



# Høgskulen på Vestlandet

## MSB210: Masteroppgave

MSB210-01-2023-VÅR-FLOWassign

### Predefinert informasjon

<b>Startdato:</b>	10-03-2023 12:00 CET	<b>Termin:</b>	2023 VÅR
<b>Sluttdato:</b>	22-05-2023 14:00 CEST	<b>Vurderingsform:</b>	Norsk 6-trinns skala (A-F)
<b>Eksamensform:</b>	Masteoppgave		
<b>Flowkode:</b>	203 MSB210 1 01 2023 VÅR		
<b>Intern sensor:</b>	(Anonymisert)		

### Deltaker

<b>Kandidatnr.:</b>	228
---------------------	-----

### Informasjon fra deltaker

<b>Antall ord *:</b>	24894
----------------------	-------

**Egenerklæring \*:** Ja  
**Jeg bekrefter at jeg har Ja**  
**registrert**  
**oppgavetittelen på**  
**norsk og engelsk i**  
**StudentWeb og vet at**  
**denne vil stå på**  
**vitnemålet mitt \*:**

### Gruppe

<b>Gruppenavn:</b>	(Anonymisert)
<b>Gruppenummer:</b>	11
<b>Andre medlemmer i gruppen:</b>	215

Jeg godkjenner avtalen om publisering av masteroppgaven min \*

Ja

Er masteroppgaven skrevet som del av et større forskningsprosjekt ved HVL? \*

Nei

Er masteroppgaven skrevet ved bedrift/virksomhet i næringsliv eller offentlig sektor? \*

Nei



**Høgskulen  
på Vestlandet**

# **MASTEROPPGAVE**

**Energimerking og boligpriser: En økonometrisk analyse av boligmarkedet i Bergensregionen i perioden 2019-2022**

**Energy-labeling and housing prices: An econometric analysis of the housing market in the Bergen Region in the period 2019-2022**

**Kevin Ha  
Ola Andre Olofsson**

Master of Science in Business, Regional Economics and Innovation  
Fakultet for økonomi og samfunnsvitenskap (FØS)

Veileder: Professor Liv Aileen Osland

Innleveringsdato: 22.05.2022

Vi bekrefter at arbeidet er selvstendig utarbeidet, og at referanser/kildehenvisninger til alle kilder som er brukt i arbeidet er oppgitt, jf. Forskrift om studium og eksamen ved Høgskulen på Vestlandet, § 12-1.

## Forord

Masterutredningen utgjør det avsluttende arbeidet på det toårige masterprogrammet Økonomi og administrasjon (Master of Science in Business), ved Høgskulen på Vestlandet. Avhandlingen er skrevet innenfor fordypningen Regional økonomi og innovasjon, i perioden januar - mai 2023, og har et omfang på 30 studiepoeng.

Det er passende å uttrykke takknemlighet overfor dem som har bidratt med støtte, slik at prosjektet ikke bare kom i gang, men også nærmer seg fullføring. I løpet av prosessen har vi tilegnet oss en dypere forståelse om boligmarkedet og energimerkes betydning. Vi drar også lærdom av hvordan data innhentes, bearbeides, analyseres, tolkes og presenteres. Oppgaveskrivingen har vært utfordrende, men spennende og svært lærerik. En spesiell anerkjennelse går til Kommunal- og distriktsdepartementet (KDD), som har styrket vår forskning gjennom deres generøse stipend.

Vi vil rette en stor takk til vår veileder, Liv Aileen Osland ved Institutt for Økonomi og Administrasjon ved Høgskulen på Vestlandet, for å ha inspirert oss og vært til stor hjelp for å kunne belyse forskningsspørsmålene på en god måte. Vi er stolte av å presentere denne utredningen som en avslutning av vår master i økonomi og administrasjon.

Mange personer har under oppgaveskrivingen vært behjelpelige på områdene oppgaven handler om. Vi ønsker å med bakgrunn i dette å takke følgende: Erlend Gjerde og Mads Martin Torp fra Ambita for tilgang til datamateriale, Andreas Jensen og Dag Martin Sundelius i Eiendomsverdi AS for tilgang til datamateriale, Trond Ivar Bøhn for tilgang til Enova sin API, og Professor Inge Thorsen og høgskolelektor Arnstein Gjestland for faglig og teknisk bistand.

Til slutt vil vi takke venner og kollegaer for inspirasjon, diskusjoner og gjennomlesning av oppgaven. Vi vil gjerne gi hverandre et klapp på skulderen for et fantastisk samarbeid som har vært fylt med effektivitet, herlig kommunikasjon og en forståelse for hverandres liv også utenfor studenttilværelsen.

Regresjonene og estimatene i denne avhandlingen er utført i RStudio. R-skript er tilgjengelig på forespørsel. Eventuelle feil eller unøyaktigheter er utelukkende vårt ansvar.



---

Kevin Ha



---

Ola Andre Olofsson

## Sammendrag

Denne masteroppgaven har studert forholdet mellom energimerking og boligpriser i Bergensregionen fra 2019-2022. En økonometrisk tilnærming var benyttet for å identifisere om energimerking kan bidra til å forklare variasjon i boligprisene. Det primære datasettet var ervervet fra Ambita gjennom personlig finansiering. Tilleggsdata var tilveiebrakt etter avtaler med ENOVA og Eiendomsverdi.

Hovedproblemstillingen har vært å undersøke om og i hvilken grad energimerking påvirker boligpriser. Vår empiriske undersøkelse finner ambivalente resultater med hensyn til effekten av energimerking på kvadratmeterprisen. Verken OLS-modellene eller de romlige modellene kunne påvise en signifikant relasjon mellom energimerker og kvadratmeterpris, mens oppvarmingskarakterene viser en signifikant effekt i flere modeller, inkludert de romlige.

Undersøkelser av forholdet mellom energimerking og boligpriser, både før og etter strømpriskrisen inntraff vinteren 2021/2022, indikerte heller ingen signifikant innvirkning av energikarakter. Videre antyder funnene at oppvarmingskarakteren innenfor energimerket kan ha mer betydning for boligprisene enn energikarakteren selv.

Resultatene fra studien vår viser at sammenhengen mellom energimerking og boligpriser er komplekst sammensatt, og ikke signifikant i vårt fokusområde. Videre forskning med mer detaljerte og omfattende data vil være nødvendig for å gi en mer omfattende forståelse av dette emnet, spesielt med tanke på romlige effekter og under påvirkning av store eksterne hendelser som pandemier og strømpriskriser.

*Nøkkelord:* Energimerking, boligpris, energieffektive boliger, romlig økonometri, hedonisk teori

## Abstract

This master's thesis has studied the relationship between energy labeling and housing prices in the Bergen region from 2019-2022. An econometric approach was used to identify whether energy labeling can help explain variation in housing prices. The primary dataset was acquired from Ambita through personal financing. Additional data was provided following agreements with ENOVA and Eiendomsverdi.

The main issue has been to investigate whether and to what extent energy labeling affects housing prices. Our empirical study generates ambivalent results with respect to the effect of energy labeling on the price per square meter. Neither the OLS models nor the spatial models could detect a significant relationship between energy labels and square meter price, while heating ratings show a significant effect in several models, including the spatial ones.

Investigations into the relationship between energy labeling and housing prices, both before and after the energy price crisis occurred in the winter of 2021/2022, also indicated no significant impact of energy rating. Furthermore, the findings suggest that the heating rating within the energy label may be more important for housing prices than the energy rating itself.

The results from our study show that the relationship between energy labeling and housing prices is complex, and not significant in our focus area. Further research with more detailed and comprehensive data will be necessary to provide a more comprehensive understanding of this subject, especially considering spatial effects and under the influence of major external events such as pandemics and energy price crises.

*Key words:* Energy Performance Certificates, housing price, energy-efficient houses, spatial econometrics, hedonic theory

# Innholdsfortegnelse

<b>Forord</b> .....	2
<b>Sammendrag (1 s.)</b> .....	2
<b>Abstract</b> .....	4
<b>Figur- og tabelloversikt</b> .....	7
<b>1.0 Introduksjon</b> .....	7
1.1 Generelt om boligmarkedet i Bergensregionen.....	8
1.2 Problemstilling.....	8
1.3 Oppgavens struktur og metode.....	10
<b>2.0 Bakgrunn - kommuneoversikt</b> .....	12
2.1 Befolkning, topografi og bystruktur.....	12
2.2 Energimerkeordning i Norge.....	16
2.3 Strømpriskrisen vinteren 2021-2022.....	20
<b>3.0 Teori- og litteraturgjennomgang</b> .....	22
3.1 Litteratur og teori om boligmarked i Norge og Vestlandet.....	22
3.1.1 Kjennetegn på det norske boligmarkedet.....	22
3.1.2 Kjennetegn på boligobjekter som et gode.....	23
3.1.3 Tidligere forskning.....	23
3.2 Hedonisk teori.....	27
3.2.1 Likevekt på etterspørselssiden av markedet.....	31
Budfunksjonen.....	32
3.2.2 Likevekt på tilbudssiden av markedet.....	34
Offerfunksjonen.....	35
3.2.3 Markedslikevekt.....	36
3.3 Geografisk avgrensning.....	38
3.4 Det teoretiske grunnlaget for regresjonsanalyse.....	39
3.4.1 Regresjonsanalyse: Anvendelse og tolkning.....	39
3.4.2 Statistiske verktøy for tolkning av resultatene.....	41
3.5 Romlig økonometri.....	42
3.5.1 Spillover-effekter.....	42
3.5.2 k-nærmeste naboer (k-NN).....	45
3.5.3 Forventningsskjev estimator.....	45
3.6 Spatial Lag Model (SAR).....	46
3.7 Spatial Error Model (SEM).....	47
3.8 Spatial Durbin Model (SDM).....	48
3.8 Spatial lag of X-modellen (SLX).....	49
3.9 Testprosedyrer.....	50

<b>4.0 Metode</b> .....	<b>52</b>
4.1 Datagrunnlag i vår undersøkelse.....	52
<b>5.0 Analyse</b> .....	<b>59</b>
5.1 Deskriptiv statistikk.....	59
5.2 Utvalg.....	61
5.3 Hypoteser.....	62
5.4 Valg av funksjonsform, variabler og modeller.....	63
<b>6.0 Resultater</b> .....	<b>68</b>
6.1 Preliminære funn.....	69
6.2 Konkluderende modeller.....	73
<b>7.0 Diskusjon</b> .....	<b>77</b>
7.1 Drøftelse av resultater.....	77
7.2 Metodologiske begrensninger.....	79
7.3 Strømpriskrisen vinteren 2021/22.....	81
<b>8.0 Konklusjon</b> .....	<b>85</b>
<b>9.0 Videre forskning</b> .....	<b>87</b>
<b>Referanser</b> .....	<b>88</b>
<b>Appendiks</b> .....	<b>98</b>
A1. 07221: Prisindeks for brukte boliger, etter kvartal.....	98
A2. 09817: Innvandrere uten norskfødte med innvandrerforeldre.....	98
A3. 09364: Kraftpriser i sluttbrukermarkedet.....	99
A4. 03014: Konsumprisindeks, total.....	99
A5. Gauss-Markov-teoremet - BLUE.....	100
A6. Deskriptiv statistikk.....	101
A7. Variable Importance Plots.....	104
A8. Korrelasjonsmatrise med varmekart.....	105
A9. Oversikt av observasjoner; energimerking.....	106
A10. Moran's plot - SEM.....	107
A11. Moran's plot - SLX.....	109
A12. Moran's plot - SDM.....	111
A13. Moran's plot - SAR.....	113
A14. Konkluderende modell (SEM med $K = 3$ ).....	115



## Figur- og tabelloversikt

Tabell 1. Oversikt over det demografiske fokusområdet	12
Tabell 2. Utviklingen av antall registrerte energiattester i Bergensregionen.	19
Tabell 3. Antall observasjoner oppsummert.	53
Tabell 4. Deskriptiv statistikk for ulike boligtyper	59
Tabell 5. Gjennomsnittspriser	59
Tabell 6. Deskriptiv statistikk av andre sentrale variabler	60
Tabell 7. Medianpris	60
Tabell 8. Median på boligattributter	61
Tabell 9. Deskriptiv statistikk for relevante variabler	61
Tabell 10. Deskriptiv statistikk for beregnet levert energi (kWh/m <sup>2</sup> )	62
Tabell 11. Variabelliste	65
Tabell 12. Referansekategorier	65
Tabell 13. Oversikt over fire beste modeller, synkende fra lavest AIC-verdi.	73
Tabell 14. SEM-modell med $K = 3$ . Energikarakterer, og regresjonskoeffisienter (avhengig variabel: logaritmen til kvadratmeterpris)	74
Figur 1. Selvlaget oversiktskart over kommunene med befolkningstall	13
Figur 2. Energimerket.	17
Figur 3. Antall attester fordelt på energikarakter.	20
Figur 4. Den hedoniske prisen og implisitte prisskjema for attributt $z_1$	31
Figur 5. Landeiers offerkurver.	36
Figur 6: Markedslikevekt.	37
Figur 7. Diagram over sammenslåinger av datasett	56
Figur 8. Diagram over variabelens viktighet.	64
Figur 9. Prisutviklingen i siste kvartal 2021	82
Figur 10. Prisutviklingen i første kvartal 2022.	83
Figur 11. Gjennomsnittlig kvadratmeterpris med strømpriskrisen	84

# 1.0 Introduksjon

I dette kapitlet vil vi belyse sentrale temaer som har relevans for vår problemstilling, med særlig vekt på aspekter ved boligmarkedet og prisdannelsen i dette markedet. Dette inkluderer en gjennomgang av konteksten relatert til boligmarkedet i Bergensregionen.

## 1.1 Generelt om boligmarkedet i Bergensregionen

Boligmarkedet i Bergen kommune har vært preget av sterk etterspørsel og økende priser de siste årene (Statistisk sentralbyrå, 2023; Krogsveen, 2023), se appendiks A1. Dette skyldes blant annet at byen er attraktiv å bo i, med vakker natur, rikt kulturliv og gode studiemuligheter for studenter. Den økte kostnaden forbundet med bolig, samt utfordringene ved å identifisere en passende og kostnadseffektiv bolig, er blitt en merkbar realitet for mange. Samtidig har dette også skapt muligheter for boligeiere og andre aktører på boligmarkedet. Bergen er også kjennetegnet ved en høy sysselsettingsgrad og en konstant befolkningsvekst, faktorer som utvilsomt bidrar til den økte etterspørselen etter boliger. I Bergen sentrum er det kun mulig å bosette seg i en brøkdel av sirkelen som omkranser byen på grunn av topografien.

I tillegg har det vært en økning i antall innvandrere til området (Statistisk sentralbyrå, 2023), referert i appendiks A2. En mulig effekt av denne innvandringen kan være en ytterligere stimulering av boligetterspørselen. En slik utvikling kan delvis tilskrives en voksende preferanse for leiligheter, særlig blant unge og studenter, kombinert med en mangel på tilgjengelige boliger i området

Boligmarkedet i Bergensområdet kan variere noe fra kommune til kommune, og det kan være forskjeller i prisnivå og tilbud av boliger avhengig av hvor i området man leter. I denne masteravhandlingen vil vi foreta en systematisk analyse av boligmarkedet i følgende kommuner i Bergensområdet: Bergen, Samnanger, Bjørnafjorden, Øygarden, Askøy, Vaksdal, Osterøy, Alver og Austrheim. For enkelhets skyld vil vi heretter referere til disse kommunene samlet som *Bergensregionen*. Med bakgrunn i dette anser vi det hensiktsmessig å gi korte beskrivelser om hver av kommunene.

Vi vil videre se på hvordan boligmarkedet i disse kommunene har utviklet seg de siste fire årene, og hvilke faktorer som har påvirket etterspørselen etter, og prisene på boliger i området.

## 1.2 Problemstilling

Boligsamvirket og byggenæringen har sentrale roller i håndteringen av noen vesentlige utfordringer: klimaendringer (Islam & Winkel, 2017, s. 1-2). Selv om det bygges mye, og bærekraftig, vil 80% av bygningsmassen i Bergen om 50 år allerede være bygd. Derfor er det viktig å redusere energibruken i eksisterende bygg, spesielt fra 60- og 70-tallet, som trenger oppgradering ifølge Klima- og energihandlingsplanen for Bergen som ble vedtatt av bystyret 21. september 2016. Energimerking kan være et effektivt verktøy for å øke bevisstheten om energiforbruk og løsninger for forbedring av energieffektiviteten, og dermed redusere totalt energiforbruk og klimagassutslipp. Dette tiltaket kan også potensielt øke verdien av boligen og gjøre den mer attraktiv for potensielle kjøpere, som et ledd i en mer bærekraftig og rettferdig boligpolitikk.

I henhold til Norges klimaforpliktelser og Bergens klimastrategi 2022-2030 (2023), krever den nødvendige elektrifiseringen av transport og industri økt energiproduksjon. Aunedi, Ortega og Green (2022, s. 3-8) viser at lokal energiproduksjon, som solceller, kan forbedre energiforbrukets fleksibilitet, redusere belastningen på det globale nettverket, senke strømkostnadene og tillate eksport ved overproduksjon. Dette er generelt positivt for boliger, klima, miljø og huseiere. Energimerking kan med bakgrunn i dette fremme bevissthet om lokale og bærekraftige energiløsninger og bidra til å redusere avhengigheten av fossil energi.

Samlet sett kan energimerking av boliger bidra til reduksjon av klimagassutslipp, økning i boligverdier og dekning av et økende kraftbehov i en framtid med potensielt kraftunderskudd (Sem et al., 2020, s. 29). Dette bør tas i betraktning når boligpriser i Bergen og andre byer vurderes i fremtiden. Energimerking gir potensielle boligkjøpere innsikt i boligens energistandard og anslåtte månedlige strømutfgifter, noe som kan påvirke både boligpriser og kjøpsbeslutninger. Et lavere energiforbruk resulterer i reduserte månedlige strømkostnader for boligeiere, noe som kan være en avgjørende faktor for kjøpere ved å redusere deres samlede boligutfgifter og bidra til økonomisk trygghet. I tillegg er lavere energiforbruk gunstig for miljøet, ettersom det reduserer boligsektorens karbonavtrykk og hjelper i kampen mot klimaendringer.

Økonomiske insentiver spiller også en betydelig rolle i denne sammenhengen. Statlige subsidier og skattefordeler kan stimulere boligeiere til å utføre energieffektiviseringsarbeid på eiendommene sine. Dette kan potensielt senke kostnadene ved å forbedre energieffektiviteten i boliger, noe som igjen kan bidra til å motivere flere boligeiere til å ta slike skritt, særlig dersom strømprisene er høye. På tampen av innleveringsfristen for denne masteroppgaven, lanserte Bergen kommune et økonomisk tilskudd for installasjon av varmepumper i private boliger. Formålet er å kutte

energibruken i Bergens bygg og boliger (Bergen kommune, 2023).

Med bakgrunn i dette, vil vi undersøke om det er en sammenheng mellom energimerking og boligpriser, og om dette varierer mellom ulike typer boliger og områder i Bergen. Vi håper å kunne bidra til økt forståelse for hvordan energimerking kan påvirke boligpriser i Bergen, og om dette kan bidra til å øke bevisstheten om energieffektivitet og redusere klimagassutslippene i byen.

Analysen i sin helhet skal bidra til å besvare vår overordnede problemstilling formulert som:

**Bidrar energimerking til å forklare variasjoner i boligpris, og hvor stor er i så tilfelle den økonomiske betydningen av energimerking?**

Denne problemstillingen bærer betydelig viktighet av flere grunner. Den har først og fremst betydning for boligeiere, utleiere og kjøpere i området, gitt at det tilbyr en dypere forståelse av hvordan energimerking kan påvirke boligprisene i en by med et mangfoldig boligtilbud som Bergen. Overgående er denne utredning av betydning som et empirisk bidrag også. Det gir en unik mulighet til å anvende, evaluere og teste teoretiske konsepter i en konkret og praktisk sammenheng, og på den måten bidra til eksisterende kunnskap i den vitenskapelige litteraturen om dette emnet eller bidra til å fylle eventuelle kunnskapshull. Med dette som bakgrunn, presenterer vi følgende forskningsspørsmål for å kaste lys over vår overordnede problemstilling:

- Hvor stor er sammenhengen mellom energimerking og boligpris, og er denne signifikant?
- Hva er den gjennomsnittlige energikarakteren, og hvordan har den endret seg over tid?
- Hvordan påvirker ulike energimerker (A-G) boligprisene, og er det noen merker som har en sterkere innvirkning på prisen enn andre?

For å undersøke forskningsspørsmålet vil vi benytte flere kilder for å drøfte sammenhengen mellom boligprisene i Bergen i over tid, og boligens energimerking. Dette historiske perspektivet omfatter årene 2019-2022, og er preget av både koronapandemien, og det medier og politikere omtaler som en strømpriskrise i kjølvannet av de høye strømprisene vinteren 2021-2022.

### **1.3 Oppgavens struktur og metode**

I denne oppgaven vil vi først presentere det teoretiske rammeverket og gi en oversikt over relevant bakgrunnsinformasjon. Deretter vil vi presentere datamaterialet og tilhørende deskriptiv statistikk, samt gi en detaljert fremstilling av vår metodologiske tilnærming, inkludert valg av

forskningsdesign og datainnsamlings- og behandlingsprosedyrer.

Videre vil vi presentere og diskutere resultatene fra dataanalysene i forhold til problemstillingen og de formulerte hypotesene. Til slutt vil vi konkludere på bakgrunn av våre funn, og komme med anbefalinger for videre forskningsarbeid.

### **Avgrensning**

Forskningen er avgrenset til ni kommuner i Vestland for å studere regionale trender. Analysen omfatter kun selveierboliger, ekskluderer næringsbygg og ignorerer leiemarkedet. Resultatene vil dermed primært være relevante for det private boligkjøps- og salgsmarkedet for selveierboliger

### **Implikasjoner**

Studiens implikasjoner kan understreke viktigheten av økt kompetanse om energimerking blant boligmarkedets aktører. Bedre forståelse av energimerkingens effekt på boligpriser kan øke bevisstheten om energieffektivitet. Energimerking kan også forsterke bærekraftig utvikling av boligsektoren, et stadig mer prominent tema. Funnene kan gi myndighetene, både lokalt i Bergen og nasjonalt, bedre innsikt i tiltak som kan styrke boligmarkedets bærekraftighet.

## 2.0 Bakgrunn - kommuneoversikt

I dette kapitlet i masterutredningen vil vi gi en grov oversikt over kommunene som er av interesse for denne studien. Formålet er å gi en grunnleggende forståelse av de spesifikke kommunene, deres geografiske plassering, demografiske sammensetning, topografi, bystruktur og andre relevante aspekter. Denne informasjonen er for å sette resten av forskningen i kontekst og forstå de unike utfordringene og mulighetene i hver kommune.

### 2.1 Befolkning, topografi og bystruktur

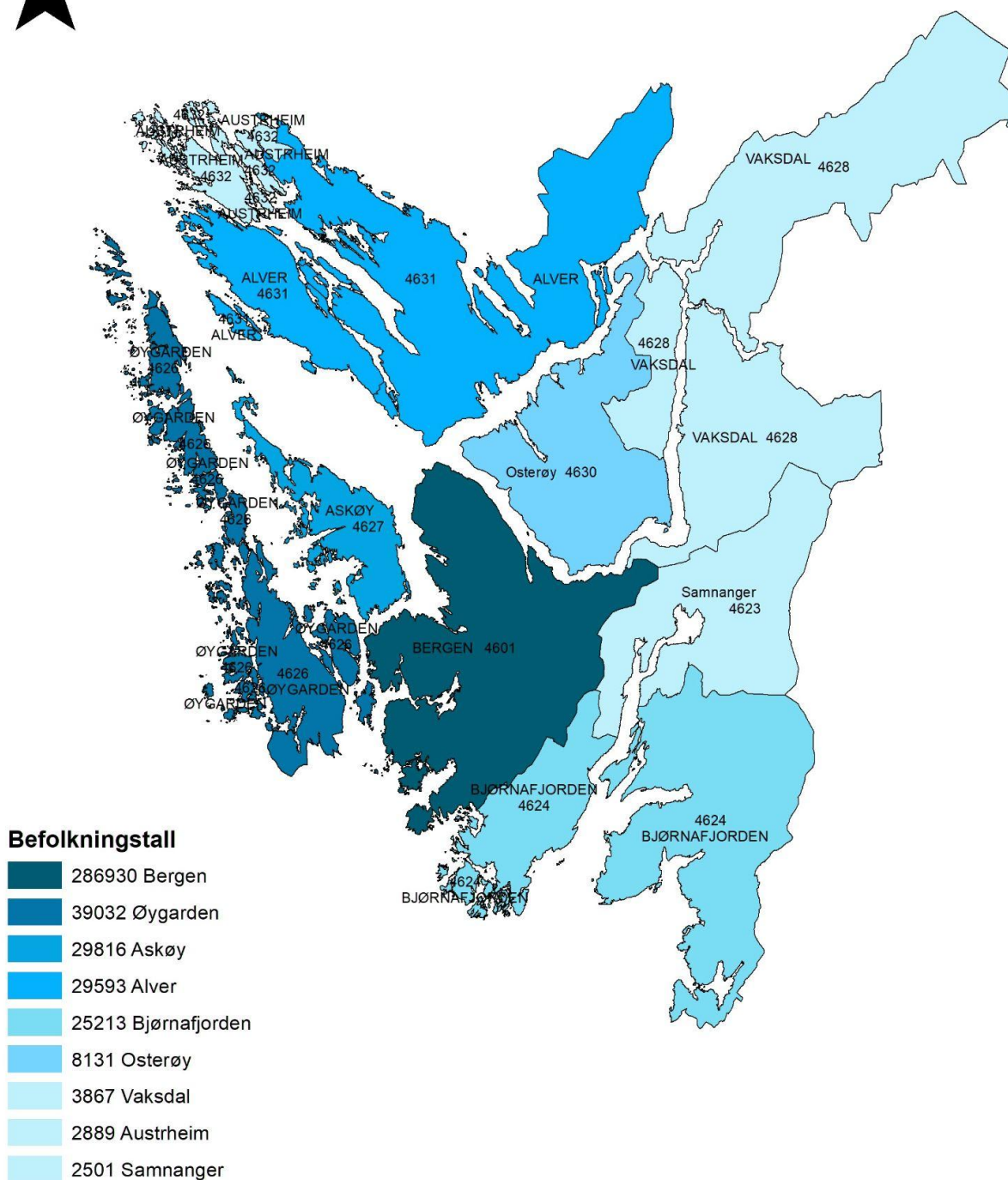
Tabell 1. Oversikt over det demografiske fokusområdet

Kommunennummer og navn (2022)	Areal av tettsted (km <sup>2</sup> )	04861: Befolkningen som bor i tettsteder	07459: Befolkning i kommunen
K-4601 Bergen	98,3	280 856	286 930
K-4623 Samnanger	0,9	1 066	2 501
K-4624 Bjørnafjorden	13,2	19 187	25 213
K-4626 Øygarden	21,1	28 776	39 032
K-4627 Askøy	17,8	26 660	29 816
K-4628 Vaksdal	1,8	2 834	3 867
K-4630 Osterøy	3,4	4 129	8 131
K-4631 Alver	11,7	17 276	29 593
K-4632 Austrheim	1,6	1 090	2 889

Data er fra «04861: Areal og befolkning i tettsteder (K) 2000 - 2022» og «07221: Prisindeks for brukte boliger, etter statistikkvariabel, region, boligtype og kvartal» av Statistisk sentralbyrå, 2023, (<https://www.ssb.no/statbank/table/04861>; <https://www.ssb.no/statbank/table/07221>). Gjengitt i henhold til Norsk lisens for offentlige data (NLOD).

Tabell 07459 viser befolkningen i kommunen, mens tabell 04861 viser befolkningen som bor i tettsteder innenfor kommunen. Ifølge SSB (2023) defineres et tettsted som en samling bygninger med minst 200 innbyggere, der avstanden mellom bygningene generelt ikke skal være mer enn 50 meter. En stor andel av befolkningen i de studerte kommunene, inkludert over halvparten i Austrheim, bor i områder utenfor tettsteder.

Figur 1. Selvlaget oversiktskart over kommunene med befolkningstall



I de neste avsnittene vil vi kort redegjøre for de tre største kommunene i areal og befolkning, og avslutningsvis kort beskrive resterende kommuner.

## **Bergen kommune**

Bergen kommune er en av de mest folkerike kommunene i Norge, og er kjent for sin vakre beliggenhet mellom fjellene og havet. Kommunen har en lang historie som en viktig havneby og handelsstad, og er i dag en viktig kultur- og næringsby i landet.

Bergen kommune har ifølge Bergen kommune (2022) et folketall på 286.930 per 1. januar 2022. Byen består av åtte bydeler, og rundt 244.000 av kommunens innbyggere bor utenfor sentrum. Terrenget gjør at bydelene er avgrenset fra hverandre, og mange har en sterk lokal tilknytning til sin bydel (Bergen kommune, 2022). Byen, som ofte omtales som Vestlandets hovedstad, har et areal på 465.3 kvadratkilometer og er omgitt av syv fjell.

Bergen kommune er preget av en variert topografi med fjell, daler, og hav. Fjellene rundt byen gir en vakker ramme, men kan også skape utfordringer for byutviklingen, som for eksempel begrensninger på tilgjengelig areal og begrensninger i infrastruktur. Vollset (2007) argumenterer for at den begrensede plassen i bysenteret førte til en økt bosetning i utkanten av sentrum, noe som resulterer i lenger gåavstand til byens kjerne og økt etterspørsel etter organisert transport. Bergen kommune har forsøkt å møte utfordringene knyttet til topografien gjennom å blant annet utvikle et godt kollektivtilbud hvor Bybanen har vært hovedsatsingen de siste årene.

## **Askøy kommune**

Askøy kommune er en øykommune som ligger vest for Bergen by. De fleste innbyggerne på Askøy bor i sørlige deler av kommunen og i tettbygde områder. 88 prosent av innbyggerne bor innenfor tettsteder, slik det er definert av Statistisk sentralbyrå (Askøy kommune, 2020, s. 11). Det er registrert 11 837 boliger på Askøy per 1. januar 2020. 68% av disse er eneboliger, 18% er rekkehus, 9% er horisontaldelte tomannsboliger og 4% er leilighetsbygg. Resten er andre typer bygg som sykehjem, bofellesskap og lignende. Askøy har en lavere andel leilighetsbygg enn nabokommunene (Askøy kommune, 2020, s. 13).

Askøy kommune, kjennetegnet av variert topografi med fjell, skoger, vann og øyer, byr på både utfordringer og muligheter for byutvikling, infrastruktur og unike boligområder med storslått natur og utsikt. Offentlig transport som buss, ferje og tog gir tilgang til Bergen sentrum og andre kommunedeler. Imidlertid byr topografien og begrensede veier på utfordringer for biltrafikk og parkering, selv om tiltak er iverksatt for å forbedre parkeringsmuligheter.



## Øygarden kommune

Øygarden kommune har omtrent 39 000 innbyggere med 2 800 yrkesaktive, og er den nest største kommunen i Vestland fylke. Kommunen representerer et område som er preget av høy folkevekst, et stadig voksende bosteds- og industriområde, samt et innholdsrikt kommunesenter på Straume, Sotra Kystby. Arealet er på omtrent 315 km<sup>2</sup>, og er sammensatt av flere hundre små og store øyer som etter hvert er og blir bundet sammen av broer (Øygarden kommune, 2022).

Kommunens unike topografi og velutviklede transportinfrastruktur påvirker boligprisene, og til tross for trafikk- og tilgjengelighetsutfordringer, arbeider kommunen aktivt med forbedringer. Det nye Sotrasambandet (Statens vegvesen, 2023) med en planlagt kollektivterminal ved Storavatnet og ferdigstillelse tidligst i 2023, vil endre transporttilbudet i vestkorridoren og Vest-Bergen. Dette sambandet vil knytte Øygarden og Bergen tettere sammen og potensielt redusere rushtidspendlingen med 20 minutter. Ifølge Gjestland et al. (2012) vil dette øke *tilgjengeligheten*, noe som igjen vil medføre en positiv *kapitalisering* i prisene på berørte eiendommer. Dette fører til to hovedspørsmål: *når* kapitaliseringen vil finne sted og *hvordan* økt tilgjengelighet skal måles. Den empiriske litteraturen gir ingen klar enighet om tidspunktet for kapitalisering, men basert på teori og under antakelse av fullkomment rasjonelle forventninger, kan det hevdes at full kapitalisering vil inntreffe så snart prosjektet blir vedtatt, eller kanskje til og med før, dersom folk forutser den positive innvirkningen. Gjestland et al. (2012) forventet at Trekantsambandet ville øke Bømlo-priser til nivåer lik Stord, men fant ingen signifikante forskjeller. I vår utredningen kan etableringen av Sotrasambandet potensielt endre den hedoniske prisfunksjonen. Det er viktig å understreke i oppgaven at dette sambandet ikke opprettes i løpet av studieperioden vår, noe som er en fordelaktig for analysen. Gjestland et al. (2012) sine analyser antyder at endringer i transportnettverket påvirker eiendomspriser først etter etablering. Derfor vil effekter fra Sotrasambandet på prisene sannsynligvis bli mer synlige etter opprettelsen, og derfor vil vår analyse ikke bli påvirket.

## Distriktskommuner

Distriktskommuner kjennetegnes ved et lite arbeidsmarked eller lang reiseavstand til større arbeidsmarkeder og tjenester. Næringslivet i distriktene har lengre avstand til sine markeder og har færre leverandører, finansinstitusjoner og kompetansemiljø lokalt enn næringslivet i mer sentrale områder (Meld. St. 5 (2019–2020), s. 25). Geografisk beliggenhet og lokale forhold har en stor

innvirkning på boligmarkedet, og denne påvirkningen varierer betydelig fra sted til sted. Denne variasjonen er spesielt tydelig når vi sammenligner urbane og distriktsområder, preget av ulike sosioøkonomiske og geografiske forhold, som igjen kan ha innvirkning på boligprisene.

*Samnanger* er en kommune i Midthordland i Vestland fylke. Den grenser i nord til Vaksdal, i øst mot Kvam, i sør mot Bjørnafjorden, og i vest mot Bergen. Samnanger kommune består av landskapstyper fra fjord til høyfjell, og har mye skog og daler.

*Bjørnafjorden* er en kommune i Midthordland i Vestland fylke som ble etablert 1. januar 2020 ved sammenslåing av kommunene Os og Fusa. Kommunen ligger ved Bjørnafjorden, et landskap bestående av fjorder, fjell og fosser - 20 minutter fra Bergen.

*Vaksdal* ligger mellom Bergen og Voss kommuner, øst for Osterøy kommune, men den vestligste delen av Vaksdal ligger på Osterøyas nordøstlige side. Det er gode pendlemuligheter mellom Vaksdal og Bergen (40 min.).

*Osterøy* ligger på innlandsøyen Osterøy, som er delt mellom kommunene Osterøy og Vaksdal. Kommunen kjennetegnes blant annet av småindustri, landbruk og fiskeoppdrett-/foredling. Nesten halvparten av de bosatte og yrkesaktive har arbeid utenfor kommunen.

*Alver* ble etablert 1. januar 2020 da de tidligere kommunene Lindås, Radøy og Meland ble slått sammen. Kommunen består av alt fra trange fjorder til åpne hav, og sysselsettingen består i stor grad av olje- og gassrelaterte yrker, i tillegg til matproduksjon. Kommunen har også betydelig innpendling fra Bergen, og nabokommuner.

*Austrheim* består av øyer ut mot Nordsjøen, samt nordvestspissen av Lindåshalvøya på fastlandet. Store deler av sysselsettingen i kommunen består av industri, og mer enn halvparten av yrkesaktive har arbeidsplasser utenfor kommunen.

## **2.2 Energimerkeordning i Norge**

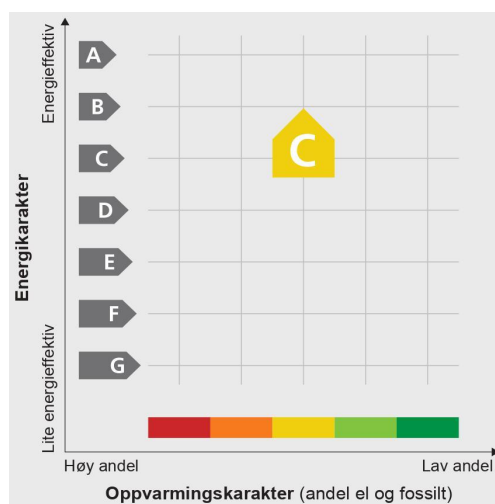
Oppgradering av energieffektiviteten i bygninger er et viktig fokus for industrialiserte land som ønsker å oppnå bærekraftig utvikling. Prosessen er imidlertid treg. Selv om det finnes flere kostnadseffektive energibesparende tiltak tilgjengelig for boligeiere, utnyttes ikke deres potensial for energibesparelser (Curtin og Maguire, 2011).

Energimerkingen av bygninger skal sikre bygningseiere, leietakere og andre aktører i

eiendomsmarkedet grunnleggende informasjon om energitilstanden til bygningen. Bygningens energitilstand vil slik bli en del av beslutningsgrunnlaget ved kjøp og leie (Ot.prp. nr. 24 (2008-2009), s. 9). Energimerkeordningen krever at boligeiere skal presentere en energiattest ved salg og utleie. Denne attesten kan hentes ved å gi informasjon til det nettbaserte Energimerkesystemet. Boligeier kan enten registrere informasjonen selv eller be noen andre om å gjøre det for dem.

*Energiattest med energimerke* er påkrevd for alle boliger som selges fra 1. juli 2010 for å øke bevisstheten om energibruk og energieffektivitet. Målet er å øke bevisstheten om energibruk, ulike oppvarmingsløsninger og løsninger som kan gjøre boligen mer energieffektiv (Enova SF, 2009). *Energimerket* består av en *energikarakter* og en *oppvarmingskarakter*.

**Figur 2.** Energimerket.



Fra «Publikasjoner, presentasjoner og dokumenter», av Enova SF, 2009. (<https://www.energimerking.no/no/energimerking-bygg/om-energimerkesystemet-og-regelverket/publikasjoner-og-dokumenter/>).

Energimerket på en bolig reflekterer den energien som trengs for å opprettholde innertemperaturen og den benyttede energikilden. Energimerking gjøres gjennom en teknisk vurdering av boligens isolering, vinduer, og oppvarmingssystemer. Energimerket gir en indikasjon på boligens energieffektivitet, med A som det mest effektive, noe som resulterer i lavere strømutfgifter. Boliger bygget før 2010 har vanligvis et energimerke gul/grønn C.

*Energikarakteren* er basert på den forventede mengden strøm som vil bli brukt ved normalt forbruk. En bolig med karakter A vil altså ha en lavere strømregning enn en med karakter G, alt annet likt.

*Oppvarmingskarakteren* måler bygningens oppvarmingssystems miljøvennlighet og bærekraft, på en femdelt skala fra grønt (fornybar energi, ingen direkte bruk av elektrisitet, olje eller gass) til rødt

(tradisjonelle, mindre miljøvennlige energikilder som elektrisitet, olje eller gass).

Energimerking har blitt et sentralt tema i både miljø- og økonomi sammenhenger de siste årene på grunn av boligsektorens betydelige energiforbruk. Bøeng (2014) indikerer at nesten 80% av husholdningenes energiforbruk i Norge i 2012 gikk til strøm, hovedsakelig til oppvarming. Det totale forbruket på 20 230 kWh per husholdning representerte omtrent 20% av landets totale energiforbruk det året. Norges vassdrags- og energidirektorat (NVE) rapporterte i 2018 at husholdningene, inkludert jordbruk, sto for ca. 35% av det totale energiforbruket (Backe, 2020).

