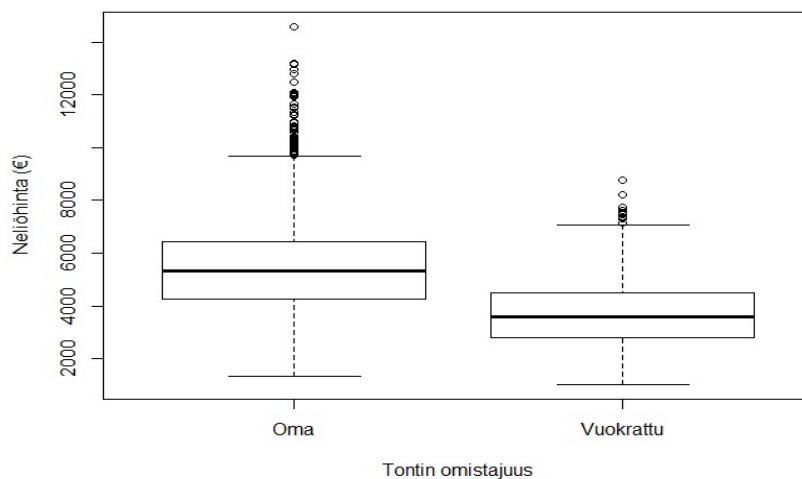
Kuvio 3.7. Laatikkokuva asuntojen neliöhinnosta kiinteistöjen kunnan mukaan.

Aineiston keskimääräinen hoitovastike per neliö oli 4,38 euroa. Hoitovastikkeiden hajonta on melko suurta, mikä näkyy myös kuviossa 3.8. Poikkeavia havaintoja esiintyy myös melko paljon. Kuviosta ei silmämääräisesti pysty havaitsemaan mahdollista yhteyttä hoitovastikkeen ja neliöhinnan välillä.



Kuvio 3.8. Neliöhinnan ja hoitovastikkeen välinen hajontakuva.

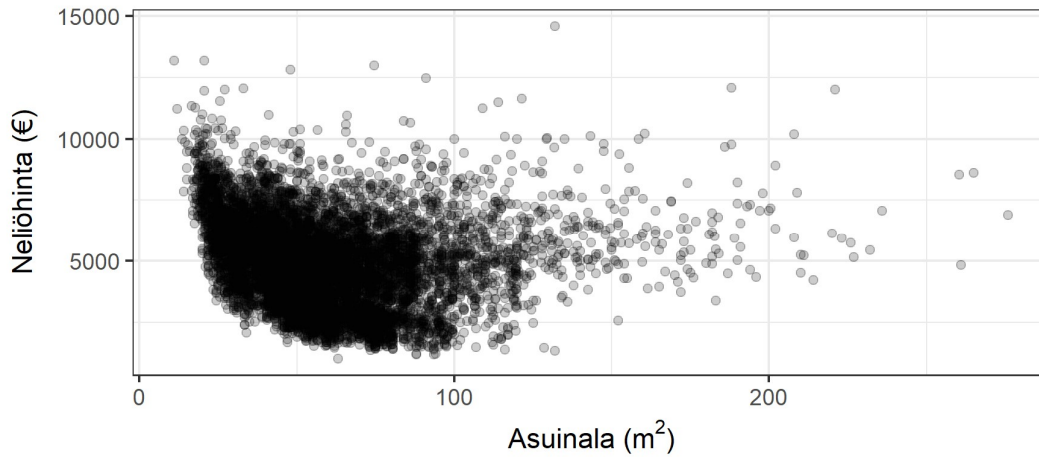
Aineiston asunnoista noin 72 prosenttia sijaitsee omassa omistuksessa olevalla tontilla ja 28 prosenttia puolestaan vuokratontilla. Vuokratontilla olevien asuntojen mediaanineliöhinta oli noin 3 591 euroa ja omalla tontilla olevilla puolestaan noin 5 345 euroa. Kuvion 3.9 perusteella tontin omistajuus näkyy selvästi asuntojen neliöhinnoissa.



Kuvio 3.9. Laatikkokuva asuntojen neliöhinnoista tontin omistajuuden mukaan.

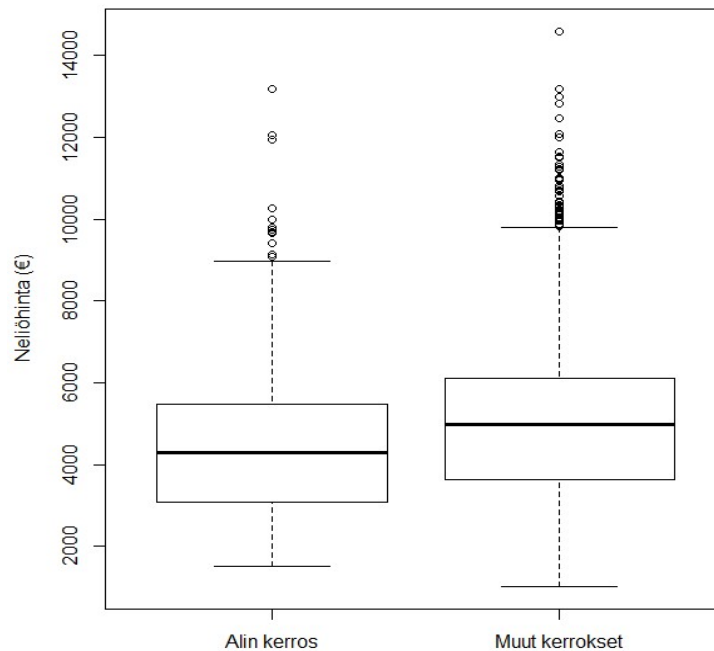
3.2.3 Asunnon ominaisuuksia mallintavat muuttujat

Aineiston asuntojen asuinalan keskiarvo on noin 59 neliometriä ja mediaani 54 neliometriä. Asuinalan 95. persentiili on 107 neliometriä, mutta kuvioista 3.10 nähdään, että aineistossa on runsaasti tätä huomattavasti suurempia asuntoja, eli hajonta on melko suurta. Kuvioista nähdään, että yksittäisten asuinneliöiden arvo hieman laskee asuinalan kasvaessa.



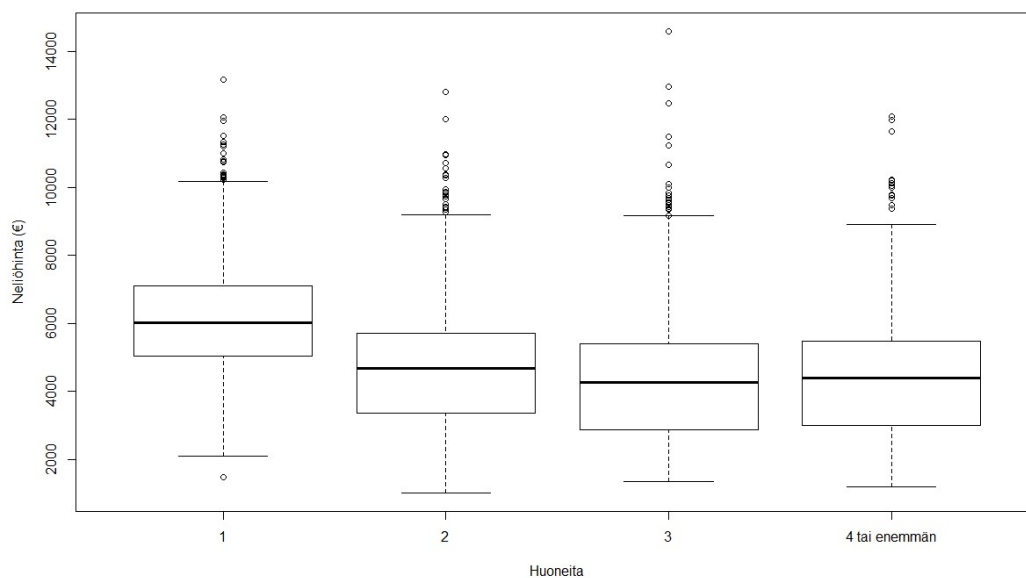
Kuvio 3.10. Neliöhinnan ja asuinalan välinen hajontakuvio.

Asuinkerroksen hintavaikutuksen kuvaamista vaikeuttaa sen voimakas riippuvuus rakennusvuodesta. Esimerkiksi 50-luvulla rakennetut rakennukset ovat harvoin kymmenen kerroksisia. Tämän vuoksi tässä tutkielmassa on rajattu asuinkerroksen tarkastelu siihen, että sijaitseeko asunto ensimmäisessä kerroksessa vai ylemmissä. Noin 19 prosenttia aineiston asunnoista sijaitsee ensimmäisessä asuinkerroksessa. Ensimmäisessä asuinkerroksessa olevien asuntojen mediaanihinta oli noin 4 286 euroa ja ylemmissä kerroksissa sijaitsevien 4 980, eli ero on huomattava, mikä näkyy selvästi kuviossa 3.11. Alimman kerroksen asuntojen neliöhinnat olivat keskimäärin noin 574 euroa matalammat.



Kuvio 3.11. Laatikkokuviot asuntojen neliöhinnasta sen mukaan sijaitseeko asunto alimmassa kerroksessa vai ylempänä.

Aineiston asuntojen huonelukumäärää tarkasteltaessa yleisimpiä huoneistotyyppisiä ovat kaksiot, jotka muodostavat noin 44 prosenttia aineistosta. Empiiristä osiota varten asunnot, joissa huoneita on neljä tai enemmän on yhdistetty yhdeksi luokaksi. Keskimäärin neliöhinnat ovat korkeimmat yksiöissä, mikä näkyy kuviossa 3.12. Asiaan kuitenkin vaikuttanee voimakkaasti taustamuuttujana asuntojen koot. Jotta huonelukumäärän todellinen vaikutus saadaan arvioitua empiirisessä osiossa, täytyy huoneiston koon olla otettuna huomioon. Edellisessä pääluvussa olevan oletuksen mukaan huonelukumäärällä on päinvastainen vaikutus asunnon arvoon, kuin mitä kuvio 3.12 antaa ymmärtää, kun taustamuuttujien vaikutusta ei ole vakioitu.

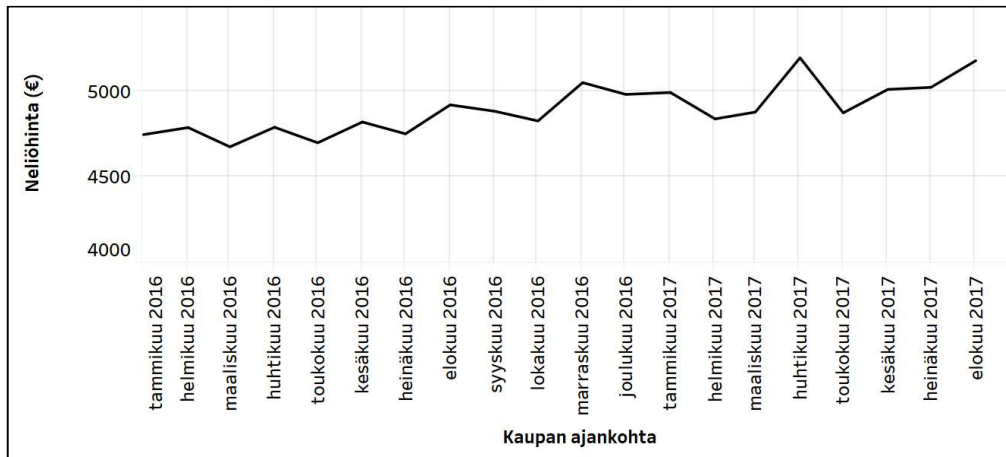


Kuvio 3.12. Laatikkokuvio asuntojen neliöhinnoista huonelukumäärän mukaan.

3.2.4 Hintakehitys aineistossa

Kuten johdantoluvussa mainittiin, on asuntojen hinnoilla ollut nouseva trendi Helsingissä. Kuviossa 3.13 on aineiston keskineliöhinnat eri kuukausina. Myös kuviossa on nähtävissä nouseva trendi – aineiston neliöhinnat ovat nousseet keskimäärin noin 23 euroa kuukaudessa ja keskimääräinen kuukausimuutos on ollut 0,5 prosenttia. Tutkielman empiirisessä osiossa kaupan ajankohtaa käytetään taustamuuttujana, jolla

pyritään vakioimaan ajankohdan vaikutus aineiston sisällä. Tällä pyritään vähentämään ajankohdasta johtuvaa vaihtelua ja siten vähentämään mallin harhaisuutta.



Kuvio 3.13. Hintojen ajallinen kehitys aineistossa.

4 LINEAARINEN REGRESSIOMALLI

Tämän tutkielman empiirisessä osiossa pyritään muodostamaan lineaarinen regressiomalli, jonka avulla kuvataan tutkielman aineiston asuntojen hinnanmuodostusta. Lineaarista regressioanalyysia käytetään yhden selitettävän muuttujan y ja yhden tai usean selittävän muuttujan, x_1, \dots, x_k , välisen lineaarisen yhteyden mallintamiseen tai selittämiseen. Kun $k = 1$, eli mallissa on vain yksi selittävä muuttuja, kutsutaan sitä yksinkertaiseksi lineaariseksi regressioksi. Kun selittäviä muuttujia on enemmän kuin yksi, kutsutaan mallia usean selittäjän lineaariseksi regressioksi. Regressioanalyysillä voidaan muun muassa ennustaa selitettävän muuttujan, eli vastemuuttujan, arvoja tiedossa olevilla selittävien muuttujien arvoilla tai arvioida selittävien muuttujien vaikutusta – tai niiden välisiä yhteyksiä – vastemuuttujaan. Lineaarinen regressiomalli kirjoitetaan muotoon

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \epsilon, \quad (4.1)$$

jossa β_0 on vakiotermi, β_k regressiokerroin selittävälle muuttujalle x_k ja ϵ on mallin virhetermi. Matriisimuodossa malli kirjoitetaan:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad (4.2)$$

jossa

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \dots \\ \beta_k \end{bmatrix}, \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \dots \\ \epsilon_n \end{bmatrix}.$$

Lineaariseen regressiomalliin liittyy useita oletuksia. Mallin selittävien muuttujien tulee kuvata riittävästi selitettävän muuttujan vaihtelua ja niitä tulee olla riittävästi. Lisäksi mallin tulee sopia niihin yksiköihin, joihin sitä sovelletaan, toisin sanoen poikkeavat havainnot täytyy ottaa huomioon. Selitettävän muuttujan ja regressoreiden välinen yhteys tulee olla likimain lineaarinen. Virhiden tulee olla riippumattomia. Virhetermit ovat homoskedastisia, eli virhetermien varianssi on kaikilla sama. Lisäksi oletetaan, että yksi-

kään \mathbf{X} :n sarake ei ole lineaarinen kombinaatio muista sarakkeista. Edellä olevien standardioletuksien lisäksi usein tehdään lisäoletus virhetermien normaalisuudesta. Oletuksien toteutumista tutkitaan regressiodiagnostiikan avulla. (Faraway 2002.)

