



Kundevurderinger og hotellenes bærekraftsertifiseringer: En analyse av deres påvirkning på norske hotellrompriser

*En empirisk studie av hotellrompriser
ved hjelp av multippel regresjonsanalyse*

Håvard Jacobsen Høgestøl og Martin Kildal

Veileder: Kenneth Fjell

Masterutredning i økonomi og administrasjon

Hovedprofil: Finansiell økonomi og Økonomisk styring

NORGES HANDELSHØYSKOLE

Dette selvstendige arbeidet er gjennomført som ledd i masterstudiet i økonomi- og administrasjon ved Norges Handelshøyskole og godkjent som sådan. Godkjenningen innebærer ikke at Høyskolen eller sensorer inntår for de metoder som er anvendt, resultater som er kommet frem eller konklusjoner som er trukket i arbeidet.

Forord

Gjennom denne oppgaven har vi fått en dypere forståelse av faktorene som påvirker hotellrompriser, og vi har sett hvor sammensatt det norske hotellmarkedet er. Vi synes det har vært et spennende tema å undersøke, samtidig som det har vist seg å være både krevende og komplekst. Ved innlevering av denne masterutredningen avsluttes vår studieperiode ved Norges Handelshøyskole.

Vi vil benytte denne muligheten til å uttrykke vår største takknemlighet til Kenneth Fjell, vår veileder ved Institutt for regnskap, revisjon og rettsvitenskap ved NHH. Takk for din rike kunnskap, dine kontinuerlige tilbakemeldinger og forslag gjennom arbeidet med oppgaven. Samtidig ønsker vi å takke hverandre for et godt samarbeid gjennom denne perioden.

Norges Handelshøyskole

Bergen, 1. juni 2023

Håvard Jacobsen Høgestøl

Martin Kildal

Sammendrag

Denne masteroppgaven har fokusert på hvordan norske hotellrompriser påvirkes av både kundevurderinger og hotellenes bærekraftsertifiseringer. I dagens digitale samfunn er tilbakemeldinger fra tidligere gjester lett tilgjengelige og synlige for potensielle kunder på plattformer som Booking.com. Dette gir potensielle kjøpere muligheten til å basere sine hotellvalg på tilbakemeldingene fra tidligere gjester. Samtidig blir hotellenes eventuelle bærekraftsertifisering synliggjort på slike tredjepartskanaler som Booking.com. Basert på disse observasjonene har vi formulert følgende problemstilling:

Hvordan påvirkes hotellromprisene i det norske markedet av kundevurderinger og hotellenes bærekraftsertifiseringer?

For å svare på dette innhentet vi data fra Booking.com ved hjelp av web-skraping. Dette ga oss raskt et utvalg på 720 hoteller. Denne svært effektive formen for datainnsamling gjorde at vi kunne utføre en tversnittundersøkelse med stor grad av reliabilitet, i et marked som preges av meget dynamisk prissetting.

Å analysere hele landet som ett samlet marked viste seg å være utfordrende, og vi så betydelige forskjeller mellom hvilke variabler som var viktige i ulike regioner. Derfor valgte vi å dele markedet inn i to dimensjoner: byhoteller og distriktshoteller. Inndelingen skiller vår studie fra tidligere publikasjoner om emnet, og representerer således en ny metode for hvordan man kan undersøke et nasjonalt hotellmarked.

Ved å se på alle hotellene i hver dimensjon, i én samlet modell, konstaterte vi at segmenttilhørighet (stjernenivå) var en viktig faktor, og at hele 94 % av våre observasjoner ligger innenfor segmentnivå tre og fire. Dette gjorde at vi utarbeidet en modell som fordelte hotellene på segmentnivå. Resultatene indikerer at hva som påvirker hotellromprisene varierer betydelig, avhengig av både markedstilhørighet og segmentnivå. Innenfor hvert segment er det kun noen få variabler som faktisk påvirker hotellromprisene. For eksempel viser det seg at beliggenhet er den eneste faktoren som har en signifikant effekt på prisen for byhoteller i segment tre, og at en økt score for attributter som fasiliteter, komfort, personale, renhold og WiFi, ikke gir en høyere hotellrompris. Dette funnet kan forklares med at kundene forventer en viss standard innenfor hvert segment. Dette er et gjennomgående mønster for alle segmenter og

tyder på at segmenttilhørighet er den viktigste indikatoren på hva et hotell tilbyr. Hotellene må imidlertid ikke undervurdere faktorer som ikke direkte påvirker romprisen, da de fortsatt betraktes som grunnleggende goder, selv om de ikke differensierer prisene mellom hotellene.