Andelen boliger som selges med energikarakter A, B eller C øker stadig ifølge Eiendomsverdi (2019; 2022), og dette skyldes flere faktorer. Stadig flere huseiere har iverksatt tiltak for energieffektivisering som resulterer i bedre energiklassifisering. I tillegg, boliger som er konstruert i samsvar med gjeldende byggeforskrifter oppfyller automatisk kravene for en høyere energikarakter. Dette medfører at en økende andel av boligene har en god energiklassifisering, og er i en utvikling i tråd med hva grønne lån insentiver til. Med dette i betraktning, vil vi nå vende oppmerksomheten mot den konkrete utviklingen av antall registrerte energiattester i Bergensregionen. Tabell 2 gir en oversikt over utviklingen de siste årene.

**Tabell 2.** Utviklingen av antall registrerte energiattester i Bergensregionen.

År	Jan	Feb	Mar	Apr	Mai	Jun	Jul	Aug	Sep	Okt	Nov	Des	Total
2009	0	0	0	0	0	0	0	0	0	0	0	18	<b>18</b>
2010	47	17	34	16	20	81	225	310	388	297	228	86	<b>1749</b>
2011	234	218	320	233	0	0	0	0	0	0	0	0	<b>1005</b>
2012	263	301	348	348	381	257	173	330	298	284	178	98	<b>3259</b>
2013	245	232	264	336	310	268	151	298	282	261	181	132	<b>2960</b>
2014	207	165	0	273	388	437	255	449	379	441	255	186	<b>3435</b>
2015	316	404	565	477	596	512	331	421	409	366	20	181	<b>4598</b>
2016	311	445	428	521	565	416	282	584	532	395	222	146	<b>4847</b>
2017	397	409	529	438	638	474	326	451	379	467	427	198	<b>5133</b>
2018	403	385	385	494	594	487	305	487	438	572	284	230	<b>5064</b>
2019	372	366	458	492	628	450	286	460	427	425	329	229	<b>4922</b>
2020	362	382	635	583	909	708	343	588	1138	813	358	179	<b>6998</b>
2021	334	390	704	629	1001	1020	427	459	437	353	329	170	<b>6253</b>
2022	339	434	812	646	389	0	2	299	569	489	406	246	<b>4631</b>

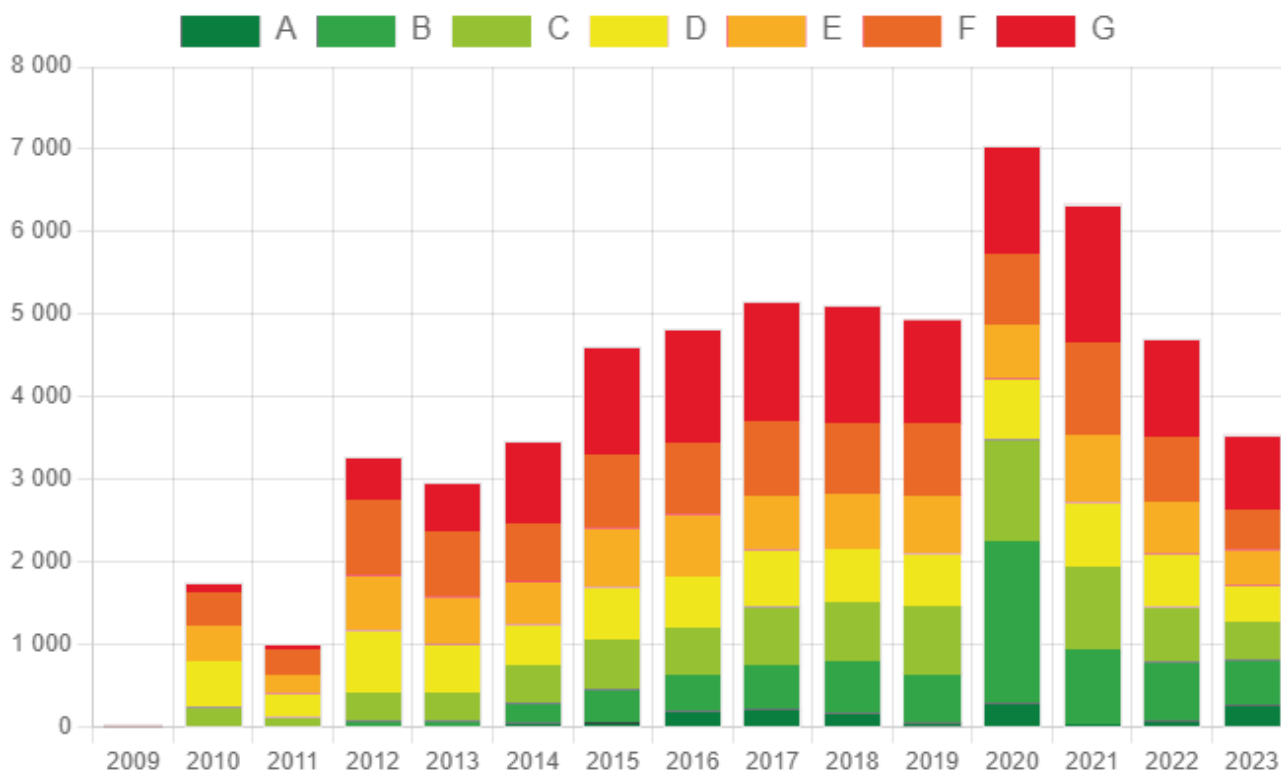
Data er fra «Antall attester over tid» av Enova Energimerking, 2023.

(<https://portal.ems.enova.no/statistikk/over-tid?geography=Vestland&kommune=4631&kommune=4627&kommune=4601&kommune=4632&kommune=4624&kommune=4630&kommune=4623&kommune=4618&kommune=4628&kommune=4626>). Gjengitt i henhold til Norsk lisens for offentlige data (NLOD).

Tabellen viser en betydelig økning i registrerte energiattester over tid, fra 18 i 2009 til et toppunkt på 6 998 i 2020. Selv om antallet sank noe i 2021, var det fortsatt høyt, med 6 253 attester. Vi observerer en sesongmessig variasjon, med en tendens til økning i vår- og sommermånedene, og en nedgang mot slutten av året. Dette mønsteret har vært konstant, til tross for noen unntak

Videre presenteres fordelingen av energiattester fordelt på energikarakter fra 2009 til 2023 i Bergensregionen. Grafen gir en oversikt over utviklingen av energikarakteren i løpet av disse årene, og fremhever en økende andel av høyere energikarakterer.

**Figur 3.** Antall attester fordelt på energikarakter.



Fra «Attester etter energikarakter og oppvarmingsmerke», av [Enova Energimerking](#), 2023.

(<https://portal.ems.enova.no/statistikk/energikarakter?buildingCategory=Smahus&buildingCategory=Boligblokker&geography=Vestland&kommune=4631&kommune=4627&kommune=4601&kommune=4632&kommune=4624&kommune=4630&kommune=4623&kommune=4618&kommune=4626>).

Fra 2009 til 2023 kan vi se en tydelig økning i antall boliger med energikarakter A, B og C. I 2009 var andelen av boliger med karakter A, B eller C ekstremt lav (0,01 %). Men i løpet av de neste årene ser vi en betydelig økning, spesielt for energikarakter B og C. I 2020 er det en bemerkelsesverdig økning i karakter A (0,50 %) og B (3,44 %), mens andelen av karakter C også forblir høy (2,08 %). Samtidig ser vi en nedgang i andelen av boliger med lavere energikarakter (F og G) fra 2012 til 2023. For eksempel, andelen av boliger med energikarakter G har sunket fra 2,47 % i 2018 til 1,18 % i 2023. Dette kan forklares med at fokus på energieffektivitet har hatt en betydelig innvirkning på boligmarkedet.

Det er viktig å merke seg at selv om andelen av boliger med høyere energikarakter (A, B og C) har økt betydelig, utgjør de fortsatt en mindre andel sammenlignet med lavere energikarakterer (D, E, F og G). Totalt sett har 2,29 % av boligene energikarakter A, 12,35 % har energikarakter B og 13,78 % har energikarakter C, mens 14,75 % har energikarakter D, 14,55 % har energikarakter E, 17,93 % har energikarakter F og 24,35 % har energikarakter G.

Grafen illustrerer den generelle trenden mot økt energieffektivitet og gir et klart bilde av hvordan

boligmarkedet har utviklet seg i forhold til energiklassifisering i løpet av de siste årene.

## 2.3 Strømpriskrisen vinteren 2021-2022

I løpet av perioden 1993-2021 økte det samlede strømforbruket i norske husholdninger med 28 prosent (Tangeland et al., 2022). Vinteren 2021, preget av lave temperaturer, resulterte i et historisk høyt strømforbruk i norske husholdninger, ettersom en stor andel av husholdningene bruker elektrisitet som primærkilde for oppvarming av boligen sin, ved hjelp av enheter som elektriske ovner, varmekabler, og/eller sentralvarmesystemer som drives av elektrisitet (Tangeland et al., 2022). Det økte forbruket kom sammen med økte strømpriser i landet vinteren 2021-2022. Prisøkningen er et resultat av flere faktorer som lav fyllingsgrad i vannmagasinene. Andre faktorer er høy gasspris og CO<sub>2</sub>-avgift i Europa, økt etterspørsel etter strøm i nabolandene, tørr sommer og høst, og rask utfasing av kull og kjernekraft.

En midlertidig *strømstøtteordning* ble etablert desember 2021 for å hjelpe husholdninger å håndtere de ekstraordinære høye strømprisene. Ordningen fungerer slik at når gjennomsnittlig spotpris for måneden går over 70 øre/kWh i det prisområdet husholdningen tilhører, vil staten dekke en andel av forbruket over dette gjennomsnittet. Ordningen har blitt styrket og forlenget flere ganger, og regjeringen foreslo nylig at utbetalingen av strømstøtte skal baseres på strømprisen time for time, at støttegraden blir økt til 90 prosent over 70 øre hele året, og at ordningen skal forlenges ut hele 2024. En husholdning kan få dekket opp til 5 000 kilowattimer per måned, og støttegraden gjelder selve strømprisen. I tillegg må husholdningene betale avgifter og nettleie (Regjeringen.no, 2023).

Strømpriskrisen har og har hatt konsekvenser for både næringslivet og husholdningene i Norge, med særlig innvirkning på Vestlandet, gitt regionens avhengighet av vannkraft. En spesifikk konsekvens som vil bli undersøkt i denne studien, er den potensielle negative effekten på boligprisene i Bergen. Økte strømpriser, særlig i vintermånedene, kan medføre at boligeiere står overfor høyere utgifter til både oppvarming og belysning. Dette kan i sin tur redusere deres betalingssevne og dermed også etterspørselen etter boliger. I tillegg kan høye strømpriser føre til lavere lønnsomhet og konkurransevne for industrien i Bergen, noe som kan påvirke sysselsettingen og inntektsnivået i byen. Dette kan igjen redusere boligetterterspørselen ytterligere.

## 3.0 Teori- og litteraturgjennomgang

I denne delen av oppgaven presenteres den teoretiske forankringen, som vil bidra til å belyse boligprising, energimerking og økonometriske modeller hver for seg, og i relasjon til hverandre. Formålet med teoridelen er å gi vekt og substans til den påfølgende analysen og diskusjonen, og vil i sum belyse oppgavens overordnede problemstilling og relaterte forskningsspørsmål.

### 3.1 Litteratur og teori om boligmarkedet i Norge og Vestlandet

I dette kapitlet vil vi beskrive boligmarkedet med forankring i litteratur, teori og statistikk.

#### 3.1.1 Kjennetegn på det norske boligmarkedet

Det norske boligmarkedet kjennetegnes av en beskjeden befolkning på om lag 5,5 millioner innbyggere og en befolkningstetthet på 14-15 individer per kvadratkilometer, noe som plasserer Norge blant verdens minst tettbefolkede nasjoner (Meld. St. 18 (2016–2017)).

Oslo, som er Norges største by, huser omtrent 700 000 mennesker, mens Bergen, Trondheim og Stavanger følger deretter med henholdsvis omtrent 280 000, 210 000 og 145 000 innbyggere. Til tross for Norges begrensede størrelse og befolkning, er boligmarkedet i landet sterkt konkurransepreget enn i andre land, på grunn av de rådende markedskreftene som styrer allokeringen av boliger. Om man finner egnede boliger avhenger i stor grad av tilbudet, som kan variere geografisk. Ytterligere særtrekk inkluderer et kjølig klima, som kan påvirke boligkonstruksjon og isolasjon og dermed også drifts- og eierskapskostnader.

Oppsummert er det norske boligmarkedet preget av en liten befolkning med høy eierandel (Sandlie og Sørvoll, 2017; Karlberg og Victorin 2004), samt enkelte utfordringer knyttet til befolkningssammensetning og klimatiske forhold. Det er avgjørende å vurdere disse faktorene når man analyserer boligpriser og tilgjengelighet i Norge.

Den norske boligmodellen er preget av et ønske om at folk skal kunne kjøpe og eie sitt eget hjem. Dette har ført til den høye eierandel i Norge - blant de høyeste i OECD (Sandlie og Sørvoll, 2017; Karlberg og Victorin 2004).

Et viktig kjennetegn ved den norske boligmodellen er at andelen av husholdningene som eier bolig er omtrent 60-70% av alle boliger i Norge er privateid, det vil si at de er direkte eid av enkeltpersoner (Sandlie og Sørvoll, 2017; Karlberg og Victorin 2004). Dette er et høyt tall



sammenlignet med andre land, og viser at eierskap av bolig er viktig for mange nordmenn.

I tillegg til direkte eierskap, finnes også andre typer eierskap i Norge. 11-14% av boliger i Norge er indirekte eid (Sandlie og Sørvoll, 2017; Karlberg og Victorin 2004), det vil si at eierskapet er delt mellom flere personer eller organisasjoner. Dette kan for eksempel være sameie, borettslag og aksjeleiligheter.

### 3.1.2 Kjennetegn på boligobjekter som et gode

Ifølge Oxley (2004, s. 31) finnes det *eksplisitte* og *implisitte* boligmarkeder. Vi velger å definere boliger som heterogene goder og vil dermed anvende en *hedonisk modell*. Målet med oppgaven er å gi en dybdeanalyse av *implisitte priser*, med særlig fokus på energimerking.

Bolig er et eksempel på et heterogent gode. Et slikt gode består ifølge Day (2001) av en blanding av attributter som, selv om de varierer i en rekke egenskaper, er så tett relatert at de anses som å være et gode. Totalprisen som betales for en bolig i eiendomsmarkedet vil bli bestemt av de spesifikke strukturelle egenskapene til boligen, omgivelsene og beliggenheten. Blant disse egenskapene vil vi særlig fokusere på energimerking ved den aktuelle boligen. Derfor vil vi forvente at eiendommer med lav energimerking vil ha lavere priser enn tilsvarende boliger isolert sett.

I en grunnleggende artikkel foreslo Rosen (1974) en modell for markedets atferd som beskrev hvordan markedene for differensierte goder fungerer. Modellen som Rosen presenterte gir det teoretiske grunnlaget for hedonisk verdsettelse og vil være tema for den neste delen.

Day (2001) skriver videre at prisene i et bestemt boligmarked blir fastsatt av interaksjonen mellom tilbud og etterspørsel. Markedet når enighet om et sett med priser for de forskjellige boligtypene som tilfredstiller markedet, og tilbudet er lik etterspørsel. Prisene kan oppsummeres ved hjelp av en *hedonisk prisfunksjon*, som er en likevektsrelasjon, som viser sammenhengen mellom totalpris på boligens egenskaper og deres implisitte priser. I neste kapittel vil vi redegjøre for den hedoniske teorien som ligger bak.

### 3.1.3 Tidligere forskning

Energimerking av boliger har fremstått som et sentralt instrument for å øke bevisstgjøring rundt energieffektivitet og redusere energikonsum i Norge. Til tross for dette, finnes det begrenset forskning som belyser hvordan energimerking påvirker boligprisene i Bergensregionen. Flere

årsaker kan ligge til grunn for dette, slik som at dette er et nyere forskningsfelt eller at det eksisterer intrikate sammenhenger mellom energimerking og boligpriser. Uavhengig av årsak, bidrar dette til å øke interessen for å utforske temaet ytterligere. Videre presenteres noen sentrale funn.

### **On the economics of energy labels in the housing market**

Brounen og Kok (2011) i Nederland er blant de første som utfører en undersøkelse om innføringen av energimerking ved hjelp av den hedoniske metoden. I studien deres, ble det oppdaget en positiv korrelasjon mellom boligpris og energimerket. Den nederlandske ordningen for energimerking, som ligner den norske, bruker også karakterene A-G. Studien konkluderte med at det er en sammenheng mellom karakterene, hvor karakter A gir høyest prisøkning. Forskerne analyserte boligtransaksjoner for omtrent 32 000 merkede boliger og dokumenterte en positiv sammenheng mellom energieffektiviteten til en bolig og dens transaksjonspris. Ved å anvende Heckmans to-trinns metode, fant de at en bolig med et grønt merke selges med en premie på 3,6% sammenlignet med tilsvarende boliger med uten grønne merker. Denne premien varierer avhengig av merkets resultat, og beregninger tyder på at denne variasjonen delvis kan forklares med det underliggende energiforbruket til boligen.

### **Energy performance certificates - Informing the informed or the indifferent?**

I vår utredning tar vi utgangspunkt i Olaussen et al. (2017) sin artikkel hvor de undersøkte om energimerking har noen betydning for prisen på en bolig når den skal selges. De sammenlignet data fra årene før energimerkingen ble innført i Norge i 2010, med data fra årene etter, og fulgte opp utviklingen senest i 2019. Olaussen et al. (2017) tror at nyrenoverte bad vil ha en større innvirkning på pris enn energimerkingen. De mener også at faktorer som sjøutsikt, et barnevennlig område eller nærhet til butikker vil ha større betydning for prisen på en bolig. Derimot tror de at grønne boliglån kan bidra til at intensjonen bak energimerkingen vil bli oppfylt i større grad.

Forfatterne forsøker å etablere en kausal sammenheng i studien ved å bruke en fasteffektmodell. Dette er et kraftig verktøy i økonometri som kontrollerer for tidsspesifikke uendrede variabler, altså de faktorene som ikke endrer seg over tid, eller endrer seg svært sakte. I dette tilfellet bruker forfatterne det for å kontrollere for variabler som boligtype, boligstørrelse, bydel og alder. For å gjennomføre dette benytter de seg av et datasett som inneholder flere transaksjoner for samme bolig

i perioden 2000 til 2014, med unntak av 2010 da energimerking ble innført. Ved å gjøre dette, kan de kontrollere for faktorer som er unike for hver bolig, men som ikke endrer seg over tid, som estetisk utseende, lokasjon og boligens standard.

Hjertet av deres kausale argument ligger i hvordan de skiller prisvirkningen av disse tidsspesifikke uendrede faktorene fra prisvirkningen forbundet med innføringen av energimerking. Ved å fjerne påvirkningen av alle disse tidsspesifikke uendrede faktorene (det vil si de faste effektene), isolerer de innvirkningen av energimerking på boligpriser. De nevner også den klare tidstrenden i boligprisdata, noe som indikerer behovet for å ta hensyn til endringer i makroøkonomiske variabler som påvirker alle boliger over tid.

Resultatene deres viste at det ikke var noen merkbar endring i betydningen av energimerking for salg av boliger, eller at innføringen av lovpålagt energimerking ikke hadde innvirkning på boligprisene. Derimot tror de at grønne boliglån kan bidra til at intensjonen bak energimerkingen vil bli oppfylt i større grad.

### **Real Estate Price Formation: Energy Performance Certificates and the Role of Real Estate Agents**

Olaussen et al. (2021) analyserte boligsalgsdata fra Oslo, og tok hensyn til faktorer som salgspris, utlysingspris, dato, adresse, bydel, størrelse, type bolig, byggeår, og energimerker. De benyttet en hedonisk modell for å analysere dataene, og trakk sammenligninger til tidligere studier av Brounen og Kok (2011) og Olaussen et al. (2017).

Blant de mest fremtredende funnene var betydningen av variabler som bygningens alder, beliggenhet, år og størrelse, som alle viste seg å være signifikante på 1% nivå i både før og etter innføringen av energimerking i 2010. Energimerkene viste imidlertid en mindre klar sammenheng med boligens prisantydning enn forventet. Merkene B, C og D viste signifikans, men A og E gjorde det ikke. Dette kan være på grunn av et lite utvalg av boliger med A-merke i etter innføringen, og ingen med A-merke i før innføringen.

Interessant nok viste studien at innføringen av energimerking ikke synes å ha endret eiendomsmeglernes verdsettelse av boliger nevneverdig. Selv om høyere energimerker generelt er forbundet med høyere prisantydninger, tyder resultatene på at eiendomsmeglere hadde evnen til å justere prisantydningen til boligens faktiske energistandard både etter og før innføringen av

energimerker. Studien konkluderer med at, til tross for energimerkets betydning i kjøpsbeslutninger, formidler eiendomsmeglere denne informasjonen uavhengig av energimerking, noe som indikerer at slike merker ikke tilfører ekstra informasjon til eiendomsmarkedet. Dette er noe de har klart å gjøre selv før introduksjonen av energisertifikater.

Statistiske studier som Brounen og Kok (2011), Fuerst et al. (2015) og Jensen et al. (2016), viser en positiv korrelasjon mellom energimerking av boliger og transaksjonspriser Olaussen, Oust og Solstad (2021, s. 7-8). I tillegg har Hyland et al. (2013), Bio Intelligence Service et al. (2013) og Marmolejo-Duarte og Chen (2018) funnet lignende resultater ved å bruke utlysingspriser istedenfor transaksjonspriser. På den andre siden kommer meningsbasert forskning, som Murphy (2014), Amecke (2012), Laine (2011) og Backhaus et al. (2011), til motsatt konklusjon.

Ifølge Olaussen et al. (2021, s. 8) bygger Olaussen et al. (2017) bro mellom de statistiske studiene og meningsbasert forskning ved å forbedre spesifikasjonen av den hedoniske regresjonsanalysen og anvende en fast effektmodell. Dette viser at resultatene fra den statistiske analysen samstemmer med dem fra meningsbasert forskning. Studien deres tyder på at kjøpere enten ignorerer energimerkingen på kjøpstidspunktet, eller at de er informert om boligens energiytelse gjennom andre kanaler. I begge tilfeller blir energimerkingen overflødig.

### **Eiendomsverdi sine analyser**

Eiendomsverdi AS er et selskap som overvåker og registrerer aktivitet og utvikling i de norske eiendomsmarkedene. Eiendomsverdi leverer blant annet faglige innlegg som baserer seg på boligprisstatistikken, som de selv leverer på oppdrag av Eiendom Norge. Serien heter «fakta om boligmarkedet,» og ble originalt publisert på eiendomsverdi.no.

I mai 2019 publiserte Eiendomsverdi AS en artikkel med en regresjonsanalyse og undersøkte om energieffektive boliger selges til en høyere pris enn de mindre energieffektive boligene. Analysen tok utgangspunkt i 417 000 omsetninger med energikarakter fra 2014 til 2019. Analyseresultatene viste en klar trend hvor boliger med lavere energikarakter selges i gjennomsnitt til lavere salgspris. Effekten av energikarakterene er også større for enebolig enn leilighet. For en enebolig var det en rabatt å gå fra karakter C til F på 325 000 kroner mens den samme rabatten var 118 000 kroner for leiligheter

Selv om modellen i stor grad er utfyllende, er det umulig å kontrollere for innvendig standard for

eksempel. Artikkelen skrevet av Olausen et al. (2017), fant at boliger med høy energikarakter allerede hadde et prispåslag før merkeordningen ble innført. Dette kan bety at boligene skiller seg fra hverandre med andre ikke observerte faktorer som kan forklare prisforskjellene enn energikarakterene. På den andre siden kan det like gjerne bety at kjøpere innser hvilke boliger som er energieffektive uten hjelp av energikarakteren. Begge deler tyder på at energieffektive boliger ikke opplevde endring i prising ved innføringen av energimerker.

I mai 2022 publiserte de en oppfølging av artikkelen nevnt over fordi strømprisene i Norge økte fra sommeren 2021, noe en også kan se tydelig i appendiks A3. Man kunne anta at høyere strømkostnader ville føre til en sterkere prispremie for boliger med lavere strømforbruk. For å undersøke om dette faktisk har skjedd, analyserte de utviklingen i prispremien over tid og var spesielt interessert i å se på effekten før og etter strømprisøkningen sommeren 2021. Det Eiendomsverdi søkte å identifisere, var merkostnaden for boliger med energikarakter A, B eller C, sammenlignet med boliger som har en lavere energikarakter.

Resultatene fra Eiendomsverdi viser at det er en prispremie for boliger med energikarakter A, B eller C sammenlignet med boliger med lavere energikarakter. Imidlertid er det ikke funnet noen betydelig endring i denne premien *over tid*, heller ikke det siste året etter økningen i strømprisene fra sommeren 2021. Dette indikerer at den relative prisingen ikke har blitt påvirket av økte strømpriser, eller at det tar tid før husholdninger og boligkjøpere tar høyde for de økte strømprisene og effekten dette har på husholdningsbudsjettet.

### 3.2 Hedonisk teori

I økonomifaget blir hedonisk metode ofte anvendt i studier av boligmarkedet. Metoden knyttes til konseptet heterogene goder, hvor det antas at de ulike godene som besitter disse egenskapene eller attributtene, gir glede eller nytte. I sammenheng med bolig, kan disse attributtene for eksempel være størrelse, sentrumsnærhet eller tilgjengelighet til friluftsområder (Osland, 2001, s. 1).

Den hedoniske metoden kan brukes til å bryte ned boligprisene til ulike parametrene som påvirker prisen, for å kunne måle påvirkningen de har på boligprisene. Formålet med modellen er å gjennomføre regresjonsanalyser på boligkarakteristikker, for å kunne fremstille implisitte priser på de individuelle karakteristikker (Malpezzi, 2002, s. 2). Den hedoniske metoden tar dermed utgangspunkt i et boligmarked der boliger blir ansett som et heterogent gode. At et gode er heterogent betyr at det er unikt, og en vil ikke kunne finne to heterogene goder som er identiske, og

nytten en vil få av to heterogene goder vil være ulik (Osland, 2001, s. 1).

Anvendelsen av hedonisk metode er fordelaktig i undersøkelser av boligmarkedet fordi den tar i betraktning hvordan individuelle, subjektive vurderinger av boligens kvaliteter påvirker prisene, noe som muliggjør en mer presis verdivurdering sammenlignet med alternative metoder. I denne forbindelse fremhever Osland (2001, s. 1) relevansen og betydningen av den hedoniske metoden. Hun påpeker at boliger utgjør den mest betydelige investeringen for mange mennesker gjennom deres levetid, og dessuten representerer den mest betydelige andelen av mange individers formue.

Rosen (1974, s. 34) gir en mer teknisk definisjon av hedoniske priser, som han beskriver som implisitte priser av egenskaper som blir avslørt for økonomiske aktører gjennom observerte priser av differensierte produkter og de spesifiserte egenskapene tilknyttet dem. Denne definisjonen understreker metodens evne til å avsløre implisitte verdier knyttet til ulike boligeenskaper.

Det er viktig å merke seg at den hedoniske modellen opererer under følgende forutsetninger i forbindelse med boligmarkedet generelt:

1. Boliger er betraktet som *heterogene* goder. Dette innebærer at selv om det overordnede godet (boligen) er heterogent, består det av flere homogene attributter eller egenskaper.
2. En annen antagelse er at boliger innebærer betydelige *produksjonskostnader*. Med andre ord, konstruksjonen av nye boliger krever betydelig økonomisk investering.
3. Videre antas det at boliger har en lang teknisk *levetid*, hvilket betyr at de er designet og bygget for å vare over lengre tid.
4. Boliger er ansett som *immobile* goder. Dette betyr at de er faste i posisjon, og at deres geografiske plassering eller romlige egenskaper blir en integrert del av godet. Derfor vil beliggenheten, sammen med de fysiske egenskapene til boligen, påvirke prisen.
5. Til slutt antar den hedoniske modellen at boliger er *udelelige*. Dette innebærer at en bolig, i sin helhet, ikke kan deles og selges i separate deler.

Disse forutsetningene danner grunnlaget for den hedoniske prismodellen og dens anvendelse i studier av boligmarkedet. I tillegg til disse grunnleggende elementene, er det essensielt å også ta hensyn til den unike strukturen i boligmarkedet, som selv er bundet av egne antagelser:

Boligmarkedet opererer under antagelsen om markedets ufullkommenheter, som skaper en tilstand av *asymmetrisk informasjon* (Osland, 2001, s. 2-3). Dette har konsekvenser for flere aspekter av markedet, inkludert informasjonskostnader, søkekostnader, transaksjonskostnader, og flyttekostnader. Dette betyr at det finnes en rekke faktorer som hindrer et boligmarked i å fungere

som et perfekt konkurransedyktig marked. I et perfekt marked ville alle kjøpere og selgere ha fullstendig og umiddelbar tilgang til all informasjon, det ville ikke være noen transaksjonskostnader, og det ville være mange kjøpere og selgere slik at ingen enkeltaktører kan påvirke prisen. I praksis er imidlertid boligmarkedet langt fra dette idealet.

En av disse ufullkommenhetene er *asymmetrisk informasjon*, som oppstår når en part i en transaksjon har mer eller bedre informasjon enn den andre. For eksempel kan en selger av en bolig kjenne til skjulte problemer med boligen som kjøperen ikke har kunnskap om. Dette kan skape en ubalanse i forhandlingsstyrken mellom partene, og potensielt føre til uønskede resultater, som for eksempel en høyere pris enn hva boligen egentlig er verdt.

Asymmetriske informasjon og andre ufullkommenheter kan også øke ulike kostnader forbundet med boligmarkedet:

- *Informasjonskostnader* er kostnadene forbundet med å innhente informasjon om tilgjengelige boliger, deres priser, kvalitet og så videre.
- *Søkekostnader* er kostnadene forbundet med å lete etter en passende bolig.
- *Transaksjonskostnader* er kostnadene forbundet med selve kjøpsprosessen, for eksempel kostnader til eiendomsmeglere, advokater, bankgebyrer og lignende.
- *Flyttekostnader* er kostnadene forbundet med å flytte fra en bolig til en annen.

Videre er boligmarkedet underlagt *offentlig regulering*, styrt og implementert av myndighetene. Dette innebærer at det eksisterer spesifikke, obligatoriske regler og retningslinjer som styrer markedet, inkludert reguleringer knyttet til landbruk og beskatning. Disse faktorene, i tillegg til de tidligere nevnte, er avgjørende for en effektiv og presis anvendelse av den hedoniske prismodellen i boligmarkedsstudier. Det har imidlertid vist seg å være utfordrende å utføre analyser på boligmarkedet under de gitte forutsetningene ovenfor. Osland (2001, s. 3) refererer til Alonso (1964), som fremmer forenklete forutsetninger for boligmarkedsmodellen:

1. *Et homogent område, med kun ett bysenter*: Regionen antas å være homogen med bare ett sentralt punkt, som et bysentrum. Her påvirker kun avstanden til sentrum boligprisene, siden geografiske forskjeller ikke eksisterer innenfor regionen.
2. *All aktivitet konsentrert i bysenteret*: Alle viktige økonomiske, sosiale og kulturelle aktiviteter er samlet i bysentrum. Dette gjør sentrum til det mest tiltrekkende området for innbyggerne og kan påvirke boligpreferansene.
3. *Avstand til sentrum korrelerer med økt reisetid*: Jo lenger unna sentrum boligen er, desto lenger blir reisetiden. Boliger nærmere sentrum kan være mer attraktive på grunn av kortere reisetid

og derfor høyere priset.

4. *Boliger betraktes som et homogent gode, fremfor et heterogent:* Boliger antas å være like og utbyttbare, selv om de i realiteten varierer i størrelse, kvalitet, beliggenhet og lignende, noe som gjør dem til et heterogent gode.
5. *Det eksisterer et fritt marked for boligkjøp:* Det antas ingen restriksjoner på kjøp eller salg av boliger, og priser bestemmes av markedets tilbud og etterspørsel. Dette overser potensielle reguleringer og restriksjoner som kan påvirke boligmarkedet.

Den hedoniske teoriens hovedmålsetting er å undersøke hvordan prisfunksjonen for et heterogent gode kan forklares som et resultat av interaksjonen mellom tilbydere og etterspørere i markedet. Denne interaksjonen er unik i forhold til tradisjonelle analyser av tilbud og etterspørsel, da den hedoniske prisfunksjonen hvor Osland (2001, s. 3) viser en omhylling av etterspørernes «budfunksjoner» og tilbydernes «offerfunksjoner» for hvert enkelt attributt.

Rosen (1974) studerte hedonisk prisfunksjon for heterogene goder generelt, men hans modell kan krediteres som grunnlaget for videreutviklingen av modellen som i dag er anvendt innen forskning på boligmarkedet med hensyn til hedonisk prisfunksjon. Denne modellen bygd på Rosens antagelser er da følgende:

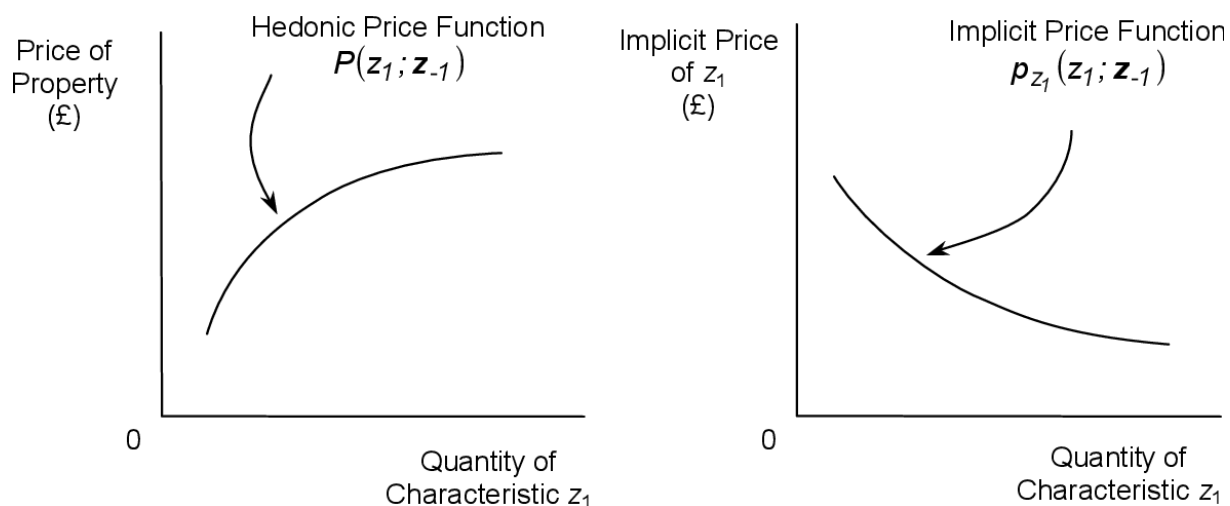
1. *En stor varebeholdning (stock)*, hvor varen er tilgjengelige boliger på markedet.
2. *Markedet er befolket av mange små aktører*, som hverken individuelt eller kollektivt har noen innflytelse på markedsforholdene eller prisene i markedet.
3. *Ingen markedsfriksjoner.* Det betyr at alle kostnader knyttet til transaksjoner, flyttin- og søking er lik null, noe som bidrar til et problemfritt og effektivt marked.
4. Det er *perfekt informasjonstilgang* på pris og kvalitet hos begge sider av markedet, både for kjøpere og selgere. Dette reduserer informasjonsasymmetri.
5. *Konsumentene kan ikke skille attributter fra det heterogene godet*, i dette tilfellet boliger. Det vil si at husholdningene ikke har mulighet til å skille ut enkelte attributter fra det heterogene godet (boligen).
6. *Utvalget av boliger på kort sikt er bestemt*, som betyr at kjøpere kan kun kjøpe boliger som er tilgjengelige på markedet.
7. *Marginalprisene* på boliger er dynamiske og endrer seg over tid.
8. *Marginalprisen* på en attributt kan avhenge av mengden av andre attributter (Day 2001, s. 22).
9. Vi ser bort fra den tradisjonelle supply-and-demand -modellen med tilbud og etterspørsel. I stedet tar denne modellen utgangspunkt i bud- og tilbudsfunksjoner av boligkarakteristikker.



Osland (2001, s. 2) viser i sin artikkel at den hedoniske prisfunksjonen som blir presentert av Rosen (1974) er en likevektsfunksjon i et marked, som er omringet/omhyllet av mange bud-, og tilbudskurver fra kunder og selger (Quigley, 1982).

Figur 4 illustrerer en hedonisk prisfunksjon og dens implisitte pris, ved et gitt boligkarakteristikk:

**Figur 4.** Den hedoniske prisen og implisitte prisskjema for attributt  $z_1$



Fra The theory of hedonic markets: obtaining welfare measures for changes in environmental quality using hedonic market data (s. 9), av B. H., Day, 2001. Centre for Social and Economic Research on the Global Environment (CSERGE): London, UK.

### 3.2.1 Likevekt på etterspørselssiden av markedet

Husholdningene tilpasser seg for å maksimere nytte under en ikke-lineær budsjettbegrensning, som kan sees på etterspørselssiden (Rosen, 1974, s. 38-41).

Maksimer nytten:

$$U_j = (Z, X, \alpha_j)$$

gitt budsjettrestriksjonen

$$Y_j = X + P(Z).$$

- $U$  er nytten for enhver husholdning  $j$  som er definert av:
- $Z$  som representerer en vektor av attributter for boligen
- $X$  som representerer en vektor av alle konsumvarer som ikke er boliger (annet konsum)

- Prisen for  $X$  er fastsatt til 1.
- $Y_j$  som betegner inntekten målt i enheter av  $X$  for husholdning  $j$ .
- $\alpha_j$  som er en vektor av parametere som beskriver preferansene.

Hver husholdning kjøper kun én bolig, som betraktes som et konsumgode. Nyttefunksjonen  $U_j$  er antatt å være strengt konkav. Denne teorien antar at både første- og andreordensderiverte av prisfunksjonen  $P(Z)$  finnes, men har ikke et bestemt fortegn. Det antas også at den har en form som sikrer en unik løsning for nyttemaksimeringsproblemet. Denne teorien forutsetter også at individene er homo economicus, eller rasjonelle aktører som vil velge det alternativet som gir dem mest nytte. Homo economicus er ifølge Mill referert i Persky (1995, s. 231) den perfekte rasjonelle økonomiske mannen som har fullstendig informasjon om konsekvensene av sine handlinger, og velger alltid handlinger som gir maksimerer økonomisk fordel eller nytte. Persky (1995, s. 231) poengterer at Mill definerer homo economicus som en person som utelukkende ønsker å eie rikdom og som er i stand til å bedømme den sammenlignende effektiviteten til midler for å oppnå det målet.

Inntekten til husholdning  $j$ , representert ved  $Y$ , tilsvarer totalbeløpet som brukes på konsumvarer utenom fritidsbolig ( $X$ ) samt kostnaden for fritidsbolig  $P(Z)$ . Ved å sette den marginale substitusjonsraten mellom  $Z_i$  og  $X$  lik den partiellderiverte av prisfunksjonen med hensyn til boligattributtene, kan man finne punktet som representerer maksimal nytte for en husholdning:

$$\frac{\partial P}{\partial Z_i} = \frac{\partial U_j}{\partial Z_i} / \frac{\partial U_j}{\partial X}$$

hvor  $i = 1, \dots, n$ . Punktet der nytten for husholdningen er høyest kalles også et *optimum*.