4.1 Pienimmän neliösumman menetelmä

Mallin parametrien estimointiin voidaan käyttää pienimmän neliösumman menetelmää (PNS-menetelmä). PNS-menetelmässä estimoidaan parametrit regressiokertoimien vektorille $\hat{\beta}$, joka minimoi residuaalien neliösumman:

$$S(\beta) = (\mathbf{y} - \mathbf{X}\beta)'(\mathbf{y} - \mathbf{X}\beta). \quad (4.3)$$

Mikä voidaan esittää edelleen muodossa:

$$\begin{aligned} S(\beta) &= \mathbf{y}'\mathbf{y} - \beta'\mathbf{X}'\mathbf{y} - \mathbf{y}'\mathbf{X}\beta + \beta'\mathbf{X}'\mathbf{X}\beta \\ &= \mathbf{y}'\mathbf{y} - 2\beta'\mathbf{X}'\mathbf{y} + \beta'\mathbf{X}'\mathbf{X}\beta. \end{aligned} \quad (4.4)$$

Virhetermien neliösumman $S(\beta)$ minimi on sen derivaatan nollakohdassa β :n suhteen:

$$\frac{\partial S}{\partial \beta} = -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\beta = \mathbf{0}, \quad (4.5)$$

josta saadaan estimoidut parametrit, ratkaisemalla yhtälö β :n suhteen. Yhtälö voidaan sieventää muotoon

$$\mathbf{X}'\mathbf{X}\beta = \mathbf{X}'\mathbf{y}. \quad (4.6)$$

Yhtälö 4.6 on pienimmän neliösumman normaaliyhtälö. Normaaliyhtälön ratkaisemiseksi kerrotaan molemmat puolet vasemmalta $\mathbf{X}'\mathbf{X}$:n käänteismatriisilla, millä saadaan β :n pienimmän neliösumman estimaattori:

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}. \quad (4.7)$$

Käänteismatriisi $(\mathbf{X}'\mathbf{X})^{-1}$ on aina olemassa, jos yksikään \mathbf{X} :n sarake ei ole lineaarinen kombinaatio muista sarakkeista. Estimoidun mallin sovitearvojen vektori $\hat{\mathbf{y}}$ voidaan kirjoittaa muotoon

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\beta} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \mathbf{H}\mathbf{y}, \quad (4.8)$$

jossa $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ on $n \times n$ matriisi, jota kutsutaan hattumatriisiksi. (Montgomery, Peck & Vining 2012.) Farawayn (2002) mukaan $\hat{\beta}$ on paras lineaarinen harhaton estimaattori, kun mallin oletukset ovat voimassa. Montgomery et la. (2012) tarkentavat, että puhuttaessa ”parhaasta” viitataan minimoituun varianssiin.

4.2 Regressiokertoimien ja mallin merkitsevyys sekä mallin selitysaste

Regressiomallin merkitsevyyden testaamiseksi tutkitaan, selittävätkö mallin selittäjät yhdessä selitettävän muuttujan vaihtelua. Eli onko ainakin yhdellä mallin selittäjällä merkitsevä yhteys selitettävään muuttujaan. Testin hypoteesit voidaan kirjoittaa muotoon:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1: \beta_j \neq 0 \text{ ainakin yhdellä } j: n \text{ arvolla, } j = 1, 2, \dots, k.$$

Jos nollahypoteesi hylätään, ainakin yksi selittävistä muuttujista selittää vastemuuttujan arvojen vaihtelua. Hypoteesien testaamiseksi kokonaisneliösumma jaetaan kahteen osaan, mikä voidaan kirjoittaa:

$$SS_T = SS_R + SS_{Res}.$$

SS_T on kokonaisneliösumma, joka kuvaa vastemuuttujan arvojen kokonaisvaihtelua, SS_R on regressioneliösumma, joka kuvaa sen osuuden vastemuuttujan vaihtelusta, jonka regressiomalli selittää ja SS_{Res} on jäännöseliösumma, joka kuvaa sen osuuden, mitä malli ei ole pystynyt selittämään. Neliösummien kaavat voidaan kirjoittaa muotoon:

$$SS_R = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2, \quad (4.9)$$

$$SS_{Res} = \sum_{i=1}^n (y_i - \hat{y}_i)^2. \quad (4.10)$$

Hypoteesien testausta varten lasketaan testisuure:

$$F = \frac{SS_R/k}{SS_{Res}/(n-k-1)} \sim F_{k, n-k-1} \quad (4.11)$$

Testisuure noudattaa Fisherin F-jakaumaa vapausastein k ja $n-k-1$ nollahypoteesin ollessa voimassa. Tässä k on selittäjien lukumäärä ja n havaintojen lukumäärä. Eli nollahypoteesi

voidaan testata F-testin avulla. Nollahypoteesi hylätään riskitasolla α , jos laskettu testisuure F on suurempi kuin Fisherin F-jakauman $(1 - \alpha)$ -fraktiilipiste vapausastein k ja $n-k-1$ ($F_{\alpha; k, n-k-1}$). (Montgomery et al. 2012; Weisberg 2005.)

Neliösummien arvoja voidaan käyttää edelleen regressiomallin selityksasteen (R^2) laskemiseen, jolla voidaan tutkia mallin sopivuutta. Selityksaste kuvaa, kuinka suuren osan malli selittää vastemuuttujan kokonaisvaihtelusta. Selityksasteen kaava voidaan kirjoittaa muotoon:

$$R^2 = \frac{SS_R}{SS_T} = 1 - \frac{SS_{Res}}{SS_T} \quad (4.12)$$

(Fidell & Tabachnick 2014). Eri malleja vertailtaessa käytetään usein korjattua selityksastetta (R^2_{Adj}), koska R^2 kasvaa aina, kun malliin lisätään muuttujia, riippumatta siitä, että selittävätkö uudet muuttujat todellisuudessa vastemuuttujaa. Korjatun selityksasteen kaava kirjoitetaan muotoon:

$$R^2_{Adj} = 1 - \frac{SS_{Res}/(n-k)}{SS_T/(n-1)} \quad (4.13)$$

missä $SS_{Res}/(n-k)$ on keskineliövirhe ja $SS_T/(n-1)$ on vakioinen riippumatta siitä, montako selittäjää mallissa on. Lisättäessä malliin uusi muuttuja korjatun selityksasteen arvo kasvaa ainoastaan, jos lisätty muuttuja pienentää keskineliövirhettä. Korjattu selityksaste siis rankaisee hyödyttömien muuttujien lisäämisestä. (Montgomery et al. 2012.)

Jos malli todetaan merkitseväksi, voidaan sen ja selityksasteen lisäksi tutkia mallin selittävien muuttujien merkitsevyyttä mallissa. Tällöin hypoteesit voidaan kirjoittaa muotoon:

$$H_0: \beta_j = 0$$

$$H_1: \beta_j \neq 0.$$

Nollahypoteesin testaamiseksi lasketaan testisuure:

$$t = \frac{\hat{\beta}_j}{s(\hat{\beta}_j)} \sim t_{n-k-1}, \quad (4.14)$$

jossa $s(\hat{\beta}_j)$ on parametrin β_j estimaatin hajonta. Testisuure noudattaa Studentin t-jakaumaa vapausastein $n-k-1$ nollahypoteesin ollessa voimassa, jolloin nollahypoteesi voidaan testata t-testin avulla. Nollahypoteesi hylätään riskitasolla α , jos lasketun t -arvon itseisarvo on suurempi kuin Studentin t-jakauman $(1 - \alpha/2)$ -fraktiilipiste vapausastein $n-k-1$ ($t_{\frac{\alpha}{2}; n-k-1}$). Jos nollahypoteesi jää voimaan, muuttuja x_j ei mallissa selitä vastemuuttujaa merkitsevästi, jolloin se voidaan poistaa mallista. Huomioitavaa on, että tämä testi testaa selittäjän merkitsevyyttä yhdessä muiden selittäjien kanssa, joten sen estimoitu parametri riippuu myös muista selittäjistä. Tämän vuoksi parametrien t -arvot tulee laskea aina uudelleen, kun malliin lisätään tai mallista poistetaan selittävä muuttuja.

4.3 Muuttujien muunnokset

Regressiomalleja rakennettaessa usein päädytään tekemään erilaisia muuttujien muunnoksia sekä mahdollisesti myös lisäämään muunnoksia malliin mallissa jo olevien selittäjien lisäksi. Muuttujien muunnoksien avulla muun muassa saadaan luokittelevat muuttajat eli faktorit halutessa mukaan malliin, varmistettua mallin oletuksien täyttyminen esimerkiksi stabiloimalla varianssia tai lineaarisoinnalla mallia ja muutettua mallin parametrejä mallin tulkinnan kannalta mielekkäämmiksi. Montgomery et al. (2012) huomauttavat, että optimaalisessa tilanteessa muunnokset tekisi tieteen tekijä, jolla on substanssietämystä, mutta monissa tilanteissa siihen ei ole kuitenkaan mahdollisuutta, jolloin muunnokset voidaan valita intuitiivisesti tai analyttisten päätelmien avulla. Seuraavaksi käydään läpi tässä tutkielmassa käytettäviä muunnoksia.

Jotta kategorisia muuttujia voidaan käyttää regressiomalleissa, täytyy olla tapa esittää niiden arvoja numeerisesti. Tämä saavutetaan muodostamalla dummy-muuttujia kategorisista muuttujista. Dummy-muuttuja on muuttuja, joka saa arvon 0 tai 1 riippuen siitä, onko havainnossa dummy-muuttujan kuvaama ominaisuus. Yhdestä kategorisesta muuttujasta muodostettavien dummy-muuttujien lukumäärä on yksi vähemmän kuin kyseisen muuttujan eri luokkien lukumäärä. Tällöin yksi luokka toimii niin sanottuna vertailuryhmänä. Tarkastellaan esimerkkiä, jossa kategorisena selittäjänä on koulutustaso, jossa on kolme eri luokkaa (kandidaatti, maisteri, tohtori). Tällöin muodostettavat dummy-muuttajat voisivat olla muotoa:

$$D_{maisteri} = 1, \text{ jos maisteri, muuten } 0,$$

$$D_{tohtori} = 1, \text{ jos tohtori, muuten } 0.$$

Tässä tilanteessa vertailuryhmä olisi ”kandidaatti”, johon havainto kuuluu, jos molemmat yllä olevat dummy-muuttujat saa arvon 0. (Montgomery et al. 2012.)

Potenssimuunnos on usein käytetty muunnos, jossa muuttujan arvo muunnetaan johonkin sen potenssiin. Potenssimuunnosten avulla voidaan stabiloida varianssia ja linearisoida mallia. Esimerkiksi puolikkaalla potenssilla eli neliöjuurimuunnoksella voidaan vähentää suurien lukujen vaikutusta. Esimerkiksi, jos muuttujana on vuositulot, todellisuudessa vuositulojen ero 20 000 ja 30 000 euron välillä tuntunee suuremmalta kuin ero esimerkiksi 100 000 ja 110 000 euron välillä. Selitettävän ja selittävän muuttujan yhteys voi olla lineaarisen sijaan eksponentiaalinen, joka on mahdollista linearisoida esimerkiksi korottamalla selittäjä kolmanteen potenssiin. Muuttujien väliset riippuvuudet voivat olla myös joskus polynomin muotoisia, jolloin niiden välistä yhteyttä voidaan kuvata n:nneen asteen polynomin avulla. Esimerkiksi malliin voidaan sisällyttää selittäjä x_1 ja sen neliö x_1^2 , joiden avulla voidaan sovittaa malliin vastemuuttujan ja selittäjän x_1 välinen toisen asteen polynomin muotoinen riippuvuus. (Weisberg 2005.) Faraway (2002) huomauttaa, että polynomin avulla riippuvuutta selitettäessä polynomin merkitsevyyttä tulisi tarkastella polynomin korkeimman asteen termin mukaan, eikä alempia termejä saisi poistaa, vaikka ne olisi parametrien testausten mukaan tilastollisesti ei-merkitseviä, koska polynomin eri termit luonnollisesti riippuvat toistensa arvoista.

Myös logaritmuunnos on suosittu muunnos, jonka avulla voidaan saada mallia linearisoitua ja virhetermien varianssia stabiloitua. Weisberg (2005) mainitsee kirjassaan peukalosäännöksi, jos muuttujan arvot vaihtelevat yli sadan yksikön verran, sen korvaaminen logaritmoidulla arvolla mallissa on todennäköisesti hyödyllistä. Lisäksi logaritmuunnoksesta voi olla hyötyä, jos halutaan mallin parametrit tulkinnan kannalta mielekkäämmiksi. Esimerkiksi, jos halutaan tutkia selittäjien suhteellista vaikutusta vastemuuttujaan, on vastemuuttujasta tehty logaritmuunnos hyödyllinen. Tällöin mallin tulkinta muuttuu siten, että selittävän muuttujan x_k muuttuessa yksikön verran vastemuuttuja muuttuu mallin mukaan $100 \times \beta_k$ prosenttia.

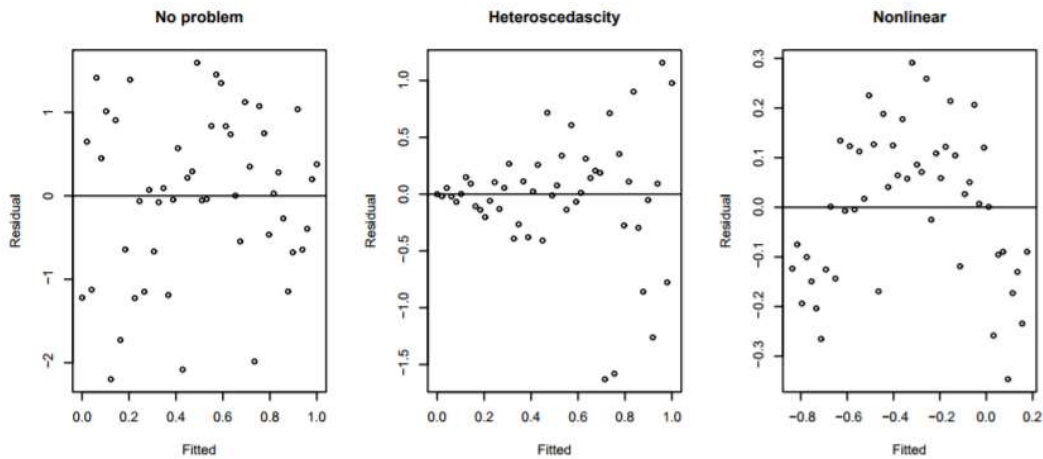
4.4 Regressiodiagnostiikka

Regressiodiagnostiikkaa käytetään mallin sovittamisen jälkeen selvittämään, soveltuuko muodostettu malli kuvaamaan aineistoa ja täyttääkö se riittävän hyvin regressiomallin oletukset. Diagnostiikka perustuu pääosin mallin jäännöksiin eli residuaalien tutkimiseen, jonka on todettu olevan tehokas tapa tutkia, kuinka hyvin malli sovittaa aineiston ja tarkistaa mallin oletukset. Residuaalien lisäksi tarkastellaan niin sanottuja vaikutusvaltaisia tai muulla tavalla poikkeavia havaintoja. Karkeat rikkomukset mallin oletuksien kohdalla voivat tuottaa epästabiilin mallin, jossa eri otoksella voisi tulla täysin erilaisia tuloksia ja johtopäätelmiä. (Weisberg 2005.)

Mallin residuaalit saadaan yhtälöstä:

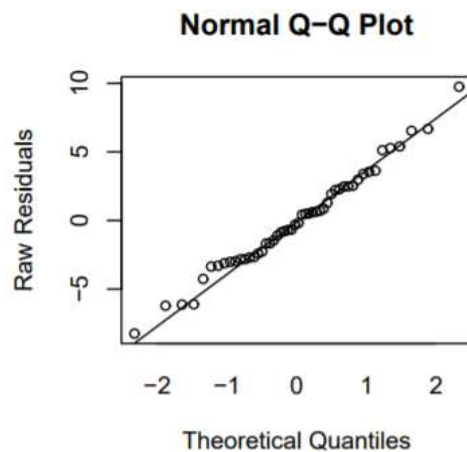
$$e_i = y_i - \hat{y}_i, i = 1, 2, \dots, n, \quad (4.15)$$

jossa y_i on havaittu arvo ja \hat{y}_i on vastaava sovitettu arvo. Montgomery et al. (2012) mainitsevat, että heidän kokemuksensa mukaan regressiomallin residuaalien tilastolliset testaukset eivät ole laajasti käytössä. Sen sijaan he huomauttavat, että käytännön tilanteissa useimmiten residuaalidiagrammit ovat informatiivisempia kuin vastaavat tilastolliset testit. Tehokas tapa residuaalien tarkasteluun on piirtää ne kuvaajaan suhteessa mallin sovitettuihin arvoihin, missä y-akselilla on residuaalien arvot ja x-akselilla sovitetut arvot. Hyvin muodostetun regressiomallin tulisi muodostaa vyömäisen pisteparven, joka on vaakasuora ja sen leveys y-akselin suunnassa tulisi olla vakio x-akselin suhteen. (Cook & Weisberg 1999.) Kuviossa 4.1 on esimerkkinä kolme residuaalidiagrammia, joissa y-akselilla on residuaalien arvot ja x-akselilla sovitetut arvot. Ensimmäisen diagrammissa on pisteparvi, jonka perusteella mallin muodostuksessa ei ole ollut ongelmia. Toisessa diagrammissa on havaittavissa heteroskedastisuutta, eli residuaalien varianssi ei ole vakioinen, vaan se kasvaa sovitteen arvojen kasvaessa. Kolmannesta diagrammista puolestaan voi havaita epälinearisuutta, kun pisteparvi ei ole vaakasuora vaan muodostaa käyrän.



Kuvio 4.1. Kolme esimerkkiä residuaalidiagrammeista (Faraway 2002).

Residuaalien normaalisuuden tarkistamiseen hyödyllinen tapa on kvantiilikuvio (Q-Q plot), josta nähdään esimerkki kuviossa 4.2. Residuaalit ovat likimäärin normaalijakautuneita, jos ne muodostavat likimain suoran. Kuviossa y-koordinaatit ovat mallin residuaalit ja x-koordinaatit $N(0,1)$ -jakauman odotetut kvantiilit.



Kuvio 4.2. Esimerkki kvantiilikuviosta (Faraway 2002).

Edellä kuvatut menetelmät ovat tehokkaita mallin harhattomuuden, homoskedastisuuden ja normaalisuuden tarkasteluun. Yksittäisten havaintojen diagnostiikkaan suositellaan käytettäväksi skaalattuja residuaaleja, jotka ovat hyödyllisiä etsittäessä havaintoja, jotka

ovat poikkeavia, tai äärimmäisten arvoja, jotka toisin sanoen ovat havaintoja, jotka eroavat jollakin tavalla muusta aineistosta. Hyvä vaihtoehto tähän on tarkastella studentisoituja residuaaleja, jotka saadaan laskettua kaavalla:

$$r_i = \frac{e_i}{\sqrt{\left(\frac{SS_{Res}}{n-k-1}\right)(1-h_{ii})}}, i = 1, 2, \dots, n \quad (4.16)$$

jossa h_{ii} on hattumatriisin \mathbf{H} i. havainnon diagonaalialkio. Hattumatriisin diagonaalialkiot ovat niin sanottuja vipulukuja (leverage). Jos havainnon vipuluku on muihin verrattuna selvästi suurempi, niin kyseinen havainto on syrjässä muihin havaintoihin nähden selittävien muuttujien arvojen suhteen. Studentisoidut residuaalit ottavat siten huomioon sekä suuret residuaalit että suuret vipuluvut. Havainnot, joilla on suuri residuaali ja vipuluku, ovat todennäköisesti vaikutusvaltaisia, eli niillä on merkittävä vaikutus regressio-kertoimiin. (Montgomery et al. 2012.)

Havaintojen vaikutusvaltaisuuden mittaamisessa tutkitaan, kuinka paljon havainnon poistaminen vaikuttaisi malliin. Tähän yhtenä menetelmänä on Cookin etäisyys, jonka avulla mitataan yksittäisten havaintojen vaikutus regressiomallin parametreihin. Cookin etäisyyden yhtälö voidaan kirjoittaa muotoon:

$$D_i = \frac{1}{k\left(\frac{SS_{Res}}{n-k-1}\right)} \sum_{j=1}^n (\hat{y}_{(i)j} - \hat{y}_j)^2, \quad (4.17)$$

missä $\hat{y}_{(i)j}$ on j. havainnon sovitettu arvo mallilla, joka on estimoitu ilman i. havaintoa. Hyödyntäen studentisoitua residuaalia Cookin etäisyys saadaan myös yhtälöstä:

$$D_i = \frac{r_i^2}{k} \times \frac{h_{ii}}{1-h_{ii}}. \quad (4.18)$$

(Cook & Weisberg 1999.) Jos yksittäisien havaintojen studentisoitu residuaali ja/tai Cookin etäisyys on selvästi suurempi kuin yleisesti, kannattaa kyseiset havainnot ottaa erikoistarkasteluun ja tarvittaessa muodostaa malli ilman kyseisiä havaintoja.

4.5 Mallin valinta

Regressioanalyysin onnistumiseen kuuluu olennaisena osana oikean mallin löytäminen. Pyrkimyksenä on pystyä selittämään selitettävän muuttujan vaihtelua mahdollisimman tarkasti käyttäen kuitenkin mahdollisimman vähän selittäjiä. Toisin sanoen pyritään ottamaan malliin vain relevantit selittäjät mukaan, jotta malli pysyisi riittävän yksinkertaisena. Selittäjien rajaamisen taustalla voi myös olla talouteen tai aikaan liittyvät seikat: selittävien muuttujien arvojen mittaaminen voi olla kallista ja aikaa vievää, jolloin niiden priorisointi on perusteltua. Selittäjien rajaamisella voidaan parantaa mallin parametrien tulkittavuutta, kun puolestaan mallissa mukana olevat turhat selittäjät lisäävät kohinaa muihin kiinnostuksen kohteina oleviin mallin suureisiin.

Tutkittavan ilmiön ja aihealueen tuntemus on mallin muodostuksen kannalta tärkeää ja relevantit muuttujat saattavat olla hyvin tiedossa etukäteen. Lisäksi toisinaan voi olla tärkeää sisällyttää tutkimukseen mukaan aihealueen teorian kannalta tärkeitä muuttujia, vaikka ne eivät välttämättä mallissa olisi tilastollisesti merkitseviä. Selittävien muuttujien määrän rajaamisesta luovutaan joskus, kun halutaan pelkästään monimutkaisesta ennustuksesta mahdollisimman tarkka, eikä olla kiinnostuneita mallin parametrien tulkinnasta tai muista ominaisuuksista.

Selittävien muuttujien valintaan on olemassa monia erilaisia menetelmiä. Automaattisista mallinvalintamenetelmistä esimerkkejä ovat lisäävä valinta, poistava valinta ja täydellinen haku. Lisäävässä valinnassa käytetään seuraavaa strategiaa:

- 1) Lähdetään liikkeelle mallista, jossa ei ole ollenkaan selittäviä muuttujia.
- 2) Testataan jokaisen muuttujan lisäämistä malliin erikseen ja valitaan lisättäväksi muuttuja, jonka p-arvo on pienin ja alle halutun tason α .
- 3) Lisäämistä jatketaan niin kauan, kunnes ei ole lisättävissä muuttujia edellä mainituilla ehdoilla.

Poistava valinta lienee olemassa olevista mallinvalintamenetelmistä yksinkertaisin. Poistava valinta etenee seuraavasti, kunnes malli on valmis:

- 1) Muodostetaan malli, jossa on kaikki selittävät muuttujat.
- 2) Estimoidaan malli siinä olevilla selittäjillä.
- 3) Testataan selittäjien merkitsevyytensä valitulla tasolla α .

- 4) Jos mallin kaikkien selittäjien p-arvot ovat alle valitun tason, on malli valmis.
- 5) Poistetaan mallista selittäjä, jonka p-arvo on suurin.
- 6) Palataan vaiheeseen (2).

Täydellisessä haussa muodostetaan kaikki mahdolliset mallit, mitä selittäjien eri kombinaatiot voivat muodostaa. Näistä malleista valitaan jonkin mallinvalintakriteerin perusteella, esimerkiksi korjatun selitysasteen, yksi tai muutama parasta mallia jatkotarkasteluun. (Faraway 2002.)

Faraway (2002) mainitsee, että samasta aineistosta voidaan saada muodostettua useita erilaisia malleja johtuen erilaisista mallinvalintamenetelmistä tai prosessin eri osien (muunnokset, diagnostiikka, ...) erilaisesta suoritusjärjestyksestä johtuen, vaikka ei tehtäisi mitään selvästi väärin. Tästä syystä johtuen Faraway (2002) korostaa toisen tai kolmannen riippumattoman analyysin tuomaa lisäarvoa.

5 AINEISTON ANALYYSI

Tässä luvussa mallinnetaan asuntojen neliöhintoja muodostamalla regressiomallit pienimmän neliösumman menetelmällä hyödyntäen kolmannessa luvussa kuvattua aineistoa. Ensin kuvataan, miten muodostetaan niin sanottu päämalli, jolla pyritään selittämään neliöhintojen vaihtelua mahdollisimman hyvin. Tämän jälkeen kuvataan lyhyesti lisäksi kaksi muuta mallia, joiden avulla pyritään kuvaamaan lisää sijainnin vaikutusta neliöhintoihin keskustaetäisyyden ja -matka-ajan avulla. Mallien esittelyn jälkeen kuvataan mallien selittäjien estimoituja vaikutuksia neliöhintoihin.

5.1 Päämalli asunnon velattomalle neliöhinnalle

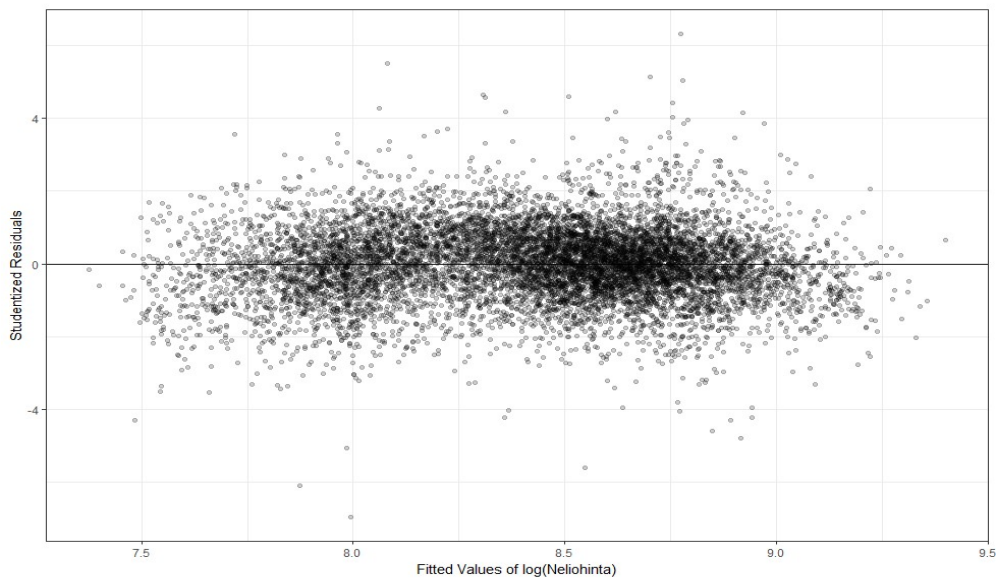
Tässä osiossa muodostetaan päämalli asunnon velattomalle neliöhinnalle. Ensin kuvataan alustavan mallin selittävät muuttujat ja muunnokset, minkä jälkeen tutkitaan mallin regressiodiagnostiikkaa. Diagnostiikan perusteella tehdään tarvittaessa muutoksia malliin sen parantamista varten. Lopuksi esitellään lopullinen malli.

Sovitettaessa aineistoon regressiomallia tehdään selitettävänä muuttujana olevalle asunnon velattomalle neliöhinnalle logaritmimuunnos, millä on linearisoiva vaikutus ja muuttaa regressiokertoimia siten, että vaikutukset tulkitaan suhteellisina absoluuttisten sijaan. Taulukosta 5 nähdään mallin sovittamisessa hyödynnetyt selittävät muuttujat. Muuttujista ”asuinala” ja ”ikä” on lisätty myös toisen ja kolmannen asteen termit, jotta malli kuvaisi niiden epälineaarisen ja -monotonisen yhteyden neliöhintaan. ”Asuinala”, ”ikä”, ”hoitovastike per neliö” ja kahdesta ensimmäisestä tehtävät muunnokset ovat mallissa jatkuvia muuttujia ja loput ovat luokittelevia muuttujia, joita käytetään dummy-muuttujina. Ennen mallin sovittamista postinumerot, joilla on alle kymmenen havaintoa, on yhdistetty lähellä olevaan postinumeroalueeseen.