Helningen av prisfunksjonen på dette punktet, som er gitt ved optimal mengde, er videre interessant å se på.

Med dette ser vi at de implisitte (hedoniske) prisene for hver boligattributt  $Z_i$ . Dette er hvordan endringen i salgsprisen på boligen er påvirket av endringen i mengden av boligattributt  $Z_i$ , for eksempel hvordan en høyere energikarakter påvirker salgsprisen på boligen. Oppgavens mål er å estimere disse implisitte prisene.

### Budfunksjonen

Budfunksjonen representerer tilbudene fra kunder og etterspørselen, og er viktig for å kunne forklare en likevekt i markedet for heterogene goder. Funksjonen blir forklart gjennom hvilken

betalingsvillighet kunder har for de ulike boligene som er på markedet, samt attributtene til boligene, når nytten og inntekten til konsumenten er konstant. Funksjonen er nøkkelen til å forstå hvordan konsumentene verdsetter forskjellige attributter i heterogene gode, som boliger. Budfunksjonen definerer maksimal betalingsvillighet for samlinger av attributter ( $Z$ ). Med bakgrunn i dette, er budrentekurven altså en indifferenskurve. Dette kan brukes til å studere sammenhengen mellom flere ulike mulige kombinasjoner av attributter ved en eller flere boliger, og de subjektive prisene i tillegg til markedsprisen (Osland, 2001, s. 9). En attributtvektor er summen av alle attributtene som inngår i godet. En spesifikk bolig vil med andre ord ha en attributtvektor som tilsvarer summen av alle egenskapene eller attributtene som boligen besitter. Budfunksjonen kan altså beskrives som en funksjon av de ulike attributtene i boligen, og hvordan konsumentene verdsetter disse. Uttrykket for maksimal betalingsvillighet, når nytte og inntekt er konstant, er gitt ved:

$$\theta_j = \theta(Z, Y_j, U_j, \alpha),$$

hvor  $\theta$  representerer den maksimale betalingsvilligheten til konsumenten. Rosen (1974) introduserer også en *marginal budfunksjon*. Denne funksjonen viser den marginale betalingsviljen for en økning i mengden av et gitt attributt i boligen. Dette gir en mer detaljert forståelse av hvordan konsumentene verdsetter ulike attributter i boliger. Med en marginal budfunksjon, kan man beregne hva konsumentene er villig til å betale for en endring eller økning i energikarakter. Ved å sammenligne den vanlige budfunksjonen med den marginale budfunksjonen, kan man forstå hvordan konsumentene verdsetter ulike attributter i et gitt gode og hva de er villige til å betale for endringer i disse attributtene. Dette gir en mer detaljert forståelse av hvordan prisene på heterogene gode som boliger, formes av konsumentenes preferanser og betalingsvilje. Det er viktig å påpeke at vi kun estimerer de implisitte prisene for de boligattributtene, og ikke tar hensyn til absolutte priser. Den marginale budfunksjonen er gitt ved:

$$\frac{\partial \theta_j}{\partial Z_i} > 0$$

som ifølge Osland (2001, s. 5) kan tolkes som maksimal betalingsvillighet for en partiell økning i et boligattributt. Hvis nyttefunksjonen er strengt konkav kan det vises at

$$\frac{\partial^2 \theta_j}{Z_i^2} < 0$$

(Rothenberg et al. 1991, sitert i Osland, 2001, s. 5). Dette betyr at betalingsvilligheten øker med økende boligattributt, men til en avtagende grad. Med andre ord, jo mer vi øker attributtet, jo

mindre økning vil vi se i betalingsvilligheten til potensielle kjøpere.

### 3.2.2 Likevekt på tilbudssiden av markedet

Teorien beskriver hvordan mange små bedrifter konkurrerer om å produsere boliger som passer kundene best, på en profittmaksimerende måte, og hvordan dette påvirker prisene.

Profittfunksjonen for hvert enkelt produsent defineres som

$$\pi = M \cdot P(Z) - C(M; Z; \beta),$$

hvor  $M$  representerer et firmas tilbud av hus (antall hus) som tilsvarer en spesifikk attributvektor  $Z$ . Inntektsfunksjonen er ikke-lineær og definert av antall hus  $M$  multiplisert med den eksogene gitte hedoniske prisfunksjonen  $P(Z)$ .  $C$  er kostnadsfunksjonen som øker konvekst med antall hus  $M$ , og  $\beta$  er en vektor av skiftparametere som representerer for eksempel faktorpriser eller produksjonsteknologi for et enkelt firma.

Bedriftene er heterogene og varierer i sine kostnader. Noen har lavere kostnader når de produserer trehus, andre når de produserer eneboliger, og igjen andre når de produserer flermannsboliger osv. Dette betyr at bedriftene velger å produsere de boligtypene som gir dem mest profitt, noe som påvirker tilbudet av boliger på markedet.

Likevekt på tilbudssiden av markedet oppnås når det ikke er noen bedrifter som kan øke profitten ved å øke eller redusere produksjonen. Ved likevekt vil alle bedriftene ha nådd et punkt hvor deres profitt er maksimal, og tilbudet av boliger vil være tilpasset etterspørselen. Prisene på boliger vil dermed være bestemt av kundenes preferanser og deres evne til å betale for boliger med de egenskapene de ønsker.

Rosen (1974) sin hedoniske pristeori gir en god forståelse av hvordan boligpriser er påvirket av deres individuelle egenskaper, og hvordan konkurransen mellom bedrifter påvirker tilbudet av boliger. Teorien gir også en forståelse av hvordan likevekt oppnås på tilbudssiden av markedet.

### Profittmaksimering

$$\frac{\partial P}{\partial Z_i} = \frac{\partial C}{\partial Z_i} / M,$$

hvor  $i = 1, \dots, n$  viser at for å oppnå maksimal profitt, bør bedriftene velge en sammensetning av

boligattributter som gjør at den implisitte prisen for en gitt attributt er lik marginalkostnadene per hus ved en delvis økning i mengden boligattributter.

$$P(Z) = \frac{\partial C}{\partial M}$$

viser at for å oppnå likevekt, bør bedriftene produsere et antall boliger slik at marginalinntekten, gitt av prisen på boligen, er lik marginalkostnadene ved produksjon av en bolig til.

### Offerfunksjonen

Rosen (1974) demonstrerte at den hedoniske prisfunksjonen representerer en likevektsforhold ved å være en omhylling av både konsumenters *budfunksjoner* og produsenters *offerfunksjoner* (Quigley, 1982), som nevnt i Osland (2001, s. 2). Med bakgrunn i dette, er begrepet offerfunksjonen en sentral komponent i denne teorien. Den beskriver den laveste prisen man må selge for, slik at man oppnår et gitt profittnivå :  $\pi^*$ .

Selve profittfunksjonen kan uttrykkes ved (Osland, 2001):

$$\pi = M * P(Z) - C(M, Z, \beta).$$

Offerfunksjonen kan beregnes ved å løse for P (den laveste prisen man er villig til å selge for) på et gitt profittnivå, ved å bruke formelen:

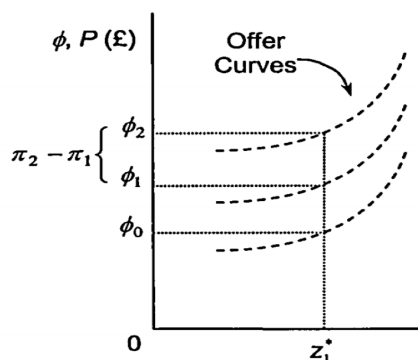
$$\Phi = \Phi (Z, \beta, \pi^*).$$

Ligningen presenterer her en funksjon som beskriver tilbudssiden i relasjonen mellom pris og karakteristikk for boliger. Den gir en indikasjon på den minste prisen produsentene er villige til å akseptere for å tilby boliger med et gitt sett med karakteristikk ( $Z$ ), et konstant profittnivå ( $\pi^*$ ) og et optimalt antall boliger som skal produseres.

Teorien sier at fortjenesten er konstant langs den stiplede offerkurven, og jo høyere offerkurve, jo høyere profitt. Dette betyr at jo høyere kostnader ved å produsere et produkt, jo høyere pris vil bedriften være villig til å akseptere for det.

Bedrifter maksimerer profitten når offerkurven tangerer priskurven. Dette er minimumsprisen firmaet er villig til å akseptere for å tilby et produkt med denne mengden av attributtene. For eksempel, hvis en bedrift produserer boliger, vil de være villige til å akseptere en lavere pris for en bolig med lavere energikarakter.

Figur 5. Landeiers offerkurver.

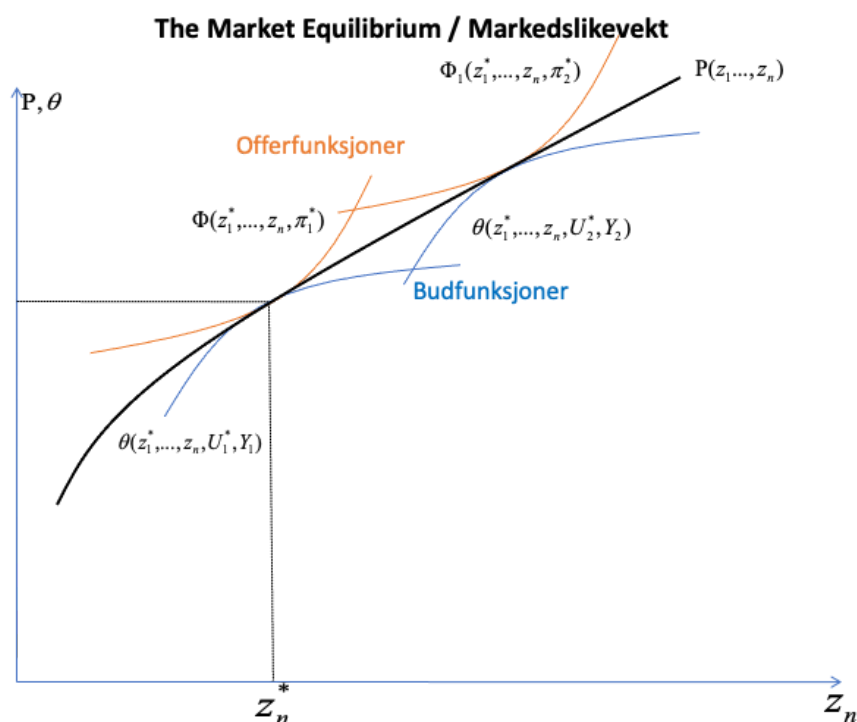


Fra The theory of hedonic markets: Obtaining welfare measures for changes in environmental quality using hedonic market data. (s. 42), av Day, B. H., 2001. Centre for Social and Economic Research on the Global Environment (CSERGE): London, UK.

### 3.2.3 Markedslikevekt

Rosen (1974) etablerte i sin hedoniske pristeori at prisfunksjonen formes av både kjøp- og salgskurver. Dette møtet mellom kjøper og selger på markedet resulterer i en hedonisk priskurve, som illustrerer både salgstilbudet og kjøpervilligheten. Kurvenes skjæringspunkt representerer markedets likevekt. Teorien foreslår videre at denne prisfunksjonen er en *likevektsligning*, formet av optimalisering fra alle markedsaktører. Ved interaksjonen mellom kjøper og selger blir en akseptabel pris bestemt, basert på husholdningenes maksimale betalingsvillighet og selgers tilbud, som etablerer markedslikevekt.

Figur 6. Markedslikevekt.



Bearbeidet fra Den hedonistiske metoden og estimering av attributtpriser , av Osland, L., 2001, s. 10.

*Husholdninger* maksimerer nytten ved å finne den lavest oppnåelige *budkurven* som tangerer den gitte hedoniske prisfunksjonen. Dette betyr at husholdningene ønsker å kjøpe et gitt produkt til lavest mulig pris, og at å by på et annet produkt med andre egenskaper vil ikke øke nytten.

*Produsenter*, derimot, optimaliserer fortjeneste ved å finne høyeste *offerkurve* som berører samme prisfunksjon, og streber etter å selge produkter til maksimal pris, der tillegg av andre egenskaper gir lavere fortjeneste. Når disse aktørene er koblet på bestemte punkter i prisfunksjonen, oppnår vi *likevekt*, hvor ingen forbedrer sin posisjon ved å velge ulike bunter med attributter, og prisfunksjonen rydder markedet.

Variasjoner i befolkning, inntekt, kostnader, attributter kan endre likevektsprisfunksjonen. For eksempel, befolkningsøkning kan øke etterspørsel og pris, akkurat som høyere produksjonskostnader kan drive opp produktprisen. Dermed er likevektsprisfunksjonen dynamisk og tilpasses markedssituasjonen. For å detektere brudd på grunn av uventede økonomiske sjokk i vår dataserie, bruker vi Chow-testen (Chow, 1960; Hill, Griffiths og Judge, 2001).

Med bakgrunn i kapittel 3.2, kan vi oppsummere ved at den hedoniske pristeorien hjelper oss med å forstå hvordan priser på boliger og andre heterogene varer formes av konsumentpreferanser og

betalingsvilje, og hvordan disse faktorene påvirker markedslivekten. En av de viktigste implikasjonene er at priser på heterogene varer, som boliger, ikke bare forklares av tradisjonelle tilbud og etterspørsel modeller. Den hedoniske prisfunksjonen baserer seg på ideen om at det er de spesifikke egenskapene til et gode som bestemmer prisen. Boligkarakteristikker er ofte referert til i boligmarkedet, der likevektsfunksjonen er basert på tilbud- og offerfunksjoner, avhengig av boligkarakteristikker og kjøpers budsjett. Derfor kan endringer i etterspørsel etter boliger påvirke prisene på boliger med bestemte egenskaper mer enn andre.

### 3.3 Geografisk avgrensning

I dette kapitlet vil vi redegjøre for den geografiske avgrensningen for vår økonometriske analyse av boligprisene i Bergen. I del 2.0 Bakgrunn er det generell informasjon om beliggenhet og befolkningstall i vårt fokusområde. I denne delen vil vi diskutere hvordan vi har valgt ut de geografiske områdene som skal inkluderes i analysen.

Juvkam (2002, s. 24) definerer en bo- og arbeidsmarkedsregion som en funksjonell region som samhandler internt, typisk mellom et senter og et omegn. Da regionen kan være begrenset av omegns tilknytning til senteret, i tillegg kan det være geografiske begrensninger, Juvkam ser derimot på et område som kan bestemme regionens størrelse, som er pendling, dvs. mobiliteten til innbyggerne og reisetid. Videre er det hensiktsmessig å ta hensyn til reisetid og reiseavstand, når en bruker pendling som et utgangspunkt for inndeling av region (Juvkam, 2002, s. 24- 26). SSBs interne notat (Byfuglien & Holm, 1988, s. 3) viser til en pendling tidsramme mellom kommuner på inntil 45 minutter, som Juvkam (2002, s.29) også refererer til, men viser til at pendling er også avhengig av tilknytningen mellom omegn og sentrum.

Juvkam (2002, s. 33-34) refererer til boken til Fosse og Selstad fra 1997 for NIBR11, som omhandler inndelingen av innbyggere i ulike regioner. Denne inndelingen er senere revidert i NIBR-rapporten fra 2013 (Juvkam, 2013, s. 19).

I denne oppdaterte rapporten blir kommunene gruppert i seks kategorier, rangert fra én til seks, på grunnlag av innbyggertallet; fra null til over 50 000. I både NIBR-rapportene fra 2002 og 2013 observerer vi en inndeling av bo- og arbeidsregion Bergen som inkluderer det vi i kapittel 1.1 omtaler som Bergensregionen. Disse kommunene er valgt på bakgrunn av deres størrelse, geografisk nærhet til hverandre, deres funksjonelle forhold i form av pendling og arbeidsmarked, samt deres felles historie som deler av den tidligere definerte bo- og arbeidsregionen Bergen.



### 3.4 Det teoretiske grunnlaget for regresjonsanalyse

Dette kapitlet konsentrerer seg om regresjonsanalyse, en sentral komponent i vår masteroppgave, definert av Johannessen et al. (2020, s. 333) som en metode for å undersøke endringer i en avhengig variabels verdier i forhold til en eller flere uavhengige variabler

For å kunne utvikle en generell formel, anta  $\beta_0$  som verdien til den avhengige variabelen når den uavhengige variabelen har verdien 0, og  $\beta_1$  som hvor mye den avhengige variabelen endrer seg når den uavhengige variabelen øker med én verdi.  $u$  representerer et feilledd, som tar høyde for det som ikke er inkludert i modellen. Den generelle formen for likningen blir:

$$P = \beta_0 + \beta_1 X + u.$$

#### 3.4.1 Regresjonsanalyse: Anvendelse og tolkning

Vi tar utgangspunkt i OLS (ordinary least squares) for å finne den lineære regresjonsmodellen som minimerer summen av kvadratene av residualene, det vil si det minste kvadrerte avviket mellom de faktiske verdiene og de estimerte verdiene. Dette gjør det mulig å estimere koeffisientene i regresjonsmodellen og å predike verdier av avhengige variabler ut fra verdiene av uavhengige variabler.

Vår studie benytter seg av en multiple linear regression (MLR)-modell:

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \varepsilon_i,$$

som analyserer flere uavhengige variabler, i motsetning til den enklere 'simple linear regression' (SLR)-modellen som begrenser analysen til to variabler. MLR gir en dypere forståelse av den avhengige variabelen, og tillater kontroll av flere påvirkende faktorer. Dette bidrar til å forklare variabiliteten i den avhengige variabelen og skaper en mer robust modell (Wooldridge, 2019, s. 66). Fem antakelser må hensyntas ved en MLR-modell, ifølge Wooldridge (2019, s. 80-88), gjengitt i appendiks A5.

*Heteroskedastisitet* og *homoskedastisitet* beskriver variansen i en økonometrisk modell. Heteroskedastisitet oppstår når variansen til feilleddet varierer (Wooldridge, 2009, s. 53), mens

homoskedastisitet viser til situasjonen der variansen er konstant. Heteroskedastisitet innebærer at variansen av avvikene fra modellen ikke er konstant, noe som kan påvirke standardfeilene til de estimerte parametrene og dermed redusere påliteligheten av hypotesetesting for disse parameterne. Dersom viktige variabler er utelatt og de korrelerer med det som studeres, kan heteroskedastisitet i verste fall gi forventningsskjevne resultater.

*Multikollinearitet* defineres av Wooldridge (2019, s. 90) som en høy, men ikke perfekt korrelasjon mellom to eller flere uavhengige variabler. Dette betyr at en stor del av variasjonen i en uavhengig variabel kan forklares av de andre uavhengige variablene (Wooldridge, 2019, s. 90). En høy grad av korrelasjon eller linearitet mellom variablene kan føre til store variasjoner i estimeringen av helningen for OLS-kurven, altså en stor  $Var(\hat{\beta}_j)$ . (Wooldridge, 2019, s. 90-91). Multikollinearitet bryter imidlertid ikke med noen av forutsetningene for MLR i appendiks A5, og trenger ikke å være et problem i seg selv (Wooldridge, 2019, s. 90-91). Når en skal teste for multikollinearitet kan en gjennomføre en VIF-test (Variance Inflation Factor-test), som ser på statistikk for de individuelle koeffisientene. En VIF-verdi over 10 regnes ofte som en terskelverdi for når multikollinearitet kan ansees som et problem (Wooldridge, 2019, s. 92).

Modellene presentert hittil antar implisitt uavhengighet mellom forskjellige enheters utfall, uten rom for spillovereffekter. Dette kan være en betydelig begrensning, særlig når man arbeider med romlige data, ifølge Elhorst og Vega (2015, s. 344) slik som vi vil gjøre i denne utredningen.

## **Funksjonell form og logistiske regresjoner**

For å kunne stole på at en regresjonsanalyse gir korrekte resultater, er det nødvendig at visse forutsetninger er oppfylt. Når det kommer til logistisk regresjon, er de mest sentrale forutsetningene at forholdet mellom variablene har riktig funksjonell form, at individuelle observasjoner er uavhengige av hverandre - det vil si at feilleddsvariasjonen er uavhengig - og at de uavhengige variablene ikke er korrelert med hverandre. Dette innebærer at det ikke bør finnes underliggende variabler i analysen som både påvirker den avhengige variabelen og samtidig er korrelert med en annen uavhengig variabel (Skog, 2004, s. 379-397; Eikemo & Clausen, 2012, s. 145).

Valget av funksjonell form i hedonisk prisfunksjon for logistisk regresjon er avgjørende, men uten teoretisk veiledning. Valget er derfor et empirisk spørsmål som krever grundig dataanalyse. Den valgte formen er sannsynligvis ustabil over tid og kan ikke forventes å dekke mer enn ett enkelt

markedsområde. Den funksjonelle formen er sannsynligvis ikke lineær, og den mest typiske formen er en kombinasjon av log-log og log-lin funksjoner. For eksempel kan alle kontinuerlige variabler være i naturlig logaritmisk form (ln-form), mens dummyvariabler opprettholdes i lineær form. Dette innebærer at det er avgjørende å være oppmerksom på hvordan variablene transformeres og integreres i modellen for å sikre en så korrekt og informativ analyse som mulig. En log-lineær modell kan uttrykkes som følger:

$$\ln(y_i) = \beta_0 + \beta_1 x_1 + \varepsilon_i,$$

og tolkes som at 1 enhet økning i  $x$  fører til  $100 \cdot \beta_1$  økning i  $y$ , mens en log-log-modell,

$$\ln(y_i) = \beta_0 + \beta_1 \ln(x_1) + \varepsilon_i,$$

kan tolkes som at 1 enhet økning i  $x$  fører til  $\beta_1$  % økning i  $y$ .

### 3.4.2 Statistiske verktøy for tolkning av resultatene

Dette underkapitlet fokuserer på de statistiske verktøyene som er anvendt for å tolke resultatene fra vår analyse.

*Korrelasjonskoeffisienten*, en indikator på lineær avhengighet variabel mellom -1 og 1, reflekterer variablenes samvariasjon (Wooldridge, 2019, s. 699). Korrelasjon måles typisk ved Pearsons produkt-moment korrelasjonskoeffisient (Johannessen et al., 2020, s. 319). Positiv korrelasjon indikerer at en økning i en variabel medfører en tilsvarende økning i den andre, og vice versa for negativ korrelasjon (Fugleberg & Kristianslund, 1995, s. 22). Denne gir innsikt i graden av lineær avhengighet mellom to variabler.

En *t-test* benyttes for å avgjøre signifikant forskjell mellom gjennomsnittet av to grupper med potensielt felles egenskaper. Man observerer t-statistikk, t-fordelingsverdier og frihetsgrader for å bestemme statistisk signifikans. T-testen kan enten være dobbeltsidig eller enkeltsidig, basert på om man søker forskjeller i spesifikke retninger eller mulighet for positive og negative forskjeller. Resultatene uttrykkes som t-verdier, som sammenlignes med kritiske grenseverdier for å bestemme signifikans (Kritzman, 1994, s. 20).

*P-verdier* gir informasjon om sannsynligheten for å observere resultater som er like eller mer ekstreme enn de vi faktisk observerte, antatt at nullhypotesen er sann. Nullhypotesen representerer den grunnleggende antagelsen om at ingen effekt eller forskjell finnes i dataene. P-verdier blir ofte brukt til å avgjøre om resultatene fra en statistisk test er *signifikante* eller *ikke*. Statistisk signifikans

angir hvilket signifikansnivå man kan bruke for å forkaste nullhypotesen vår (Wooldridge, 2009, s. 128). Dersom p-verdien er lav (vanligvis  $> 0,05$ ), anses nullhypotesen som forkastelig. Ulempene er at de ikke gir en absolutt sannsynlighet for at resultatene er korrekte, ifølge Greenland et al. (2016). De gir også ikke informasjon om størrelsen på effekten eller forskjellen vi undersøker. Bruk av p-verdier i kombinasjon med andre metoder, som konfidensintervaller, gir en mer omfattende resultatanalyse. Greenland et al. (2016) påpeker en ofte begått feil; å redusere p-verdier til en binær klassifisering av 'signifikant' eller 'ikke-signifikant', noe som overser den kontinuerlige naturen og kan føre til misforståelser eller feiltolkninger av resultater.

### 3.5 Romlig økonometri

Konseptet *romlig økonometri* ble ifølge Anselin (1988, s. 7) først introdusert av den belgiske økonomen Jean Paelinck i artikkelen *Spatial Econometrics* fra 1978. Videre definerer Luc Anselin romlig økonometri som et forskningsfelt som fokuserer på romlige aspekter av data, som krever en annen tilnærming enn standard økonometriske metoder. Romlige data har ifølge (Chi & Zhu, 2019, s. 5) en form for geografisk tilknytning. Når det gjelder romlig avhengighet, antar man at den avhengige variabelen  $Y$  i en prisfunksjon kan være avhengig av andre  $Y$ -verdier i nærheten. Ved tilstedeværelse av romlig *heterogenitet* eller *autokorrelasjon*, er det residualene som påvirkes av residualene i nærheten. Disse avhengige variablene kan også påvirkes av nærliggende faktorer, som kalles en *spillover-effekt*. Eksempelvis kan en bolig med høy energimerking kan ha en positiv effekt på boligprisene i nærheten, da det kan øke verdien av eller øke etterspørselen etter boliger i området.

Dette er ifølge Elhorst og Vega (2017) ett av to hovedinteresser i litteraturen om romlig økonomi og romlig økonometri. Den andre interessen er såkalte *spatial lags*. Et spatial lag måler effekten av den avhengige variabelen ( $Y$ ), de forklarende variablene ( $X$ ) eller feilledet ( $u$ ) observert i andre tverrsnittsenheter  $j$  enn enhet  $i$  på den avhengige variabelen i enhet  $i$ . En romlig spillover-effekt er definert som den marginale effekten av en endring i en forklarende variabel i en bestemt tverrsnittsenhet på de avhengige variabelverdiene i en annen enhet, og er avledet fra den reduserte formen av en romlig økonometrisk modell.

#### 3.5.1 Spillover-effekter

Brueckner (2003) forklarer at spillover-effekter oppstår når en persons valg, som for eksempel

vedlikehold av en bolig, ikke bare påvirker nytten for den enkelte, men også for nabolaget. Dette innebærer at et velholdt boligområde kan gi fordeler som går utover den enkelte beboers nytte, for eksempel ved å øke velvære og heve boligprisene i området. På den andre siden kan etablering av et nytt vannkraftverk i et boligområde resultere i økt støy, noe som kan påvirke både velvære og salgspriser på boliger negativt. Anselin (1988, s. 8) påpeker generelt at effekten avtar med økende geografisk avstand, det vil si at objekter som er nærmere hverandre har en tendens til å påvirke hverandre mer enn de som er lenger unna.

LeSage (2014) definerer to hovedkategorier av spillover-effekter observert i økonometriske analyser: *globale* og *lokale* spillovers.

*Globale romlige spillovers* karakteriseres av deres omfattende effekter, der en endring ved en gitt sted (i) resulterer i omfattende endringer i hele regionen eller markedet, som igjen har en tilbakevirkende effekt på utgangspunktet (i). Det som skjer i Bergen smitter over på alle andre steder og vil gi spillover-effekter til Bergen i retur. Andre eksempler kan være endringer i regjeringens energi- og klimapolitikk, økninger i energipriser, eller globale trender mot økt miljøbevissthet, som alle kan påvirke interessen for miljøvennlige boliger.

*Lokale spillover-effekter*, derimot, er geografisk begrensede i sin innflytelse. I disse tilfellene medfører en handling på et bestemt sted endringer i verdier eller nytte i umiddelbar nærhet, men uten å stråle ut til større geografiske områder. For eksempel, en bolig i Bergen med høy energimerking kan potensielt ha en positiv innvirkning på nærliggende boligpriser. Dette kan være et resultat av en økt sannsynlighet for høyere salgspris, reduserte energiutgifter, og/eller forbedret inneklima. Spillover-effekten begrenset til lokalområdet - det som skjer i Bergen, forblir i Bergen og påvirker kun nærliggende områder.

Budskapet fra LeSage (2014) er at de fleste romlige spillover-effekter er lokale. En formell definisjon av romlige spillovers, som gitt av LeSage (2014, s. 3), er:

$$\frac{\partial y_i}{\partial x_j} \neq 0 \quad (i \neq j) \text{ hvor } i \text{ og } j \text{ er ulike lokasjoner}$$

### **Romlige virkninger**

Når romlige prosesser inkluderes i hedoniske modeller, er det ikke lenger mulig å redusere de implisitte prisene for miljøattributter til de estimerte verdiene av koeffisienten  $\beta_k$  (Maslianskaia-Pautrel og Baumont, 2016). Den implisitte prisen for miljøattributter legger til to

effekter: en *direkte* effekt knyttet til observasjon  $i$  og en *indirekte* effekt som kommer fra nærliggende observasjoner ( $j$ ). Dette involverer en romlig diffusjonsprosess som kommer fra naboer (*lokale spillovere*) og kan til og med bli styrket av en romlig multiplikatoreffekt som kommer fra tilbakekoblingseffekter på alle observasjonene som er spredt over området (*globale spillovere*). Totalverdien av den implisitte prisen beregnes både med de estimerte verdiene av forklarende parametere og de estimerte verdiene av romlige parametere. Ettersom effekten av endringer i en forklaringsvariabel varierer mellom alle regioner, introduserer Pace og LeSage (2006) sitert i (LeSage, 2008, s. 34) et sammendragsmål som fanger opp disse ulike påvirkningene og som kan benyttes for å beregne gjennomsnittseffekten av disse endringene for samtlige institusjoner.

Virkningene kan være positive eller negative og kan være av varierende styrke og varighet. De defineres som hvor mye et utfall ( $Y$ ) endres ved en gitt liten endring i en variabel ( $x$ ).

Man kan skille mellom direkte, indirekte og total påvirkning.

Den direkte påvirkningen er definert som:

$$\frac{\partial y_i}{\partial x_i}$$

hvor utfallet/virkningen ( $y$ ) på lokasjon  $i$ , endres på grunn av en liten endring i en variabel ( $x$ ) på samme sted.

Den indirekte påvirkningen kan defineres ved å benytte den krysspartielle deriverte, angitt som

$$\frac{\partial y_i}{\partial x_j} (i \neq j).$$

I denne konteksten, utgjør endringer i ( $x$ ) på punkt ( $i$ ), som påvirker posisjoner utenfor ( $i$ ), et romlig 'spillover', og det eksisterer ( $n$ ) slike effekter, hvor ( $n$ ) tilsvarer antallet observasjoner. For å beregne de samlede gjennomsnittlige effektene (direkte, indirekte og total), summeres effektene over alle observasjoner, hvorpå summen divideres med ( $n$ ). Slik genererer vi en kvantifisert representasjon av de respektive gjennomsnittlige effektene.

$$\text{Average direct impact: } ADI = \frac{1}{n} \sum_{i=1}^n \frac{\partial y_i}{\partial x_i}$$

$$\text{Average total impact: } ATI = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial y_j}{\partial x_i}$$

$$\text{Indirect impact} = ATI - ADI$$

Ved romlig autokorrelasjon kan korreksjon oppnås ved å innlemme relevante variabler, eller alternativt gjennom romlig regresjonsanalyse med estimatorene som er justert for spatial effekter

(Florax og Van der Vlist, 2003). Avhengig av resultatene fra testen, vil vi vurdere en *spatial lag modell* (SAR) eller en *spatial error modell* (SEM) (Revelli, 2006).

### 3.5.2 k-nærmeste naboer (k-NN)

Når man utfører romlige analyser, er det viktig å definere nabolagsstrukturene mellom geografiske enheter. Nabolag kan beskrives som en samling av stedbaserte egenskaper forbundet med klynger av boliger, noen ganger i forbindelse med annen arealbruk (Galster 2003). Når disse stedbaserte egenskapene er til stede, oppfattes og verdsettes, reflekteres og kapitaliseres egenskapene i boligprisene.

K-nærmeste naboer (k-NN) er en teknikk for å bestemme *punktdatas* nabolagsstruktur (Chi og Zhu, 2008). Den identifiserer (k) nærmeste naboer for hver observasjon, der (k) er forhåndsbestemt, basert på beregnede avstander mellom punkter. Anselin (2002) påpeker at k-NN adresserer utfordringen med avstandsbaserte romlige vektmatriser som genererer for mange naboer i urbane og for få i landlige områder. Altman (1992) beskriver metoden som en enkel og populær ikke-parametrisk statistisk teknikk for å estimere funksjonsverdier basert på observasjoner i et datasett. Hovedideen bak k-NN er å finne et sett av k nærmeste naboer til et gitt datapunkt, og deretter estimere funksjonsverdien eller klassen for dette datapunktet basert på naboene. Altman (1992) forklarer hvordan man kan velge *k* i k-NN metoden, og hvordan valget av *k* påvirker modellens ytelse, herunder forventningsskjevhet og varians. Metoden baseres på prinsippet om at observasjoner nærmere hverandre i datasettet har mer likhet i deres avhengige variabler. For å utføre en k-NN-regresjon, velges et antall nærmeste naboer (k), og avhengige variabler for et nytt datapunkt estimeres ved å beregne et veid gjennomsnitt av avhengige variabler til de k nærmeste naboene i datasettet. Metoden ansees spesielt nyttig i romlige analyser der avstand er en relevant faktor, for eksempel i klyngedannelse, romlig interpolering eller romlig regresjon.

### 3.5.3 Forventningsskjev estimator

Å utelate relevante variabler er ifølge Mauro (1990, s. 315) den vanskeligste typen spesifikasjonsfeil å oppdage, og den med de mest alvorlige konsekvensene. Å utelate en relevant variabel vil øke anslagene for standardfeilene til de inkluderte variablene, noe som vil forårsake at testene av signifikans for disse prediktorer blir unøyaktige. Samtidig er det umulig å inkludere alle relevante variabler i en regresjonsligning. Som et resultat av dette er det uunngåelig å utelate relevante variabler (Clarke, 2005, s. 349). Når den utelatte variabelen er korrelert med de inkluderte

variablene, vil anslagene for regresjonskoeffisientene for disse variablene være skjeve.

Weisstein (1995) definerer en estimator som en matematisk prosedyre som anvendes for å beregne en ukjent parameter, som for eksempel effekten av en boligs energikarakter på dens pris. Dersom estimatoren systematisk avviker fra den sanne parameterverdien, er den *forventningsskjev*. En forventningsrett estimator sikrer derimot at estimatets gjennomsnittsverdi korresponderer med parameterens sanne verdi. Forventningsskjevhet kan oppstå dersom utvalget begrenses til boliger i et spesifikt område med en unormal fordeling av energikarakterer. Dette vil resultere i en estimator som ikke nøyaktig reflekterer den generelle sammenhengen mellom boligpriser og energikarakterer.

### 3.6 Spatial Lag Model (SAR)

Spatial lag modellen (SAR) er en økonometrisk metode som brukes til å undersøke sammenhengen mellom geografiske områder og deres økonomiske aktivitet (Anselin, 1988). Modellen hensyntar geografiske forhold som kan påvirke økonomien.

Anselin (1988) sin romlige modellvariant; en romlig lagmodell med *lag*  $Y$ :

$$Y = \rho WY + X\beta + u.$$

- $\rho$  er en romlig autoregressiv koeffisient
- $W$  er en  $n \times n$  romlig vektmatrise som definerer nabolagsstrukturen
- $u$  er en  $n \times 1$  vektor av feilledd
- De romlige spillovereffektene kommer fra  $\rho \cdot W \cdot Y$ .

SAR er hensiktsmessig når salgsprisen på et hus reagerer direkte på salgsprisen på nabohusene. Vi må også anta at spillovereffektene er globale når vi bruker denne modellen.

Merk at  $P$  i uttrykket  $P = \rho WP + X\beta + u$  vises på begge sider av likhetstegnet, og likningen må derfor løses for  $P$ . Beregningen av total påvirkning innebærer både  $X\beta$  og  $WP$ .

$$P = \rho WP + X\beta + u$$

$$P = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1} u$$

I SAR beregnes både den direkte og indirekte totale påvirkningen, og det kan vises at den gjennomsnittlige direkte effekten av modellen er:

$$\frac{\partial y_i}{\partial x_i} = (I - \rho W)^{-1} \beta$$



I den hedoniske prismodellen, er dette den implisitte prisen på karakteristikk  $x$ , forutsatt at vi har estimert en romlig lagmodell, men merk at ikke kan  $\beta$  tolkes slik at den gjenspeiler de vanlige lineære regresjonshellingsestimatene (f.eks. partielle effekter, eller implisitte priser som i den lineære hedoniske boligprismodellen). Når en bruker SAR-modellen må man bruke impacts-funksjonen for å hente ut de indirekte og direkte virkningene.

### 3.7 Spatial Error Model (SEM)

En romlig feilmodell kan benyttes dersom LM-tester indikerer at vi har romlig feilavhengighet i modellen. Dette kan oppstå dersom restleddet  $\varepsilon$  inkluderer utelatte ikke-signifikante variabler som er romlig avhengige (Brueckner, 2003).

$$Y = X\beta + u, \text{ hvor } u = \lambda \cdot Wu + \varepsilon$$

- Vi har lagt til et romlig lag ( $Wu$ ) av residualen/feilleddet  $\varepsilon$
- $W$  er en  $n \times n$  romlig vektmatrise som spesifiserer de definerte naboene.
- $u$  er en  $n \times 1$  vektor for feilleddet  $\varepsilon$ .
- $\varepsilon$  er en  $n \times 1$  vektor av feilleddet  $\varepsilon$  som er normalfordelt og uavhengig distribuert.

SEM bør vurderes man har romlig autokorrelasjon i residualene eller utelatte variabler som er romlig korrelert. Styrken av den romlige korrelasjonen i residualene er bestemt av størrelsen på  $\lambda$  (absoluttverdi). Høyere verdier av  $\lambda$  indikerer sterkere romlig autokorrelasjon i residualene.  $\lambda$ -verdien ligger mellom 0 og 1 (absoluttverdi), gitt en spesifikk spesifisering av  $W$  (radstandardisert) (LeSage, 1997).

SEM og OLS-modellen er nestet, og hvis  $\lambda = 0$ , får vi en OLS-modell ettersom SEM-modellen inneholder alle vilkårene i OLS-modellen i tillegg til det romlig laggede feilleddet.

Å benytte SEM er fordelaktig for å analysere boligpriser og geografiske faktorer. For det første kan det være romlige mønstre i feilleddet på grunn av utelatte tilfeldige faktorer som er romlig korrelert, noe som ofte refereres til som romlig autokorrelasjon som en *nuisance* - at autokorrelasjonen ikke kan tillegges en substansiell tolkning, i kontrast til autokorrelasjon relatert til den avhengige variabelen, som kan ha en mer substansiell tolkning. Boligpriser er sterkt påvirket av geografi, og det kan være utfordrende å hensynte alle relevante faktorer.

Til slutt er det viktig å ha en substansiell tolkning av resultatene, som forteller oss hva funnene sier om den virkelige verden, i stedet for bare å fokusere på regresjonskoeffisienter. Ved å bruke SEM kan man oppnå en bedre forståelse av hvordan geografiske faktorer påvirker boligpriser og vurdere de komplekse romlige mønstrene som kan være til stede.

Bruk av OLS i stedet for SEM, som egentlig ville vært korrekt modellvalg, vil undergrave tilliten til koeffisientestimatene. Selv om dette ikke nødvendigvis medfører en betydelig komplikasjon, indikerer betydelige avvik i koeffisientestimatene at noe kan være feilspesifisert i modellen.

### **3.8 Spatial Durbin Model (SDM)**

Omkring starten av 2000-tallet økte interessen for å inkludere romlige effekter (Lundberg, 2006). Tidligere romlige økonometriske modellene var hovedsakelig basert på én type romlig interaksjonseffekt, enten gjennom en romlig lagmodell eller en romlig feilmodell. Disse modellene inneholdt enten en romlig lag-variabel eller en romlig autoregressiv prosess i feilledet.