Taulukko 5. Mallin (1) selittävät muuttujat.

Selittävä muuttuja
<i>Asuinala</i>
<i>Asuinala²</i>
<i>Asuinala³</i>
<i>Kiinteistön kunto</i>
<i>Tontin omistajuus</i>
<i>Hoitovastike per neliö</i>
<i>Ikä</i>
<i>Ikä²</i>
<i>Ikä³</i>
<i>Alin kerros</i>
<i>Huoneita</i>
<i>Kaupun ajankohta</i>
<i>Postinumero</i>

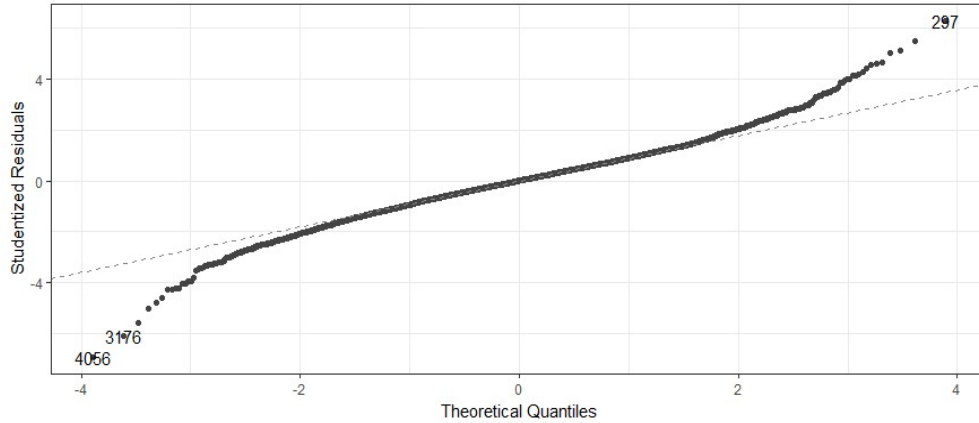
Taulukon 5 muuttujia käytettäessä selittäjinä saadaan logaritmoidulle neliöhinnalle sovitettua alustava regressiomalli, jonka selitysaste on 88,9 prosenttia, eli selittävät muuttujat selittävät logaritmoidun neliöhinnan vaihtelusta noin 89 prosenttia. Kuvion 5.1 perusteella malli vaikuttaa melko homoskedastiselta, eikä epälinearisuutta ole havaittavissa kuvion perusteella. Itseisarvoltaan suurimmat residuaalit on hyvä ottaa erikseen tarkasteluun. Näitä näyttäisi olevan aineiston kokoon nähden varsin maltillisesti.



Kuvio 5.1. Mallin (1) studentisoidut residuaalit sovitettujen arvojen suhteen.

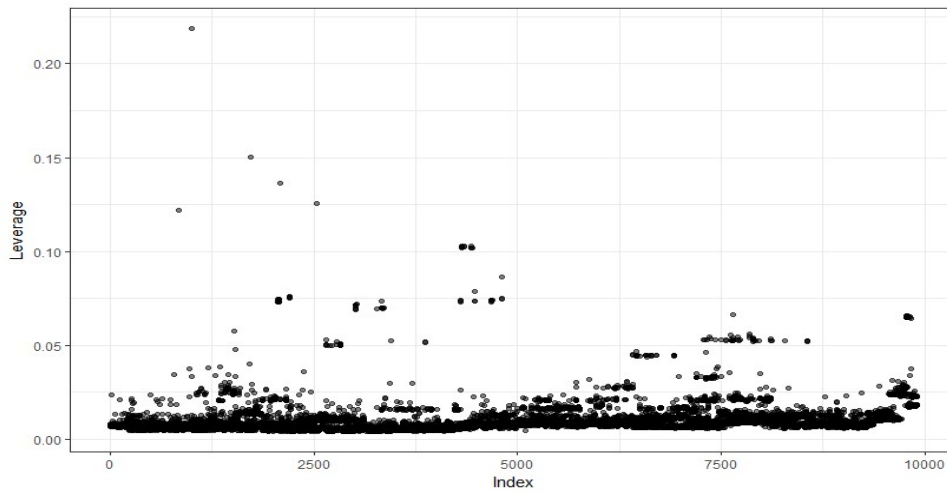
Kuvion 5.2 perusteella studentisoidut residuaalit ovat normaalijakaumaa vahvähäntäisemmästä jakaumasta. Itseisarvoltaan suurimpien residuaalien arvot ovat niiden kvanttilien odotettuja arvoja jonkin verran suurempia, minkä vuoksi jakauman hännät ovat ole-

tettua paksummat. Tarkemmasta tarkastelusta selviää, että poikkeuksellisen suuret studentisoidut residuaalit tulevat kohteille, joiden velaton hinta on ollut selvästi poikkeuksellisen suuri tai pieni.



Kuvio 5.2. Kvantiilikuvio mallin (1) studentisoiduista residuaaleista.

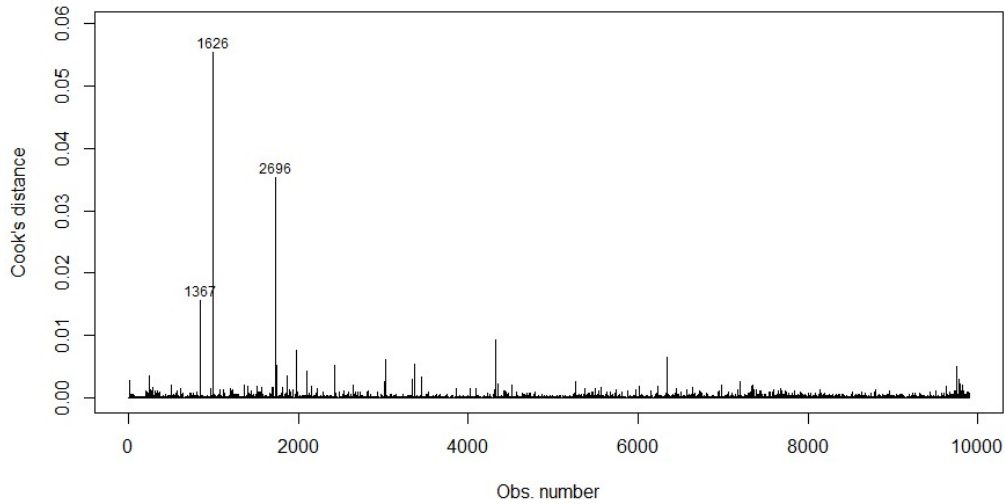
Kuviosta 5.3 nähdään havaintojen vipuluvut mallissa (1). Nähdään, että aineistosta löytyy havaintoja, joiden vipuluvut poikkeavat selvästi, eli ne ovat selittävien muuttujien suhteen syrjässä muihin havaintoihin nähden. Poikkeavat vipuluvut otetaan erikseen tarkasteluun, mikä auttaa paljastamaan mallin ja/tai aineiston hyvytyteen liittyviä asioita.



Kuvio 5.3. Havaintojen vipuluvut mallissa (1).

Lisäksi tutkitaan havaintojen vaikutusvaltaisuutta ja poikkeavuutta laskemalla Cookin etäisyydet havainnoille. Kuviosta 5.4 nähdään, että kolmen havainnon Cookin etäisyydet ovat huomattavasti suurempia muuhun aineistoon nähden. Niiden lisäksi tarkastellaan

myös muita havaintoja, joiden Cookin etäisyydet poikkeavat huomattavasti keskimääräisestä tasosta.



Kuvio 5.4. Havaintojen Cookin etäisyydet.

Tarkasteltaessa poikkeavia residuaaleja, vipulukuja ja Cookin etäisyyksiä nousee usein esille kohteita, jotka ovat hyvin vanhoja tai asuinaltaan poikkeuksellisen suuria – varsinkin erittäin suurten kohteiden kohdalla edellä mainitut tunnusluvut ovat huomattavasti keskimääräistä tasoa korkeammat. Tarkemman tarkastelun perusteella malli ylisovittaa iän ja varsinkin asuinalan vaikutuksen niiden suurilla arvoilla. Muuttujaa ”ikä” muunnetaan siten, että se saa arvon 100, jos kohteen ikä on sata vuotta tai enemmän. Muuttujaa ”asuinala” puolestaan muokataan siten, että se saa arvon 120, jos kohteen asuinala on 120 tai enemmän. Lisäksi muunnetun muuttujan ”asuinala” sijaan käytetään mallin muodostamisessa apumuuttujaa, jonka arvot ovat muunnettu muuttuja ”asuinala” vähennettynä luvulla 11. Tämän apumuuttujan käyttäminen ei itsessään muuta mallin toimivuutta, mutta helpottaa asuinalaan liittyvien regressioparametrien tulkitsemista, kun asunnon referenssikokona mallissa on täten 11 neliometriä (aineiston minimi). Sen sijaan, että referenssikokona olisi nolla neliometriä, mikä ei arvona ole todellisuudessa mahdollinen. Muunnosten jälkeen malli sovitetaan uudelleen. Osa mallista on nähtävillä taulukossa 6, ja liitteestä C löytyy koko malli sisältäen myös vakiotermin sekä postinumeroalueiden regressiokertoimet.

Taulukko 6. Sovitettu regressiomalli (1).

<i>Dependent variable:</i>	
log(Neliöhinta)	
Asuinala - 11	-0.02542198*** (0.0008196579)
(Asuinala - 11) ²	0.00027206*** (0.00001525247)
(Asuinala - 11) ³	-0.0000009177137*** (0.00000008353553)
Kiinteistön kunto: Tyydyttävä	0.08415202*** (0.006598253)
Kiinteistön kunto: Hyvä	0.184245*** (0.006502972)
Kiinteistön kunto: Erinomainen	0.2172594*** (0.01153948)
Kiinteistön kunto: Uusi	0.2059082*** (0.01029274)
Tontin omistajuus: Vuokrattu	-0.04170708*** (0.004246858)
Hoitovastike per neliö (€)	-0.01381239*** (0.001459299)
Ikä	-0.0132997*** (0.0006352476)
Ikä ²	0.0001674645*** (0.00001429135)
Ikä ³	-0.0000005552989*** (0.00000008903838)
Kerros: Alin	-0.03128393*** (0.003390344)
Huoneita: 2	0.03121159*** (0.005686906)
Huoneita: 3	0.05401325*** (0.008043051)
Huoneita: 4 tai enemmän	0.01633077 (0.01034577)
Pvm: 2016-02	0.01651909** (0.008285337)
Pvm: 2016-03	0.006865484 (0.008348761)
Pvm: 2016-04	0.02683272*** (0.008286704)
Pvm: 2016-05	0.03077974*** (0.00829173)
Pvm: 2016-06	0.02290448*** (0.00817587)
Pvm: 2016-07	0.03309258*** (0.009068635)
Pvm: 2016-08	0.03401857*** (0.008458457)
Pvm: 2016-09	0.03846425*** (0.0082391)
Pvm: 2016-10	0.03933082*** (0.008366595)
Pvm: 2016-11	0.03738475*** (0.008198579)
Pvm: 2016-12	0.04431445*** (0.008925722)
Pvm: 2017-01	0.0480807*** (0.008767572)
Pvm: 2017-02	0.04919527*** (0.008358947)
Pvm: 2017-03	0.0629785*** (0.007974398)
Pvm: 2017-04	0.07406198*** (0.00835245)
Pvm: 2017-05	0.08072626*** (0.008313289)
Pvm: 2017-06	0.07957667*** (0.008136912)
Pvm: 2017-07	0.07874524*** (0.008556635)
Pvm: 2017-08	0.08140489*** (0.01763745)
Observations	9,901
R ²	0.8907134
Adjusted R ²	0.8895419
Residual Std. Error	0.128467 (df = 9795)
F Statistic	760.3025*** (df = 105; 9795)

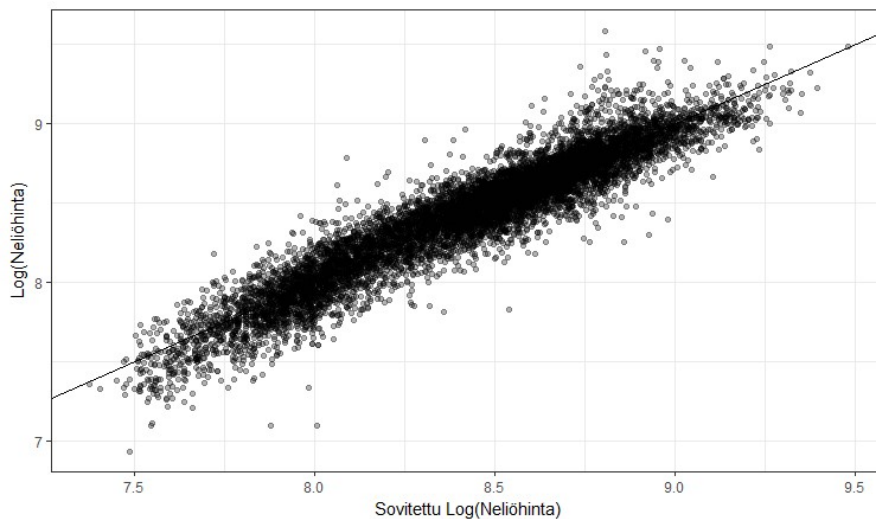
Note:

*p<0.1; **p<0.05; ***p<0.01

Taulukosta 6 nähdään postinumeroalueita lukuun ottamatta mallin selittäjien regressio-kertoimet ja kyseisten parametrien estimaattien hajonnat. Selitysaste on noin 89,1 prosenttia, jota voidaan pitää melko korkeana. Kiinteistön kuntoa kuvatessa vertailuluokkana on luokka ”huono”, tontin omistajuuden kohdalla vertailuluokkana on ”oma tontti”, asuinkerroksen vertailuluokkana on muut kerrokset kuin alin, huonelukumäärän vaikutusta verrataan yksiöihin ja myyntikuukauden vertailukuukausi on tammikuu 2016. Lisäksi postinumeroalueiden vertailuluokkana on postinumeroalue 00790.

Valmiille mallille tehdystä regressiodiagnostiikasta selviää, että kymmenen suurinta vipulukua saavaa kohdetta sijaitsevat kaikki 00830 postinumeroalueella, eli Tammisalossa. Kyseisen alueen kohteiden välillä on hintaeroja, joita malli ei pysty selittämään, eikä aineistossa ilmene selkeitä virheitä. Edellä tehtyjen muunnosten myötä Cookin etäisyyksissä ei ole enää yhtä poikkeavia yksittäisiä arvoja kuin aiemmassa mallin versiossa. Lopullisen mallin regressiodiagnostiikka on nähtävillä liitteessä D.

Kuviosta 5.5 nähdään, että todellisten logaritmoitujen neliöhintojen ja mallin ennustamien arvojen riippuvuus on hyvin lineaarinen ja voimakas. Kuvio vahvistaa mallin kuvaavan tehokkaasti logaritmoidun neliöhinnan vaihtelua siihen valittujen selittävien muuttujien avulla, eli pienimmän neliösumman menetelmällä sovitettu regressiomalli on toimiva tapa tutkia asuntojen hintoja.



Kuvio 5.5. Logaritmoidun neliöhinnan vastearvot sovitetten suhteen.

5.2 Lisämallit asunnon velattomalle neliöhinnalle

Tässä aliluvussa kuvataan lyhyesti edellä olevan päämallin lisäksi kaksia mallia, joilla pyritään kuvaamaan keskustaetäisyyden ja -matka-ajan yhteyttä asuntojen velattomiin neliöhintoihin. Malleihin on sovellettu samoja muunnoksia kuin edellä ja postinumeron tilalla malleissa on selittäjinä keskustaetäisyys kilometreinä ja matka-aika keskustaan minuutteina, joka on laskettu keskiarvona matka-ajasta autolla ja julkisella liikennevälineellä. Lisäksi matka-aikaa kuvaavaa muuttujaa muunnetaan siten, että se saa arvon 10, jos matka-aika on vähemmän kuin kymmenen minuuttia. Alle kymmenen minuutin matka-ajan päässä olevien kohteiden kohdalla ei liene mielekästä tarkastella liikkumista keskustaan autolla tai julkisilla, kun todennäköisempi vaihtoehto liikkumiselle lienee esimerkiksi kulkeminen kävellen tai pyörällä. Lisäksi näin lyhyillä matka-ajoilla mittausharhan riski on suurempi, johtuen esimerkiksi pysäköintipaikan etsinnästä tai kulkuneuvon odottamisesta pysäkillä. Edellä mainituista selittäjistä on mukana malleissa myös toisen ja kolmannen asteen termit.

Taulukko 7. Sovitettut lisämallit.

	<i>Dependent variable:</i>	
	Malli (2)	Malli (3)
	log(Neliöhinta)	
Asuinala - 11	-0.02611192*** (0.001236516)	-0.02497404*** (0.001015561)
(Asuinala - 11) ²	0.0002736951*** (0.00002296329)	0.0002527122*** (0.00001886309)
(Asuinala - 11) ³	-0.0000007825484*** (0.0000001256826)	-0.0000007347205*** (0.0000001032267)
Keskusta_min	0.1938214*** (0.008441604)	
Keskusta_min ²	-0.01058981*** (0.0004206327)	
Keskusta_min ³	0.0001577013*** (0.000006587723)	
Keskusta_km		-0.02164429*** (0.004598685)
Keskusta_km ²		-0.008234211*** (0.0006706561)
Keskusta_km ³		0.0004259881*** (0.00002797144)
Kiinteistön kunto: Tyydyttävä	0.08604666*** (0.009992506)	0.08680118*** (0.008210099)
Kiinteistön kunto: Hyvä	0.2002548*** (0.009845062)	0.194936*** (0.008086652)
Kiinteistön kunto: Erinomainen	0.2538597*** (0.01745674)	0.2454905*** (0.01433428)
Kiinteistön kunto: Uusi	0.1626505*** (0.01508661)	0.2063718*** (0.01250108)
Tontin omistajuus: Vuokrattu	-0.1131526*** (0.005030902)	-0.0718522*** (0.004254958)
Hoitovastike per neliö (€)	-0.02049356*** (0.002125844)	-0.01717859*** (0.001753748)
Ikä	-0.02326167*** (0.0008185176)	-0.01791697*** (0.0006888386)
Ikä ²	0.0003876547*** (0.0000182698)	0.0002600829*** (0.00001548514)
Ikä ³	-0.000001704799*** (0.0000001137028)	-0.000001056119*** (0.00000009700386)
Kerros: Alin	-0.04096193*** (0.005100829)	-0.03174463*** (0.004188366)

Huoneita: 2	0.05436043*** (0.008554291)	0.05357389*** (0.007023606)
Huoneita: 3	0.06795138*** (0.01201423)	0.08415964*** (0.009875275)
Huoneita: 4 tai enemmän	0.009806281 (0.01544744)	0.0400837*** (0.01271695)
Pvm: 2016-02	0.01995446 (0.01254508)	0.01448849 (0.01030445)
Pvm: 2016-03	-0.0009536472 (0.01264357)	-0.0003743476 (0.01038461)
Pvm: 2016-04	0.02612404** (0.01254234)	0.02853403*** (0.01030305)
Pvm: 2016-05	0.02505006** (0.01254363)	0.0305853*** (0.01030264)
Pvm: 2016-06	0.01782424 (0.01236921)	0.02554789** (0.01016417)
Pvm: 2016-07	0.0200429 (0.01370642)	0.02750945** (0.01125862)
Pvm: 2016-08	0.03180657** (0.01280345)	0.03263271*** (0.01051599)
Pvm: 2016-09	0.02035147 (0.01246949)	0.032176*** (0.01024287)
Pvm: 2016-10	0.03854969*** (0.01266154)	0.04152407*** (0.0103997)
Pvm: 2016-11	0.03819216*** (0.01240486)	0.0393748*** (0.01018809)
Pvm: 2016-12	0.05402986*** (0.01350809)	0.04993524*** (0.01109465)
Pvm: 2017-01	0.04625186*** (0.0132689)	0.0455236*** (0.01089809)
Pvm: 2017-02	0.05353599*** (0.0126509)	0.05022942*** (0.01039075)
Pvm: 2017-03	0.05386163*** (0.01205255)	0.05723949*** (0.00990059)
Pvm: 2017-04	0.07401234*** (0.01265327)	0.07600406*** (0.010391)
Pvm: 2017-05	0.07436637*** (0.01259242)	0.07590616*** (0.01034101)
Pvm: 2017-06	0.09046712*** (0.01231681)	0.0878129*** (0.01011496)
Pvm: 2017-07	0.0779541*** (0.01294358)	0.07560523*** (0.01063325)
Pvm: 2017-08	0.08011513*** (0.0266781)	0.08119296*** (0.02191138)
Constant	8.363764*** (0.05802032)	9.537346*** (0.02213039)
Observations	9,901	9,901
R ²	0.7461567	0.8287634
Adjusted R ²	0.7451786	0.8281036
Residual Std. Error (df = 9862)	0.1951241	0.1602604
F Statistic (df = 38; 9862)	762.8616***	1,256.074***

Note:

*p<0.1; **p<0.05; ***p<0.01

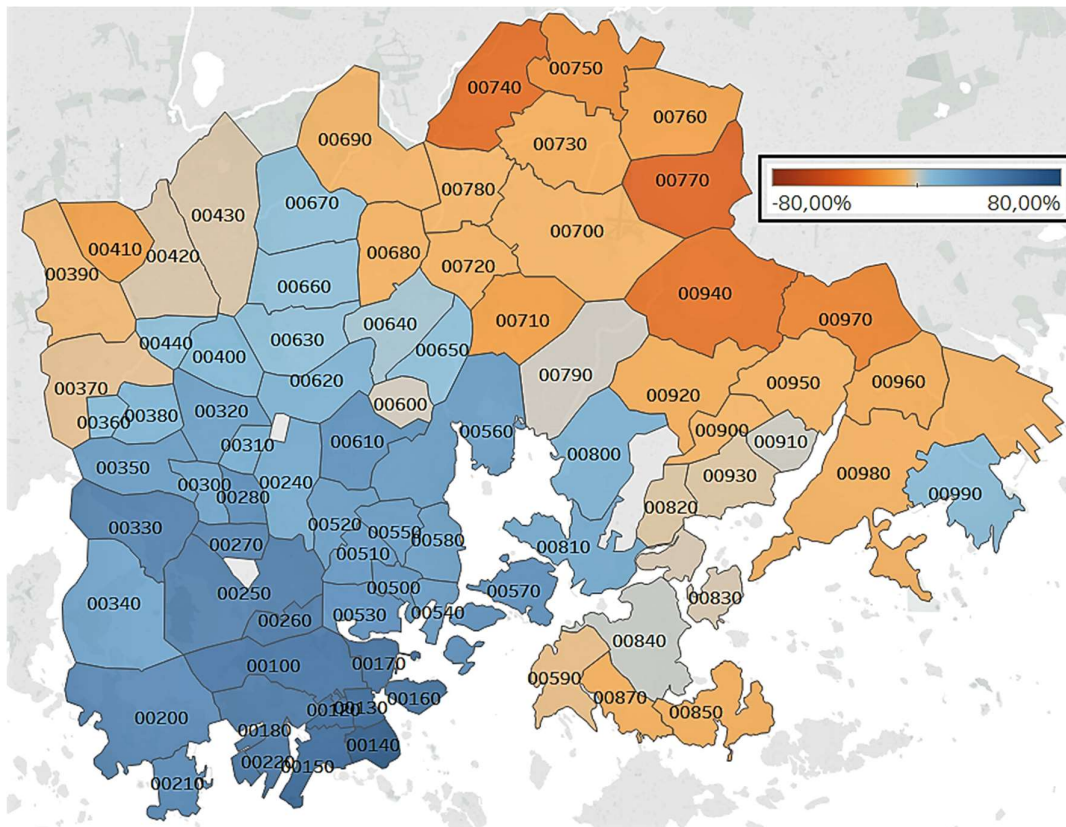
Taulukossa 7 nähdään sovitetut lisämallit. Mallien selitysvoima on heikompi kuin päämallissa, kun mallien selitysasteet ovat noin 74,6 ja 82,9 prosenttia. Matka-ajan selitysvoimaa heikentäneen edellä mainittu seikka, että se ei ota huomioon pysäköintipaikan etsintään tai kulkuneuvon odottamiseen mahdollisesti kuluvaan aikaan. Keskustaetäisyyttä ja matka-aikaa kuvaavat selittäjät ja niiden toisen sekä kolmannen asteen termit ovat maleissa tilastollisesti merkitseviä selittäjiä.

5.3 Estimaattien kuvailu

Tässä aliluvussa kuvataan selittäjien vaikutusta neliöhintoihin regressioparametrien estimaattien perusteella. Kuvailussa hyödynnetään pääosin mallin (1) tuloksia, mutta keskus-taetäisyyden ja matkustusajan vaikutusta kuvataan lisämallien avulla.

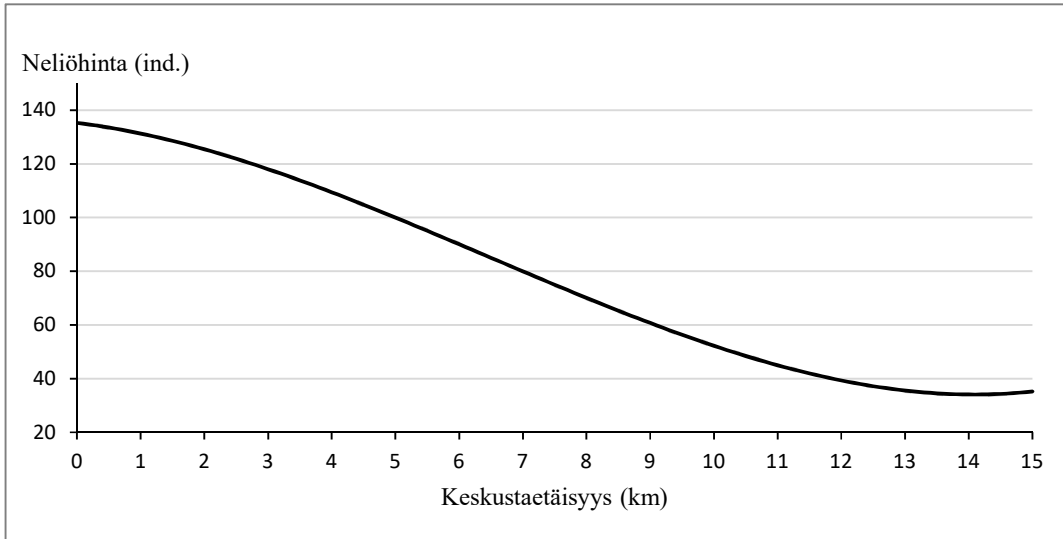
5.3.1 Sijaintitekijät

Postinumeroalueiden regressiokertoimien perusteella asuinalueella on erittäin suuri vaikutus asuntojen neliöhintoihin Helsingissä. Vertailuluokkana on Viikki (00790), jonka keskimääräiset neliöhinnat ovat lähellä Helsingin keskitasoa. Kuviosta 5.1 nähdään, miten eri postinumeroalueiden vaikutus hintoihin jakautuu maantieteellisesti. Regressioker-toimet kuvaavat alueiden hintavaikutusta, kun muut selittäjät on vakioitu. Esimerkiksi Sörnäisissä sijaitsevien kaksioiden neliöhinta oli keskimäärin noin 1,3 prosenttia enemmän kuin Munkkiniemen kaksioiden, mutta mallin mukaan Munkkiniemi asuinalueena nostaa asuntojen arvoa keskimäärin noin 8,4 prosenttia Sörnäisiin verrattuna.