LeSage og Pace (2009) introduserte den romlige Durbin-modellen (SDM) som overgår tidligere modeller i flere aspekter. SDM begrenser ikke størrelsen på potensielle romlige spillovere, og kan generere ulike globale eller lokale effekter for forskjellige forklaringsvariabler. I tillegg produserer den forventningsrette koeffisientestimater selv om den datagenereringsprosessen er en romlig lag- eller feilmodell.

En romlig lagmodell (SLM, ofte SAR - spatial autoregressiv) innebærer en romlig prosess i responsvariabelen (LeSage og Pace, 2009). Durbin-modeller inkluderer romlig laggede uavhengige variabler til uavhengige variabler inkludert i den romlige modellen. SDM er egnet for å håndtere estimater påvirket av romlige effekter, og for å gi estimater for direkte og indirekte virkninger. Ifølge LeSage and Pace (2009) er SDM mer robust mot ulike romlige relaterte feilspesifiseringer. SDM kan avdekke romlige effekter og gi ytterligere informasjon via direkte og indirekte romlige spillovere. Denne tilnærmingen er motivert av tilstedeværelsen av signifikante uforklarte romlige effekter i residualene (Anselin, 2002).

Det er få studier på boligmarkedet som benytter SDM. Eksempler på studier er Brasington og Hite (2005), Osland (2010) og Fernandez-Aviles et al. (2012). Det har vært en generell mangel på forståelse for effektberegning, og resultatene har ofte blitt feiltolket (LeSage og Fischer, 2008; Fischer et al., 2009). Ifølge Elhorst (2010, s. 26), bør en moderne tilnærming til romlig økonometri også vurdere SDM. I nyere tid har det vært en oppsving i bruken av SDM, delvis grunnet

utviklingen av relevant programvare. Dette markerer et viktig skritt fremover i vår forståelse og anvendelse av romlig analyse i boligmarkedsstudier, og understreker betydningen av å teste SDM i vår metodiske tilnærming.

Elhorst (2010) gir en oversikt over de mest relevante romlige økonometriske modellene. Han argumenterer for at SDM er den eneste modellen som gir parameterestimer som ikke er forventningsskjev, og gir riktige standardfeil, selv når den virkelige datagenereringsprosessen er en av de andre nevnte romlige regresjonsmodellene, hvor alle parametrene er identifiserbare (Elhorst, 2010, s. 14). Dette er i samsvar med LeSage og Pace (2009) som viser at SDM fanger datagenereringsprosessen, selv når relevante romlige relaterte variabler utelates.

$$P = \rho WP + X\beta_0 + \rho WX\beta_1 + \varepsilon$$

- $\rho$  er en romlig autoregressiv koeffisient
- $P$  er en vektor av observasjoner av priser
- $X$  er en matrise av observasjoner av uavhengige variabler
- $W$  er en  $n \times n$  romlig vektmatrise som spesifiserer spesifisere antatt nabolagsstruktur for observasjonene.

Uttrykket reflekterer en hypotese om at modellen inkluderer en romlig lagging av den avhengige variabelen, i tillegg til en romlig lagging av alle uavhengige variabler. Romlig lagmodell inneholder en romlig lagging av den avhengige variabelen bare, og romlig errormodell inneholder en romlig lagging av feilleddet. Det er mulig å vise at SDM inkluderer de to mer vanlig brukte modellene, romlig lag- eller feilmodell (Osland, 2010).

I en SDM, representerer rho-parameteren graden av romlig avhengighet. Denne parameteren, mellom -1 og 1, reflekterer hvorvidt endringer i en avhengig variabel på en gitt plassering påvirker tilsvarende variabler i de omkringliggende stedene. En positiv rho indikerer positiv romlig autokorrelasjon, som impliserer at liknende verdier er geografisk konsentrert, mens en negativ rho indikerer ulike verdier har en tendens til å opptre i nærheten av hverandre.

### 3.8 Spatial lag of X-modellen (SLX)

Det romlige lagget til X-modellen er en lineær regresjonsmodell utvidet til å inkludere forklarende variabler observert på tverrsnittsenheter i nabolaget, og er den enkleste romlige økonometriske modellen som produserer fleksible romlige spillover-effekter.

Modellen har altså lokale spillovers, og inkluderer bare de stedene som er koblet i henhold til  $W$ . Dette er ofte den mest fornuftige antagelsen ifølge LeSage (2014).

$$Y = X\beta + WX\theta + u$$

Tolkningen av de estimerte koeffisientene er rett frem:

$$\text{Direkte virkninger: } \frac{\partial y_i}{\partial x_i} = \beta_k \qquad \text{Indirekte virkninger: } \frac{\partial y_i}{\partial x_i} = \theta$$

Det siste begrepet, krysspartiellderivert = indirekte innvirkning, er spillovers som faller på alle naboer (LeSage, 2014). Den totale effekten er summen av disse direkte og indirekte virkningene. I situasjoner hvor romlig avhengighet er svak, antyder LeSage (2014, s. 16) at den best tilpassede modellspesifikasjonen kan være SLX. Dette innebærer at hverken SDM eller Spatial Durbin Error Model (SDEM) nødvendigvis representerer den sanne modellen.

### 3.9 Testprosedyrer

Når man skal avgjøre hvilken av de lineære modellene som har mest forklaringskraft benyttes vanligvis  $R^2$ . Den gir et uttrykk på hvor stor del av variasjonen i den avhengige variabelen som kan forklares av de inkluderte uavhengige variablene, og varierer mellom 0-1. (Wooldridge, 2009).  $R^2$  Det er ingen minimumskrav for  $R^2$ -verdien, og det er viktig å presisere at lav  $R^2$  ikke nødvendigvis indikerer en svak modell.  $R^2$  øker alltid når flere uavhengige variabler inkluderes i modellen (Wooldridge, 2009). Derfor bør man være forsiktig med å tolke dette målet, men heller bruke justert  $R^2$  som korrigerer for mengden forklaringsvariabler (Wooldridge, 2009).

#### Goodness of fit

I tillegg til å vurdere  $R^2$ , kan man gjøre sammenlikning mellom den lineære modellspesifikasjonen (OLS), spatial lag modellen (SAR) og en modell som er satt opp med en spatial error tilnærming (SEM). Dette gjøres for å sammenlikne hvilken av de tre modellspesifikasjonene som har best forklaringskraft. For å sammenligne modellene benyttes verdier for blant annet log likelihood, Akaike informasjonskriterium (Akaike, 1974) og det nært beslektede Bayesianske informasjonskriterium (BIC) (Schwarz, 1978). De to sistnevnte mål er basert på likelihoodfunksjonen, men straffer et høyt antall parametere.

Loglikelihood er en statistisk teknikk som optimerer parameterne for å maksimere sannsynligheten for de observerte data (Millar, 2011). Høyere verdier indikerer bedre tilpasning. AIC-verdier, derimot, tar hensyn til antall estimerte parametre, og det anbefales å velge modellen med de laveste verdiene (Chi & Zhu, 2020, s. 66). Ifølge dem er AIC basert på likelihood-verdier og tar hensyn til kompleksiteten og straffer komplekse modeller. Dette tilsier at lavere AIC verdier representerer en bedre balanse mellom modelltilpasning og kompleksitet. Formelen uttrykkes som

$$AIC = 2 \cdot k - 2 \cdot \ln(\hat{L}),$$

hvor  $k$  = antall parametre. Dersom vi har mange parametre (høy  $k$ ), straffes man fordi modellen blir for kompleks og unøyaktig. Mens AIC straffer modeller kun basert på antall variabler, inkluderer BIC også utvalgsstørrelse (Christophersen, 2018, s. 112). På denne måten straffer BIC hardere enn AIC, og foretrekker vanligvis mindre modeller.

### **Morans's I og Lagrange-multiplikator (LM) tester**

Moran's I (Moran, 1948) er et korrelasjonsmål, der en ser på hvordan en enkelt variabel samvarierer i en geografi. Det er det mest brukte målet for **spatial autokorrelasjon** ifølge Chi og Zhu (2008). Moran's I på -1 indikerer negativ avhengighet, 0 en tilfeldig fordeling (uavhengighet), og +1 positiv romlig avhengighet (Haining, 2003, s. 75-80) Moran's I er nyttig for å oppdage direkte romlig diffusjon i en enkelt variabel, og enda mer nyttig som et diagnostisk verktøy. Ved å estimere Moran's I for residualene av en modell som inkluderer en romlig avhengig Y-variabel, er det mulig å bestemme om uavhengige variabler kan forklare den romlige avhengigheten i Y, og hvor mye som fortsatt gjenstår å bli forklart (Ward og Gleditsch, 2008).

Siden Moran's I testen ikke skiller mellom hva som er den mest sannsynlige kilden til spatial avhengighet så har Anselin m.fl (1996) foreslått at man i tillegg til Moran's I testen også burde benytte to tester basert på Lagrange Multiplikator (LM) prinsippet. Henholdsvis LM og robust LM test. Dette skal gi en klar indikasjon på om vi har en spatial error avhengighet, spatial lag avhengighet eller begge deler.

## 4.0 Metode

I denne seksjonen av undersøkelsen vil vi presentere metodiske valg som er gjort. Vi vil presentere det valgte forskningsdesignet og -metoden, i tillegg til å beskrive hvordan datainnsamlingsprosessen har foregått og hvordan dataene er behandlet etterpå.

### 4.1 Datagrunnlag i vår undersøkelse

For å analysere det norske eiendomsmarkedet, benyttes sekundærdata hentet fra flere kilder, og vi anvender en *kvantitativ* tilnærming. Disse dataene er opprinnelig samlet inn for andre formål, og representerer derfor informasjon som allerede eksisterer (Gripsrud et al., 2010, s. 19, 57-59). De er uttrykt gjennom tall og mengdeenheter, og kan kvantifiseres (Gripsrud et al., 2004). De dataene vi benytter utgjør ikke en totaltelling over alle solgte boliger, men er basert på salg som er formidlet av eiendomsmeglere og annonsert gjennom Finn.no. Selv om sekundærdata er mer tilgjengelig og enkelt å tilegne seg, er det ingen garanti for at disse dataene vil være direkte relevante for vår problemstilling, ettersom de opprinnelig ble innhentet med en annen hensikt (Saunders et al., 2009, s. 268-288). Et eksempel på dette er selskapet Eiendomsverdi AS, som har automatiserte prosedyrer for å ekskludere transaksjoner som ikke er relatert til bolig- eller fritidsboligsalg (for eksempel salg av tomter, garasjer, etc.), noe som skiller seg fra praksisen til Ambita.

Foreliggende utredning er basert på data fra fire hovedkilder: Enova, Kartverket, Eiendomsverdi og Ambita. De neste sidene redegjør for hvordan de ulike kildene er behandlet.

#### **Enova SF**

Enova har utviklet et programmeringsgrensesnitt (API) som gir interessenter, inkludert banker, tilgang til Energimerkestatistikken, noe vi har fått tilgang til gjennom en inngått avtale på bakgrunn av utredningens formål. Data fra 2010 ble innhentet og bearbeidet ved hjelp av RStudio og Excel. Datasettet inkluderte boligens detaljer på lokasjon som kommune, poststed, matrikkel, samt informasjon om energieffektivitet, energikarakterer og energikilder, men ikke koordinater.

De dataene som ble stilt til rådighet for oss omfattet en oversikt over bygningene som var knyttet til matrikkelen for sikker identifikasjon, mens bygninger som hadde gjennomgått energimerking uten en slik matrikkelforbindelse ikke var tilgjengelige. Som et resultat inkluderte oversikten kun

omkring 50% av alle energimerkede boliger og omkring 75% av de energimerkede næringsbyggene. Enova arbeider kontinuerlig for å knytte en større andel av bygningene og deres respektive energimerkeregistreringer til matrikkelen, og dette arbeidet er ennå ikke fullført.

Utvalget vi fikk tilgang til fra Enova hadde totalt 5424 observasjoner i 2019, 5576 observasjoner i 2020, 4310 observasjoner i 2021 og 4675 observasjoner i 2022. Disse observasjonene var imidlertid landsomfattende, så det var nødvendig å isolere data for Bergensregionen. Etter å ha filtrert på kommunenivå, satt vi igjen med henholdsvis 375, 365, 342 og 341 observasjoner for årene 2019-2022. Med data fra 2010 ender vi opp med totalt 3880 observasjoner for Bergensregionen.

**Tabell 3.** Antall observasjoner oppsummert.

<b>Årstall</b>	<b>Observasjoner, bergensregionen</b>	<b>Observasjoner, nasjonalt</b>
2010-2018	2457	31 250
2019-2022	1423	19 985
Sum	3 880	51 235

Datasettet vårt, med 24 variabler fra individuelle matrikler til energikarakterer, møtte tre utfordringer: behov for strukturerte matrikkeldata, tilstedeværelsen av irrelevante variabler, og NA-verdier. For å håndtere dette, brukte vi RStudio til å konsolidere matrikler til en strukturert form, ekskluderte irrelevante variabler og utelot kolonner med utelukkende NA-verdier for å sikre analysens integritet.

Variabelen fossilandelen viste seg å være noe problematisk for vårt formål, ettersom hele kolonnen bestod av NA-verdier som ikke kunne benyttes. Vi besluttet derfor å fjerne denne kolonnen, ikke fordi den var irrelevant, men fordi vi for øyeblikket ikke har mulighet til å innhente og anvende slik informasjon. Etter å ha fjernet disse variablene, sitter vi igjen med ENOVA-datasettet vårt, som er konstruert på grunnlag av datasettene vi fikk tilgang til fra ENOVA.

Gitt vårt ønske om å utføre både en romlig analyse og en konvensjonell regresjonsanalyse i tråd med den hedoniske teorien, bestemte vi oss for å koble disse dataene til en spesifikk geografisk lokasjon, nemlig eiendommene i de undersøkte kommunene. Følgelig valgte vi å integrere vårt datasett for Bergen og omegn, som er basert på ENOVA sine datasett, med geografisk data fra Kartverket.

## Kartverket

Vi brukte derfor datasett fra Kartverket for å knytte ENOVAs data til eiendommer. Datasettet fra Kartverket inkluderte informasjon om grunneiendommer og fester, men ikke adressens knytning til matrikkelnummer på seksjonsnivå.

Datasettet fra Kartverket gav oss også tre problemstillinger. Den første var at dataene fra Kartverket innebar hele Vestlandet, mens vi bare var interessert i Bergen og omegn. Den andre problemstillingen var at vi ikke hadde seksjonsnummer, og dermed ikke kunne slå sammen datasettet vårt som vi bygde fra Enova sine data, med datasettet fra Kartverket. Enovas datasett hadde 1 423 observasjoner i 2019-2022, mens datasettet fra Kartverket hadde 297 351 observasjoner, som vi måtte kombinere.

Vi delte prosessen inn i flere steg for å håndtere ovennevnte problemstillinger. Datasettet fra Kartverket hadde kommunenummer, og dette utnyttet vi med å først filtrere ut de ni kommunene vi var interessert i, og vi satt igjen med 149 387 observasjoner for disse kommunene.

Manglende seksjonsnummer ble løst ved å slå sammen de fire første matriklene fra Kartverket sine data, med samme formelle oppsett som vi satt opp i Enovas datasett. Vi brukte RStudio til å midlertidig ta ut seksjonsnummeret fra Enovas datasett og slo sammen datasettene i et midlertidig datasett, for å slå dette sammen med et nytt permanent datasett, med hensyn på gateadresse. Et matrikkelnummer på fire eiendomsidentifikatorer er såpass unikt at vi kunne kombinere dette og gateadressen for å lage et datasett hvor vi matcher unike småhus og boligblokker fra Enova og Kartverket sammen. Etter dette slo vi sammen de gjenværende observasjonene med matchende seksjonsnummer.

Vi har dermed et datasett med unike observasjoner fra datasettet vi lagde med grunnlag fra Enova slått sammen med posisjonsdata og koordinater fra Kartverket. Dette datasettet gir oss dermed 3317 unike observasjoner i vårt fokusområde, med alle fem eiendomsidentifikatorene i et formelt oppsett og koordinater som vi kan bruke videre i våres analyser.

Et vesentlig formål med denne oppgaven er å gjennomføre en romlig analyse på dataene vi får samlet inn. Vi anser Bergen som regionssenteret i vår studie, og koordinatene til Bergen er utgangspunktet når man måler distanse til sentrum (CBD) i kilometer.



## Eiendomsverdi

Eiendomsverdi er et selskap som gir en oversikt over aktivitet og utvikling i boligmarkedet, og som inneholder informasjon om boliger i Norge som er eller har vært til salgs .

For å undersøke utviklingen i prisene på boliger i Bergensområdet, ble det oversendt data om boliger som ble solgt i perioden 2019-2022. Fokuset var på leiligheter, eneboliger, tomannsboliger og rekkehus, og kun boliger som var selveiet ble registrert for å unngå eventuelle variasjoner i pris som kan følge av ulike eierformer. Alle salg fra 2019 og frem til 2022, samt det siste salget før 2023, ble registrert, og for hvert salg ble det innsamlet opplysninger om boligens salgpris, boligtype, bydel, BRA, tomtestørrelse, byggeår, etasje, antall soverom, totalpris og salgsmåned. Datasettet inneholdt ikke informasjon som koordinater eller matrikelnumre.

## Ambita

Ambita er et eiendomsteknologiselskap spesialisert på eiendomsdata og tilbyr ulike typer tjenester for levering av eiendomsrelatert informasjon, elektronisk tinglysning, analysetjenester, karttjenester og mer. Vi har i forbindelse med foreliggende utredning kjøpt en engangsløst på kr 6 250 NOK av eiendomsdata som omfatter en oversikt over omsatte eiendommer og borettsandeler i Norge, med en historikk tilbake til 1993 for eiendommer og 2007 for borettsandeler. Datasettet inneholdt alt fra lokaliseringsdata som matrikel, kommune og poststed, koordinater og lignende, til data rundt boligkarakteristikker. I tillegg inneholdt det også data om bygningenes juridiske status, eierforhold og planstatus, salgstype, samt salgs- dato og pris, mm.

Kostnaden for leveransen er dekket av undertegnede, med bakgrunn i egen motivasjon. Det er viktig å merke seg at *fellesgjeld* ikke er inkludert i kjøpesummen for borettsandeler. Dermed er alle andelsleiligheter ekskludert fra analysen. Til vårt formål har vi kun beholdt matrikkeldata for både eiendommer med selveier som eierform, omsetningstransaksjoner, geografisk lokalisering (stadfesting) i en historikk for perioden fra 2019 til og med 2022, i våre fokusområder.

Det datasettet vi ervervet fra Ambita var av betydelig omfang, bestående av nærmere 300 000 observasjoner og 121 variabler. En rekke av disse variablene var overflødige for vår analyse, enten fordi de var irrelevante eller fordi de ikke kunne anvendes i vår metodikk. Følgelig besluttet vi å eliminere slike variabler ved hjelp av programvaren RStudio. Vår studie fokuserer på boligprisene ved kjøp, og samme bolig opptrådte ofte flere ganger i disse observasjonene. For å unngå duplikater

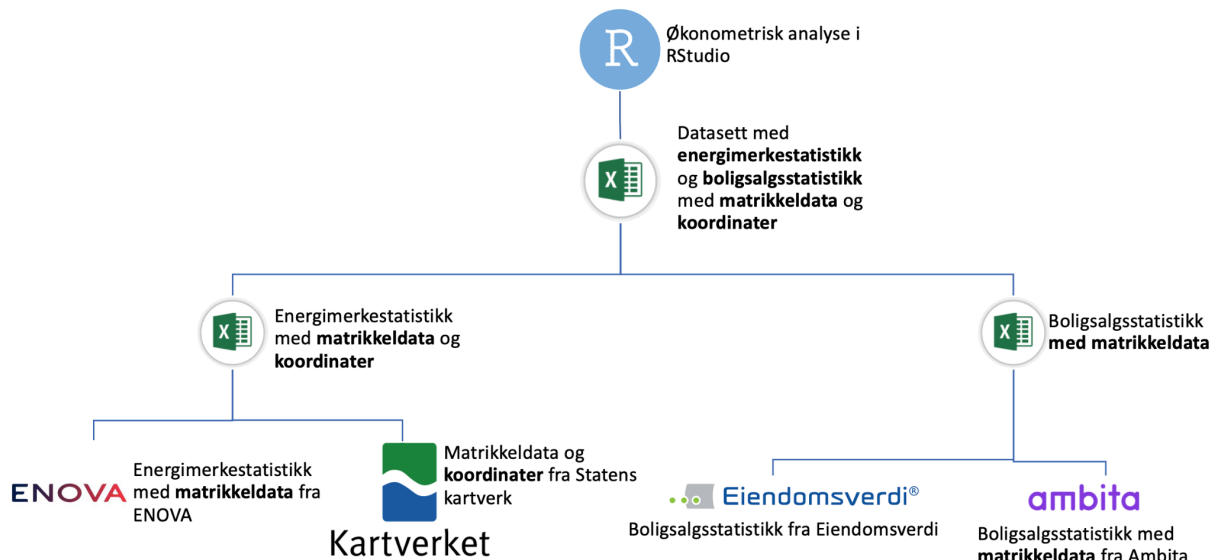
og bevare relevansen, besluttet vi å fjerne observasjonene fra boligselgers ståsted.

Etter gjennomført rensing av Ambita AS-datasettet, inneholdt det 123 000 observasjoner og 60 variabler. Imidlertid var det nødvendig å konsolidere dette datasettet med vårt datasett fra Enova for å inkludere energikarakterer for boliger solgt i perioden 2019-2022.

## Datasettsammenslåing

Før sammenslåingen av datasettene, var datarensing avgjørende, hvor vi fjernet duplikater, korrigererte feil, standardiserte formatet (for eksempel datoer og tall) og eliminerte overflødige data. Vi identifiserte matrikkelnummeret (knr, gnr, bnr, snr, fnr) som felles nøkkel for sammenslåing.

**Figur 7.** Diagram over sammenslåinger av datasett



Etter integreringen med Enova og Ambita datasettene, møtte vi et problem med manglende seksjonsnummer for eneboliger. Dette hindret RStudio i å differensiere individuelle boligenheter. Løsningen var identifisering av en unik variabel som en distinkt identifikator for hver boenhet, hvor vi konsentrerte oss om variablene 'dokument ID ved salg' og 'intern ID for boenheter'. Disse variablene antas å være unike, og ved å implementere disse i datasettet vårt, var vi i stand til å differensiere mellom forskjellige salg av forskjellige boliger. Dette resulterte i et datasett med 2964 observasjoner. For å rette fokus mot et åpent marked, filtrerte vi datasettet ytterligere for å inkludere bare frie salg av boliger. Dette valget er forankret i teorien om hedoniske modeller, som antar at boligmarkedet opererer som et åpent marked. Etter denne filtreringsprosessen, bestod datasettet vårt av totalt 2798 observasjoner.

I lys av den gjennomførte analysen oppdaget vi flere feil og manglende registreringer knyttet til kritiske variabler som antall rom, toaletter, og kjøkken. Disse utilstrekkelige observasjonene kunne vi ikke inkludere i studien vår, og de ble derfor ekskludert fra den videre undersøkelsen. Som tidligere omtalt var datasettet vårt ikke delt mellom boliger med og uten fellesgjeld, noe som nødvendiggjorde ytterligere opprydning. Etter gjennomføringen av disse justeringene, stod vi igjen med 2118 observasjoner relatert til energikarakterer i Bergensregionen.

Etter å ha fullført de nevnte stegene, konverterte vi true/false-verdier til binære tall, da datasettet anvendte disse om hverandre. Videre opprettet vi en ny variabel for å beregne boligens alder ved å sammenligne salgsåret med byggeåret. Endelig ble det lagt til en kvartalsvariabel for å lettere kunne identifisere tidspunktet for salg, og for å analysere potensielle variasjoner i priser og energiklassifiseringer over ulike perioder av året. Etter gjennomføringen av disse trinnene, ble datasettet vurdert som egnet for videre analyse.

## **Datasettbeskrivelse**

Det endelige datasettet som vi skal anvende i analysen, en kombinasjon av flere datasett fra Ambita, Eiendomsverdi, Enova og Kartverket, omfatter 2118 observasjoner i 2019-2022, og 66 variabler. Datasettet omfatter variabler relatert til energi, boligattributter og romlige målinger.

*Salgspris* er kritisk i vår analyse, da hele poenget er å analysere energikarakterers innvirkning på boligprisene. Salgsprisene er justert med konsumprisindeks (KPI) hvor indeksen settes til 108,4 i basisåret 2018, og justert med årsendringene i henhold til Statistisk sentralbyrå sin tabell 03014 i appendiks A4. *Salgsår* og *kvartal* er relevant for å se når boligen er blitt solgt, og hvordan prisen har utviklet seg over tid, også gjennom det inneværende året. Vi kan også analysere om boliger solgt i samme år og kvartal som er relativt like, men ulike energikarakter, har noen betydelig prisforskjell. *Energikarakterer*, *oppvarmingskarakter*, *levert energi* og *byggematerialer* er også helt sentralt for oppgaven vår. Med tilgang til alle disse dataene kan vi også finne om det er korrelasjon mellom en eller flere av disse variablene.

*Byggeår* og *alderen* til boligen er relevant for analysen da vi kan undersøke hvordan salgspris og energikarakter samhandler med alderen til boligen. Da det kan oppstå korrelasjoner her mellom variabler, så blir denne variabelen enda mer relevant. Rubin (1993) påpeker en negativ alderseffekt i

boligpriser som skyldes preferanser for nyere hus, og at hus avskrives over tid. Malpezzi et al. (1987) anslår betydelige depresieringseffekter i størrelsesorden 0,9% til 0,3% som støtter dette funnet. Godman og Thibodeau (1995) og Osland (2013) understreker viktigheten av å være forsiktig med spesifiseringen av aldersvariabelen i en økonometrisk analyse. De foreslår at alder bør inkluderes som et kvadrert ledd for å fange opp mulige ikke-lineære sammenhenger mellom alder og boligpriser. I tillegg påpeker Osland (2013) at aldersvariabelen kan fange opp mange (uobserverte) trekk ved nabolaget, noe som gjør den til en viktig variabel for å forstå sammenhengen mellom boligpriser og energimerking i Bergensregionen. Det er verdt å merke seg at Enova inkluderer byggeåret eksplisitt i beregningen av energikarakter. For å isolere effekten av energieffektivitet, må vi kontrollere for faktorer som påvirker prisen og varierer blant forskjellige energikarakterer.

*Det helhetlige arealet* på boligen er relevant da det kan forventes en sammenheng mellom boligstørrelse og energikarakter, samt salgsprisen. Når vi undersøker hvordan størrelsen på boligen slår ut på dette, så blir det naturlig å ta med antall, bad og WC i boligen, da det er sannsynlig med en sterk positiv korrelasjon mellom størrelse og antall rom. Vi må dermed ha muligheten til å undersøke dette og eventuelt utelukke det.

*Koordinater og matrikkel* lar oss finne ut hvor i Bergensregionen boligen ligger, som er relevant for salgsprisen. Vi kan også muligens bruke matrikkelen for å undersøke om noen områder har en større eller mindre andel av ulike energikarakterer. Koordinater lar oss også undersøke om det er noen romlige effekter som påvirker prisen på boliger i et område.

Til tross for et relativt lavt antall på 2118 observasjoner, gir datasettet muligheten til å utforske sammenhengen mellom energimerking og boligpriser. Med mulighet for å konsentrere oss om bestemte variabler, geografiske områder, bygningsmaterialer og lignende faktorer, muliggjør datasettet et spekter av analytiske tilnærminger.

## 5.0 Analyse

I dette analysekapitlet innleder vi med en presentasjon av deskriptiv statistikk (5.1) for å undersøke hovedtendenser i datasettet. Videre utforskes utvalget (5.2). Våre hypoteser presenteres i (5.3), og til slutt diskuterer vi valget av funksjonsform (5.4), inkludert vurdering av ulike modeller.

### 5.1 Deskriptiv statistikk

I denne delen av utredningen fokuserer vi på de vanligste beskrivende statistikkene. *Gjennomsnittet* gir en indikasjon på sentral tendens i dataene.

**Tabell 4.** Deskriptiv statistikk for ulike boligtyper

<b>Prisstatistikk for ulike boligtyper for hele området</b>	<b>Gjennomsnittspris</b>	<b>Medianpris</b>	<b>Minimumspris</b>	<b>Maksimumspris</b>
Annet	2 839 059 kr	2 700 000 kr	1 410 000 kr	8 300 000 kr
Blokkleilighet	3 397 789 kr	3 140 000 kr	1 000 000 kr	16 300 000 kr
Frittliggende enebolig	5 153 386 kr	4 750 000 kr	1 042 112 kr	14 750 000 kr
Rekkehus/kjede	3 958 478 kr	3 445 000 kr	1 105 000 kr	16 700 000 kr
Tomannsbolig	4 079 924 kr	3 650 000 kr	1 300 000 kr	11 200 000 kr
Uoppgitt	3 341 522 kr	3 075 000 kr	1 100 000 kr	11 000 000 kr

**Tabell 5.** Gjennomsnittspriser

<b>Gjennomsnittsprisen for boliger i hele området vårt</b>		<b>Gjennomsnittsprisen for boliger i Bergen kommune</b>		<b>Gjennomsnittsprisen for utenfor Bergen kommune</b>	
2019	3 465 167 kr	2019	3 364 026 kr	2019	2 864 445 kr
2020	3 429 641 kr	2020	3 565 745 kr	2020	2 893 181 kr
2021	3 655 071 kr	2021	3 758 074 kr	2021	3 202 842 kr
2022	3 817 127 kr	2022	3 899 831 kr	2022	3 472 524 kr
Gjennomsnitt	3 583 772 kr	Gjennomsnitt	3 711 160 kr	Gjennomsnitt	3 081 863 kr

Energikarakteren, et resultat av beregnet levert energi under normal bruk til boligen eller bygningen, følger standarden NS 3031. Karakteren representerer energimengden som tilføres bygningen utenfra ved normal bruk, som i de fleste tilfeller kan betraktes som den energimengden som må kjøpes. For å gi en oversikt over energiforbruket, presenterer vi deskriptiv statistikk av beregnet levert energi, målt i kilowattimer per kvadratmeter (kWh/m<sup>2</sup>), for årene 2019 til 2022.

**Tabell 6.** Deskriptiv statistikk av andre sentrale variabler

<b>Beregnet levert energi (kWh/m<sup>2</sup>)</b>	<b>Gjennomsnitt</b>	<b>Median</b>	<b>Gjennomsnittlig boligalder (år)</b>		<b>Gjennomsnittlig km fra solgte boliger til Bergen sentrum (CBD)</b>	
2019	198	145	2019	28	2019	8,35
2020	210	150	2020	34	2020	7,49
2021	233	170	2021	39	2021	7,31
2022	233	160	2022	39	2022	7,48
Gjennomsnitt	216	157	Gjennomsnitt	35	Gjennomsnitt	7,66

Fra 2019 til 2022 ser vi en økning i beregnet levert energi (kWh/m<sup>2</sup>), noe som tyder på et høyere energiforbruk per kvadratmeter i boligene. Samtidig viser dataene at gjennomsnittlig boligalder også har økt over denne perioden. En mulig tolkning av denne trenden kan være at eldre boliger generelt kan ha dårligere isolasjon og dermed kreve mer energi for oppvarming. Dette ville føre til en positiv korrelasjon mellom boligalder og energiforbruk. Når det gjelder avstanden til Bergen sentrum (CBD), ser det ut til at det er en liten nedgang fra 2019 til 2021, før det øker litt igjen i 2022. Det er mindre tydelig hvordan dette kan korrelere med energiforbruket, da det er mange andre faktorer som kan påvirke dette forholdet. For eksempel kan boliger nærmere sentrum være eldre og dermed mindre energieffektive slik som i Sandviken-området, men de kan også være mindre i størrelse, noe som reduserer det totale energiforbruket.

Tabell 7 og 8 gir en oversikt over medianverdiene for ulike variabler i datasettet vårt. *Median* er en robust statistisk indikator som representerer den midterste verdien i et datasett når det er ordnet numerisk. Den gir en indikasjon på den sentrale tendensen i datasettet, og er mindre påvirket av ekstreme verdier (utliggere) enn gjennomsnittet.

Tabell 7 viser medianprisen for boliger i hele området, i Bergen og utenfor Bergen for perioden 2019-2022. Det er verdt å merke seg at medianprisen generelt er høyere i Bergen sammenlignet med områdene utenfor byen. Over tid ser vi også en økning i medianprisen for boliger både i og utenfor Bergen, noe som indikerer en generell oppadgående trend i boligprisene i regionen.

**Tabell 7.** Medianpris

<b>År</b>	<b>Hele området</b>	<b>I Bergen</b>	<b>Utenfor Bergen</b>
2019	3 100 000 kr	3 260 000 kr	2 577 500 kr
2020	3 050 000 kr	3 170 000 kr	2 577 500 kr
2021	3 300 000 kr	3 400 000 kr	2 962 500 kr
2022	3 430 000 kr	3 490 000 kr	3 122 500 kr
Total	3 200 000 kr	3 305 000 kr	2 762 500 kr

Tabell 8 presenterer medianverdiene for ulike boligattributter i hele datasettet, samt for hvert år fra

2019 til 2022.

**Tabell 8.** Median på boligattributter

Variabel	Hele datasettet	2019	2020	2021	2022
Median leilighetsstørrelse (m <sup>2</sup> )	65	69	65	64	64
Median boligstørrelse (m <sup>2</sup> )	167	158	176	164	175
Median boligalder (år)	15	10	14	18	17
Median avstand til Bergen sentrum (km)	5,92	6,66	5,92	5,56	5,67
Median beregnet levert energi (kWh/m <sup>2</sup> )	157	145	150	170	160

## 5.2 Utvalg

I dette kapitlet gis en beskrivelse av utvalget som danner grunnlaget for vår forskning. Vi har konsentrert oss om fem sentrale variabler, som vi mener er avgjørende for vår studie. Det er viktig å merke seg at disse dataene kan avvike fra tidligere presenterte tabeller. Dette skyldes at tallene i tabellen nedenfor representerer verdiene i vårt spesifikt filtrerte datasett, som er begrenset til det geografiske området relevant for vår studie.

**Tabell 9.** Deskriptiv statistikk for relevante variabler

Variabel	M	SD	Min	Max
Energikarakter (A=1, G=7)	4.21	1.51	1	7
Kvadratmeterpris i hele regionen (kr)	49 812	16 524	11 889	160 500
Boligalder	23	33	0	163
Avstand til Bergen sentrum (km)	8.03	6.36	0.27	47.91
Beregnet levert energi (kWh/m <sup>2</sup> )	180.46	111.38	84.24	743.87

*M: Gjennomsnitt, SD: Standardavvik*

I tabell 9 ser vi at *energikarakteren*, som strekker seg fra 1 (A) til 7 (G), har et gjennomsnitt på 4.21 med et standardavvik på 1.51. Dette indikerer at boligene i vårt utvalg generelt har middels til høy energikarakter, og det er relativt stor variasjon i energikarakterene. *Kvadratmeterprisen* varierer betydelig, som reflektert av standardavviket på 16 524 kr, hvilket indikerer stor spredning i prisene. *Boligalderen* varierer betydelig med en spennvidde fra 0 til 163 år, noe som kan antyde en variert boligmasse. *Avstanden* til Bergen sentrum er i gjennomsnitt 8.03 km med et standardavvik på 6.36 km, noe som indikerer betydelig geografisk spredning i datasettet.

**Tabell 10.** Deskriptiv statistikk for beregnet levert energi (kWh/m<sup>2</sup>)

Beregnet levert energi (kWh/m <sup>2</sup> )	M	SD	Min	Max	M (2019)	M (2022)
Bergen kommune	192	121	91	744	178	196
Distriktskommuner	135	37	84	452	132	138
Hele regionen	180	111	84	744	167	185

*M: Gjennomsnitt, SD: Standardavvik*

I tabell 10 er det fremhevet at den gjennomsnittlige beregnede leverte energien har økt fra 2019 til 2022.

### 5.3 Hypoteser

Dette avsnittet inneholder de fundamentale hypotesene som utgjør basis for vår vitenskapelige undersøkelse, som er konstruert for å strukturere vår forskning og hjelpe oss med å besvare våre forskningsspørsmålene på en grundig måte. Basert på den overordnede problemstillingen og den teoretiske bakgrunnen, fremsetter vi følgende hypoteser som skal testes empirisk:

Nullhypotesen  $H_0$ : Energimerkingens påvirkning på boligprisene er ikke statistisk signifikant, hverken i perioden før og etter strømpriskrisen inntraff i 2021-2022

Alternativhypotesen  $H_1$ : Energimerkingens påvirkning på boligprisene er statistisk signifikant, både i perioden før og etter strømpriskrisen inntraff i 2021-2022

Dersom det er en signifikant korrelasjon mellom energimerking og pris i noen av periodene, avviser vi nullhypotesen og konkluderer med energimerkingens påvirkning på pris. Uten slik korrelasjon, kan vi ikke avvise nullhypotesen og dermed ikke fastslå energimerkingens betydning

Med bakgrunn i ovennevnte, er det også andre aspekter vi ønsker å undersøke for å få et mer helhetlig bilde. Følgelig vi formulert flere tilleggshypoteser for å utforske temaet ytterligere:

- Eldre boliger har generelt lavere energikarakter enn nyere boliger.
- Det er ingen korrelasjon mellom boligstørrelse og energimerking.
- Det er en positiv sammenheng mellom oppvarmingskarakter og energikarakter.
- Høyere energikarakter korrelerer positivt med høyere boligpris.
- En høyere oppvarmingskarakter resulterer i en høyere boligpris.
- Økt energiforbruk korrelerer positivt med høyere boligpris, grunnet høyere boligstørrelse.



## 5.4 Valg av funksjonsform, variabler og modeller

Målet med denne studien er å utføre regresjonsanalyser på boligpriser ved hjelp av flere variabler, og å gjennomføre romlige analyser for å forstå deres romlige strukturer. For å oppnå dette formålet, vil vi formulere flere hedoniske modeller og empirisk evaluere deres prognosekraft. Vår tilnærming vil fokusere på å vurdere de statistiske sammenhengene mellom ulike faktorer og boligprisene i et romlig perspektiv. For å definere den avhengige variabelen, vil vi bruke den logaritmiske transformasjonen av boligens transaksjonspris per kvadratmeter,  $\log(pm2)$  som måleenhet. Logaritmen av transaksjonsprisen per kvadratmeter forklares av tradisjonelle forklaringsvariabler som kan sammenlignes med de i Brounen og Kok (2011).

Det er verdt å merke seg at Enova inkluderer byggeåret eksplisitt i beregningen av energikarakter. For å isolere effekten av energieffektivitet, må vi kontrollere for faktorer som påvirker prisen og varierer blant forskjellige energikarakterer.

For dette formålet bruker vi regresjonsanalyse og definerer salgsprisen som en funksjon av variabler som størrelse, boligtype, byggeår og energikarakter. Dette lar oss se på den marginale effekten av energikarakteren mens vi holder de andre variablene konstant.