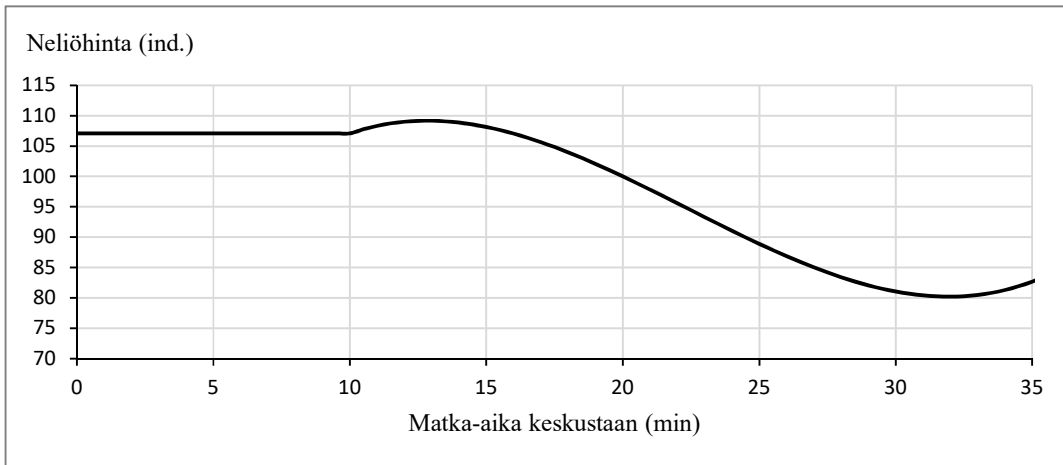
Kuvio 5.1. Postinumeroalueiden keskimääräinen vaikutus neliöhintoihin (vertailuluokkana 00790).

Kuviosta 5.2 nähdään keskustaetäisyyden mallissa (3) estimoitu vaikutus neliöhintoihin. Hinnat ovat korkeimmillaan keskusta-alueella ja laskevat selvästi siirryttäessä keskustasta kauemmaksi. Esimerkiksi kahden kilometrin päässä keskustasta sijaitsevat asunnot ovat keskimäärin noin 25,4 prosenttia arvokkaampia kuin asunnot, jotka sijaitsevat viiden kilometrin päässä keskustasta. Noin kolmentoista kilometrin jälkeen hintavaikutus tasaantuu. Keskustaetäisyyden hintavaikutus on havaittavissa myös kuviossa 5.1.



Kuvio 5.2. Keskustaetäisyyden vaikutus neliöhintaan (indeksi, 5 km = 100).

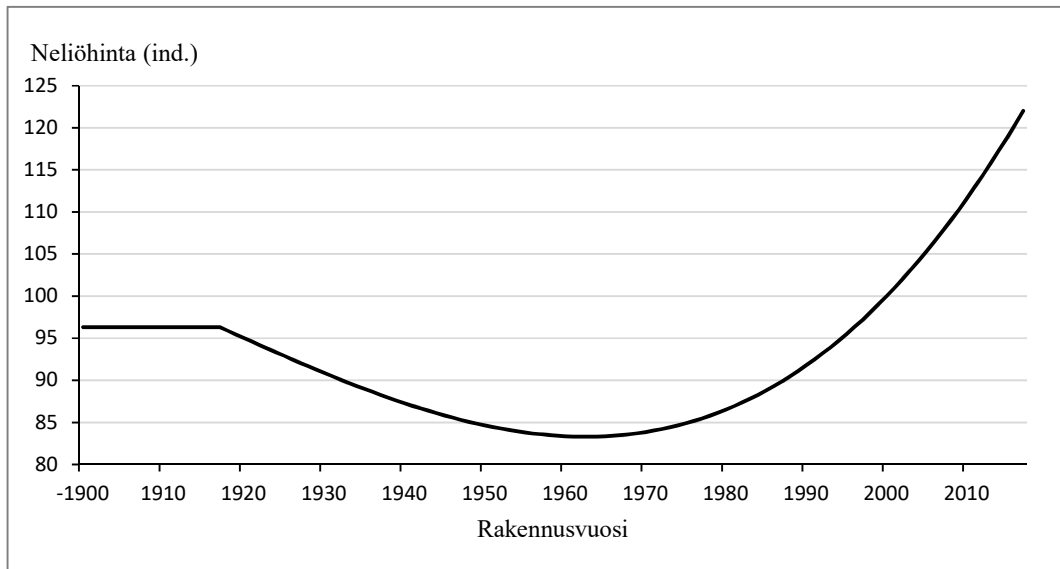
Matka-ajan estimoitu vaikutus puolestaan nähdään kuviossa 5.3. Mallin (2) perusteella nelliöhinnat maksimoituvat noin 13 minuutin matka-ajan päässä keskustasta, jonka jälkeen matka-ajan kasvulla on hintoja laskeva vaikutus, mikä on voimakkaimmillaan 17 ja 27 minuutin matka-aikojen välillä. Esimerkiksi matka-ajan kasvaessa 15 minuutista 25 minuuttiin asuntojen arvo laskee keskimäärin noin 17,8 prosenttia.



Kuvio 5.3. Matka-ajan vaikutus neliöhintaan (indeksi, 20 min = 100).

5.3.2 Taloyhtiötä ja asuinrakennusta kuvaavat tekijät

Kuviossa 5.4 on esillä iän vaikutus asuntojen neliöhintoihin. Hinnat ovat korkeimmillaan uusissa asunnoissa, minkä jälkeen hinnat laskevat voimakkaasti, enimmillään jopa 1,3 prosenttia vuodessa. Hinnat ovat matalimmillaan vuonna 1963 rakennetuissa kohteissa, eli noin 54 vuotta vanhoissa kohteissa, jota vanhemmissa kohteissa puolestaan hinnat kääntyvät nousuun.



Kuvio 5.4. Rakennusvuoden/iän vaikutus neliöhintaan (indeksi, 2000 = 100).

Hinnat ovat korkeimmillaan asunnoissa, joiden kiinteistön kunto on erinomainen, mikä nähdään taulukosta 8. Puolestaan matalimmat hinnat ovat asunnoilla, jotka ovat kunnollaan huonoja. Suurimmat erot ovat tyydyttäväkuntoisten ja sitä paremmassa kunnossa olevilla. Puolestaan uusien, hyvässä tai erinomaisessa kunnossa olevien asuntojen hintaerot ovat huomattavasti pienempiä. Esimerkiksi Erinomaisessa kunnossa olevat asunnot ovat noin 2,8 prosenttia arvokkaampia kuin hyväkuntoiset asunnot.

Taulukko 8. Kiinteistön kunnan vaikutus neliöhintaan (vertailuluokkana ”huono”).

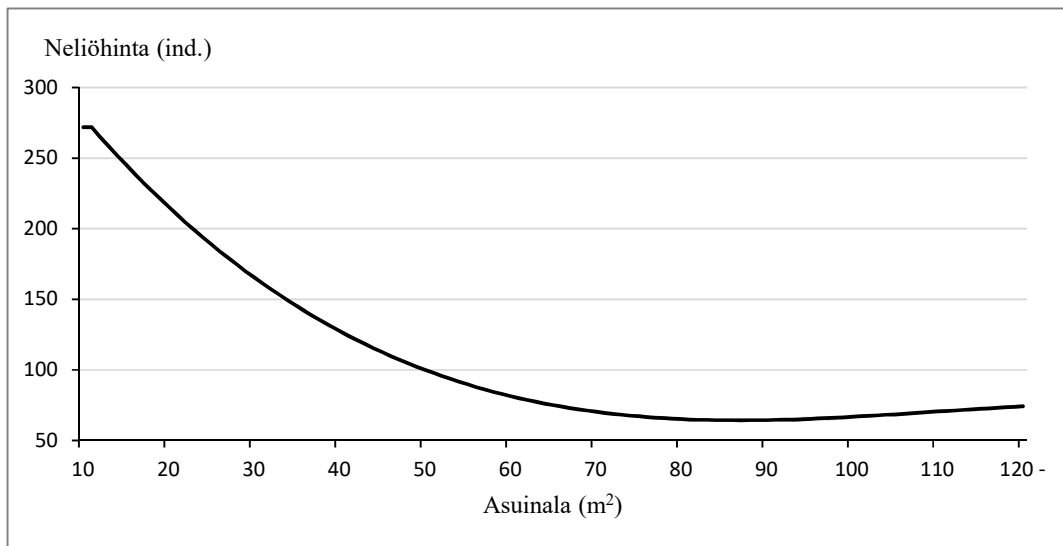
Kiinteistön kunto	Hintavaikutus
Huono	-
Tyydyttävä	+ 8,4 %
Hyvä	+ 18,4 %
Erinomainen	+ 21,7 %
Uusi	+ 20,6 %

Mallin (1) mukaan hoitovastikkeella on asunnon arvoa laskeva vaikutus, kun keskimäärin yhden euron kasvu hoitovastikkeessa asuinneliötä kohti laskee asunnon arvoa noin 1,4

prosenttia. Lisäksi omalla tontilla olevien asuntojen neliöhinnat ovat keskimäärin 4,4 prosenttia korkeammat kuin vuokratontilla olevat.

5.3.3 Asunnon ominaisuuksia kuvaavat tekijät

Pienimmillä asunnoilla on keskimäärin selvästi korkeimmat neliöhinnat, mikä nähdään kuvioista 5.5. Neliöhinnat laskevat selvästi asuinalan kasvaessa pienimmistä, mutta asuinalan vaikutus selvästi tasaantuu asuinalan kasvaessa yli noin 45 neliön. Yli 85 neliöisten asuntojen kohdalla puolestaan asuinalan kasvulla näyttää olevan hieman nostava vaikutus. Mallin (1) mukaan alimmassa kerroksessa sijaitsevien asuntojen neliöhinnat ovat keskimäärin noin 3,1 prosenttia matalampia kuin ylemmissä kerroksissa sijaitsevien.



Kuvio 5.5. Asuinalan vaikutus neliöhintaan (indeksi, 50 m² = 100).

Taulukosta 9 nähdään, että mallin (1) perusteella lisähuoneet keskimäärin nostavat neliöhintoja verrattuna asuntoihin, joissa huonelukumäärä on pienempi, mutta ovat muilta ominaisuuksiltaan vastaavanlaisia.

Taulukko 9. Huoneiden lukumäärän vaikutus neliöhintaan (vertailuluokkana yksiöt).

Huoneita	Hintavaikutus
Yksiö	-
Kaksio	+ 3,1 %
Kolmio	+ 5,4 %
Neljä huonetta tai enemmän	+ 1,6 % (*)

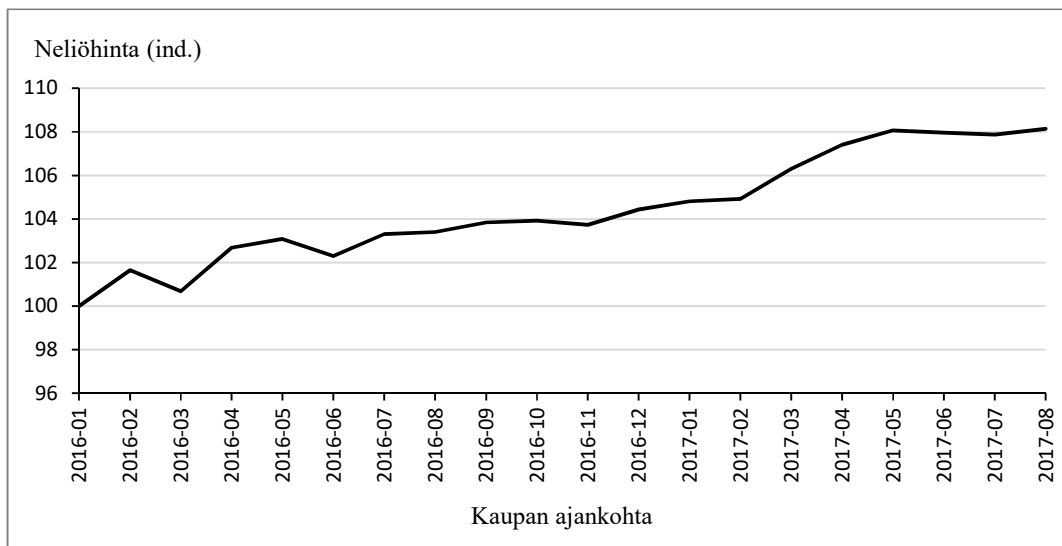
* ei tilastollisesti merkitsevä

Asuntojen, joissa on huoneita neljä tai enemmän, neliöhinnat eivät ole tilastollisesti merkitsevästi suurempia yksiöihin verrattuna. Tämän osaltaan saattaa selittää se, että löytyy

hyvin vähän vertailukelpoisia kohteita, koska keskimäärin yksiöt ovat huomattavasti pienempiä asuinalaltaan. Esimerkiksi aineiston pienin asunto, jossa oli neljä huonetta tai enemmän, on 62 neliötä, mutta yksiöitä, joiden koko on 62 neliötä tai enemmän, löytyi aineistosta vain kuusi kappaletta.

5.3.4 Ajallinen hintakehitys

Kuviosta 5.6 nähdään, kuinka asuntojen hintataso on kehittynyt keskimäärin tammikuun 2016 ja elokuun 2017 välillä mallin (1) perusteella. Kuvion perusteella helsinkiläisten kerrostaloasuntojen neliöhinnat ovat nousseet melko tasaisesti edellä mainitun ajanjakson aikana. Esimerkiksi vuosien 2016 ja 2017 tammikuuiden välisenä aikana, eli vuodessa, hintataso oli noussut keskimäärin noin 4,8 prosenttia. Koko ajanjaksona puolestaan hinnat nousivat noin 8,1 prosenttia.



Kuvio 5.6. Neliöhintojen hintakehitys ajassa (indeksi, 2016-01 = 100).

6 JOHTOPÄÄTÖKSET

Kuten johdantoluvussa todettiin, on oman asunnon hankinta useimmille elämän suurin yksittäinen hankinta. Helsingissä on koko maan korkein hintataso kerrostaloasuntomarkkinoilla. Tämän vuoksi Helsingissä on erityisen tärkeää saada mahdollisimman hyvä arvio asunnon markkina-arvosta sitä ostaessa ja myytäessä. Lisäksi asuntojen hinnat on erittäin suosittu aihe suomalaisessa mediassa ja siihen liittyvä uutisointi sekä informaatio ovat usein melko yksinkertaistettua, mikä saattaa toisinaan olla jopa harhaanjohtavaa. Tästä johtuen tämän tutkielman tavoitteena oli lineaarisen regressioanalyysin avulla mallintaa helsinkiläisten kerrostaloasuntojen hinnanmuodostusta. Mallintamalla hinnanmuodostusta pyrittiin kuvaamaan hintaan vaikuttavia tekijöitä ja niiden vaikutusten merkittävyyttä. Usean selittäjän lineaarisen regression vahvuus on, että se kuvaa mallissa olevien yksittäisten selittäjien arvojen vaihtelun vaikutusta, kun muiden selittäjien arvot ovat vakioituna. Tällä tavalla saadaan esimerkiksi eri postinumeroalueiden hintavaikutuksen arvioinnissa otettua huomioon alueiden asuntokannan erilaisuus (muun muassa iän suhteen).

Tutkielmassa estimoitiin logaritmimuunnoksen avulla asuntojen velattomille neliöhinnoille kolme regressiomallia, joista ensimmäinen toimi päämallina ja kaksi jälkimmäistä toimivat apumalleina, joilla saatiin kuvattua keskustaetäisyyden ja -matka-ajan vaikutusta. Päämallin selitysaste oli erittäin hyvä, sen selittäessä velattoman neliöhinnan vaihtelusta jopa noin 89,1 prosenttia. Apumallien selitysvoima oli heikompi päämalliin nähden vaihdettaessa sijaintia kuvaavia selittäjiä.

Odotusten mukaisesti sijainnilla on erittäin suuri vaikutus asunnon arvoon jopa pelkästään Helsingin sisäisillä markkinoilla. Päämallin perusteella eri postinumeroalueilla on arvostuksessa hyvin suuria arvoja. Mallin perusteella arvokkain alue on Kaivopuisto – Ullanlinna (00140) ja puolestaan vähiten arvostettu Jakomäki – Alppikylä (00770). Mallin mukaan sijaintina Kaivopuisto – Ullanlinna nostaa asunnon arvoa jopa noin 202 prosenttia verrattuna Jakomäki – Alppikylä -postinumeroalueeseen. Arvokkaimmat asuinalueet sijaitsevat pääosin keskustassa tai kantakaupungissa ja puolestaan matalimman arvostuksen omaavat alueet Pohjois-Helsingissä.

Vakioimalla muiden selittäjien vaikutus neliöhintoihin saatiin selville asuinalueiden arvostuksen todelliset erot. Esimerkiksi aineistossa oli kohteita 78:ltä eri postinumeroalueelta, ja keskineliöhintojen perusteella 24. arvokkain alue on Herttoniemi (00810) ja 25. arvokkain Käpylä (00610). Kyseisillä alueilla myytyjen asuntojen velattomat keskineliöhinnat olivat noin 5070 ja 4876 euroa. Mutta mallin mukaan Käpylän arvostus 19. suurin ja Herttoniemen vasta 32. suurin. Mallin mukaan Käpylä nostaa asuntojen arvoa noin 11,8 prosenttia Herttoniemeen verrattuna. Sijoitusten muutos voidaan selittää esimerkiksi iän, tontinomistajuuden ja kiinteistön kunnan avulla. Herttoniemessä sijaitsevien asuntojen keski-ikä oli noin 17 vuotta, kun puolestaan Käpylässä sijaitsevien kohteiden noin 68 vuotta. Käpylän asuinnoista jopa noin 97 prosenttia oli vuokratontilla, kun puolestaan Herttoniemessä vastaava luku oli noin 64 prosenttia. Lisäksi Käpylässä sijaitsevista asunnoista noin 45 prosenttia kuuluu kiinteistön kunnan osalta luokkaan ”huono” tai ”tyydyttävä”, kun Herttoniemeläisistä asunnoista vain noin 19 prosenttia kuului näihin luokkiin. Tämä on malliesimerkki siitä, minkälaisen hyödyn malli tarjoaa eri alueiden arvoa arvioidessa keskineliöhintojen suoran vertailun sijaan.

Korvattaessa mallin selittäjissä postinumero keskustaetäisyydellä heikkenee selityssaste 82,9 prosenttiin, eli keskustaetäisyydellä on heikompi selitysvaima kuin postinumerolla. Asuntojen hinnat ovat korkeimmillaan keskustassa ja keskustaetäisyyden kasvulla on selvästi hintoja laskeva vaikutus. Tämä kuvanee sitä, että lähempänä keskustaa pääsee parhaiten hyödyntämään kaupungistumisen tarjoamista eduista, kun palvelut, ihmiset ja taloudelliset toiminnot ovat lähellä toisiaan. Neliöhintojen ja keskustaetäisyyden välinen yhteys on melko voimakas, kun esimerkiksi kolmen kilometrin päässä keskustasta sijaitseva asunto on keskimäärin 68,4 prosenttia arvokkaampi kuin kahdeksan kilometrin päässä. Keskustaetäisyyden ollessa yli 11 kilometriä sen hintoja laskeva vaikutus heikkenee ja 14 kilometrin jälkeen hinnat jopa keskimäärin hieman nousevat. Tämä selittynee sillä, että pisimmillä etäisyyksillä saattavat naapurikuntien liikekeskusten läheisyys vaikuttaa hintoihin niitä nostavasti.

Korvattaessa edelleen mallin selittäjissä keskustaetäisyys matka-ajalla keskustaan heikkenee mallin selityssaste jälleen, ollen lopulta 74,6 prosenttia. Tähän osaselityksenä lienee se, että estimoituihin matka-aikoihin muodostuu hieman harhaa, koska ne eivät osaa ottaa huomioon mahdollista aikaa, joka kuluu esimerkiksi kulkuneuvon odottamiseen pysäkillä tai pysäköintipaikan löytämiseen. Lisäksi alle 10 minuutin matka-ajoilla ei ollut merkit-

seviä eroja, mikä selittynee osittain sillä, että auto tai julkinen liikenne ei välttämättä käytännöllisin vaihtoehto niin lyhyillä matkoilla. Mallin mukaan korkeimmat hinnat ovat keskimäärin noin 13 minuutin matka-ajan päässä keskustasta. Tämä johtunee siitä, että Helsingin arvokkaimmat asuinalueet sijaitsevat hieman ydinkeskustan ulkopuolella. Tähän selityksenä voi olla se, että niissä liikenne-, melu- ja muut vastaavat haitat ovat vähäisempiä, mutta kuitenkin pääsee nauttimaan kaupungissa asumisen hyödyistä helposti.

Matka-aikojen hintavaikutusten estimoinnin avulla voidaan esimerkiksi arvioida julkisten liikennehankkeiden vaikutuksia asuntojen arvoihin. Tuoreena esimerkkinä tästä on Lautasaaren metroasema, joka avattiin 18. marraskuuta 2017, minkä myötä keskustan saavutettavuudessa saattoi tapahtua suuriakin muutoksia eri asuntojen kohdalla – sekä nopeutumista että hidastumista, kun myös bussiliikenne muuttui. Esimerkiksi julkisten liikenneyhteyksien hidastuessa 15 minuutista 20 minuuttiin siten, että keskimääräinen matka-aika kasvaa 15 minuutista 17,5 minuuttiin, laskee mallin mukaan asuntojen arvoa noin 3,1 prosenttia.

Rakennuksen iällä on hintaan laskeva vaikutus keskimäärin noin 55 vuoteen asti, minkä jälkeen hinta kohoaa, kun ikä kasvaa. Iän hintavaikutus on melko voimakas, minkä vuoksi on tärkeää ottaa se huomioon asuntojen hintoja arvioitaessa. Esimerkiksi saatujen tulosten perusteella 20 vuotta vanha kohde on keskimäärin noin 8,1 prosenttia arvokkaampi kuin vastaavanlainen 30 vuotta vanha kohde. Iän vaikutus kuvastanee, kuinka asuinrakennusten korjaustarve ja siitä syntyvät kustannukset kertyvät vuosien saatossa, kunnes tehdään suuria remonteja, kuten linjasaneeraus, joiden myötä korjaustarve on jälleen huomattavasti pienempi, joka välittyy asunnon hintaan. Lisäksi iän hintavaikutuksesta osa selittynee eri aikakausien eroilla käytettyjen rakennustekniikoiden ja -materiaalien suhteen. Kiinteistön kunnolla on myös hyvin voimakas vaikutus hintoihin. Kuntoluokitusten ”uusi”, ”erinomainen” ja ”hyvä” hintavaikutukset eivät vaikuta eroavan keskenään suuresti, mutta seuraavan kuntoluokan ”tyydättävä” arvostus on huomattavasti matalampi näihin nähden (noin 8,4 prosenttia matalampi verrattuna luokkaan ”hyvä”). Heikoimman kuntoluokan ”huono” arvostus on vielä huomattavasti matalampi – noin 7,7 prosenttia heikompi. Lisäksi saatujen tulosten perusteella kuntoluokituksen ”erinomainen” omaavat asunnot ovat keskimäärin 0,9 prosenttia arvokkaampia kuin uudet vastaavat kohteet, mikä saattaa hieman yllättää. Tähän selityksenä voi olla, että uusiin kohteisiin voi kohdistua runsaasti erilaisia epävarmuustekijöitä. Esimerkiksi millaiseksi taloyhtiö, naapurusto ja

muut vastaavat asiat lopulta muodostuvat uusissa paikoissa. Lisäksi uudiskohteista saattaa syntyä enemmän vaivaa liittyen takuukorjauksiin ja niiden tarpeellisuuden selvittämiseen.

Estimoitujen tulosten perusteella hoitovastikkeella on hintaa laskeva vaikutus – hoitovastikkeen kasvaessa euron asuinneliötä kohti laskee neliöhinta keskimäärin noin 1,4 prosenttia. Tulos on odotusten mukainen, koska korkeampi hoitovastike nostaa asumisen kustannuksia, mikä todennäköisesti vähentää maksuhalukkuutta. Lisäksi omalla tontilla olevat asunnot ovat keskimäärin noin 4,4 prosenttia arvokkaampia vuokratontilla oleviin nähden. Tosin vaatisi tarkempaa tutkimusta selvittää kattaako hintaerot tonttien diskonttatut arvot, eli välittykö tonttien arvot kokonaisuudessaan hintoihin.

Asuntojen kokoja tarkastellessa selvästi korkeimmat neliöhinnat ovat pienimmillä asunnoilla, joiden kohdalla ”lisäneliöiden” arvo on huomattavasti korkeampi kuin suuremmissa asunnoissa. Esimerkiksi 20 neliöisten asuntojen neliöhinnat ovat keskimäärin noin 14,1 prosenttia korkeammat kuin 25 neliöisten, mutta puolestaan 60 neliöisten asuntojen neliöhinnat ovat vain noin 4,4 korkeammat verrattuna 65 neliöisiin. Asuinalan ja neliöhintojen välinen suhde kuvaa hyvin, kuinka tärkeitä yksittäiset neliöt ovat eri kokoisissa kohteissa. Asuinalan kasvaessa yli 90 neliön hintavaikutus kääntyy lievään kasvuun, mikä selittynee sillä, että suurempien asuntojen kohdalla tarjonta on kerrostalomarkkinoilla huomattavasti matalampi verrattuna pienempiin asuntoihin. Lisähuoneilla vaikuttaisi olevan arvoa nostava vaikutus saatujen tulosten perusteella. Useampi huone lisää asuntojen seinäalaa ja lisää asumisratkaisuja. Esimerkiksi 80 neliöinen kolmio on todennäköisesti suositumpi vaihtoehto lapsiperheelle tai yhteisasunnoksi verrattuna 80 neliöiseen kaksioon nähden. Myös asuinkerroksella on selvä vaikutus neliöhintaan. Alimmassa kerroksessa sijaitsevien asuntojen neliöhinnat ovat keskimäärin 3,1 prosenttia matalammat kuin ylemmissä kerroksissa sijaitsevien. Tulos on odotusten mukainen, koska alimmassa kerroksessa on todennäköisemmin heikommät näkymät, suuremmat meluhaitat ja heikompi yksityisyys.

Tutkielman tulosten perusteella helsinkiläiset kerrostaloasunnot ovat hyvin moniulotteisia hyödykkeitä, joiden hinnat määräytyvät useiden hintatekijöiden yhteisvaikutuksen perusteella. Tämä alleviivaa entisestään sitä, että asuntojen hintojen tulkinnassa ei tulisi tehdä liikaa yksinkertaistuksia, esimerkiksi tarkasteltaessa hintoja pelkästään postinumeralueen perusteella.

Tutkielmassa onnistuttiin mallintamaan asuntojen hintoja melko tarkasti. Sijaintia voitane pitää suurimpana yksittäisenä arvotekijänä asunnoille, minkä lisäksi ikä ja kunto vaikuttavat hyvin voimakkaasti – silti unohtamatta muita hintatekijöitä. Tutkielman tuloksissa näkyi lisäksi selvästi asuntomarkkinoiden suhdanteen kehitys, kun mallin mukaan hintataso on selvästi noussut tammikuun 2016 ja elokuun 2017 välillä, mikä vastaa hyvin yleistä näkemystä tilanteesta sekä peilaantuu johdonmukaisesti tilastokeskuksen lukujen kanssa.

Mallin hyvästä toimivuudesta huolimatta sen kehittämiseen ja jatkotutkimukseen on runsaasti mielenkiintoisia mahdollisuuksia. Malli voitaisiin saada entistä tarkemmaksi lisäämällä aineistoon esimerkiksi mahdollisen tiedon edellisen ja/tai seuraavan linjasaneerauksen ajankohdasta sekä useita erilaisia mikrolokaatioihin liittyviä selittäjiä, kuten etäisyys meren rantaan, lähimpään kouluun ja niin edelleen. Myös laadukkaampi tieto liittyen hissien, saunan ja parvekkeen olemassaoloon parantaisi hyvin todennäköisesti hintojen mallinnettavuutta. Tutkielman tekemisen yhteydessä selvisi, että regressiodiagnostiikka voisi olla yksi hyvin hyödyllinen aineiston laadunvalvontakeino henkilöille, jotka työskentelevät säännöllisesti asuntokauppa-aineistojen parissa. Myös yli- tai aliarvostettujen kohteiden löytämiseen voisi olla mallista sekä sen diagnostiikasta hyötyä. Yksittäisen kohteen kohdalla malli ei liene riittävä ainoaksi keinoksi määrittämään asunnon arvoa, mutta antanee hyvän perustan, josta arviota voi tarkentaa esimerkiksi ei-mitattavissa olevien hintatekijöiden tietämyksen avulla. Lisäksi malli antanee hyvän peruskäsityksen, jonka pohjalta voi siirtyä tarkempiin tutkimuksiin aiheeseen liittyen.