### Trinnvis modelleringsprosess

For å velge variabler til vår modell, har vi valgt dem som i teorien (Hårsman, 1981; Laakso, 1997) skal påvirke boligprisene, slik som antall rom, alder, bad og flere. Vi startet med den umiddelbare og relevante variabelen for vår avhengige variabel. En etter en la vi til ytterligere variabler, og systematisk vurderte deres signifikante innflytelse på den avhengige variabelen. Videre ble det utført analytiske tester for å optimalisere variabelutvalget. Utvalget er et resultat av en rekke tester som omfatter følgende:

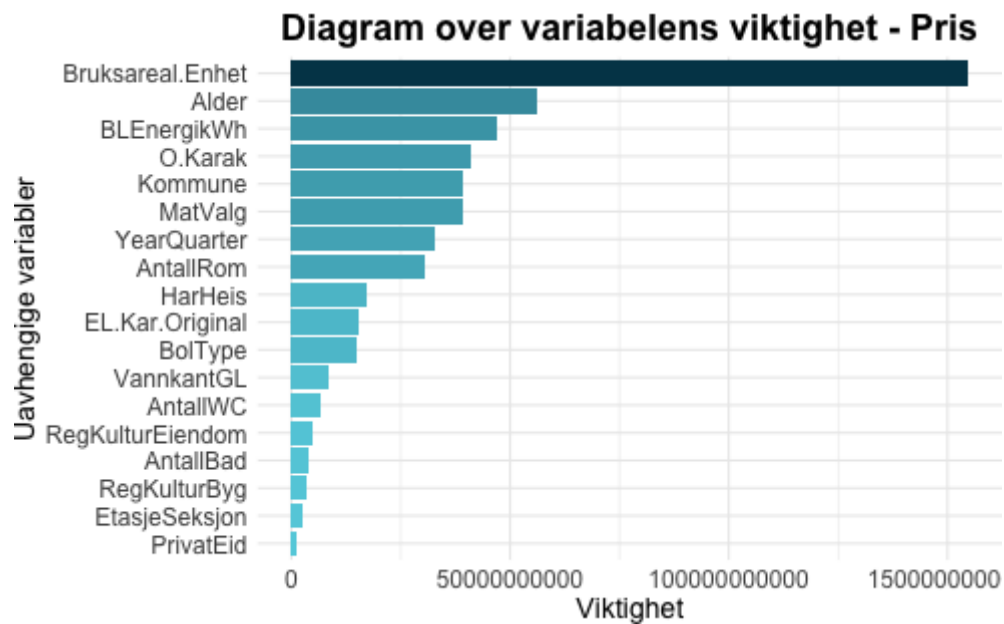
- *VIF-verdier* som angir grad av multikollinearitet
- *Korrelasjonsmatriser* for å finne variabler som er sterkt korrelert med hverandre
- *Trinnvis regresjon* for å finne en optimal kombinasjon av uavhengige variabler for vår modell. Denne metoden starter med en tom modell og legger til eller fjerner variabler basert på deres statistiske signifikans og bidrag til modellens forklaringskraft. På denne måten kan vi bygge en modell som inkluderer de mest relevante variablene for å forklare boligprisen.

Vi har også konstruert en grafisk illustrasjon som viser betydningen av forskjellige variabler i en

prediktiv modell. I diagrammet er hver variabel representert med en bar, og lengden på hver bar indikerer hvor viktig den tilsvarende variabelen er i modellen. De mest betydningsfulle variablene er de som har lengst barer og er rangert øverst i plottet.

Det er verdt å merke seg at betydningen av en variabel ikke nødvendigvis betyr at den har en stor effekt på boligprisene, men snarere at den bidrar betydelig til modellens nøyaktighet. Diagrammet er illustrert i figuren som er inkludert nedenfor.

**Figur 8.** Diagram over variabelens viktighet.



Tabell 11 viser variabelutvalget vårt, hvor \* markerer dummyvariabler

**Tabell 11.** Variabelliste

	Våre variabler (*dummy)	Beskrivelse
1	pm2	Kvadratmeterpris på boligen
2	*EL.Kar.Original	Energikarakter for boligen hvor <b>A = best</b> , og <b>G = verst</b>
3	Alder	Boligens alder
4	*O.Karak	Oppvarmingskarakter for boligen, basert på type energi (fossil, el., mm) <b>Grønn = best</b> , og <b>Rød = verst</b>
5	BEnergikWh	Beregnet levert energi totalt i kWh/m <sup>2</sup> ; den totale mengden energi er levert av et bestemt energisystem/kilde per/m <sup>2</sup> , basert på en beregning eller modell.
6	*MatValg	Materialevalg for boligens konstruksjon
7	Bruksareal.Enhet	Bruksareal (BRA): Alt areal av bruksenheter og felles deler som ligger innenfor omsluttende vegger, dvs. at for beregning av bruksareal tas mål fra inside vegg eller annen begrensing.
8	AntallRom	Antall rom i boligen
9	AntallBad	Antall bad i boligen
10	AntallWC	Antall toaletter i boligen
11	HarHeis	Indikator for om bygningen har heis
12	*YearQuarter	Årstall og kvartal for salgstidspunkt
13	dist_cbd_km	Distanse til Bergen sentrum i kilometer

Tabell 12 viser referanseverdiene til dummyvariablene, som gjennomgående er brukt.

**Tabell 12.** Referanse kategorier

Dummy	Referanse	Forventet retning	Årsak
Energikarakter	A	–	Flere signifikante verdier. Vi har heller ikke samme rasjonale/problemstilling som Olaussen et al. (2017)
Kvartal	Q4 2022	–	Siste periode. Høyest forventet boligverdi.
Oppvarmingskarakter	Grønn	–	Vi forventer en reduksjon i boligpriser ved en forverring av oppvarmingskarakter
Materialvalg	Betong	–	Bergen, dominert av betongboligblokker, antas å ha høyere priser enn omgivelsene, noe som indikerer en forventning om et negativt fortegn for andre materialer.

Modell 1:

$$\begin{aligned} \log(pm2) = & \beta_0 + \beta_1 \text{Energikarakter}_A + \beta_2 \text{Energikarakter}_B + \beta_3 \text{Energikarakter}_C + \\ & \beta_4 \text{Energikarakter}_D + \beta_5 \text{Energikarakter}_E + \beta_6 \text{Energikarakter}_F + \\ & \beta_7 \text{Energikarakter}_G + \varepsilon \end{aligned}$$

Vår første modell fremstår som enkel og rudimentær og skal kun tjene som en preliminær tilnærming for vårt videre arbeid med å utvikle mer robuste modeller. Modell 1 tar sikte på å isolere virkningen av energimerking på boligprisene. Denne modellen oppfyller imidlertid ikke de strengeste kriteriene for modellbygging, og vil kun fungere som en sammenligningsramme for å observere hvordan dette spesifikke virkemiddelet påvirker boligprisene når det kombineres med andre variabler i fremtidige modeller.

Videre for modell 2, så har vi valgt å legge til kun en variabel. I vår modell la vi til variabelen *Alder*, som er variabelen for hvor gammel boligen var da den ble solgt. Ved å legge til kun én variabel ønsker vi å se hvor stort utslag det har på modellen vår når vi kjører OLS-analysen, frem til vi finner den modellen som har størst forklaringskraft, lavest AIC, beste loglik verdi, samt beste signifikans verdier.

Modell 2:

$$\begin{aligned} \log(pm2) = & \beta_0 + \beta_1 \text{Energikarakter}_A + \beta_2 \text{Energikarakter}_B + \beta_3 \text{Energikarakter}_C + \\ & \beta_4 \text{Energikarakter}_D + \beta_5 \text{Energikarakter}_E + \beta_6 \text{Energikarakter}_F + \\ & \beta_7 \text{Energikarakter}_G + \beta_8 \log(\text{BoligAlder}) + \beta_9 \text{Oppvarmingskarakterer} \\ & + \beta_{10} \log(\text{StrømForbruk}) + \beta_{11} \text{Byggemateriale} + \varepsilon \end{aligned}$$

I de neste modellene vi lagde, la vi til én og én variabel og studerte effekten. I tillegg sjekket vi *goodness-of-fit* og sørget for å sikre at vi ikke konstruerte en inadekvat eller suboptimal modell

Vi utførte ulike tester med hensyn til korrelasjon, multikollinearitet, og lignende, for å finne korrelasjonen mellom variabler, og eventuelt utelukke variabler med multikollinearitet. Når dette var gjennomført ble hver modell testet, og den modellen som har best resultat ble brukt til videre analyse på flere hedoniske funksjoner

Modell 3 er formulert som følger:

$$\begin{aligned} \log(pm2) = & \beta_0 + \beta_1 \text{Energikarakterer} + \beta_2 \log(\text{BoligAlder}) + \beta_3 \text{OppvarmingsKarakter} \\ & + \beta_4 \log(\text{EnergiForbruk}) + \beta_5 \text{Byggemateriale} + \beta_6 \log(\text{AntallRom}) + \beta_7 \log(\text{AntallBad}) \\ & + \beta_8 \log(\text{AntallWC}) + \beta_9 \text{HarHeis} + \beta_{10} \text{SalgsKvartal} + \beta_{11} \log(\text{BoligAreal}) \\ & + \beta_{12} \log(\text{AvstandSentrum}) + \varepsilon \end{aligned}$$

Modell 4 er formulert som følger:

$$\begin{aligned} \log(pm2) = & \beta_0 + \beta_1 \text{Energikarakterer} + \beta_2 \log(\text{BoligAlder}) + \beta_3 \text{OppvarmingsKarakter} \\ & + \beta_4 \log(\text{EnergiForbruk}) + \beta_5 \text{Byggemateriale} + \beta_6 \log(\text{AntallRom}) + \beta_7 \log(\text{AntallBad}) \\ & + \beta_8 \log(\text{AntallWC}) + \beta_9 \text{HarHeis} + \beta_{10} \text{SalgsKvartal} + \beta_{11} \log(\text{BoligAreal}) \\ & + \beta_{12} \log(\text{AvstandSentrum}) + \beta_{13} \text{Kommune} + \varepsilon \end{aligned}$$

Modell 4 inkluderer en dummyvariabel for kommunene, mens Modell 3 ikke gjør det. Inkluderingen av kommune er viktig for å fange opp lokale forskjeller i prisene, og kan øke modellens forklarende kraft, kan det også føre til overjustering, noe som kan gjøre modellen mindre nøyaktig når den brukes til å forutsi priser. I tillegg, hvis det er multikollinearitet mellom variablene, kan dette påvirke modellens presisjon og stabilitet.

I løpet av vår studie, har vi utført omfattende modellering basert på et variert utvalg av antall naboer. Dette ble gjort for å identifisere den modellen som empirisk best tilpasset oppgaven. For hver modell og tilsvarende utvalg av naboer, ble det også utført en Morans I-test. Dette var for å evaluere og bekrefte den romlige autokorrelasjonen i boligprisdatabasene.

Vi har kjørt forskjellige modeller ved å variere antallet av nærmeste naboer, med det formål å identifisere den modellen som passer empirisk best til oppgaven. Det er gjort Morans I test for hver av disse modellene, hvor utvelgelsen av nabomatrise har vært et kritisk element. Vår observasjon antyder at det ikke foreligger vesentlige endringer i AIC-verdiene, som antyder at modellene demonstrerer lav sensitivitet overfor det valgte antallet naboer, K. For å dekke et bredt spekter av nabolag, har vi testet følgende alternativer på samtlige modeller:

- Tre nærmeste naboene (K = 3)
- Fem nærmeste naboene (K = 5)
- Syv nærmeste naboene (K = 7)
- Ti nærmeste naboene (K = 10)
- Femten nærmeste naboene (K = 15)

## 6.0 Resultater

Denne studien undersøker fire økonometriske modeller for boligprisestimering. Følgende inneholder avgjørelser som påvirker utredningen, foreløpige funn, og våre endelige modeller.

### **Variabelen *boligareal***

I vår studie har vi også valgt å inkludere boligarealet som en uavhengig variabel, til tross for at kvadratmeterprisen, som er vår avhengige variabel, er direkte avledet fra dette. Det er en innvending som kan oppstå, gitt at boligarealet allerede er innebygget i kvadratmeterprisen, noe som resulterer i at vi har to variabler i modellen vår som er fundamentalt knyttet til hverandre.

Vi vurderte fire modeller, med og uten boligareal, og fant at inkluderingen forbedret AIC-verdiene betydelig. Dette valget støttes også av tidligere forskning som Olausen et al. (2017). Basert på de observerte forskjellene i modellene våre og tidligere forskningsarbeid, har vi bestemt oss for å inkludere boligareal i modellene våre, til tross for det tilsynelatende paradokset det kan representere. Vi mener at denne tilnærmingen vil gi oss en mer robust og nøyaktig modell for å analysere boligprisene.

### **Modellvalg**

Analysen vår viser at SEM bare marginalt overgår SDM statistisk. Til tross for SDMs evne til å takle komplekse romlige avhengigheter, fremstår SEM, en enkel modell, ofte som førstevalg. Dette underbygger viktigheten av enkelhet og anvendbarhet ved valg av estimeringsmodell. De to andre modellene vi vurderte, SAR og SLX, hadde lavere statistisk signifikans og goodness-of-fit sammenlignet med SEM og SDM, og deres resultater er derfor ikke studert ytterligere enn det som presenteres i seksjon 6.1.

### **Vektmatriser**

Vi testet forskjellige vektmatriser, men fant ingen vesentlige forskjeller. Dermed presenteres matrisen med  $K = 3$ . I geografisk spredte boliger, spesielt i landlige kommuner, kan det være rimeligere å anta tre nærmeste naboer ( $K = 3$ ) enn ti ( $K = 10$ ).

## Geografisk avgrensning

Videre i studien vurderte vi også muligheten å inkludere hele regionen, bestående av ni kommuner, i analysen, men ekskluderte noen grunnet for få observasjoner for modellering. Vi fokuserte derfor på Bergen kommune som er et datarikt område. Resultatene herfra viste forbedret modellytelse, noe som understreker datakvalitet og volums betydning for robuste resultater.

Gitt disse funnene, har vi valgt å ikke sammenligne kommunene, men fokusert på modellering på nivået av Bergen kommune eller regionen som helhet (Bergensregionen).

## 6.1 Preliminære funn

I dette kapitlet vil vi presentere de foreløpige funnene fra vår økonometriske analyse. Vår første delproblemstilling var: *Hvor stor er sammenhengen mellom energimerking og boligpris, og er denne sammenhengen signifikant?* Resultatene fra analysen gir oss svar på dette spørsmålet. Våre beregninger viser en negativ korrelasjon på -9,8% mellom energimerking A-G (rangert som 7-1) og boligpris. Med andre ord, når energimerket forbedres, tenderer boligprisen til å falle. Denne negative korrelasjonen er statistisk signifikant på 1% nivå.

Samlet sett indikerer disse resultatene at det er en betydelig negativ sammenheng mellom energimerking og boligpris. Dette antyder at høyere energieffektivitet kan være assosiert med lavere boligpriser. Dette resultatet kan først og fremst beskrives som uventet. I det følgende presenteres essensen fra de ulike regresjonsanalysene.

### Resultater fra OLS-modell med kun energirelaterte variabler i Bergensregionen

OLS-modellen tilbyr en baseline for sammenligning med romlige modeller, og viser forholdet mellom kvadratmeterpris og energimerking i en enkel økonometrisk kontekst. Modellene er forventningsskjevne grunnet utelatelsen av viktige variabler som har en sterk korrelasjon med energikarakter. Uten mulighet til å inkludere slike variabler kan vi ha skjevhet i estimatene våre fordi de inkluderte variablene kan ha absorbert effekten på energikarakter, noe som fører til feilfortolkning av deres egentlige påvirkning.

En isolert analyse av energivariabler for Bergensregionen viser at *energikarakterer* B-G har negativ effekt på kvadratmeterpris sammenlignet med A, selv om resultatene ikke er statistisk signifikante.

*Oppvarmingskarakteren* viser lignende mønstre mot grønn som referanse, bortsett fra lysegrønn som gir økt pris, med signifikante resultater.

Variabelen for *forventet energiforbruk* viser en negativ korrelasjon med kvadratmeterprisen, men uten statistisk signifikans. *Eldre* boliger og økt brukbart *areal* gir lavere kvadratmeterpris, med signifikante resultater. Modellen forklarer rundt 65% av variasjonen i kvadratmeterprisen.

### **Resultater fra SAR-modell**

SAR-modellen viser ingen signifikante effekter av energimerking på kvadratmeterpris, hverken for Bergen alene eller regionen samlet, når boligarealet er medregnet. Oppvarmingskarakter har derimot signifikant innvirkning på boligpriser i Bergen, mens i regionen er alle unntatt lysegrønn merking signifikante. Oppvarmingskarakterene har negative koeffisienter, noe som indikerer at boligprisene har en tendens til å avta med dårligere oppvarmingskarakterer. Videre er det verdt å merke seg at boligstørrelse, alder og forventet energiforbruk er alle signifikante variabler med negative koeffisienter. Dette antyder at kvadratmeterprisene påvirkes negativt av større boligareal, økt alder på boligen, og høyere forventet energiforbruk.

SAR-modellen benytter rho for observasjoner og prediksjoner, med beregnede rho-verdier på 0,4631 og 0,4897 for tre og ti naboer i Bergen kommune, og 0,5239 og 0,5513 i Bergensregionen.  $\beta$  i SAR-modeller reflekterer ikke typiske lineære regresjonshellinger, og impacts-funksjonen kan brukes til å avlede direkte og indirekte effekter.

### **Resultater fra SEM-modell**

SEM-modellen indikerer at energikarakterene B, D og E påvirker prisen signifikant i Bergensregionen, men kun uten variabelen *boligareal*. Dette kan dog ikke sammenlignes med Olausen et al. (2017) ettersom de inkluderer variabelen *boligareal*. Ved inkludering av boligareal, blir ingen energikarakterer statistisk signifikante.

Alle oppvarmingskarakterene, unntatt lysegrønn, er signifikante for både Bergen kommune og Bergensregionen. Bortsett fra i Bergen kommune med ti nærmeste naboer, er alle oppvarmingskarakterene negative. Alder, beregnet levert energi og boligareal er statistisk signifikante med negative koeffisienter i alle SEM-modellene.



Moran I-statistikken viser liten romlig autokorrelasjon, men lambda-verdiene tyder på mulig romlig avhengighet, som SEM sannsynligvis har korrigert for.

Lambda-verdier indikerer en regresjonshelning mellom observerte og latente variabler. For Bergen var lambda 0,6811 for tre nærmeste naboer og 0,8233 for ti naboer. For regionen samlet gav modellen oss lambda lik 0,7201 for tre naboer og 0,8618 med ti naboer.

### **Resultater fra SDM-modell**

SDM-modellen indikerer at energikarakterene ikke har signifikante koeffisienter for hverken Bergen kommune eller Bergensregionen, mens oppvarmingskarakterene (minus lysegrønn) er signifikante med negative koeffisienter. Også alder, areal og forventet energiforbruk har negativ signifikans.

Lag-verdiene avslører at *energikarakterene* mangler signifikans med tre naboer, men E, F, og G er signifikante med ti naboer for Bergen kommune. For *oppvarmingskarakterene* viser analysen at lysegrønn og gul har signifikante verdier i Bergensregionen med ti naboer, mens oransje og rød viser signifikans med tre naboer. I Bergen kommune er alle *oppvarmingskarakterene* unntatt lysegrønn signifikante med ti naboer, mens bare lysegrønn og oransje er signifikante med tre naboer. I forhold til *boligareal*, *alder* og *forventet energiforbruk* i Bergen kommune, viser analysen signifikans for boligareal med tre naboer, mens alder og forventet energiforbruk er signifikante med ti naboer. *Boligarealets* signifikans bortfaller ved økning av nærmeste naboer. I Bergensregionen er boligareal signifikant med tre naboer, og med ti naboer er både boligareal og beregnet levert energi (kWh/m<sup>2</sup>) signifikante.

Moran's I-statistikk, beregnet med  $K = 10$  for vårt datasett, gir en verdi mellom -0.003 og -0.005. Denne verdien ligger tett opp mot null, noe som antyder en tilfeldig fordeling av residualer. Rho-verdiene i SDM-modellen indikerer en moderat til sterk positiv romlig avhengighet i Bergen kommune (0,5533 og 0,1895 for henholdsvis tre og ti nærmeste naboer), mens i Bergensregionen ser vi sterk positiv avhengighet (0,6288 og 0,6250). Dette antyder sterke påvirkninger på boligpriser fra nærliggende områder.

## Resultater fra SLX-modell

SLX-modellresultatene viser at energikarakterene ikke er signifikante, hverken i Bergen eller i hele regionen. Signifikante negative koeffisienter er registrert for oppvarmingskarakterene (unntatt lysegrønn) for både Bergen og hele regionen, noe som tyder på lavere boligpriser med dårligere oppvarmingskarakterer. Også alder, forventet energiforbruk og boligstørrelse er signifikante med negative koeffisienter.

Laggede verdier i SLX-modellen viser varierte resultater avhengig av antall naboer vurdert. Energikarakterene er generelt ikke signifikante med tre naboer, men noen blir signifikante med ti naboer. Oppvarmingskarakterene viser varierende resultater. Boligareal mister sin signifikans i noen tilfeller, mens alder og forventet energiforbruk opprettholder signifikansen under visse forhold.

Modellene for både Bergen kommune og Bergensregionen, viser oppvarmingskarakterene generelt signifikante direkte, indirekte og totale virkninger. Variablene alder, boligareal og forventet energiforbruk viser signifikante direkte virkninger med tre naboer, og indirekte og totale virkninger under visse forhold.

SLX-modellen antyder lokale romlige effekter på noen energikarakterer, samt oppvarmingskarakter, boligstørrelse og forventet energiforbruk. Moran's I-statistikk er konsekvent positiv og signifikant forskjellig fra null, noe som indikerer en positiv romlig autokorrelasjon i residualene, og tyder på at modellen ikke fullt ut forklarer romlig variasjon i boligpriser.

## Anselin-tilnærmingen

Basert på resultatene vi oppnådde fra de fire modellene, valgte vi å implementere Anselin-tilnærmingen ved bruk av Lagrange multiplikatorer for å bestemme den mest hensiktsmessige modellen. Denne analysen ble utført for både hele regionen og spesifikt for Bergen, og vi tok hensyn til både tre og ti nærmeste naboer. Alle Lagrange multiplikator (LM)-tester viste seg å være signifikante, og  $H_0$  forkastes for begge disse testene. Vi måtte dermed gjennomføre robuste LM-tester. For Bergen kommune, med tre nærmeste naboer, var begge testene signifikante, og  $H_0$  forkastes for begge. Når antallet nærmeste naboer økte til ti, var kun den robuste LM feil (RLMerr) testen signifikant.

For Bergensregionen samlet med tre nærmeste naboer var begge RLM-verdiene signifikante og

igjen forkastes  $H_0$  for begge robuste tester. Når vi utvidet analysen til ti nærmeste naboer, var det bare den robuste LM feil testen (RLMerr) som viste seg å være signifikant. Osland (2010, s. 303) gir en omfattende veiledning for å fatte beslutninger om modellvalg når begge robuste tester viser signifikans. Når begge tester er signifikante, er det modellen med den høyeste RLM-verdien som prefereres. I vårt spesifikke tilfelle oppnår RLMerr-verdien overlegent beste prestasjon for begge scenariene der begge robuste tester viser signifikans. Disse funnene tyder på at SEM-modellen er den mest passende for videre analyse.

## 6.2 Konkluderende modeller

Innenfor spekteret av våre mest vellykkede modeller, har vi foretatt et utvalg basert på de beste tilpasningsmålene. Det finnes en rekke teknikker for evaluering av modeller, og i denne evalueringen vil vi benytte oss av Moran's I, AIC og BIC. Disse kriteriene har veiledet oss i å identifisere fire modellene som best oppfyller våre tilpasningsmål.

**Tabell 13.** Oversikt over fire beste modeller, synkende fra lavest AIC-verdi.

Modell	K	Region	Moran's I		AIC	BIC
			Observert	Forventet		
1 - SEM	K = 3	Bergensregionen	-0,0116	-0,0005	-1496	-1269
2 - SDM	K = 10	Bergensregionen	-0,0055	-0,0005	-1469	-1033
3 - SDM	K = 10	Bergen kommune	-0,0034	-0,0006	-1131	-713
4 - SEM	K = 3	Bergen kommune	-0,0080	-0,0006	-1103	-886

Basert på tabellen ovenfor, fremgår det at vårt valg hovedsakelig står mellom Spatial Error Model (SEM) og Spatial Durbin Model (SDM), begge anvendt med enten tre eller ti nærmeste naboer. Gitt målsetningen for vår forskning, er det nødvendig å velge en modell og en nabomatrise for vår videre analyse. Vår evaluering indikerer at SEM-modellen med de tre nærmeste naboene leverer den beste ytelsen (AIC = - 1496). I tillegg er SEM spesielt godt egnet for å studere boligmarkedet med tanke på romlige effekter, siden den justerer for romlige korrelasjoner gjennom feilledet. Dette gjør SEM til et sterkt verktøy for å forstå og modellere romlige mønstre i boligpriser, og utredninger vil i hovedsak bruke SEM-modellen med de tre nærmeste naboene som grunnlag videre.

Etter å ha utforsket forskjellige modeller og tilnærminger i vår analyse, presenterer vi nå resultatene

fra vår konkluderende modell i dette kapitlet. Vi har valgt å bruke en Spatial Error Model (SEM) med  $K = 3$  for vår endelige analyse, da denne har vist seg å gi den mest pålitelige og robuste modellen basert på våre tidligere funn, som vist i delkapitlene 6.0 og 6.1.

Tabell 14 presenterer resultatene fra en Spatial Error Model (SEM) med  $K = 3$ , der energikarakterene, oppvarmingskarakterene, boligalderen, energiforbruket og boligarealet er tatt som uavhengige variabler, mens den avhengige variabelen er logaritmen til kvadratmeterprisen. Modellen er anvendt for to forskjellige områder: Bergensregionen som helhet og Bergen kommune spesifikt.

**Tabell 14.** SEM-modell med  $K = 3$ . Energikarakterer, og regresjonskoeffisienter (avhengig variabel: logaritmen til kvadratmeterpris)

	Modell 1		Modell 4	
	Bergensregionen		Bergen kommune	
	Estimat	SE	Estimat	SE
<i>Skjæringspunkt</i>	13,462***	0,180	13,213***	0,206
B	-0,065	0,056	-0,010	0,068
C	-0,015	0,059	0,042	0,070
D	-0,010	0,060	0,046	0,072
E	-0,067	0,064	-0,011	0,077
F	-0,018	0,070	0,050	0,084
G	0,070	0,076	0,139	0,092
<i>Oppvarmingskarakterer</i>				
Lysegrønn	-0,014	0,021	-0,001	0,028
Gul	-0,194***	0,035	-0,177***	0,041
Oransje	-0,168***	0,021	-0,140***	0,026
Rød	-0,101***	0,016	-0,073***	0,019
<i>Andre variabler</i>				
Alder	-0,034***	0,008	-0,042***	0,009
Beregnet levert energi (kWh/m <sup>2</sup> )	-0,109***	0,031	-0,114***	0,036
Boligareal	-0,407***	0,021	-0,393***	0,024
Lambda $\lambda$	0,720		0,681	
Antall observasjoner	2118		1679	

SE = Standardfeil, \*\*\* Signifikant på 1% nivå. \*\* Signifikant på 5% nivå. \* Signifikant på 10 % nivå.

Variablene B, C, D, E, F og G representerer de forskjellige energikarakterene med referanse til karakteren A. Ingen av disse variablene viser signifikante resultater på 10%, 5% eller 1% nivå, noe som indikerer at disse karakterene ikke har en signifikant innvirkning på kvadratmeterprisen.

I begge modellene er gul, oransje og rød signifikante på 1% nivå, noe som betyr at disse oppvarmingskarakterene har en signifikant effekt på kvadratmeterprisen sammenlignet med

referansekategorien grønn. Disse koeffisientene viser endringen i den forventede logaritmen til kvadratmeterprisen når oppvarmingskarakteren endrer seg fra grønn til den aktuelle karakteren. De negative tallene indikerer en nedgang i kvadratmeterprisen, desto verre oppvarmingskarakter boligen har - men det er marginale forskjeller mellom dem.

For å tolke koeffisientene til dummyvariablene i prosent, er det viktig å huske på at vi anvender en logaritmisk modell. Dette innebærer at når vi skifter fra referansekategorien (grønn) til en annen oppvarmingskarakter, vil det medføre en prosentvis endring i kvadratmeterprisen. Denne endringen kan beregnes ved formelen:

$$(e^{\text{koeffisient}} - 1) * 100\%,$$

som gir endringen i prosent. Sammenlignet med grønn referansekategori, antyder tallene at boliger med gul, oransje, og rød oppvarmingskarakter ville ha lavere kvadratmeterpris. Spesifikt, ville en gul oppvarmingskarakter redusere prisen med 16.22%-17.66%, en oransje med 13.02%-15.47%, og en rød med 7.05%-9.65%, avhengig av om en anvender Modell 1 eller 4.

Videre er boligens alder, beregnet levert energi (kWh/m<sup>2</sup>) og boligareal alle signifikante på 1% nivå i begge modellene. Koeffisienten for boligens alder reflekterer forandringen i den forventede logaritmen av kvadratmeterprisen for hvert år boligen eldes. En negativ koeffisient på -0,034 bekrefter at eldre boliger forventes å ha en lavere kvadratmeterpris, noe som stemmer overens med våre antakelser. Dette betyr at for hvert år en bolig eldes, vil kvadratmeterprisen forventes å redusere med omtrent 3-4%, alt annet likt.

For hver ekstra kWh/m<sup>2</sup> i beregnet levert energi, vil kvadratmeterprisen forventes å reduseres med omtrent 10%, alt annet likt. Dette er som forventet, gitt at større boliger med høyt energiforbruk typisk vil ha en lavere kvadratmeterpris. Det er også viktig å påpeke at en lavere kvadratmeterpris er assosiert med en høyere alder, høyere energiforbruk per kvadratmeter og større boligareal. Dette indikerer at eldre boliger, boliger med høyere energiforbruk og større boliger generelt har lavere priser per kvadratmeter.

I begge tilfellene er Moran's I-statistikk negativ, noe som indikerer negativ romlig autokorrelasjon i modellresidualene. Dette betyr at observasjoner som er geografisk nær hverandre, har tendens til å være mer ulike enn hva vi ville forvente ved tilfeldigheter. Selv om Moran's I-statistikk er negativ, er det ikke nok bevis til å si at det er en signifikant grad av romlig autokorrelasjon i residualene ettersom p-verdiene er henholdsvis 0.8528 og 0.9504.

Lambda-verdiene er henholdsvis 0.6811 og 0.72011, noe som indikerer en viss grad av romlig autokorrelasjon i de opprinnelige dataene. Men gitt Moran's I-testresultatene, kan vi konkludere med at SEM sannsynligvis har vært vellykket i å ta hensyn til denne romlige avhengigheten i sin modellering.

Så for å oppsummere, resultatene antyder at det er en viss grad av romlig avhengighet i de opprinnelige dataene, men SEM har sannsynligvis vært vellykket i å ta hensyn til denne avhengigheten, slik at det ikke er noen signifikant romlig autokorrelasjon i residualene.

## 7.0 Diskusjon

Denne delen av oppgaven vil ta for seg den drøftelse av resultatene, samt redegjøre for metodologiske begrensninger og avslutningsvis presentere funnene relatert til strømpriskrisen.

### 7.1 Drøftelse av resultater

#### **Romlige effekter: Observasjoner fra SLX-modellen**

Studien fokuserte på romlige effekter knyttet til energirelaterte faktorer. SEM-modellen utmerket seg, men andre modeller som SLX avdekket interessante resultater. Det ble notert romlige effekter for energikarakterene E, F og G, spesielt i Bergen. Dette gjaldt også regionalt, hvor energiklasse G viste begrensede romlige effekter. Til tross for at disse ikke betydelig påvirker pris, antyder modellene at de laveste energikarakterene kan ha marginale romlige effekter.

Oppvarmingskarakterene er også ganske interessante for oppgaven her også. Gjennomgående har vi observert en statistisk signifikant sammenheng mellom oppvarmingskarakterene og boligprisene. Våre modeller avslører også tydelige indikasjoner på romlige effekter knyttet til oppvarmingskarakterene, samt svake tegn på romlige effekter for energikarakterene.

Oppvarmingskarakterene viste en statistisk signifikant sammenheng med boligprisene og klare tegn på romlige effekter. Også alder og boligareal viste romlige effekter i Bergen kommune, men ved Modell 1 mistet alder sin signifikante effekt mens boligareal bevarte sin. Modellen indikerte også romlige effekter for energiforbruk, muligens korrelert med oppvarmingskarakterene og de laveste energikarakterene.

#### **Våre empiriske funn**

Våre empiriske funn knyttet til en potensiell prispåvirkning assosiert med energikarakterene er fortsatt tvetydig og delvis motstridende. Eksempelvis observerte Brounen og Kok (2011) en markant prispremie relatert til energimerking i sitt hedoniske datasett fra Nederland, mens andre undersøkelser, inkludert Murphy's studie fra 2014 i samme land, rapporterte kun minimale eller ubetydelige effekter. I motsetning til Brounen og Kok (2011), kunne vi imidlertid ikke reprodusere en positiv prispåvirkning. Selv i romlige modeller, fant vi ingen bevis for en prispremie assosiert

med energimerker.

Analysen vår, i likhet med Olausen et al. (2017), antyder at virkningen av energimerking på kvadratmeterprisen er usikker, ettersom energikarakterene i det store og hele ikke har noen statistisk signifikant påvirkning, mens oppvarmingskarakterene ifølge flere av modellene våre har en statistisk signifikant effekt. En potensiell forklaring av ovennevnte kan være forankret i kjøpernes persepsjon av energieffektivitet, som synes å ha en sterkere korrelasjon med oppvarmingskarakteren enn med energikarakteren. Oppvarmingskarakteren gir en mer umiddelbar og intuitiv indikasjon på boligens energieffektivitet, ved å vurdere hvorvidt boligens oppvarmingssystem er basert på fornybare energikilder, og om det kan generere varme uten direkte bruk av elektrisitet, olje eller gass. Dette kan fremstå som en mer konkret og håndgripelig indikator på energieffektivitet enn energikarakteren, som primært er basert på den forventede mengden elektrisitet som vil bli brukt fra strømmettet under normalt forbruk.

En bolig med høy oppvarmingskarakter kan bety lavere strømutfgifter, noe som kan være særlig attraktivt for potensielle kjøpere. Derav kan det være slik at oppvarmingskarakteren har en mer direkte og merkbar innvirkning på boligpriser enn den overordnede energimerkingen, ettersom kjøpere kan se mer umiddelbare fordeler (i form av reduserte strømutfgifter og økt miljøvennlighet) med en høy oppvarmingskarakter. I lys av dette resultatet er det rimelig å hevde at boligprisene i større grad påvirkes av oppvarmingskarakteren enn av den generelle energimerkingen.

Resultatene for energikarakterene fra vår analyse, sett i sammenheng med resultatene fra Olausen et al. studie (2017), er interessante. Spesielt når vi tar hensyn til strømpriskrisen vi står i, samt funnene fra Lainé (2011), som viste i hennes studie at en betydelig andel av boligkjøpere vektlegger andre faktorer mer enn energikarakterer ved boligkjøp, selv om det gir en økonomisk belastning. Appendiks A3 gir en kvartalsvis oversikt over strømprisene fra 2019 til 2022, som korresponderer med perioden for vår studie. Det fremgår tydelig at strømprisene i denne perioden i flere tilfeller var tre til fire ganger høyere enn i den forutgående perioden. For å lindre den økonomiske byrden for befolkningen, ble det innført tiltak relatert til strømpriser. Til tross for disse markante endringene i strømpriser, har vi ikke funnet statistisk signifikante bevis for at energikarakterer påvirker boligpriser. Ved første øyekast kan dette virke kontraintuitivt, gitt den sterke økningen i strømpriser og økt fokus på energieffektivitet. Imidlertid støttes våre funn av tidligere forskning, inkludert arbeidet til Lainé (2011), som ikke fant signifikante sammenhenger mellom energikarakterer og boligpriser.



## 7.2 Metodologiske begrensninger

Vår studie har støtt på en rekke metodologiske utfordringer som det er viktig å bemerke. I det følgende redegjør vi kort for dem.

### Feilspesifikasjon i regresjonsmodellen

Vår analyse avdekket for øvrig en feilspesifikasjon i den hedoniske regresjonsmodellen; koeffisientene for energikarakterene synes å reflektere noe annet enn en direkte effekt av karakterene selv. Vi står overfor problemet med en eller flere utelatte variabler som er korrelert med energikarakterene. Disse variablene omfatter:

- Registreringsstatus for bygningen som et kulturbygg.
- Typen av energikilde benyttet for oppvarming, enten fossilt brensel eller elektrisitet
- Hvilken type etasje boligen befinner seg i.

Hver av disse faktorene har potensialet til å innvirke på de endelige resultatenes integritet. Likevel er de ikke av primær relevans for denne studien, med unntak av den fossile energiandelen. Det skal bemerkes at det for dette spesifikke fokuspunktet ikke var tilgjengelig data fra ENOVA.

Det er viktig å anerkjenne at utelatelsen av disse variablene kan ha innflytelse på våre resultater. Vi anser det dermed som viktig å belyse dette for å gi leserne en helhetlig forståelse av metoden vår, og for å fremheve potensielle områder for videre forskning og analyse. Selv om disse variablene er utelatt i vår nåværende studie, kan de være verdige å utforske i fremtidige undersøkelser.

En annen mulig faktor kan være boligens estetiske utseende. Det kan tenkes at kjøpere ved kjøpstidspunktet ikke prioriterer boligens energiytelse, og i enda mindre grad energikarakterene. Lainé (2011) fant i sin studie at de fleste som kjøper bolig prioriterer faktorer som har større innvirkning, som hage og utendørsområde, beliggenhet, nabolagets karakter, og eiendommens størrelse. Hennes studie for energimerking fant at bare rundt 14% prioriterte energimerking fremfor andre faktorer ved boligkjøp i 2011, fremfor andre faktorer (Lainé, 2011, s. 46). Hennes studie viste også at selv om bare 14% prioriterte energimerking, så måtte 76% rasjonere strømmen sin, fordi de ikke hadde råd til å varme opp huset året før (Lainé, 2011, s. 7)

## **Valgfrihet i energimerking**

Vår forskning støtter seg på data hentet fra flere kilder, primært fra Ambita AS og ENOVA. Ambita har bidratt med data relatert til boligens egenskaper, mens ENOVA har levert data knyttet til energimerking. Det er imidlertid viktig å påpeke at det er frivillig å oppdatere energimerket på boligen i Norge for de som allerede eier en bolig, men må oppdatere ved salg. Dette innebærer at det er begrenset med observasjoner i noen områder. Dette utgjør en metodologisk begrensning i vår studie.

## **Mangel på tilgang til data om fellesgjeld**

Et annet aspekt som begrenser vår forskning, er mangel på tilgang til data om fellesgjeld og andelsleiligheter. Disse dataene kunne ha bidratt til en mer nyansert forståelse av boligmarkedet, og deres fravær gjør det utfordrende å skille mellom ulike typer boliger som leiligheter, eneboliger og rekkehus.

## **Begrensede observasjoner i enkelte distriktskommuner**

Videre har vi også støtt på utfordringer med begrensede observasjoner i enkelte distriktskommuner. For å sikre robustheten i vår analyse har vi valgt å utelate disse fra vår studie. Dette er et nødvendig, men uheldig kompromiss, da det begrenser representativiteten til våre funn.

## **Differensiering mellom ulike typer boliger**

Til sist, har vi også møtt på utfordringer knyttet til differensiering mellom ulike typer boliger. Vi har ikke vært i stand til å skille effektivt mellom leiligheter, eneboliger og rekkehus. Dette skyldes mangelen på detaljert informasjon i de tilgjengelige dataene, og kan potensielt påvirke tolkningen av våre resultater.

Til tross for disse begrensningene, mener vi at vår studie tilbyr verdifull innsikt i sammenhengen mellom energimerking og boligpriser. Vi anerkjenner imidlertid at videre forskning vil ha nytte av mer detaljerte og omfattende data.

## Skjevheter i datasettet fra ENOVA

I forbindelse med søknadsprosessen for tilgang til Enovas database, ble det klart at eiere i dag selv står for registreringen av energiattester, i stedet for at dette håndteres av profesjonelle organisasjoner eller offentlige instanser. Det er et viktig aspekt å være klar over at en betydelig andel av eiere av eldre boliger som har gjennomført oppgraderinger, i tillegg tenderer til å benytte en forenklet registreringsprosess for å generere attestene. Dette skyldes ofte at de oppfatter en mer detaljert registrering som overveldende og kompleks. Konsekvensen av dette valget kan være at mange eksisterende boliger som blir lagt ut på markedet, kan ha en energiattest som viser en relativt lav energikarakter. Dette kan potensielt ha en negativ innvirkning på salgsprisen for disse boligene.

## 7.3 Strømpriskrisen vinteren 2021/22

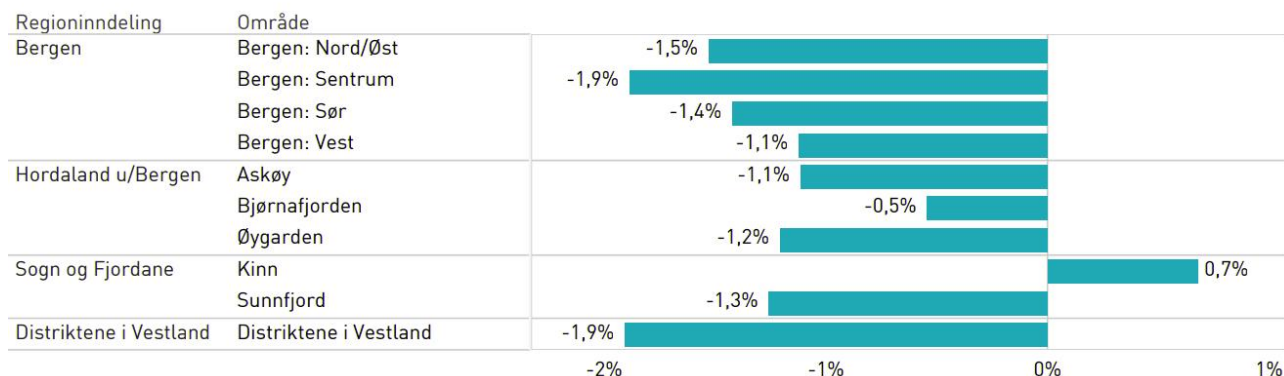
Som en del av vår studie, utførte vi også analyser av datasettet både før og etter strømpriskrisen som fant sted vinteren 2021-2022, og er fortsatt i effekt per i dag. Det var vår intensjon å undersøke potensielle endringer i kvadratmeterpris som kan ha oppstått som følge av denne betydelige hendelsen. Som illustrert i appendiks A3, har strømprisene vært ekstremt høye og er fortsatt unormalt høye i forhold til perioden før strømpriskrisen. Dette fenomenet, i lys av EUs nylig foreslåtte innføring av minstekrav til energikarakterer – et mandat som tilsier at alle boliger i nasjonen skal oppnå minimum karakter F innen 2030 og karakter E innen 2033 (Olje- og Energidepartementet, 2023) – burde intuitivt føre til at boligkjøpere legger større vekt på boligens energikarakter. For å undersøke dette, formulerte vi to hypoteser:

1. Nullhypotesen  $H_0$ : Energimerkingens påvirkning på boligprisene er ikke statistisk signifikant, hverken i perioden før og etter strømpriskrisen inntraff i 2021-2022
2. Alternativhypotesen  $H_1$ : Energimerkingens påvirkning på boligprisene er statistisk signifikant, både i perioden før og etter strømpriskrisen inntraff i 2021-2022

*Før strømpriskrisen* ble det observert at den implisitte prisen på energikarakterer ikke var statistisk signifikant. Dette antyder at energikarakteren ikke hadde en merkbar innvirkning på kvadratmeterprisen på boligene i vårt datasett. Dette resultatet kan tolkes som at forbrukerne ikke verdsetter energieffektivitet i samme grad, eller at informasjonen formidlet gjennom energikarakteren ikke ble oppfattet som relevant eller pålitelig.

**Figur 9.** Prisutviklingen i siste kvartal 2021

Kvartalsvis prisutvikling, utvalgte områder.



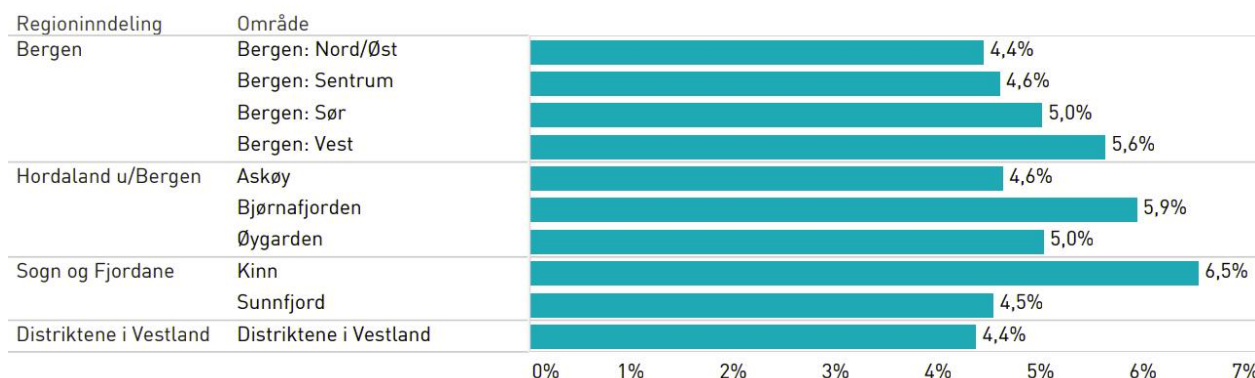
Fra «Eiendom Norge regionsrapporter K4 2021», av Eiendom Norge, 2021.

(<https://eiendommenorge.no/boligprisstatistikk/statistikkbank/rapporter/regionsrapporter/vestlandet/?article=2198#filesDownloadElement> ).

*Etter strømpriskrisen* endret imidlertid situasjonen seg. Vi observerte et skifte i sammenhengen mellom energikarakterer og kvadratmeterpris på boliger, hvor energikarakterene gikk ifra å ha negative koeffisienter, til å få positive koeffisienter. Det betyr at energimerking fikk et større positivt innslag på boligprisen per kvadratmeter. Imidlertid er det viktig å merke seg at observasjonene etter strømpriskrisen er for få til å kunne trekke klare konklusjoner. Datasettet etter strømpriskrisen inneholdt et for begrenset antall observasjoner til å gi en robust og pålitelig analyse. Selv om det er en antydning til en endring i mønsteret, oppnår vi ikke signifikant evidens på grunn av det begrensede antallet observasjoner. Dette er noe som kan være verdifullt å studere videre litt frem i tid med flere observasjoner og etter endringer i markedet. Vi erkjenner imidlertid betydningen av å forstå virkningen av slike hendelser på boligmarkedet og oppfordrer til videre forskning når et mer omfattende datasett blir tilgjengelig.

**Figur 10.** Prisutviklingen i første kvartal 2022.

Kvartalsvis prisutvikling, utvalgte områder.



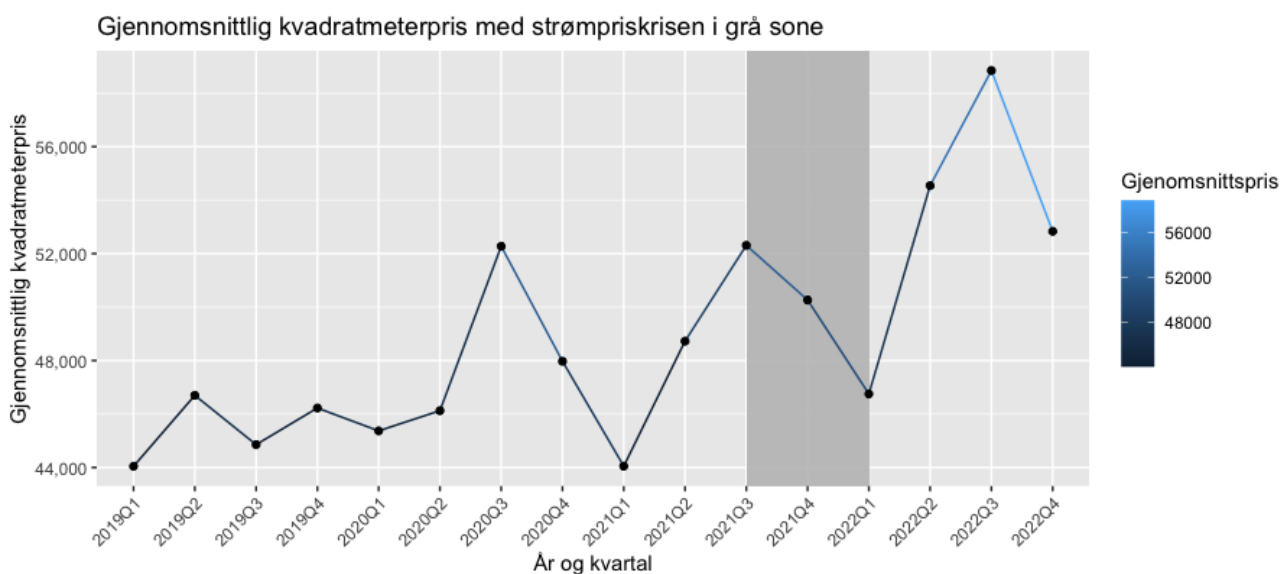
Fra «Eiendom Norge regionsrapporter K1 2022», av Eiendom Norge, 2021.

([https://eiendommorge.no/boligprisstatistikk/statistikkbank/rapporter/regionsrapporter/vestlandet/?article=2227#filesDownloadElement\\_\\_](https://eiendommorge.no/boligprisstatistikk/statistikkbank/rapporter/regionsrapporter/vestlandet/?article=2227#filesDownloadElement__)).

Disse funnene understreker behovet for videre forskning på dette området, spesielt i lys av økende strømpriser og den pågående diskusjonen om energieffektivitet i boliger. Den potensielle innflytelsen fra energikarakterer på boligpriser kan ha betydelige implikasjoner for både boligmarkedet og energipolitikken.

I våre analyser observerer vi en markert nedgang i kvadratmeterprisen når strømpriskrisen inntreffer i fjerde kvartal av 2021. Dette mønsteret endrer seg imidlertid sterkt etter første kvartal av 2022, hvor vi ser en betydelig økning i kvadratmeterprisen. Disse observasjonene er i samsvar med dataene presentert av Eiendom Norge.

**Figur 11.** Gjennomsnittlig kvadratmeterpris med strømpriskrisen



Basert på disse funnene, støtter våre data nullhypotesen ( $H_0$ ), men gir ikke tilstrekkelig bevis for å støtte alternativhypotesen ( $H_1$ ). Vi kan derfor ikke konkludere med at energimerking har en signifikant påvirkning på pris, hverken før eller etter strømpriskrisen. Vi beholder da  $H_0$ .

En viktig del av analysen var å undersøke om det var *strukturelle brudd* i datasettene våre, det vil si punkter der de underliggende mønstrene i dataene endret seg markant. Dette ble først og fremst gjort ved en t-test for å undersøke forskjeller mellom datasettene for kvadratmeterprisene før og etter strømpriskrisen. Resultatene viste statistisk signifikant forskjell mellom gjennomsnittet i de to datasettene ( $t = 9.9562$ ,  $df = 712.2$ ,  $p\text{-verdi} < 0.001$ ), som antyder betydelig forskjell i kvadratmeterpris før og etter krisen. For å forstå dynamikken i disse endringene mer detaljert, undersøkte vi også muligheten for strukturelle brudd i dataene ved hjelp av Chow-testen i RStudio. Den identifiserer to mulige strukturelle brudd før strømpriskrisen, ved observasjonene 182 og 799, henholdsvis i fjerde kvartal i 2019 og andre kvartal i 2021. Dette tyder på at det også var betydelige endringer i prisene i disse periodene, selv før krisen begynte. Etter krisen inntraff ble det identifisert et mulig brudd ved observasjon 295 i Q4 i 2022, noe som indikerer en betydelig endring i kvadratmeterprisene etter dette punktet.

Disse funnene kan indikere at strømpriskrisen kan ha hatt en betydelig innvirkning på kvadratmeterprisene, men de viser også at det har vært andre viktige endringer i markedet, selv før krisen. Dette understreker kompleksiteten i markedet, og viktigheten av å ta hensyn til flere forskjellige faktorer når man analyserer kvadratmeterprisen på boliger.

## 8.0 Konklusjon

Den sentrale problemstillingen vi har undersøkt i foreliggende utredning er:

### **Bidrar energimerking til å forklare variasjoner i boligpris, og hvor stor er i så tilfelle den økonomiske betydningen av energimerking?**

Analysen vår, lik Olausson et al. (2017), antyder at energimerkingens innvirkning på kvadratmeterprisen er inkonklusiv, da energikarakterene generelt ikke har noen statistisk signifikant effekt, mens oppvarmingskarakterene ofte viser signifikant effekter. Dette kan skyldes kjøpernes oppfatning av energieffektivitet, som ser ut til å være mer knyttet til oppvarmingskarakteren. Resultatene våre konsoliderer også tendensen som ble fremhevet av Lainé (2011), hvor andre faktorer blir vektlagt mer enn energikarakterer.

Til tross for tidligere studier som indikerer en signifikant korrelasjon mellom energimerker og kvadratmeterpris, avdekker vår anvendelse av OLS-modellen ingen signifikant sammenheng mellom disse variablene. OLS-modellen viser en statistisk signifikant korrelasjon mellom kvadratmeterprisen og oppvarmingskarakterer, hvor kvadratmeterprisen synker ved dårligere oppvarmingskarakterer. Den lysegrønne karakteren utmerker seg, ved å ha en statistisk signifikant positiv effekt på kvadratmeterprisen i Bergen kommune, sammenlignet med referansekategorien grønn. Disse funnene fremstår som statistisk signifikante, og divergerer fra den eksisterende litteraturen på feltet. Imidlertid tilfredsstillers ikke denne modellen vårt formål om å undersøke de romlige effektene.

Basert på resultatene fra våre undersøkelser, ser det ut til at spatial error model (SEM) er den mest hensiktsmessige modellen, på grunn av dens enkelhet og gode statistiske tilpasninger. Våre analyser indikerer at energimerking spiller liten til ingen statistisk rolle for boligprisene når man hensyntar romlige effekter. Derimot peker vårt studie på oppvarmingskarakterer som en signifikant betydning for boligpriser, både før og etter strømpriskrisen inntraff. Dette peker på at andre faktorer blir mer vektlagt ved boligkjøp, fremfor energikarakterer, som samstemmer med funnene til Lainé (2011).

Innledningsvis ønsket vi også å differensiere på bakgrunn av boligtyper og geografisk område, men manglende data har hindret oss fra å gjøre dette. Dette kan ha hatt en innvirkning på resultatene av analysen. Tilgang til mer informasjon kunne ha gitt oss en mer presis forståelse av hvordan energimerking påvirker boligpriser.

Svaret på våres problemstilling og forskningsspørsmål er med andre ord kompleks og sammensatt. Resultatene fra analysen trekker i retning av at energimerking har liten innvirkning på boligprisen, mens oppvarmingskarakterer har større utslagsgivende betydning for kvadratmeterpris enn energikarakteren i energimerket.



## 9.0 Videre forskning

Vi erkjenner at en masteroppgave ikke vil være fullstendig tilstrekkelig til å gi omfattende og definitive svar og konklusjoner. Snarere vil det heller tjene som et verdifullt bidrag til eksisterende kunnskap, og forhåpentligvis fungere som en katalysator for videre forskning på feltet.

Masteroppgaven søker å belyse sammenhengen mellom energimerking og boligprisen for å styrke forståelsen til kjøpere og selgere. Vi møtte på problemstillinger, primært knyttet til utilgjengelig data og manglende systemkommunikasjon. Bedre datatilgang og samarbeid vil være gunstig for fremtidig forskning.

En begrensningene i vår oppgave er et lavt antall observasjoner i perioden før og etter strømpriskrisen inntraff. Fortegnsskift i koeffisienter for energimerker kan indikere fremtidig påvirkning på pris etter strømpriskrisen, gitt flere observasjoner og markedstilpasning.

Videre forskning kan vurdere om våre resultater gjelder i andre norske byer. Vi fokuserte på Bergen, men det vil være interessant å undersøke byer som Oslo, Stavanger, Haugesund og Trondheim, særlig i Nord-Norge hvor det ofte er kaldere enn Sør-Norge i vinterhalvåret.

En interessant undersøkelse kunne være av variasjon i energikarakter mellom forskjellige bydeler. Selv om vår analyse har inkludert Bergensregionen, og Bergen by som en helhet, kunne en studie designet for å evaluere signifikante forskjeller mellom individuelle bydeler gi innsiktsfulle funn.

Vårt forskningsdesign hadde imidlertid begrensninger i differensiering mellom forskjellige typer boliger, hovedsakelig på grunn av mangel på adekvat data. Vi oppfordrer derfor fremtidige studier til å ta høyde for dette aspektet, som krever tilgang til mer detaljerte datasett. En slik utforskning vil nødvendigvis gjøre samarbeid med leverandører som er villige til å gi tilgang til datasett med tilstrekkelig detaljnivå for å kunne utvinne relevante funn.

Det er også verdifullt å undersøke hvordan boligeieres og potensielle kjøperes vurderinger av energimerking påvirker deres beslutningsprosess ved boligkjøp. Ved å kartlegge innvirkningen dette har på boligpriser, kan man få en mer nyansert forståelse av markedskreftenes prispåvirkning.

En studie av hvordan ulike typer boliger, som eneboliger, leiligheter og rekkehus, responderer forskjellig på energimerking i forhold til deres pris, vil gi en dypere forståelse av sammenhengen mellom boligtype, energimerking og pris. Ved å forstå disse sammenhengene i detalj, kan forskere bidra til en mer komplett forståelse av effekten av energimerking på boligprisene.

## Referanser

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6), 716-723.
- Alonso, A. (1964): *Location and Land Use. Toward a General Theory of Land Rent*. Harvard University Press. Cambridge, Massachusetts.
- Altman, N. S. (1992). «An introduction to kernel and nearest-neighbor nonparametric regression». *The American Statistician*. 46 (3): 175–185
- Ann-Kristin Loodtz. (2022, 12. mars). Befolkning. Bergen kommune.  
<https://www.bergen.kommune.no/omkommunen/fakta-om-bergen/befolkning/befolkning>
- Ann-Kristin Loodtz. (2022, 12. mars). Folketall per 1. januar 2022. Bergen kommune.  
<https://www.bergen.kommune.no/omkommunen/fakta-om-bergen/befolkning/folketall-per-1-januar-2022>
- Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Dordrecht.  
[https://doi.org/10.1007/978-94-015-7799-1\\_2](https://doi.org/10.1007/978-94-015-7799-1_2)
- Anselin, L. (2002). Under the hood issues in the specification and interpretation of spatial regression models. *Agricultural economics*, 27(3), 247-267.
- Askøy kommune. (2020). *Statistikk og utviklingstrekk for Askøy: Kunnskapsgrunnlag for kommunal planstrategi og folkehelseoversikt 2020-2024*.  
<https://askoy.kommune.no/plan-og-utvikling/858-kunnskapsgrunnlag-til-kommunal-planstrategi-19-11-20-1/file>
- Aunedi, M., Ortega, J.E.C. and Green, T.C. 2022. Benefits of flexibility of Smart Local Energy Systems in supporting national decarbonisation. *Energy Revolution Research Centre*, Strathclyde, UK. University of Strathclyde Publishing.
- Backe, T., G. (2020, 26. august). Dette bruker vi mest strøm på i Norge. Fortumbloggen.  
<https://blogg.fortum.no/dette-bruker-vi-mest-str%C3%B8m-p%C3%A5-i-norge>
- Bergen kommune. (2023, 16. april). *Grønn strategi - Klima og energihandlingsplan for Bergen 2015*. <https://www.bergen.kommune.no/api/rest/filer/V304556>

- Bergen kommune. (2023, 16. april). *Grønn strategi - Klimastrategi for Bergen 2022 - 2030*.  
<https://www.bergen.kommune.no/politikere-utvalg/api/fil/bk360/7186170/Gronn-strategi-Klimastrategi-for-Bergen-2022-2030>
- Bergen kommune. (2023, 19. mai). *Tilskudd til varmepumpe i privat bolig*. Innbyggerhjelpen.  
<https://www.bergen.kommune.no/innbyggerhjelpen/natur-klima-miljo/miljo-/klimateilskudd-til-varmepumpe-i-privat-bolig>
- Bishop, K. C., Kuminoff, N. V., Banzhaf, H. S., Boyle, K. J., von Gravenitz, K., Pope, J. C., ... & Timmins, C. D. (2020). Best practices for using hedonic property value models to measure willingness to pay for environmental quality. *Review of Environmental Economics and Policy*.
- Bøeng, Ann., Christin. (2014). På verdenstoppen i bruk av strøm. *Samfunnsspeilet*, 4/2014, 9.
- Brasington, D. M., & Hite, D. (2005). Demand for environmental quality: a spatial hedonic analysis. *Regional science and urban economics*, 35(1), 57-82.
- Brounen, D., & Kok, N. (2011). On the economics of energy labels in the housing market. *Journal of environmental economics and management*, 62(2), 166-179.
- Brueckner, J. K. (2003). Strategic interaction among governments: An overview of empirical studies. *International regional science review*, 26(2), 175-188.
- Byfuglien, J., Holm, S. (1988). *Inndeling i Arbeidsmarkedsregioner 1988*. SSB.  
[https://www.ssb.no/a/histstat/in/in\\_8915.pdf](https://www.ssb.no/a/histstat/in/in_8915.pdf)
- Chi, G., & Zhu, J. (2008). Spatial regression models for demographic analysis. *Population Research and Policy Review*, 27, 17-42.
- Chi, G., & Zhu, J. (2019). *Spatial regression models for the social sciences*. SAGE publications.
- Chi, G., & Zhu, J. (2020). *Spatial regression models for the social sciences*. SAGE publications.
- Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica: Journal of the Econometric Society*, 591-605.
- Christophersen, Knut-Andreas. (2018). *Introduksjon til statistisk analyse*. Gyldendal Akademisk.
- Clarke, K. A. (2005). The phantom menace: Omitted variable bias in econometric research. *Conflict*

*management and peace science*, 22(4), 341-352.

Curtin, J., & Maguire, J. (2011). *Thinking deeper: Financing options for home retrofit*. IIEA.

Day, B. H. (2001). *The theory of hedonic markets: obtaining welfare measures for changes in environmental quality using hedonic market data*. Centre for Social and Economic Research on the Global Environment (CSERGE): London, Storbritannia.

Eikemo, T.A. & Clausen, T.H. (red.) (2012) *Kvantitativ analyse med SPSS. En praktisk innføring i kvantitative analyseteknikker*. (2. utg). Tapir Akademiske Forlag. Trondheim.

Elhorst, J. P. (2010). Applied spatial econometrics: raising the bar. *Spatial economic analysis*, 5(1), 9-28.

Elhorst, J. P., & Halleck Vega, S. (2017). The SLX model: extensions and the sensitivity of spatial spillovers to W. *Papeles de Economía Española*, 152, 34-50.

Enova SF. (2009, 08. oktober). *Om energimerkeordningen*. Energimerking.no.

<https://www.energimerking.no/no/energimerking-bygg/om-energimerkesystemet-og-regelverket/>

Enova SF. (2009, 08. oktober). *Publikasjoner, presentasjoner og dokumenter*. Energimerking.no.

<https://www.energimerking.no/no/energimerking-bygg/om-energimerkesystemet-og-regelverket/publikasjoner-og-dokumenter/>

Enova SF. (2009, 13. oktober). *Karakterskalaen*. Energimerking.no.

<https://www.energimerking.no/no/energimerking-bygg/om-energimerkesystemet-og-regelverket/karakterskalaen/>

Fernández-Avilés, G., Minguez, R., & Montero, J. M. (2012). Geostatistical air pollution indexes in spatial hedonic models: the case of Madrid, Spain. *Journal of Real Estate Research*, 34(2), 243-274.

Fischer, M. M., Bartkowska, M., Riedl, A., Sardadvar, S., & Kunnert, A. (2009). The impact of human capital on regional labor productivity in Europe. *Letters in Spatial and Resource Sciences*, 2, 97-108.

Florax, R. J. G. M. & Van Der Vlist, A. J. (2003): Spatial econometric data analysis, moving beyond traditional models. *International Regional Science Review* 26(3), 223 – 243.

- Fugleberg, O & Kristianslund, I. (1995). *Innføring i regresjonsanalyse og multivariate metoder*. Bedriftsøkonomenes Forlag A/S.
- Galster, G. (2003). Neighbourhood dynamics and housing markets. I T. O'Sullivan & K. Gibb (Eds.), *Housing economics and public policy* (Kap. 9, s. 153–171). Oxford: Blackwell Science.
- Gjestland, A., McArthur, D., Osland, L., & Thorsen, I. (2012). 10. A bridge over troubled waters: valuing accessibility effects of a new bridge. *Accessibility Analysis and Transport Planning: Challenges for Europe and North America*, 173.
- Goodman, A. C., & Thibodeau, T. G. (1995). Age-related heteroskedasticity in hedonic house price equations. *Journal of Housing Research*, 25-42.
- Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *European journal of epidemiology*, 31, 337-350.
- Gripsrud, G., Olsson, U. H., & Silkoset, R. (2010). *Metode og dataanalyse* (2. utg.). Cappelen Damm.
- Gripsrud, G., Olsson, U. H., og Silkoset, R. (2004). *Metode og dataanalyse—med fokus på beslutninger i bedrifter*. Cappelen Damm.
- Gundersen, F., Juvkam, D. (2013). *Inndelinger i senterstruktur, sentralitet og BA-regioner* (NIBR-rapport 2013:1). Norsk institutt for by- og regionforskning.  
[https://www.regjeringen.no/globalassets/upload/krd/rapporter/rapporter2013/nibrrapport\\_2013\\_1.pdf](https://www.regjeringen.no/globalassets/upload/krd/rapporter/rapporter2013/nibrrapport_2013_1.pdf)
- Haining, R. (2003). *Spatial data analysis: Theory and practice*. Cambridge: Cambridge University Press.
- Halleck Vega, S., & Elhorst, J. P. (2015). The SLX model. *Journal of Regional Science*, 55(3), 339-363.
- Hårsman, B. (1981). Housing Demand Models and Housing Market Models for Regional and Local Planning. *Swedish Council for Building Research*, Document D13, Stockholm.
- Hill, R. C., Griffiths, W. E., Judge, G. G., & Reiman, M. A. (2001). *Undergraduate econometrics*

(Vol. 4). New York: Wiley.

Islam, S. N., & Winkel, J. (2017). Climate Change and Social Inequality. *Department of Economic & Social Affairs*, (DESA Working Paper No. 152), 1-2.

Jacobsen, D. I. (2005). *Hvordan gjennomføre undersøkelser? Innføring i samfunnsvitenskapelige metode* (2. utg.). Høyskoleforlaget.

Jensen, A. (2022, mai). *Markedsverdien av energieffektive boliger*. Eiendomsverdi.

Johannessen, A., Christoffersen, L. & Tufte, P. A. (2020). *Forskningsmetode for økonomisk-administrative fag*. (4.utg). Abstrakt forlag AS.

Juvkam, D. (2002). *Inndeling i bo- og arbeidsmarkedsregioner* (NIBR-rapport 2002:20). Norsk institutt for by- og regionforskning.

<https://distriktssenteret.no/wp-content/uploads/2021/04/Inndeling-i-bo-og-arbeidsmarkedsregioner.pdf>

Karlberg, B., & Victorin, A. (2004). *Housing tenures in the Nordic countries. Housing and housing policy in the Nordic countries*, 7, 57-78.

Kartverket. (2023, 15. januar). *Ofte stilte spørsmål*. Se eiendom. <https://seeiendom.kartverket.no/>

Kononuchenko, V. (2011). *Energieffektivisering av næringsbygg og konsekvenser av energimerking* (Masteroppgave, Norges miljø- og biovitenskapelige universitet, Ås). Brage NMBU vitenarkiv. <https://nmbu.brage.unit.no/nmbu-xmlui/handle/11250/188200>

Kritzman, M. (1994). What practitioners need to know... about hypothesis testing. *Financial Analysts Journal*, 50(4), 18-22.

Krogsveen. (2023, 13., april). *Prisutvikling for Bergen*. Boligprisstatistikk. <https://www.krogsveen.no/prisstatistikk/bergen-monthly>

Laakso, S. (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. *ETLA A*.

Lainé, L., & Focus, C. (2011). *As easy as EPC? Consumer views on the content and format of the energy performance certificate*. Consumer Focus. United Kingdom.

- LeSage, J. P. (1997). Regression analysis of spatial data. *Journal of Regional Analysis and Policy*, 27(1100-2016-89650), 83-94.
- LeSage, J. P. (2008). An introduction to spatial econometrics. *Revue d'économie industrielle*, (123), 19-44.
- LeSage, J. P. (2014). What regional scientists need to know about spatial econometrics.  
<https://ssrn.com/abstract=2420725>
- LeSage, J. P., & Fischer, M. M. (2008). Spatial growth regressions: model specification, estimation and interpretation. *Spatial Economic Analysis*, 3(3), 275-304.
- LeSage, J. P., & Pace, R. K. (2006). Interpreting Spatial Econometric Models [Paper presentasjon]. Regional Science Association International North American meetings, Toronto, CA.
- LeSage, J., & Pace, R. K. (2009). *Introduction to spatial econometrics*. Chapman and Hall/CRC.
- LeSage, J., & Pace, R. K. (2014). Interpreting Spatial Econometric Models. Fischer, M., Nijkamp, P. (red) *Handbook of Regional Science*. Springer, Berlin, Heidelberg.  
[https://doi.org/10.1007/978-3-642-23430-9\\_91](https://doi.org/10.1007/978-3-642-23430-9_91)
- Lincoln, Y. S. og Guba, E. G. (1985). *Naturalistic Inquiry*. SAGE Publications
- Lundberg, J. (2006). *Spatial interaction model of spillovers from locally provided public services*. *Regional Studies*, 40(6), 631-644.
- Malpezzi, S. (2002). *Hedonic Pricing Models: A Selective and Applied Review*. The Center for Urban Land Economics Research.  
[https://web.archive.org/web/20160221230431id\\_/http://down.cenet.org.cn:80/upfile/49/20072137445140.pdf](https://web.archive.org/web/20160221230431id_/http://down.cenet.org.cn:80/upfile/49/20072137445140.pdf)
- Malpezzi, S., Ozanne, L., & Thibodeau, T. G. (1987). Microeconomic estimates of housing depreciation. *Land Economics*, 63(4), 372-385.
- Maslianskaia-Pautrel\*, M., & Baumont †, C. (2016). Environmental spillovers and their impacts on housing prices: A spatial hedonic analysis. *Revue d'économie politique*, (5), 921-945.  
<https://doi.org/10.3917/redp.265.0921>
- Mauro, R. (1990). Understanding L.O.V.E. (left out variables error): A method for estimating the effects of omitted variables. *Psychological Bulletin*, 108(2), 314–329.

<https://doi.org/10.1037/0033-2909.108.2.314>

Meld. St. 18 (2016–2017) Melding til Stortinget. Berekraftige byar og sterke distrikt.

<https://www.regjeringen.no/contentassets/d15e0f19e7d9439ea5c1b08ba888bdd1/nn-no/pdfs/stm201620170018000dddpdfs.pdf>

Meld. St. 5 (2019–2020). Levende lokalsamfunn for fremtiden— Distriktsmeldingen. Kommunal- og distriktsdepartementet.

<https://www.regjeringen.no/no/dokumenter/meld.-st.-5-20192020/id2674349/?ch=3>

Millar, R. B. (2011). *Maximum likelihood estimation and inference*. John Wiley

Moran, P. A. (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society. Series B (Methodological)*, 10(2), 243-251.

Moses, J. W., & Knutsen, T. L. (2019). *Ways of knowing: Competing methodologies in social and political research* (3. utg.). Red Globe Press.

Olaussen, J. O., Oust, A., & Solstad, J. T. (2017). Energy performance certificates—Informing the informed or the indifferent?. *Energy Policy*, 111, 246-254.

Olaussen, J. O., Oust, A., & Solstad, J. T. (2021). Real Estate Price Formation: Energy Performance Certificates and the Role of Real Estate Agents. *Journal of Sustainable Real Estate*, 13(1), 1-11.

Olje- og energidepartementet. (2023). *Forslag til revidert direktiv for bygningers energiytelse* (2021/802).

<https://www.regjeringen.no/no/sub/eos-notatbasen/notatene/2022/des/forslag-til-revidert-bygningsenergidirektiv/id2959442/>

Olje- og energidepartementet. (2023, April 5). 32,6 milliardar kroner I Straumstøtte i 2022.

Regjeringa.no. Hentet 05. Mai, 2023, fra

<https://www.regjeringen.no/no/aktuelt/326-milliardar-kroner-i-straumstotte-i-2022/id2970938/>

Osland, L. (2001). Den hedonistiske metoden og estimering av attributtpriser. *Norsk økonomisk tidsskrift*, 115(1), 1-22.

Osland, L. (2010). An Application of Spatial Econometrics in Relation to Hedonic House Price



Modeling. *Journal of Real Estate Research*, 32(3), s. 303.

<https://doi.org/10.1080/10835547.2010.12091282>

Osland, L. (2013). The importance of unobserved attributes in hedonic house price models. *International Journal of Housing Markets and Analysis*.

Ot.prp. nr. 24 (2008-2009). *Om lov om endringer i lov 29. juni 1990 nr. 50 om produksjon, omforming, overføring, omsetning, fordeling og bruk av energi m.m. (energiloven)*. Olje- og energidepartementet.

<https://www.regjeringen.no/no/dokumenter/otprp-nr-24-2008-2009-/id540753/?ch=3>

Oxley, M. (2004). *Economics, planning and housing*. Macmillan International Higher Education

Paelinck, J. (1978). Spatial econometrics. *Economics Letters*, 1(1), 59-63.

Persky, J. (1995). Retrospectives: The ethology of homo economicus. *Journal of Economic Perspectives*, 9(2), 221-231.

Quigley, J.M. (1982). *Nonlinear Budget Constraints and Consumer Demand: An Application to Public Programs for Residential Housing*. *Journal of Urban Economic* 12, 177-201.

Revelli, F. (2006). Spatial interactions among governments. *Handbook of fiscal federalism*, 106-130.

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy*, 82(1), 34-55.

Rubin, G. M. (1993). Is housing age a commodity? Hedonic price estimates of unit age. *Journal of Housing Research*, 165-184.

Sandlie, H. C., & Sørvoll, J. (2017). Et velfungerende leiemarked?. *Tidsskrift for velferdsforskning*, 20(1), 45-59.

Saunders, M., Lewis, P. og Thornhill, A. (2009). *Research Methods for Business Students* (5. utg.). Pearson Education Limited.

Saunders, M., Lewis, P. og Thornhill, A. (2012). *Research methods for business students* (6. utg.). Pearson Education Limited.

Saunders, M., Lewis, P. & Thornhill, A. (2019). *Research Methods for Business Students* (8. utg.).

Pearson Education Limited.

Sem, I. V., Hole, J., Aulie, F. H., Spilde, D., Arnesen, F., & Jelsness, S. (2020). *Langsiktig kraftmarkedsanalyse 2020-2040*. NVE, Tech. Rep. 37. 29

Skog, O.-J. (2004). *Å forklare sosiale fenomener: En regresjonsbasert tilnærming* (2. utg.). Gyldendal Norsk Forlag.

Statens vegvesen. (2023, 18. april). *Rv. 555 Sotrasambandet*. Veiprosjekter.  
<https://www.vegvesen.no/vegprosjekter/riksveg/sotrabergen/>

Statistisk sentralbyrå. (2023). *04861: Areal og befolkning i tettsteder (K) 2000 - 2022*.  
<https://www.ssb.no/statbank/table/04861/>

Statistisk sentralbyrå. (2023). *07221: Prisindeks for brukte boliger, etter statistikkvariabel, region, boligtype og kvartal*. <https://www.ssb.no/statbank/table/07221/>

Statistisk sentralbyrå. (2023). *09817: Innvandrere og norskfødte med innvandrerforeldre*.  
<https://www.ssb.no/statbank/table/09817/>

Statistisk sentralbyrå. (2023). *Tettsted. Variabeldefinisjon*.  
<https://www.ssb.no/a/metadata/conceptvariable/vardok/141/nb>

Statistisk sentralbyrå. (2023, januar). *Boforhold, registerbasert*.  
<https://www.ssb.no/bygg-bolig-og-eiendom/bolig-og-boforhold/statistikk/boforhold-registerbasert>

Sundelius, D. M. (2019, mai). *Markedsverdien av energieffektive boliger*. Eiendomsverdi.

Tangeland, T., Alfnes, F., Throne-Holst, H., & Dulsrud, A. (2022). (rep.). *Bærekraftig energiforbruk: Forbrukermobilitet og-fleksibilitet i strømmarkedet*. Forbruksforskningsinstituttet SIFO.  
<https://oda.oslomet.no/oda-xmlui/bitstream/handle/11250/3052026/SIFO-rapport%2013-2022%20BFD%20B%C3%A6rekraftig%20energiforbruk.pdf?>

Vollset, M. (2007). *På sporet av bybanen*. Bodini Forlag

Ward, M. D., & K. S. Gleditsch. (2008). *Spatial regression models*. Vol. 07-155. Los Angeles, CA: Sage.

Weisstein, Eric W. (1995). Estimator. Wolfram MathWorld.

<https://mathworld.wolfram.com/Estimator.html>

Wooldridge, J. M. (2014). *Introduction to Econometrics* (Europe, Middle East and Africa).

Andover: Cengage Learning.

Wooldridge, J. M. (2019). *Introductory Econometrics. A Modern Approach* (7. utg.). Cengage

Learning.

Wooldridge, J.M. (2006). *Introductory Econometrics. A Modern Approach* (3. utg.)

Thomson/South-Western, Mason.

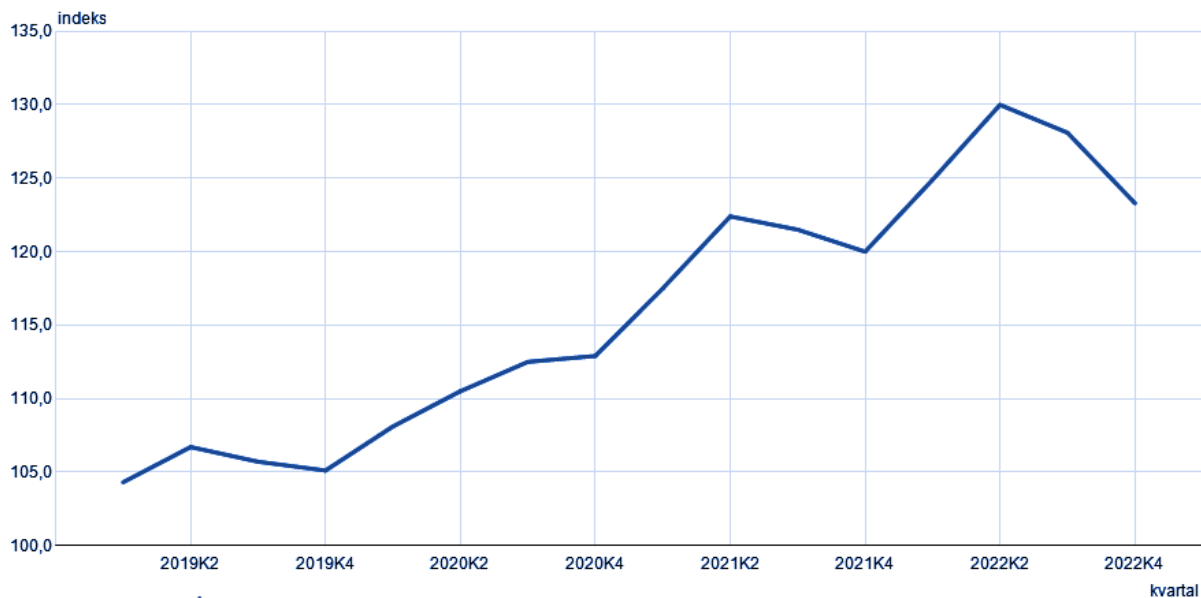
Øygarden kommune. (2022, 10. august). Om oss. Øygarden kommune.