LÄHTEET

- Avoindata.fi, 2017. *Suomalaisten rakennusten osoitteet, postinumerot ja WGS84-koordinaatit*. Väestönrekisterikeskus [viitattu 8.11.2017]. Saantitapa: <https://www.avoindata.fi/data/fi/dataset/postcodes>
- Cook, R. D. & Weisberg, S. 1999. *Applied regression including computing and graphics*. New York; Chichester: Wiley.
- Faraway, J. 2002. *Practical Regression and ANOVA using R*. Saantitapa: <https://cran.r-project.org/doc/contrib/Faraway-PRA.pdf>
- Fidell, L.S. & Tabachnick, B. G. 2014. *Using multivariate statistics*. 6. ed., Pearson new international edition. Harlow: Pearson.
- Finanssivalvonta, 2016. Asuntolainat. [viitattu 25.10.2017]. Saantitapa: <http://www.finanssivalvonta.fi/fi/Finanssiasiakas/Tuotteita/Lainat/Asuntolainat/Pages/Default.aspx>.
- Harjunen, O., Kortelainen, M. & Saarimaa, T. 2014. *Best Education Money Can Buy? Capitalization of School Quality in Finland*. Työpapereita 58, Valtion taloudellinen tutkimuskeskus VATT.
- Haulos, S., Nevala, T., Palo, M. & Sirén, M. 2015. *Kiinteistönvälittäjän käsikirja*. 8. tarkistettu painos. Helsinki: Suomen Kiinteistönvälittäjäliitto.
- Helsingin Sanomat 7.5.2017. *Asuntojen hinnat huimassa nousussa, Kallio ja Sörnäinen pomppasivat jo ohi perinteisten arvoalueiden – Katso, miten hinnat ovat muuttuneet omalla asuinalueellasi*.
- Hietala, M., Huovari, J., Kaleva, H., Lahtinen, M., Niemi, J., Ronikonmäki, NM., Vainio, T. 2015. *Asuinrakennusten korjaustarve*. PTT raportteja 251. ISBN 978-952-224-171-9.
- Jussila, I. 2017. *Miksi ihminen omistaa asunnon?* Kirjoitus Taloustutka-blogissa. Peller-von taloustutkimus PTT. [viitattu 29.10.2017].
- Kasso, M. 2016. *Kiinteistönvälitys ja -arviointi*. 2. painos. Helsinki: Talentum.

Laakso, S. & ETLA. 1997. *Urban housing prices and the demand for housing characteristics: A study on housing prices and the willingness to pay for housing characteristics and local public goods in the Helsinki Metropolitan Area*. Helsinki: Elinkeinoelämän tutkimuslaitos.

Laakso, S. & Loikkanen, H. 2004. *Kaupunkitalous: Johdatus kaupungistumiseen, kaupunkien maankäyttöön sekä yritysten ja kotitalouksien sijoittumiseen*. Helsinki: Gaudeamus.

Melo, R. & Zarruk, D. 2016. *gmapsdistance: Distance and Travel Time Between Two Points from Google Maps*. R package version 3.1. <https://github.com/rodazuero/gmapsdistance>.

Montgomery, D. C., Peck, E. A. & Vining, G. G. 2012. *Introduction to linear regression analysis*. Fifth edition. Hoboken, New Jersey: John Wiley & Sons Ltd.

Nikola, N. 2011. *The effect of pipe repairs on housing prices*. Pro gradu -tutkielma. Aalto-yliopisto, Rahoituksen laitos.

Tilastokeskus, 2015. Suomen virallinen tilasto (SVT): Kotitalouksien varallisuus [verkkójulkaisu]. ISSN=2242-3214. 2013, 2. *Oma asunto keskeisin kotitalouksien varallisuus-erä*. Helsinki: Tilastokeskus [viitattu: 10.10.2017]. Saantitapa: http://www.stat.fi/til/vtutk/2013/vtutk_2013_2015-04-01_kat_002_fi.html

Tilastokeskus, 2017a. Suomen virallinen tilasto (SVT): Osakeasuntojen hinnat [verkkójulkaisu]. ISSN=2323-878X. Helsinki: Tilastokeskus [viitattu: 14.10.2017]. Saantitapa: <http://www.stat.fi/til/ashi/index.html>

Tilastokeskus, 2017b. Suomen virallinen tilasto (SVT): Asunnot ja asuinolot [verkkójulkaisu]. ISSN=1798-6745. Helsinki: Tilastokeskus [viitattu: 17.10.2017]. Saantitapa: <http://www.stat.fi/til/asas/2016/index.html>

Tuominen, H. 2014. *Länsimetron rakentamisen vaikutukset asuinhuoneistojen arvoon*. Pro gradu -tutkielma. Aalto-yliopisto, Insinööritieteiden korkeakoulu.

Vainio, T. 2016. *Asuntotuotantotarve 2015–2040*. Teknologian tutkimuskeskus VTT Oy. ISBN 978-951-38-8395-9.

Venäläinen, I. 2013. *Putkiremontin vaikutus asunnon myyntihintaan Kuopiossa*. Pro gradu -tutkielma. Itä-Suomen yliopisto, kauppätieteiden laitos.

Verohallinto, 2017. [viitattu 26.10.2017]. Saantitapa: www.vero.fi

Weisberg, S. 2005. *Applied linear regression*. 3rd ed. Hoboken, New Jersey: Wiley.

LIITTEET

Liite A: Helsingin postinumeralueet.

Postinumeralue	Postinumeralueen nimi
00100	Helsinki Keskusta - Etu-Töölö
00120	Punavuori
00130	Kaartinkaupunki
00140	Kaivopuisto - Ullanlinna
00150	Eira - Hernesaari
00160	Katajanokka
00170	Kruununhaka
00180	Kamppi - Ruoholahti
00190	Suomenlinna
00200	Lauttasaari
00210	Vattuniemi
00220	Jätkäsaari
00230	Ilmala
00240	Länsi-Pasila
00250	Taka-Töölö
00260	Keski-Töölö
00270	Pohjois-Meilahti
00280	Ruskeasuo
00290	Meilahden sairaala-alue
00300	Pikku Huopalahti
00310	Kivihaka
00320	Etelä-Haaga
00330	Munkkiniemi
00340	Kuusisaari-Lehtisaari
00350	Munkkivuori-Niemenmäki
00360	Pajamäki
00370	Reimarla
00380	Pitäjänmäen teollisuusalue
00390	Konala
00400	Pohjois-Haaga
00410	Malminkartano
00420	Kannelmäki
00430	Maununneva
00440	Lassila
00500	Sörnäinen
00510	Etu-Vallila - Alppila
00520	Itä-Pasila
00530	Kallio
00540	Kalasadama
00550	Vallila
00560	Toukola-Vanhakaupunki
00570	Kulosaari
00580	Verkkosaari
00590	Kaitalahti
00600	Koskela-Helsinki
00610	Käpylä
00620	Metsälä-Etelä-Oulunkylä
00630	Maunula-Suursuo

00640	Oulunkylä-Patola
00650	Veräjämäki
00660	Länsi-Pakila
00670	Paloheinä
00680	Itä-Pakila
00690	Tuomarinkylä-Torpparinmäki
00700	Malmi
00710	Pihlajämäki
00720	Pukinmäki-Savela
00730	Tapanila
00740	Siltämäki
00750	Puistola
00760	Suurmetsä
00770	Jakomäki - Alppikylä
00780	Tapaninvainio
00790	Viikki
00800	Länsi-Herttoniemi
00810	Herttoniemi
00820	Roihuvuori
00830	Tammisalo
00840	Laajasalo
00850	Jollas
00860	Santahamina
00870	Etelä-Laajasalo
00880	Roihupellon teollisuusalue
00890	Itäsalmi
00900	Puotinharju
00910	Puotila
00920	Myllypuro
00930	Itäkeskus-Marjaniemi
00940	Kontula - Vesala
00950	Vartioharju
00960	Pohjois-Vuosaari
00970	Mellunmäki
00980	Etelä-Vuosaari
00990	Aurinkolahti

Liite B: Yhdistetyt postinumerot.

Todellinen postinumero	Muokattu
00430	00420
00660	00630
00670	00630
00680	00720
00690	00780
00850	00870
00220	00180

Liite C: Malli (1) kokonaisuudessaan.

	<i>Dependent variable:</i>
	log(Neliöhinta)
Asuinala - 11	-0.02542198*** (0.0008196579)
(Asuinala - 11) ²	0.00027206*** (0.00001525247)
(Asuinala - 11) ³	-0.0000009177137*** (0.00000008353553)
Kiinteistön kunto: Tyydyttävä	0.08415202*** (0.006598253)
Kiinteistön kunto: Hyvä	0.184245*** (0.006502972)
Kiinteistön kunto: Erinomainen	0.2172594*** (0.01153948)
Kiinteistön kunto: Uusi	0.2059082*** (0.01029274)
Tontin omistajuus: Vuokrattu	-0.04170708*** (0.004246858)
Hoitovastike per neliö (€)	-0.01381239*** (0.001459299)
Ikä	-0.0132997*** (0.0006352476)
Ikä ²	0.0001674645*** (0.00001429135)
Ikä ³	-0.0000005552989*** (0.00000008903838)
Kerros: Alin	-0.03128393*** (0.003390344)
Huoneita: 2	0.03121159*** (0.005686906)
Huoneita: 3	0.05401325*** (0.008043051)
Huoneita: 4 tai enemmän	0.01633077 (0.01034577)
Pvm: 2016-02	0.01651909** (0.008285337)
Pvm: 2016-03	0.006865484 (0.008348761)
Pvm: 2016-04	0.02683272*** (0.008286704)
Pvm: 2016-05	0.03077974*** (0.00829173)
Pvm: 2016-06	0.02290448*** (0.00817587)
Pvm: 2016-07	0.03309258*** (0.009068635)
Pvm: 2016-08	0.03401857*** (0.008458457)
Pvm: 2016-09	0.03846425*** (0.0082391)
Pvm: 2016-10	0.03933082*** (0.008366595)
Pvm: 2016-11	0.03738475*** (0.008198579)
Pvm: 2016-12	0.04431445*** (0.008925722)
Pvm: 2017-01	0.0480807*** (0.008767572)
Pvm: 2017-02	0.04919527*** (0.008358947)
Pvm: 2017-03	0.0629785*** (0.007974398)
Pvm: 2017-04	0.07406198*** (0.00835245)
Pvm: 2017-05	0.08072626*** (0.008313289)
Pvm: 2017-06	0.07957667*** (0.008136912)
Pvm: 2017-07	0.07874524*** (0.008556635)
Pvm: 2017-08	0.08140489*** (0.01763745)
Postinumero: 00100	0.5891752*** (0.01467845)
Postinumero: 00120	0.6502414*** (0.01636052)
Postinumero: 00130	0.6749723*** (0.02340787)
Postinumero: 00140	0.7627038*** (0.01588428)
Postinumero: 00150	0.6383855*** (0.01526579)
Postinumero: 00160	0.5869916*** (0.01848136)
Postinumero: 00170	0.6070442*** (0.01616961)

Postinúmero: 00180	0.5682591*** (0.01502687)
Postinúmero: 00200	0.45784*** (0.01484388)
Postinúmero: 00210	0.418227*** (0.01526986)
Postinúmero: 00240	0.2428807*** (0.01947728)
Postinúmero: 00250	0.517995*** (0.01536596)
Postinúmero: 00260	0.5777991*** (0.01751586)
Postinúmero: 00270	0.481977*** (0.01629443)
Postinúmero: 00280	0.421107*** (0.01986645)
Postinúmero: 00300	0.2878719*** (0.02132323)
Postinúmero: 00310	0.225224*** (0.02931797)
Postinúmero: 00320	0.2955915*** (0.01533029)
Postinúmero: 00330	0.4985254*** (0.01639079)
Postinúmero: 00340	0.2897278*** (0.03101063)
Postinúmero: 00350	0.3179335*** (0.0161762)
Postinúmero: 00360	0.1266562*** (0.02199854)
Postinúmero: 00370	-0.03715309* (0.02172241)
Postinúmero: 00380	0.1753051*** (0.01890187)
Postinúmero: 00390	-0.0626791*** (0.01832361)
Postinúmero: 00400	0.1513888*** (0.01661327)
Postinúmero: 00410	-0.1548752*** (0.01897866)
Postinúmero: 00420	-0.01855565 (0.01591694)
Postinúmero: 00440	0.101546*** (0.01823031)
Postinúmero: 00500	0.3826905*** (0.01483557)
Postinúmero: 00510	0.347353*** (0.01517499)
Postinúmero: 00520	0.3155116*** (0.01778598)
Postinúmero: 00530	0.3939815*** (0.01445725)
Postinúmero: 00540	0.3256659*** (0.03557668)
Postinúmero: 00550	0.3289964*** (0.01594335)
Postinúmero: 00560	0.3114551*** (0.01676547)
Postinúmero: 00570	0.3592611*** (0.01986786)
Postinúmero: 00580	0.3006123*** (0.03662236)
Postinúmero: 00590	-0.04592754 (0.03684158)
Postinúmero: 00600	-0.003796062 (0.02601701)
Postinúmero: 00610	0.3755274*** (0.01813706)
Postinúmero: 00620	0.1815359*** (0.02224934)
Postinúmero: 00630	0.0899461*** (0.01834616)
Postinúmero: 00640	0.04984754*** (0.01741193)
Postinúmero: 00650	0.101752*** (0.03155034)
Postinúmero: 00700	-0.07391022*** (0.01718524)
Postinúmero: 00710	-0.1683277*** (0.01599825)
Postinúmero: 00720	-0.08739458*** (0.01738653)
Postinúmero: 00730	-0.1083769*** (0.02260656)
Postinúmero: 00740	-0.3637577*** (0.0206059)
Postinúmero: 00750	-0.2373741*** (0.02255908)
Postinúmero: 00760	-0.1574606*** (0.03474072)
Postinúmero: 00770	-0.4161207*** (0.02424869)
Postinúmero: 00780	-0.07459174*** (0.02307412)
Postinúmero: 00800	0.1786397*** (0.0160668)

Postinnumero: 00810	0.2307091*** (0.01558959)
Postinnumero: 00820	-0.02252418 (0.01645332)
Postinnumero: 00830	-0.01184456 (0.0428003)
Postinnumero: 00840	0.01089761 (0.01731354)
Postinnumero: 00870	-0.1017158*** (0.02134691)
Postinnumero: 00900	-0.07322885*** (0.02023032)
Postinnumero: 00910	0.003771267 (0.01762042)
Postinnumero: 00920	-0.1139225*** (0.01861727)
Postinnumero: 00930	-0.02274522 (0.02417494)
Postinnumero: 00940	-0.3464849*** (0.01524151)
Postinnumero: 00950	-0.08264156*** (0.03145957)
Postinnumero: 00960	-0.1125817*** (0.01667269)
Postinnumero: 00970	-0.2750933*** (0.01635945)
Postinnumero: 00980	-0.09389906*** (0.01511557)
Postinnumero: 00990	0.1163371*** (0.01575599)
Constant	8.900754*** (0.0204633)
<hr/>	
Observations	9,901
R ²	0.8907134
Adjusted R ²	0.8895419
Residual Std. Error	0.128467 (df = 9795)
F Statistic	760.3025*** (df = 105; 9795)
<hr/>	
Note:	*p<0.1; **p<0.05; ***p<0.01

Liite D: Mallin (1) regressiodiagnostiikka.