<https://www.oygarden.kommune.no/politikk-og-organisasjon/organisasjon/om-oss>

## Appendiks

### A1. 07221: Prisindeks for brukte boliger, etter kvartal.

07221: Prisindeks for brukte boliger, etter kvartal. Prisindeks for brukte boliger, Bergen, Boliger i alt.



### A2. 09817: Innvandrere uten norskfødte med innvandrerforeldre

09817: Innvandrere uten norskfødte med innvandrerforeldre		2020	2021	2022
4601 Bergen	Andel av befolkningen (prosent)	15,18	15,23	15,24
4623 Samnanger	Andel av befolkningen (prosent)	8,37	8,23	8,52
4624 Bjørnafjorden	Andel av befolkningen (prosent)	11,04	11,01	11,05
4626 Øygarden	Andel av befolkningen (prosent)	10,24	10,54	10,64
4627 Askøy	Andel av befolkningen (prosent)	9,11	9,29	9,49
4628 Vaksdal	Andel av befolkningen (prosent)	9,48	9,37	9,54
4630 Osterøy	Andel av befolkningen (prosent)	9,30	9,62	9,83
4631 Alver	Andel av befolkningen (prosent)	8,78	8,95	9,10
4632 Austrheim	Andel av befolkningen (prosent)	12,47	11,57	11,39

### A3. 09364: Kraftpriser i sluttbrukermarkedet

	Kraftpris ekskl. avgifter (øre/kWh)	Kraftpris inkl. mva (øre/kWh)
2019K1	55,2	67,4
2019K2	45,6	55,7
2019K3	42,5	51,9
2019K4	46,8	57,2
2020K1	26,8	32,7
2020K2	14,9	18,2
2020K3	14,4	17,6
2020K4	22,2	27,1
2021K1	53,5	65,3
2021K2	50,9	62,2
2021K3	76,3	93,1
2021K4	108,1	132,0
2022K1	117,2	143,1
2022K2	124,2	151,6
2022K3	214,3	261,6
2022K4	169,6	207,0

### A4. 03014: Konsumprisindeks, total

TOTAL Totalindeks	Årsendring (prosent)				
	2018	2019	2020	2021	2022
	2,7	2,2	1,3	3,5	5,8

## A5. Gauss-Markov-teoremet - BLUE

Vi ønsker å oppfylle Gauss-Markovs teorems forutsetninger ved bruk av OLS, og disse vil nå bli presentert sekvensielt.

1. *Linearitet i parameterne.* I modellen er, den avhengige variabelen  $P$ , relatert til de uavhengige variabel  $X_k$  og feilleddet  $u$ ,

$$P = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + u,$$

hvor  $\beta_0, \beta_1, \dots, \beta_k$  er de ukjente parameterne av interesse, og  $u$  er et uobservert tilfeldig feilledd. Med andre ord, er forholdet mellom den avhengige og uavhengige variabelen lineær.

2. *Tilfeldig utvalg.* Det er et tilfeldig utvalg av størrelsen  $n$ ;

$$\{(x_{i1}, x_{i2}, \dots, x_{ik}, y_i) : i = 1, 2, \dots, n\}.$$

Observasjonene skal være statistisk uavhengige av hverandre.

3. *Ingen perfekt korrelasjon.* I utvalget er ingen av de uavhengige variablene konstant og det er ingen eksakt lineær sammenheng mellom de.
4. *Null-betinget gjennomsnitt.* Feilleddet  $u$  har forventet verdi lik 0, gitt alle mulige verdier av uavhengige variabler. Med andre ord;

$$E(u/x_1, x_2, \dots, x_k) = 0.$$

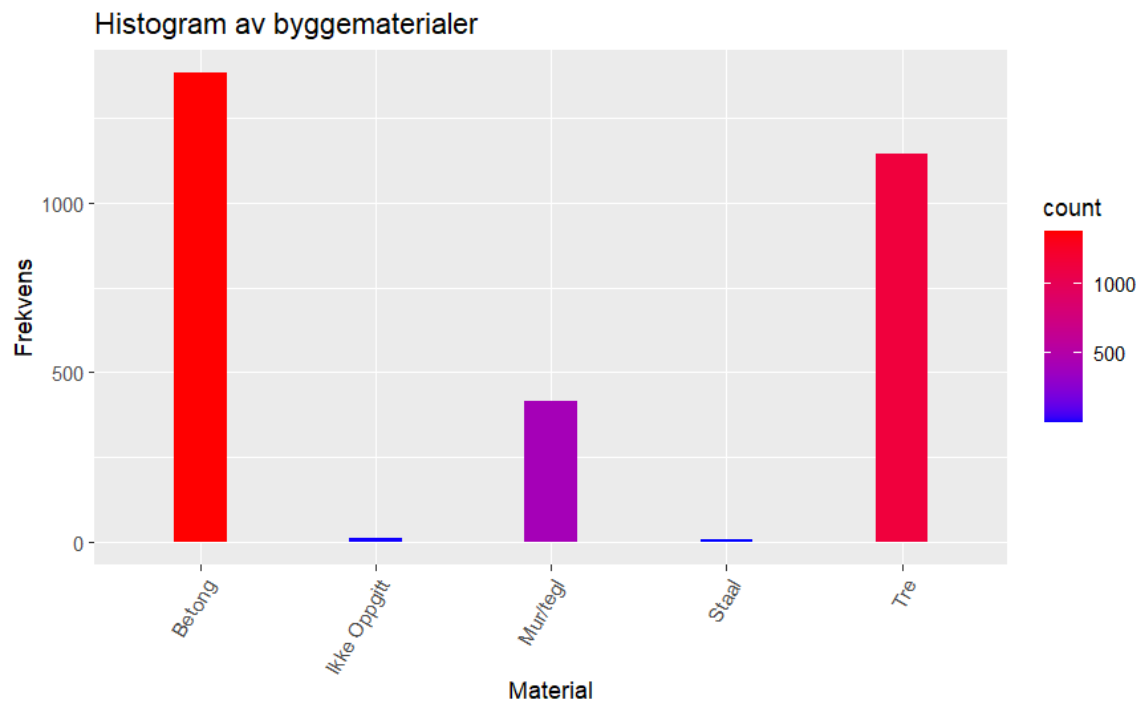
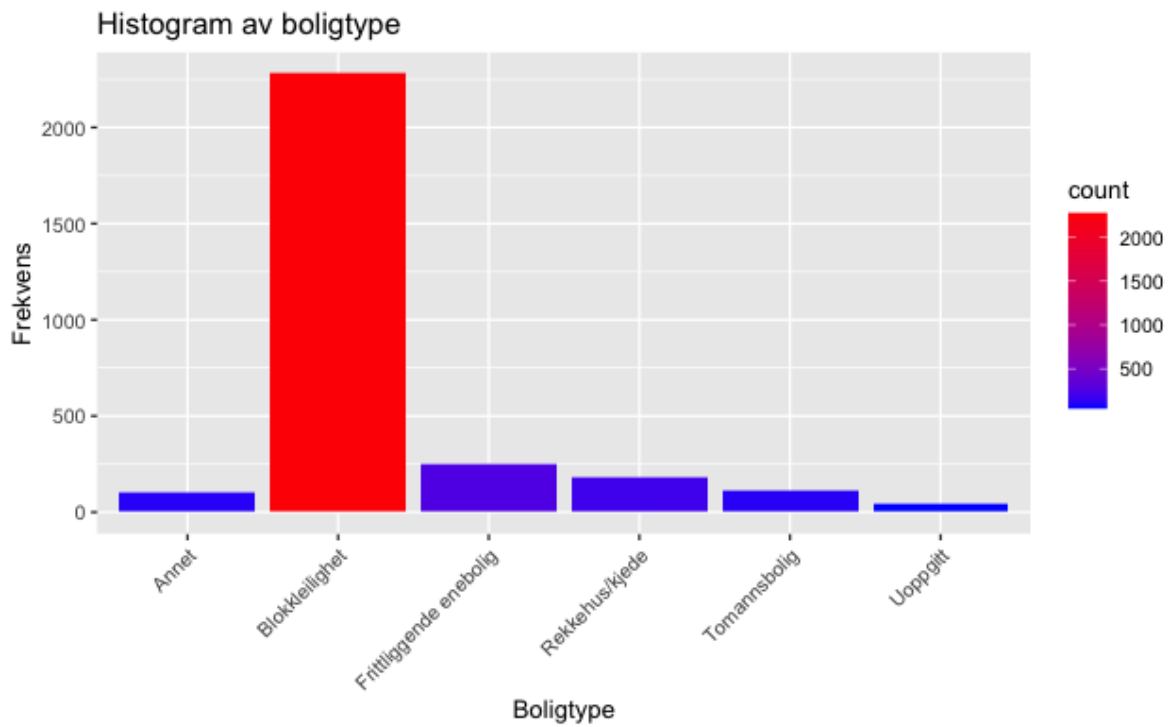
5. *Homoskedastisitet.* Feilleddet  $u$  har konstant varians gitt noen verdi av uavhengige variabler. Med andre ord,

$$\text{Var}(u/x_1, \dots, x_k) = \sigma^2.$$

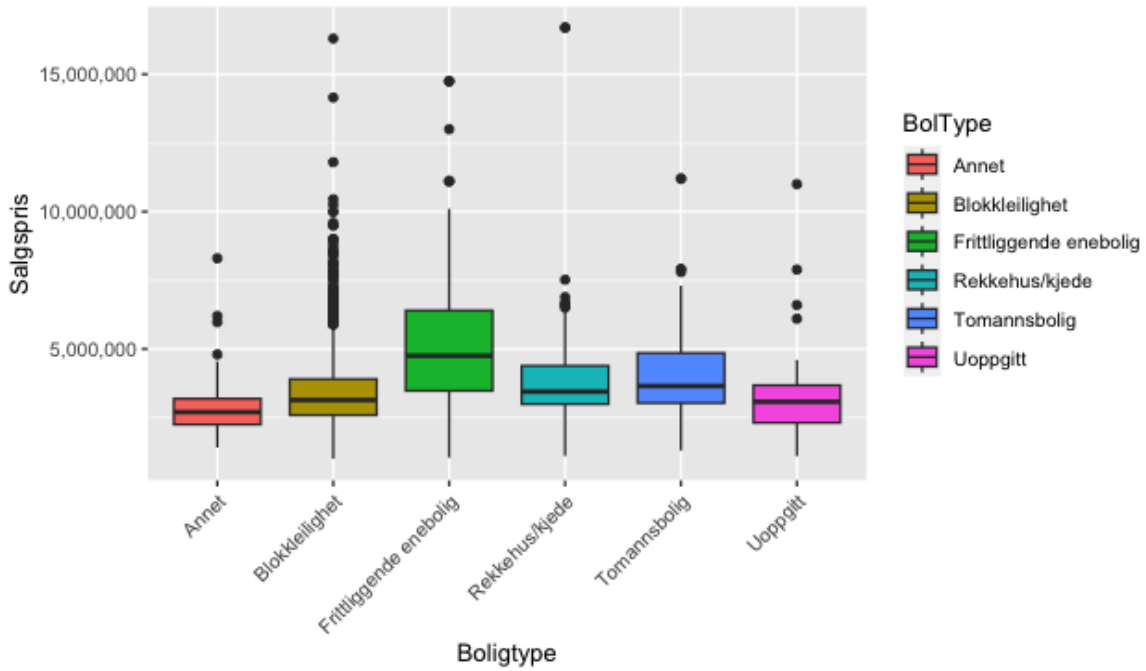
Det skal ikke være et høyt korrelasjonsnivå mellom uavhengige variabler.

Hvis alle disse antakelsene er oppfylt oppnår vi ifølge Wooldridge (2014) BLUE: Best, linear, unbiased estimators.

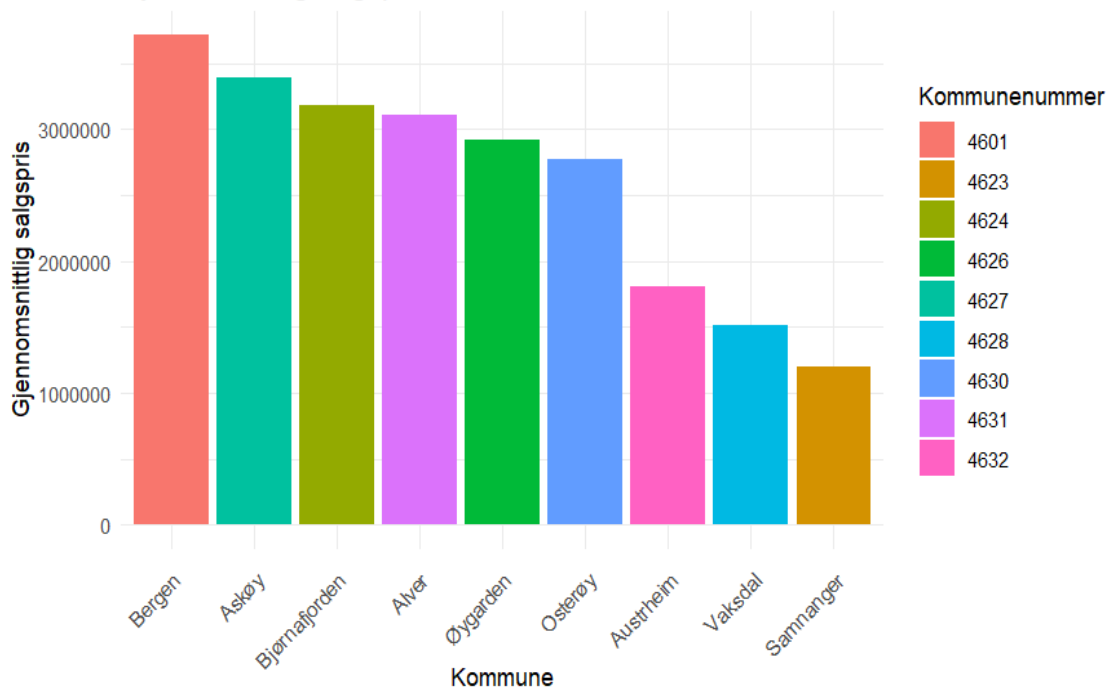
## A6. Deskriptiv statistikk



Box plot av salgspris etter boligtype

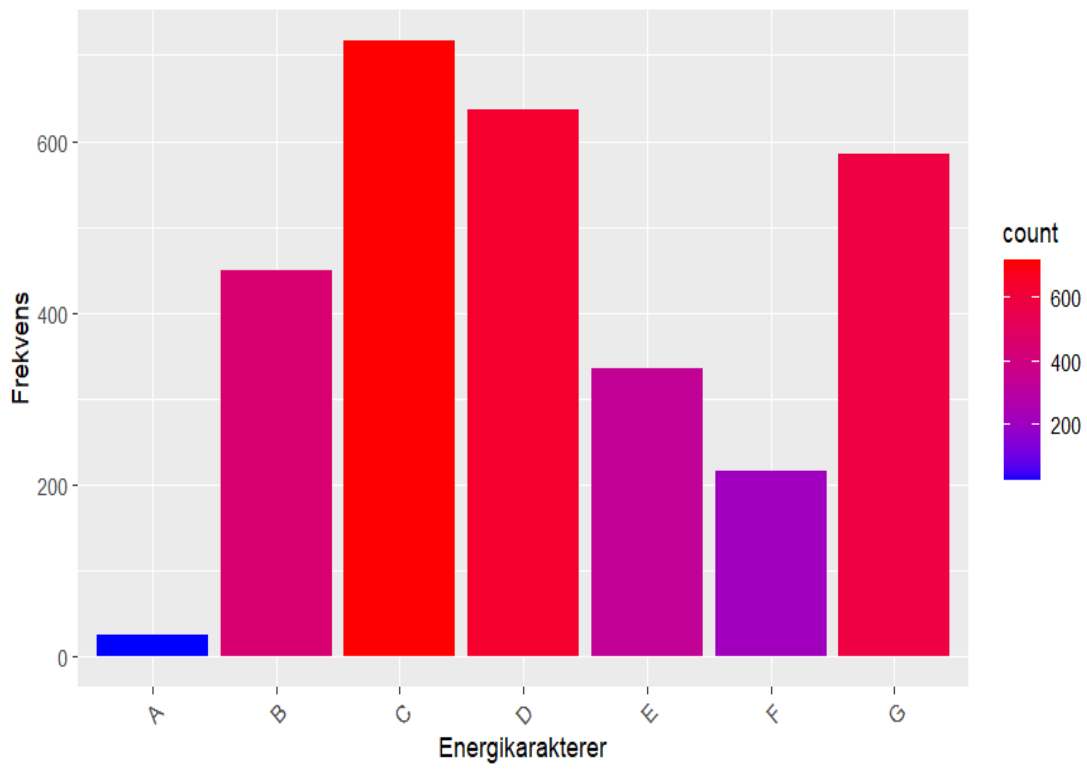


Gjennomsnittlig salgspris i ulike kommuner

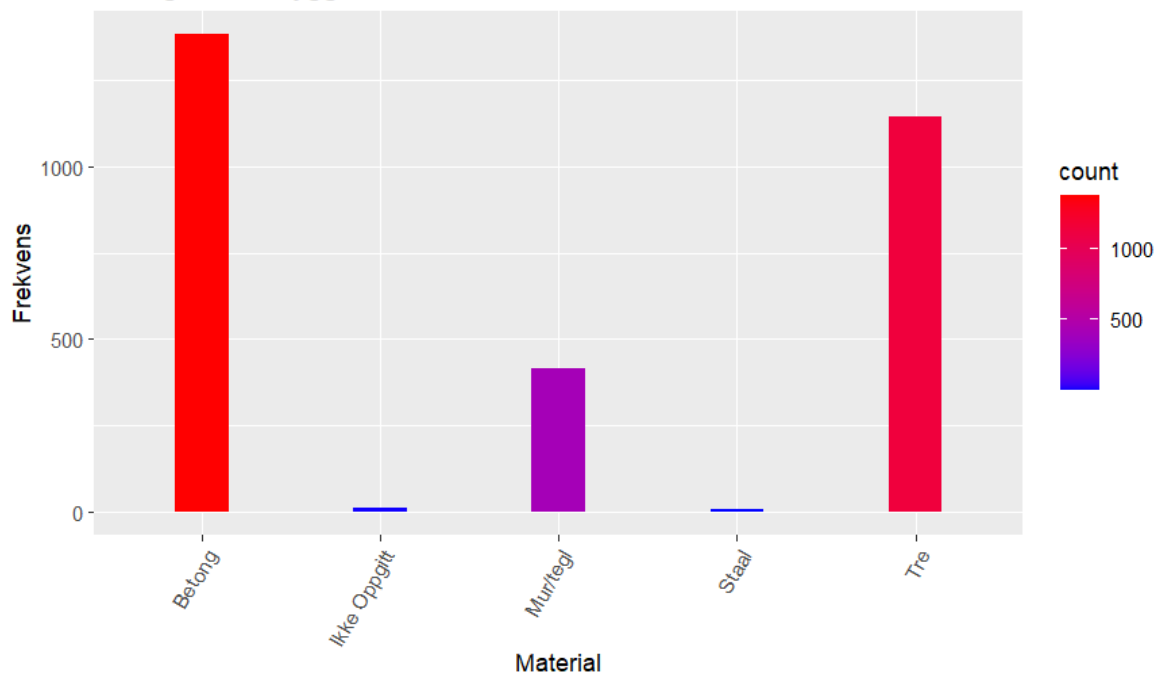




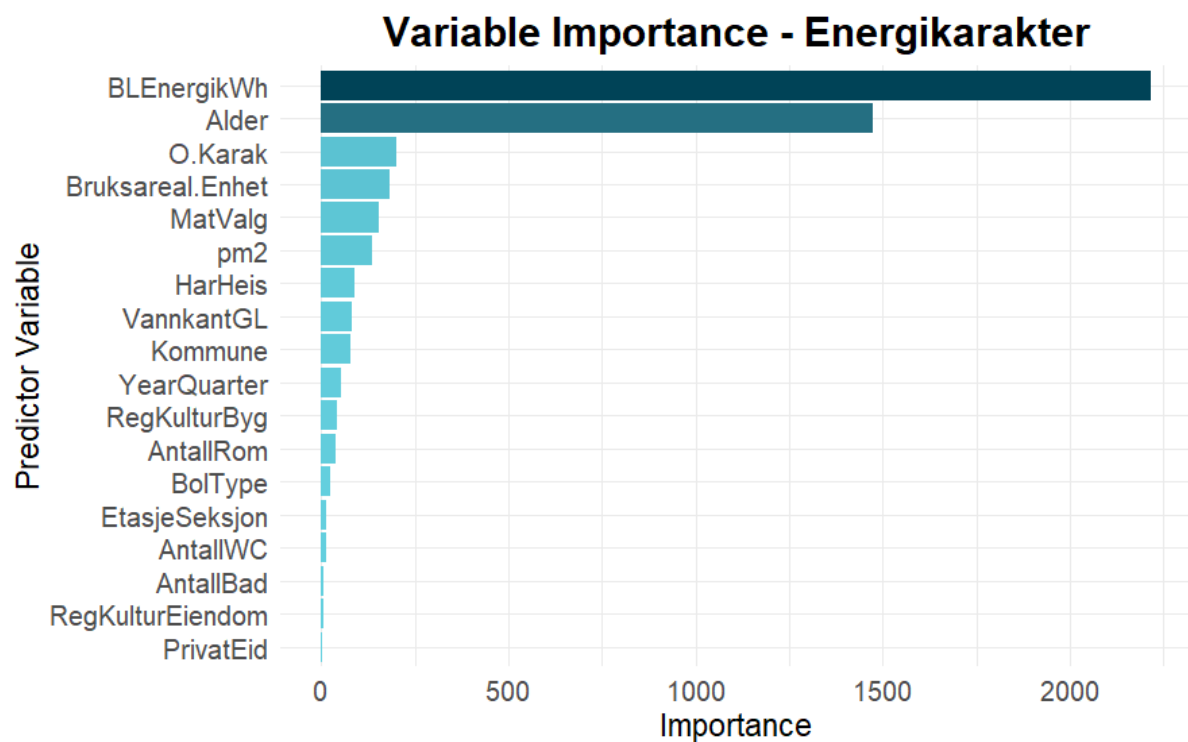
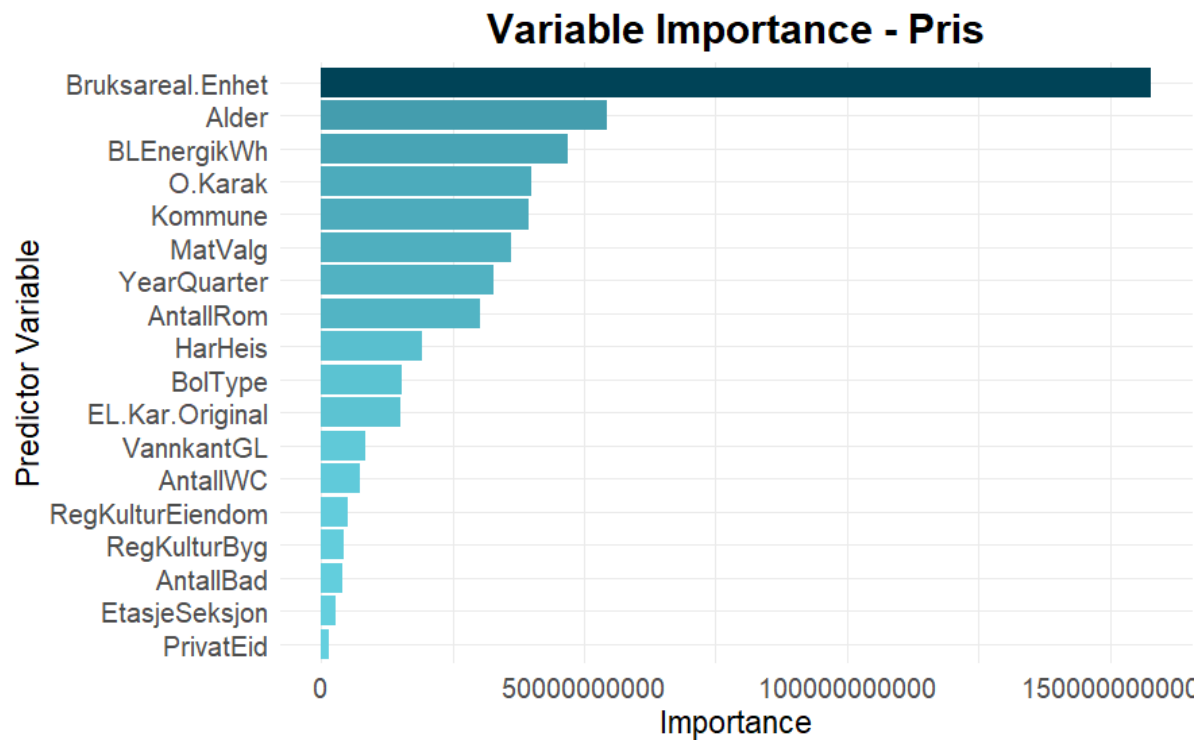
Histogram av energikarakterer



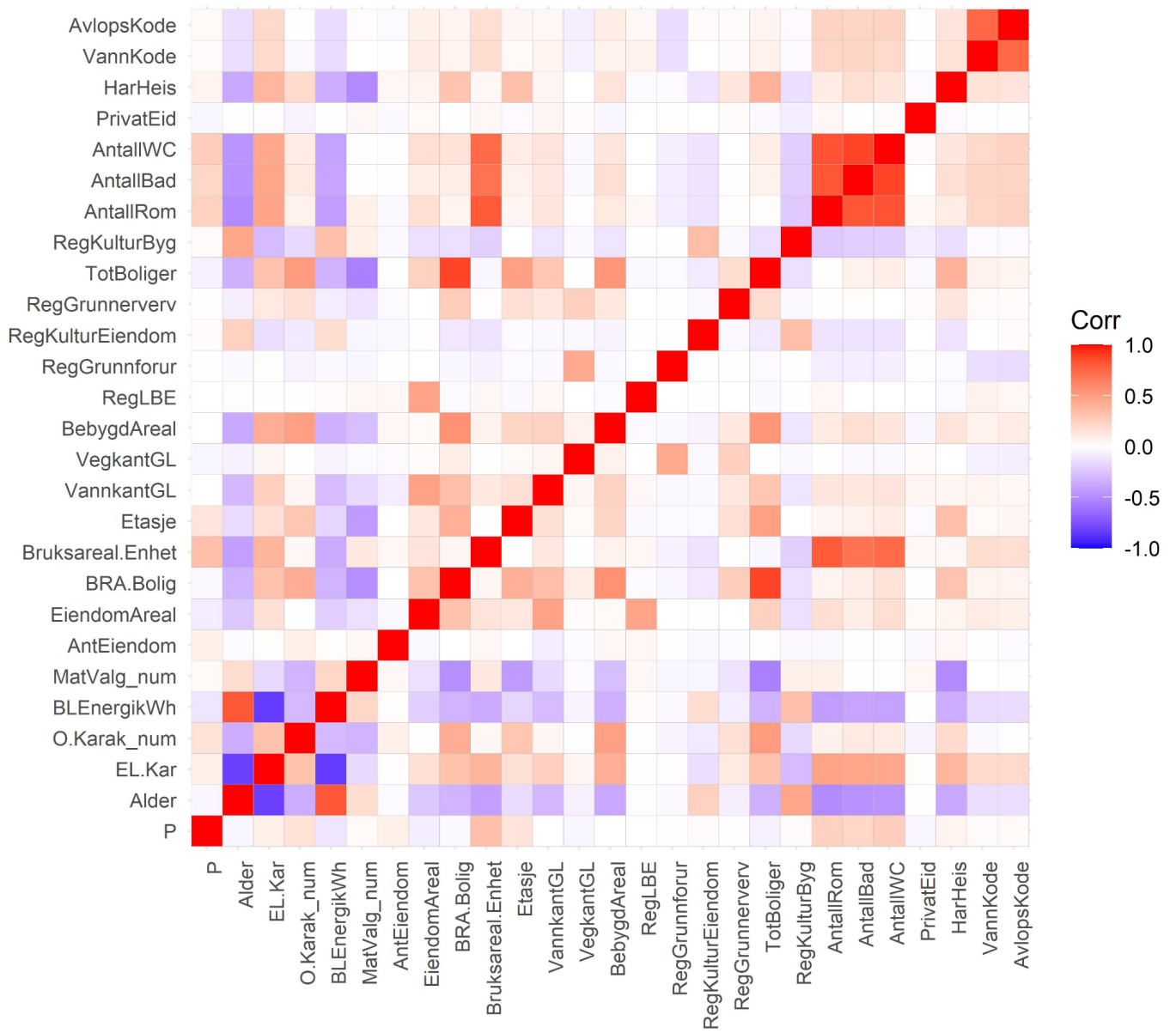
Histogram av byggematerialer



## A7. Variable Importance Plots



## A8. Korrelasjonsmatrise med varmekart



## A9. Oversikt av observasjoner; energimerking

### Oversikt over antall energikarakterer i vårt utvalg

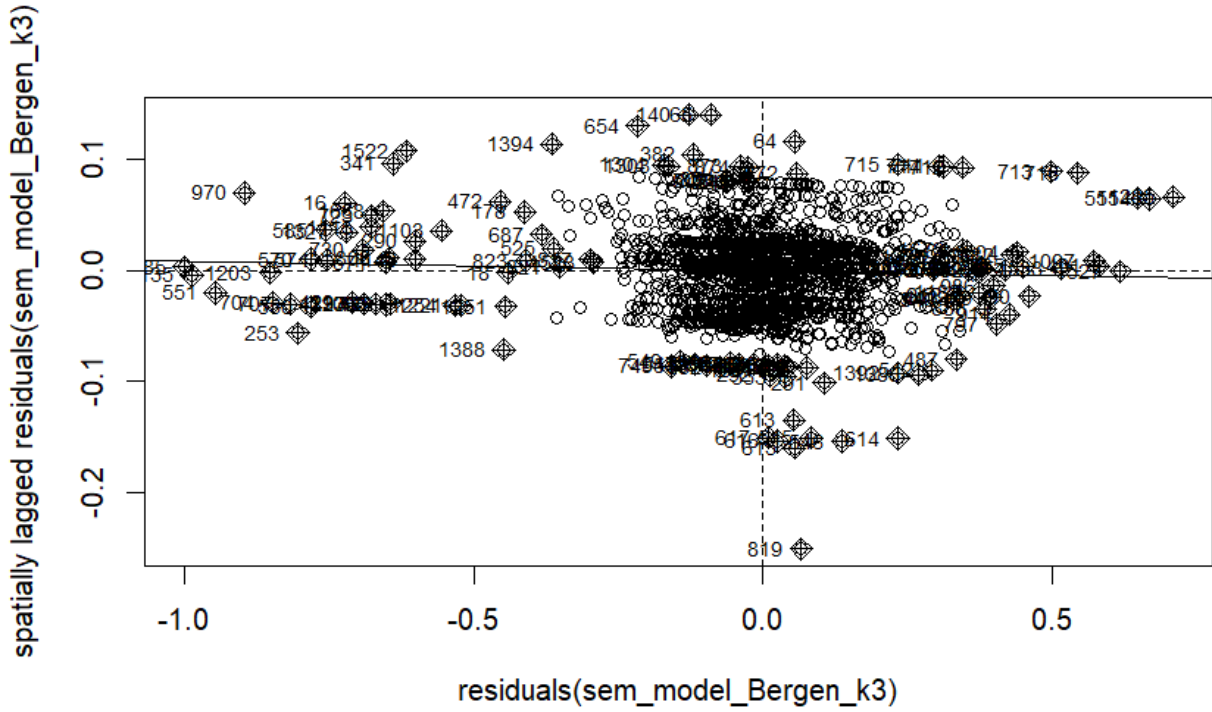
Hele perioden 2019-2022	A	B	C	D	E	F	G
<b>Bergensregionen</b>	21	398	641	542	203	91	222
<b>Bergen kommune</b>	15	244	511	407	192	89	221
<b>Før strømpriskrisen</b>							
<b>Bergensregionen</b>	19	298	476	393	141	63	158
<b>Bergen kommune</b>	15	182	377	291	132	62	158
<b>Etter strømpriskrisen</b>							
<b>Bergensregionen</b>	2	100	165	149	62	28	64
<b>Bergen kommune</b>	0	62	134	116	60	27	63

### Oversikt over antall oppvarmingskarakterer i vårt utvalg

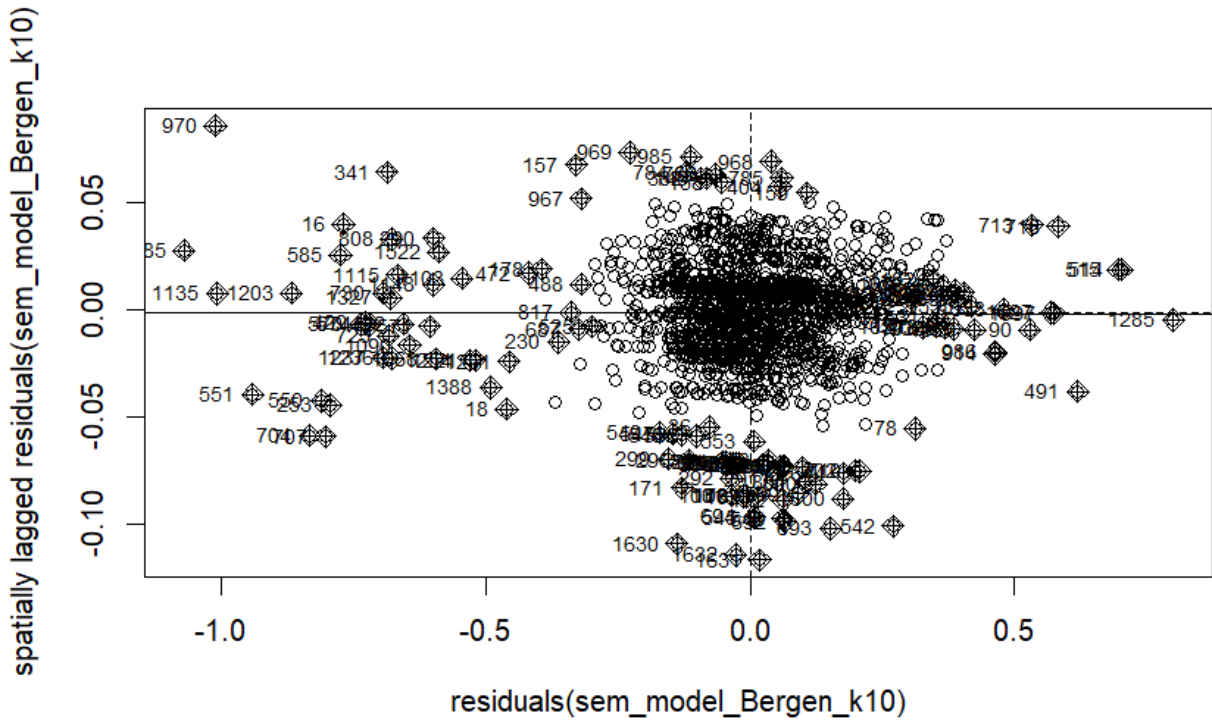
Hele perioden 2019-2022	Grønn	Lysegrønn	Gul	Oransje	Rød
<b>Bergensregionen</b>	611	106	38	293	1062
<b>Bergen kommune</b>	566	70	32	211	800
<b>Før strømpriskrisen</b>					
<b>Bergensregionen</b>	476	72	29	213	758
<b>Bergen kommune</b>	434	49	25	150	559
<b>Etter strømpriskrisen</b>					
<b>Bergensregionen</b>	143	34	9	80	304
<b>Bergen kommune</b>	132	21	7	61	241

# A10. Moran's plot - SEM

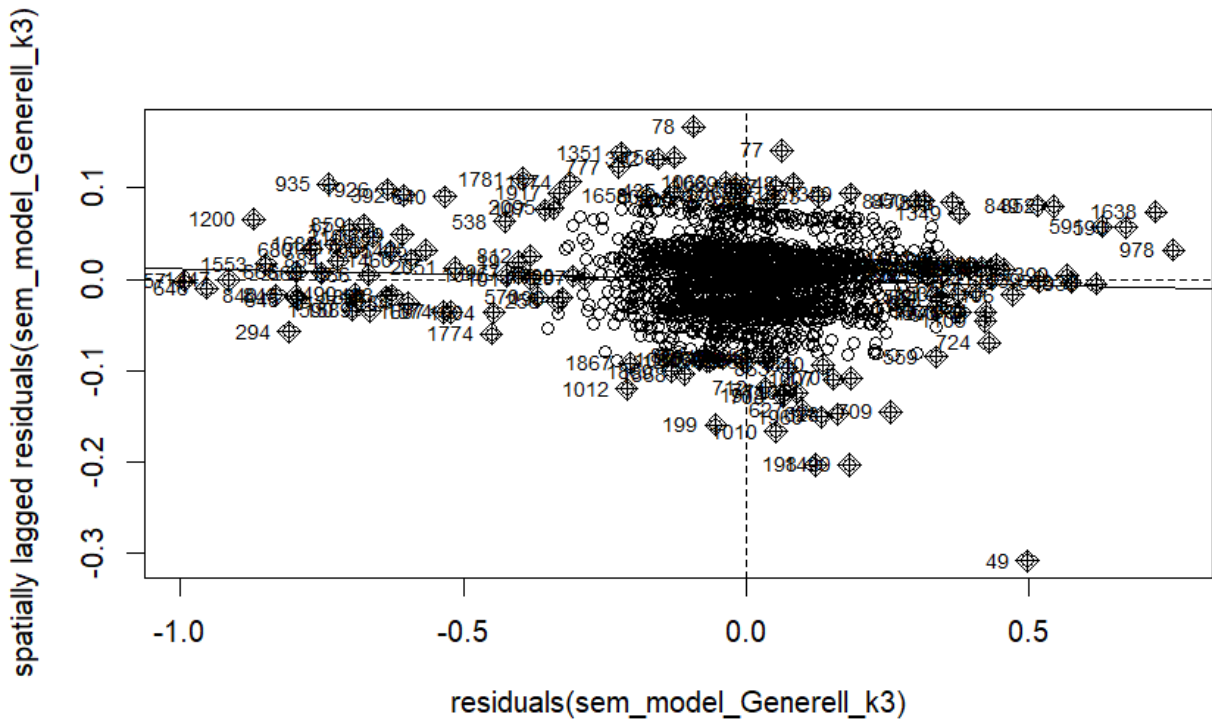
SEM, Bergen, K=3



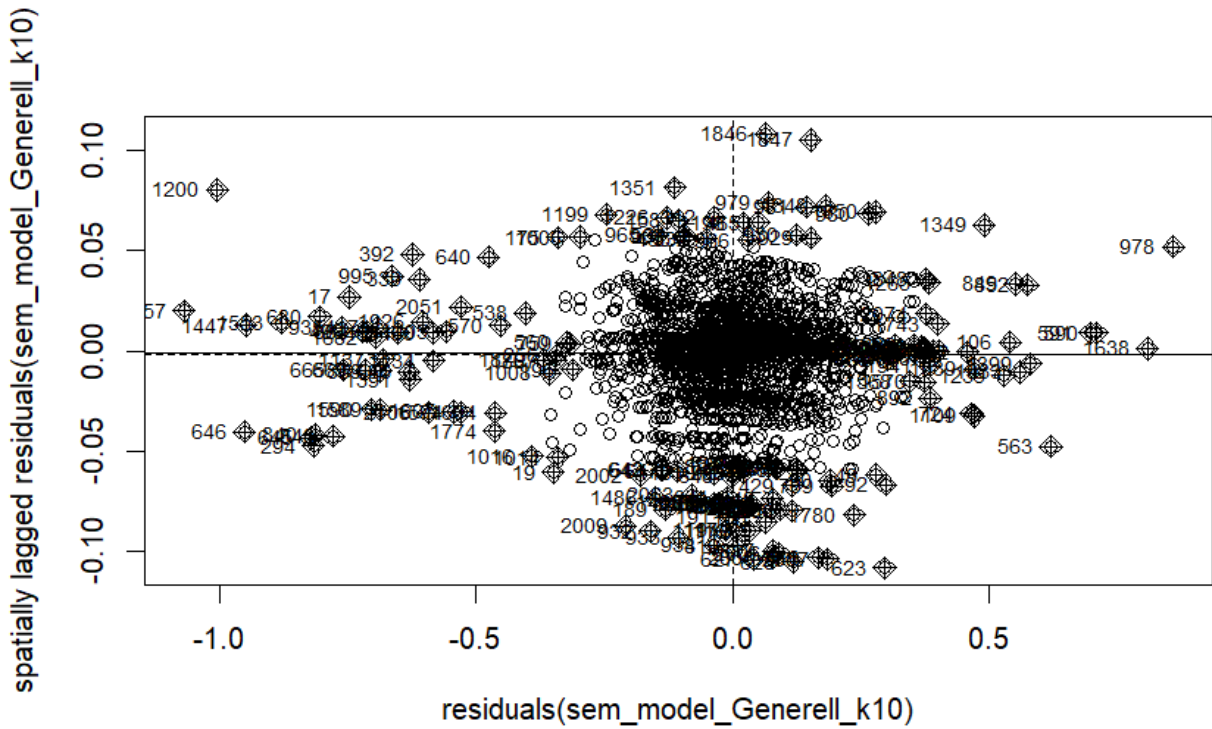
SEM, Bergen, K=10



### SEM, Bergensregionen, K=3

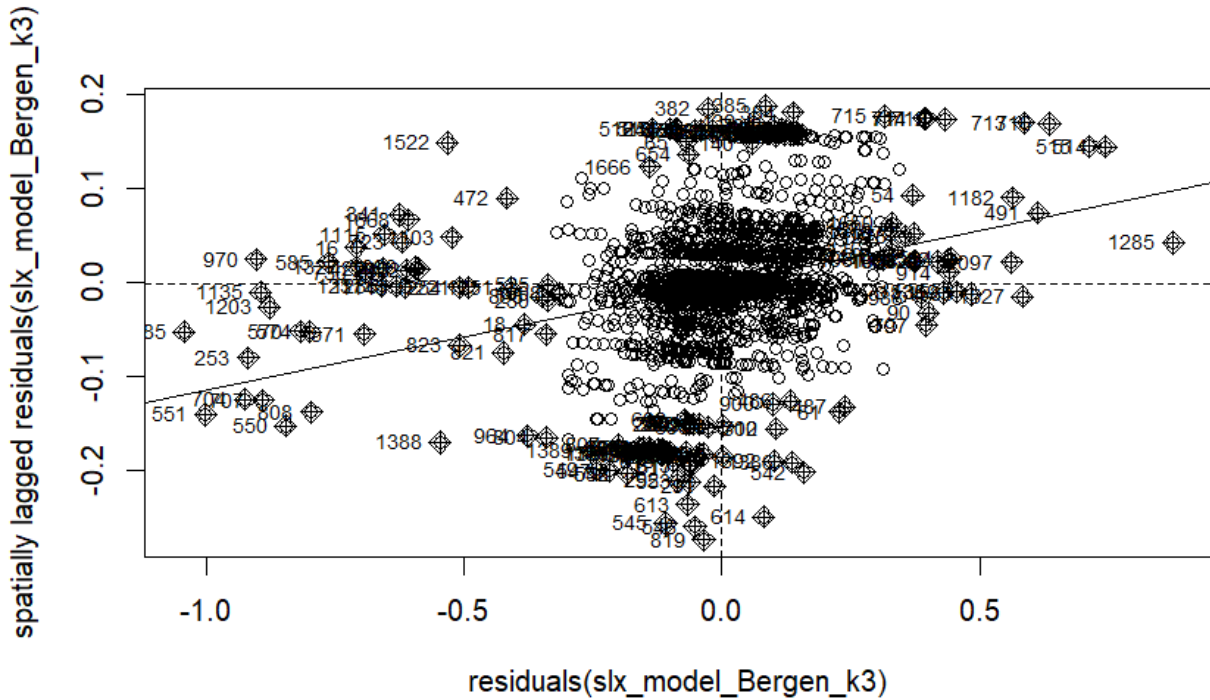


### SEM, Bergensregionen, K=10

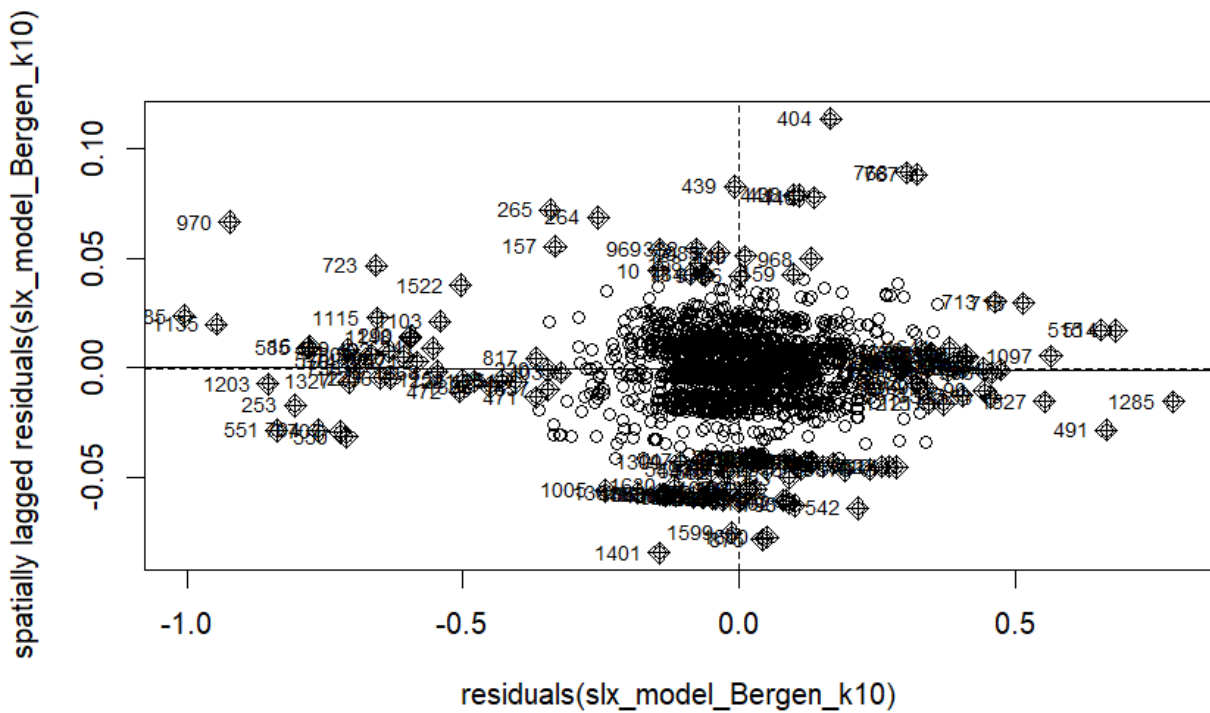


## A11. Moran's plot - SLX

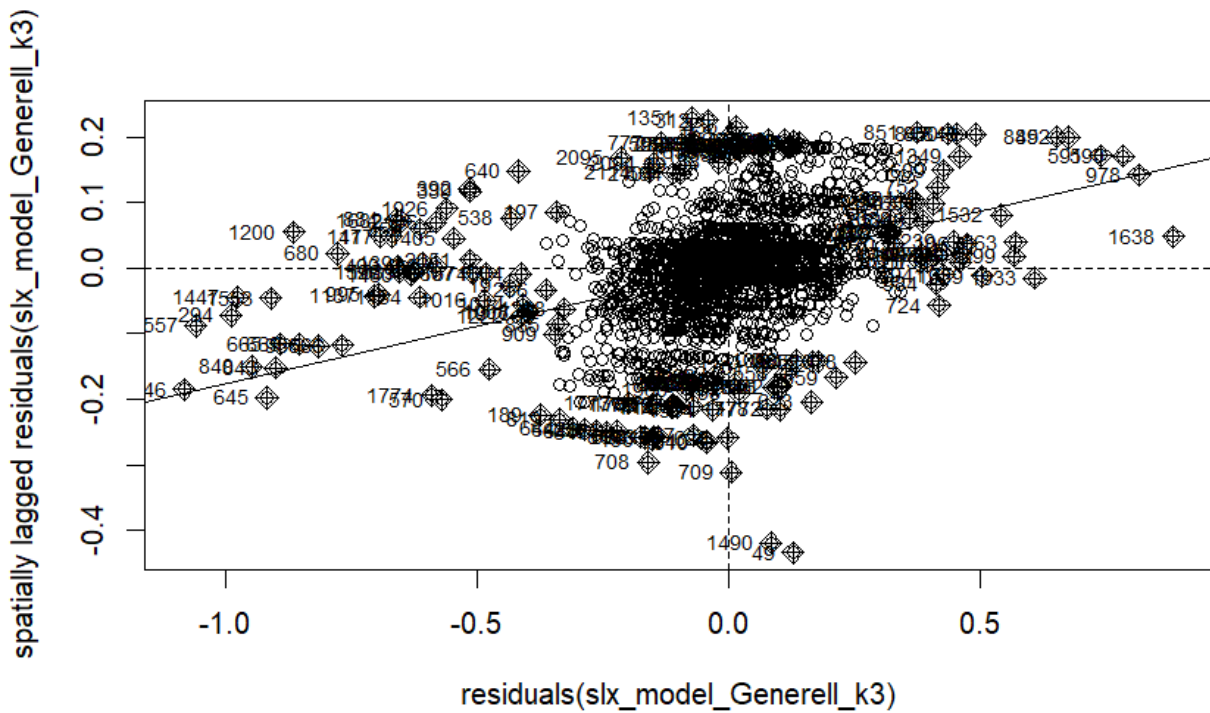
SLX, Bergen, K=3



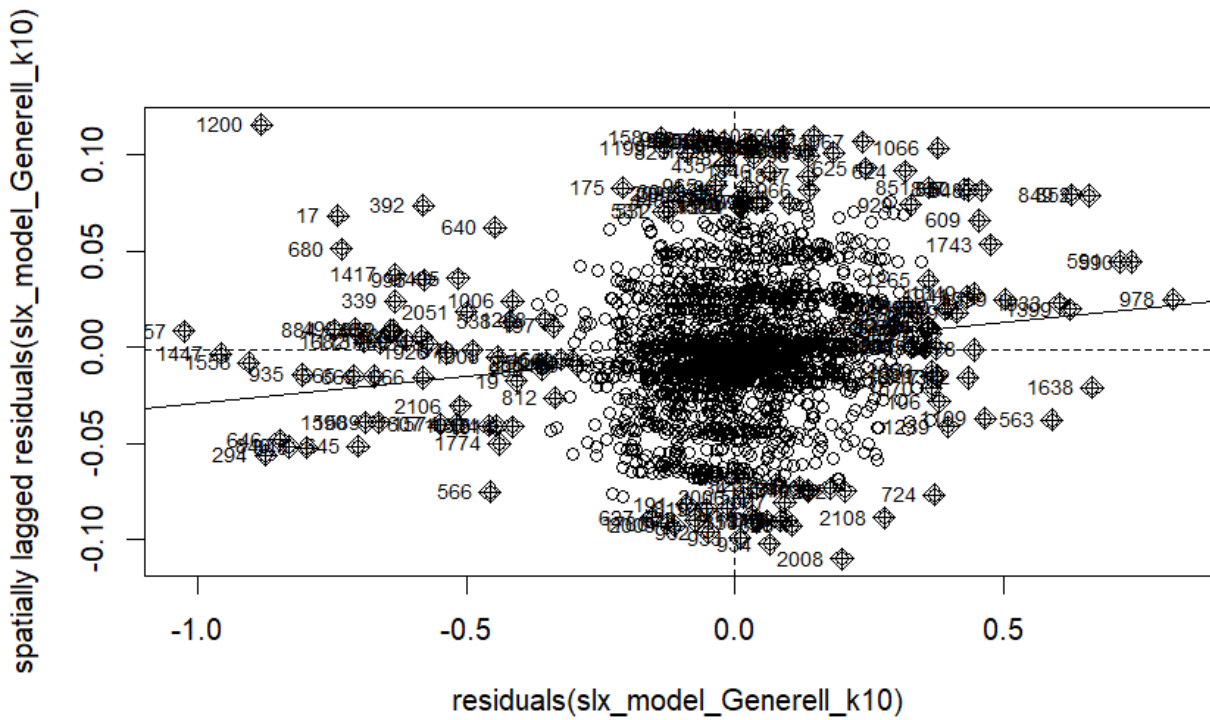
SLX, Bergen, K=10



SLX, Bergensregionen, K=3



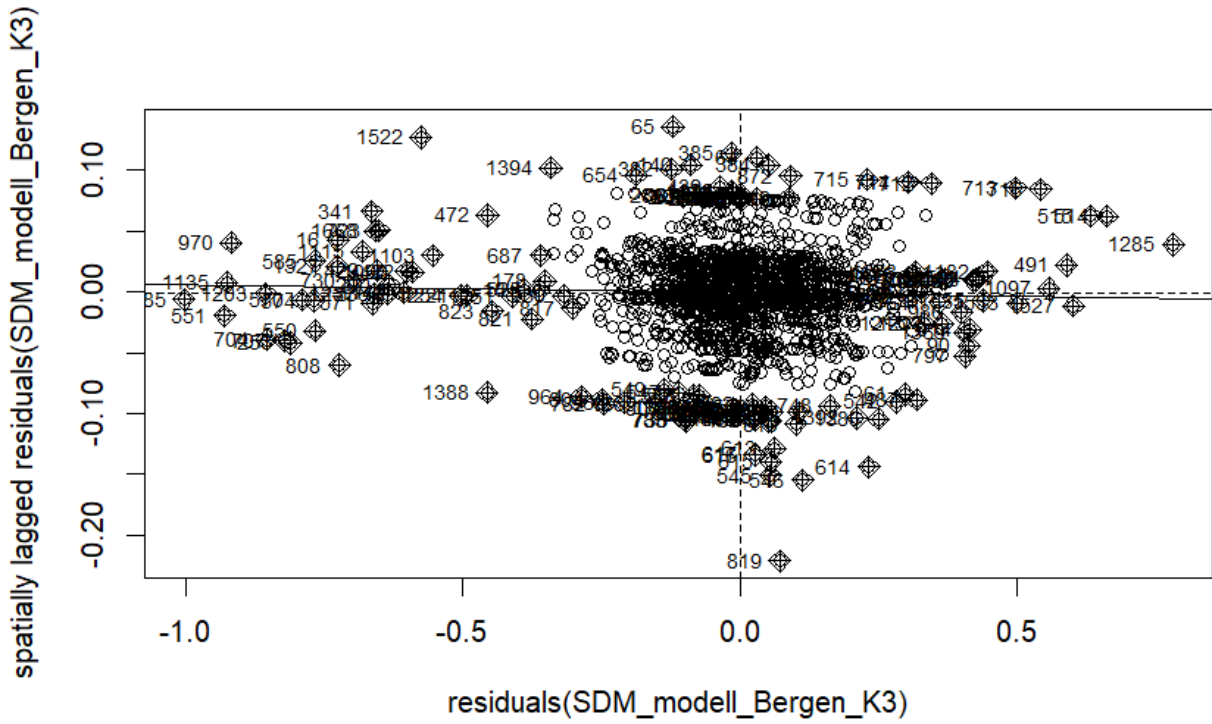
SLX, Bergensregionen, K=10



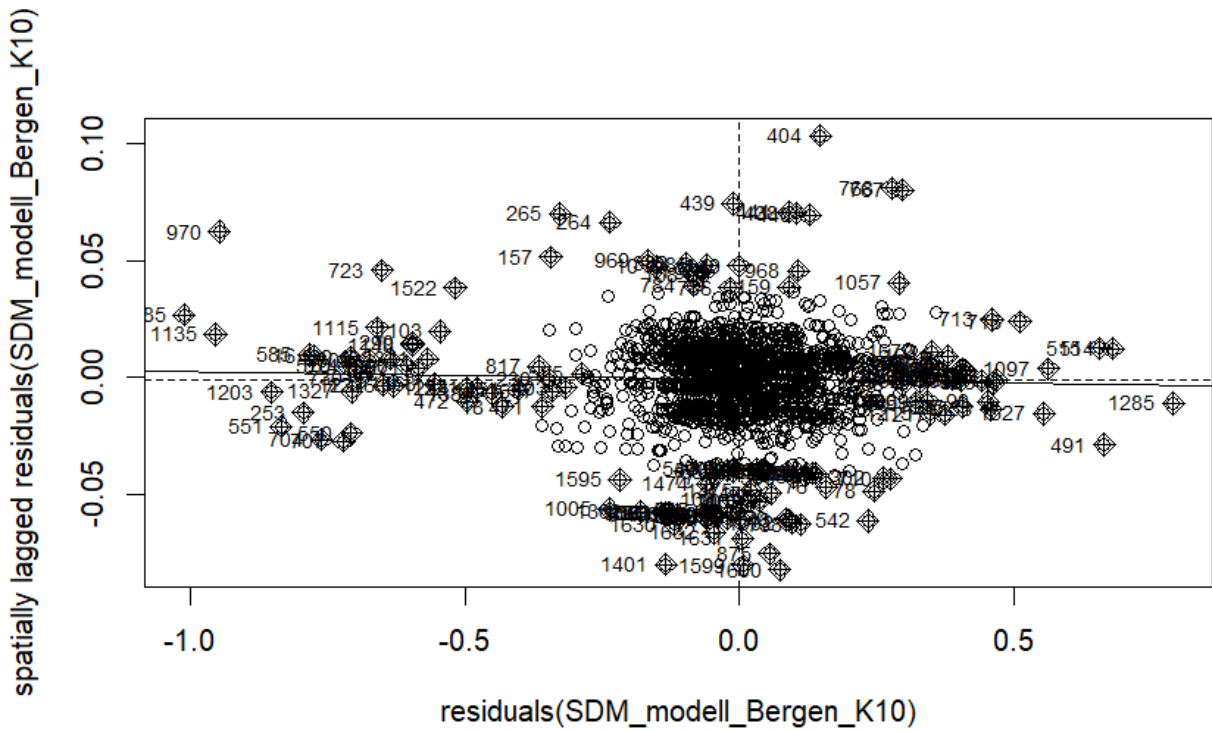


# A12. Moran's plot - SDM

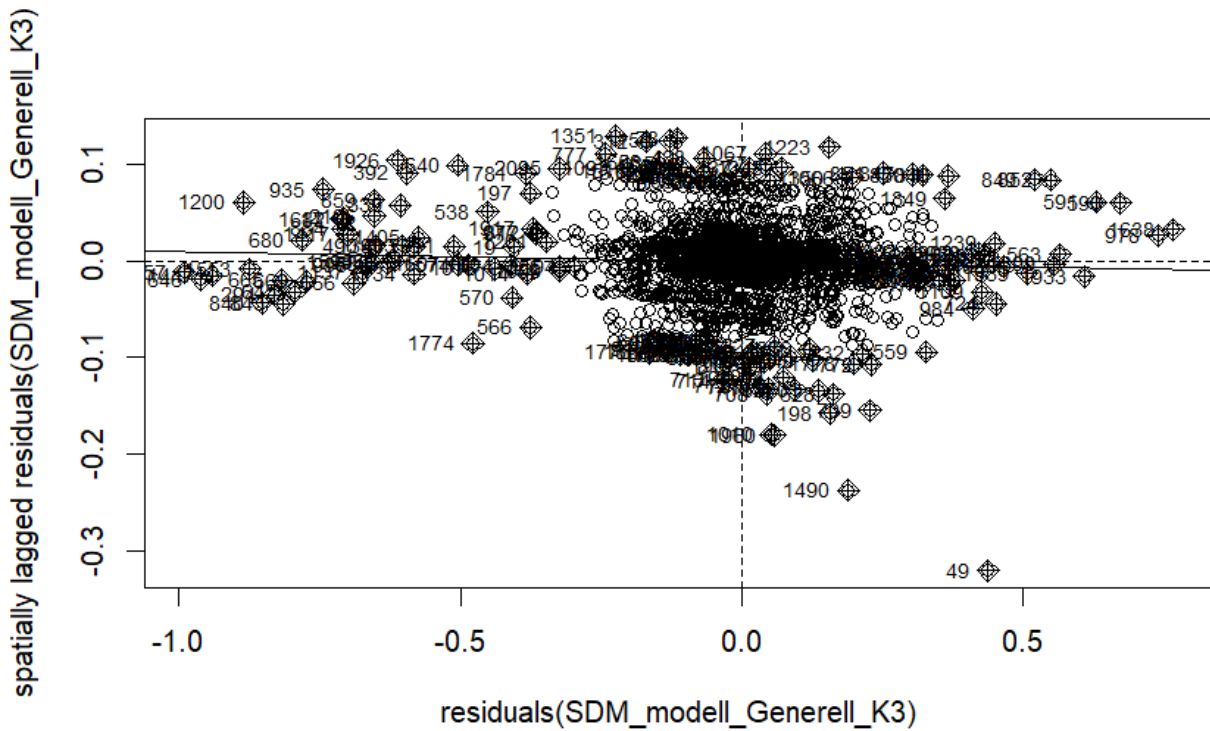
SDM, Bergen, K=3



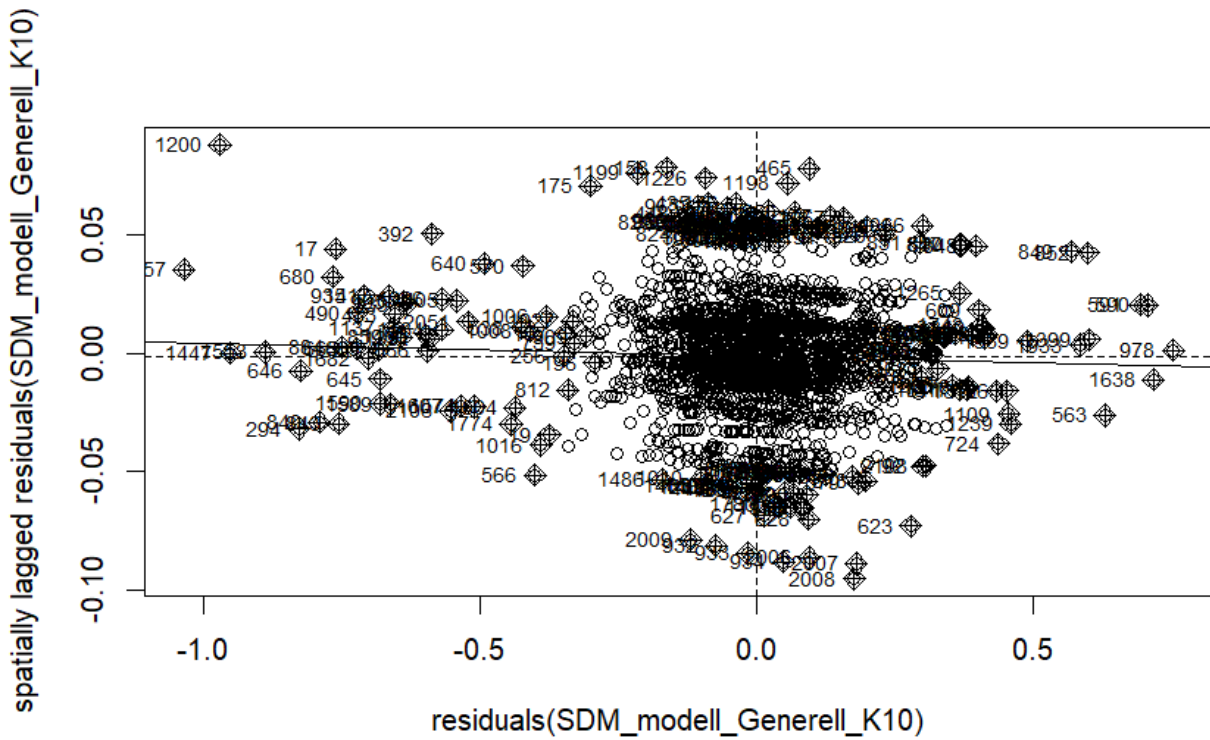
SDM, Bergen, K=10



### SDM, Bergensregionen, K=3

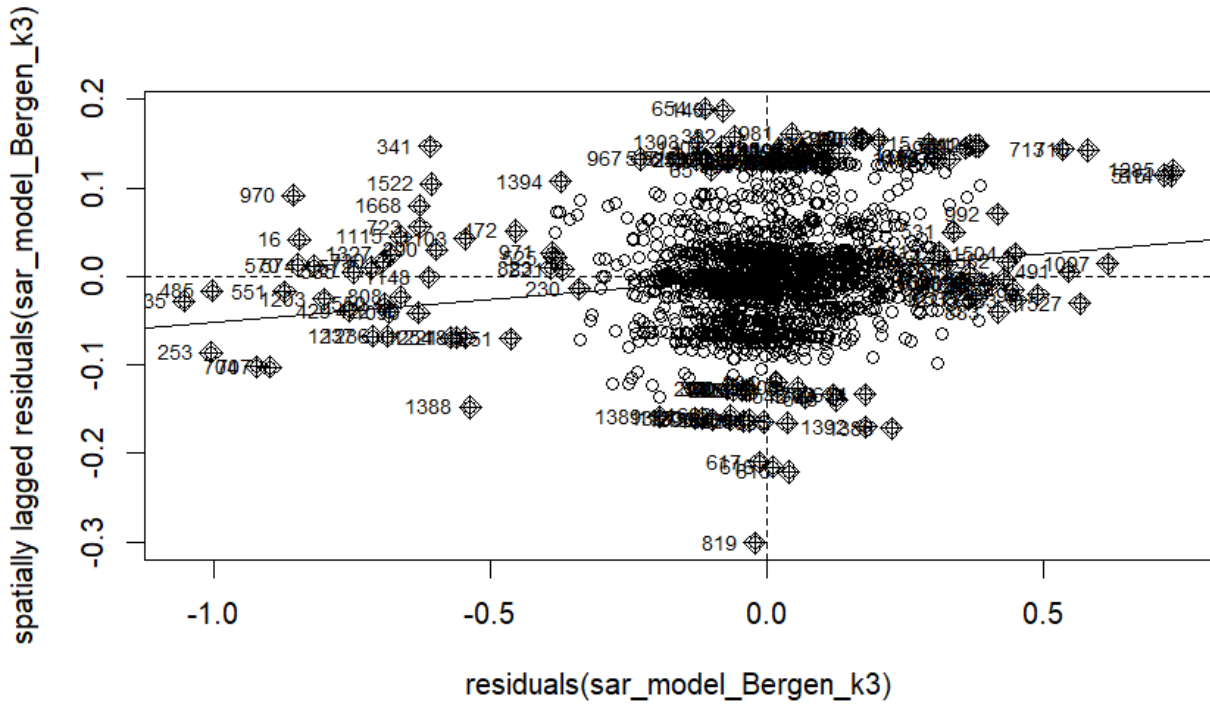


### SDM, Bergensregionen, K=10

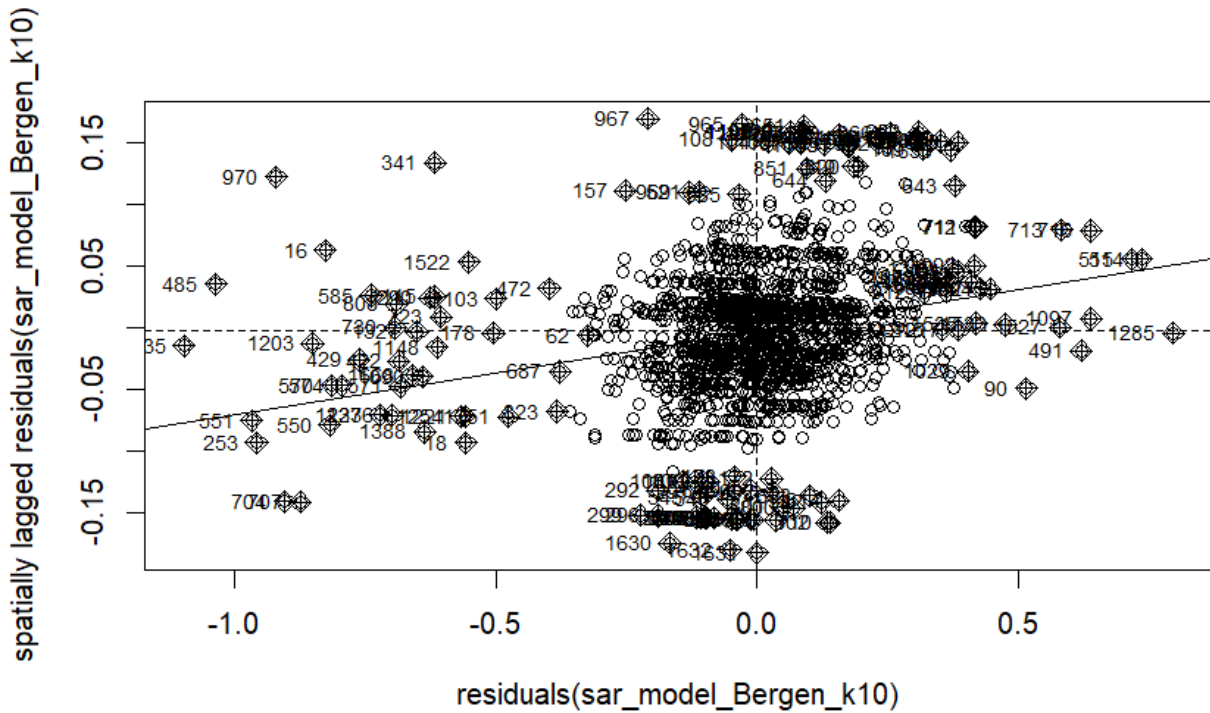


# A13. Moran's plot - SAR

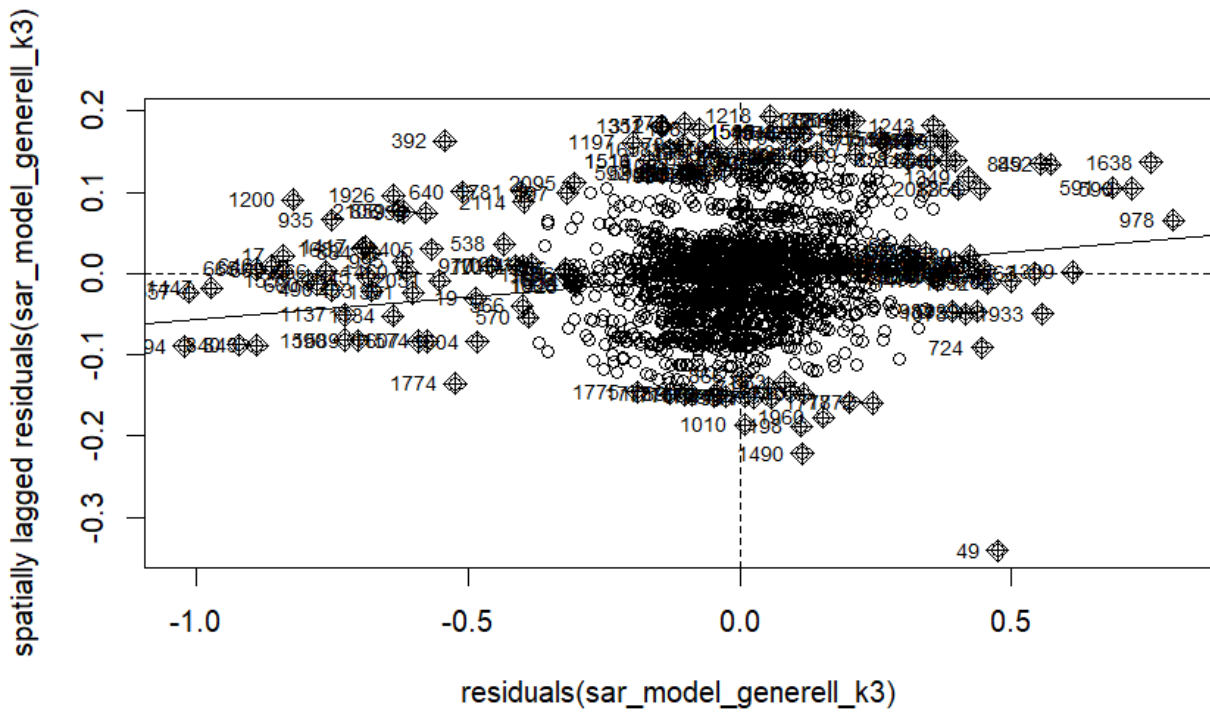
SAR, Bergen, K=3



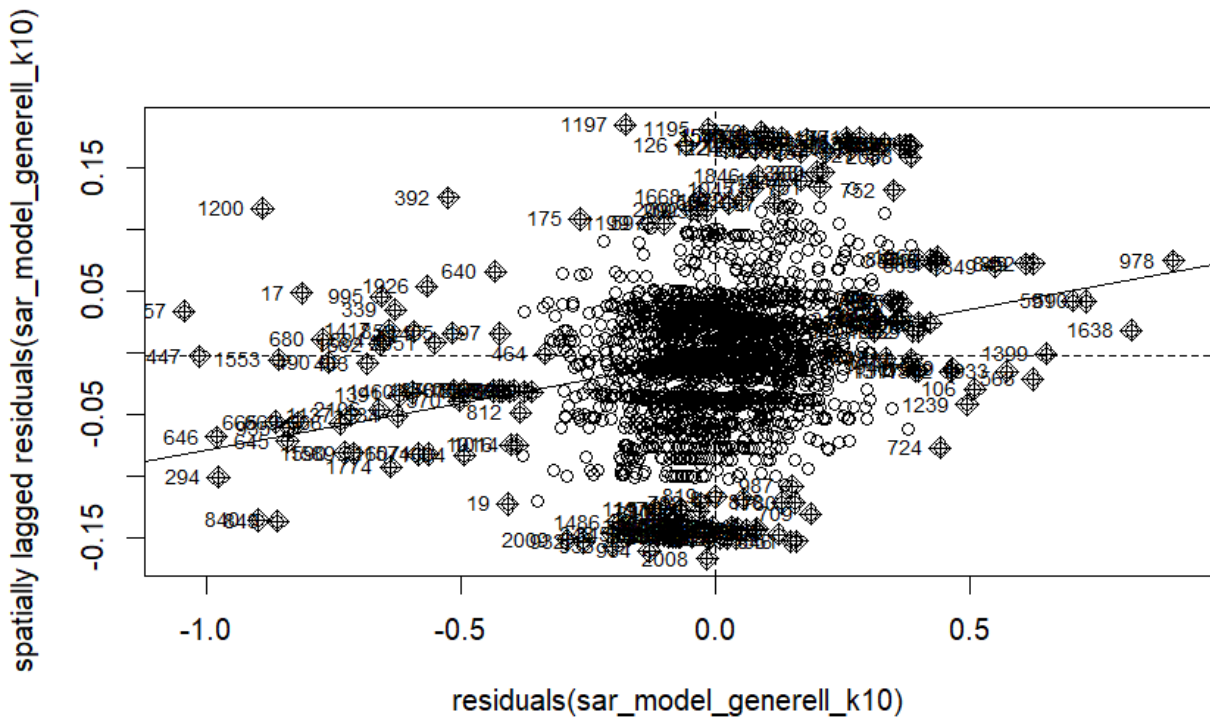
SAR, Bergen, K=10



### SAR, Bergensregionen, K=3



### SAR, Bergensregionen, K=10



## A14. Konkluderende modell (SEM med K = 3)

	Modell 1		Modell 4	
	Bergensregionen		Bergen kommune	
	Estimat	SE	Estimat	SE
<i>Skjæringspunkt</i>	13,462***	0,180	13,213***	0,206
B	-0,065	0,056	-0,010	0,068
C	-0,015	0,059	0,042	0,070
D	-0,010	0,060	0,046	0,072
E	-0,067	0,064	-0,011	0,077
F	-0,018	0,070	0,050	0,084
G	0,070	0,076	0,139	0,092
<i>Oppvarmingskarakterer</i>				
Lysegrønn	-0,014	0,021	-0,001	0,028
Gul	-0,194***	0,035	-0,177***	0,041
Oransje	-0,168***	0,021	-0,140***	0,026
Rød	-0,101***	0,016	-0,073***	0,019
<i>Byggematerialer</i>				
Ikke Oppgitt	0,028	0,062	-0,012	0,080
Mur/Tegl	0,045**	0,018	0,046**	0,020
Stål	0,032	0,068	0,022	0,072
Tre	0,001	0,013	0,009	0,016
<i>Kvartalsvariabler</i>				
2019 Q1	-0,236***	0,024	-0,268***	0,029
2019 Q2	-0,200***	0,022	-0,217***	0,026
2019 Q3	-0,238***	0,022	-0,244***	0,025
2019 Q4	-0,218***	0,022	-0,246***	0,027
2020 Q1	-0,210***	0,023	-0,218***	0,026
2020 Q2	-0,212***	0,023	-0,220***	0,027
2020 Q3	-0,178***	0,022	-0,190***	0,024
2020 Q4	-0,166***	0,023	-0,177***	0,027
2021 Q1	-0,154***	0,023	-0,173***	0,027
2021 Q2	-0,098***	0,021	-0,105***	0,024
2021 Q3	-0,095***	0,021	-0,102***	0,024
2021 Q4	-0,067***	0,023	-0,075***	0,026
2022 Q1	-0,034	0,023	-0,043	0,027
2022 Q2	0,022	0,023	0,005	0,026
2022 Q3	0,046**	0,022	0,034	0,025
<i>Andre variabler</i>				
Alder	-0,034***	0,008	-0,042***	0,009
Beregnet levert energi (kWh/m <sup>2</sup> )	-0,109***	0,031	-0,114***	0,036
Boligareal	-0,407***	0,021	-0,393***	0,024
Antall rom	0,072***	0,021	0,067***	0,025

Antall bad	0,056**	0,027	0,094***	0,031
Antall WC	0,136***	0,022	0,127***	0,025
Heis	0,034***	0,011	0,027**	0,013
Avstand til sentrum	-0,165***	0,014	-0,138***	0,015
Lambda $\lambda$	0,720		0,681	
Antall observasjoner	2118		1679	

*SE = Standardfeil, \*\*\* Signifikant på 1% nivå. \*\* Signifikant på 5% nivå. \* Signifikant på 10 % nivå.*