Et annet aspekt ved denne studien, som skiller den fra andre studier, er at den analyserer både kundevurderinger og bærekraftsertifiseringer opp mot pris. Våre funn for bærekraftsertifiseringer tyder på at hoteller som prioriter bærekraft ikke oppnår en høyere rompris, nesten uavhengig av segment- og dimensjonstilhørighet. Vi har kun identifisert betalingsvilje for bærekraftsertifisering blant kundene i segment fire i distriktene. Generelt er faktoren segmenttilhørighet viktigere enn hvorvidt et hotell er bærekraftsertifisert.

Oppgaven konkluderer med at *hotellromprisene i det norske markedet blir påvirket av kundevurderinger og hotellenes bærekraftsertifiseringer*, men bare til en viss grad. Det norske hotellmarkedet er komplekst, og det er vanskelig å generalisere i hvilken grad hvert attributt påvirker hotellromprisen. Dette avhenger av hotellenes segmenttilhørighet og hvor hotellet er plassert.

Innholdsfortegnelse

1	Innledning	1
1.1	Problemstilling og avgrensning	1
1.2	Oversikt av oppgaven	2
2	Generell bransjeinformasjon	3
2.1	Det norske hotellmarkedet	3
2.1.1	Største hotellkjeder	4
2.2	Prisvariasjonene i markedet	4
2.3	Klassifisering av hotellene	5
2.3.1	Segmentrangering av hotellene	5
2.3.2	Bærekraftsertifisering	6
2.4	Tredjepartskanaler (OTA).....	6
2.4.1	Kundevurderinger på Booking.com.....	8
2.5	Oppsummering av generell bransjeinformasjon	8
3	Teoretisk rammeverk	9
3.1	Den hedoniske prismodellen.....	9
3.1.1	Implementering av den hedoniske pristeorien	9
3.1.2	Utledning av den hedoniske prismodellen	10
3.1.3	Begrensninger ved hedonisk pristeori.....	16
3.1.4	Oppsummering av hedonisk pristeori	16
3.2	Hedoniske prisstudier av hotellmarkeder.....	17
3.2.1	Segmentnivå.....	17
3.2.2	Beliggenhet	18
3.2.3	Fasiliteter.....	19
3.2.4	Komfort.....	20
3.2.5	Personale	20
3.2.6	Renhold	21
3.2.7	WiFi	22
3.2.8	Bærekraftsertifisering	23
3.2.9	Oppsummering av variablene	24
3.3	Halo-effekten i tilbakemeldingene.....	24
3.4	Konklusjon av teori og formulering av forskningsspørsmål.....	25

4	Data og datainnsamling	26
4.1	Utvalg.....	26
4.2	Innhenting av data.....	27
4.2.1	Etikk ved Web-skraping	28
5	Metode	30
5.1	Forskningstilnærming, studiedesign og kvalitetssikring.....	30
5.1.1	Forskningstilnærming	30
5.1.2	Studiedesign.....	30
5.1.3	Kvantitativ metode.....	31
5.1.4	Reliabilitet.....	31
5.1.4	Validitet.....	32
5.2	Multippel regresjon.....	33
5.2.1	Minste kvadraters forutsetninger for modellen.....	34
5.2.2	Andre forutsetninger	36
5.2.3	En nærmere undersøkelse av residualene	38
5.2.4	Korrelasjon.....	41
5.2.5	Autokorrelasjon.....	41
5.2.6	Tilpasning av regresjonsmodellene.....	41
5.2.7	Modellens forklaringsgrad	41
5.3	Oppsummering av metode	42
6	Presentasjon av data	43
6.1	Geografisk plassering av hotellene	43
6.2	Fordeling på segmentnivå.....	44
7	Analyse og diskusjon	48
7.1	Inndeling av markedet.....	48
7.2	Korrelasjonsanalyse	50
7.2	Modell 1	51
7.2.1	Residualer	53
7.2.2	VIF-analyse.....	55
7.2.3	Attributteneffekter	56
7.2.4	Oppsummering av Modell 1	62
7.3	Modell 2	63
7.3.1	Residualer	65

7.3.2	VIF-analyse.....	67
7.3.3	Attributtenes effekter	68
7.3.4	Oppsummering av Modell 2	73
8	Konklusjon og fremtidig arbeid	76
8.1	Hovedfunn.....	76
8.2	Hva resultatene bidrar med	78
8.3	Begrensninger og ekstern validitet.....	79
8.4	Forslag til fremtidig arbeid	80
	Referanseliste.....	82
	Appendiks	88
A1	Generell bransjeinformasjon	88
A5	Metode.....	89
A5.1	Multipel regresjon.....	89
A6	Presentasjon av data	90
A6.1	Geografisk inndeling av data	90
A6.2	Deskriptiv statistikk	91
A7	Analyse og diskusjon.....	93
A7.1	Alternative regresjonsmodeller.....	93
A7.2	Statistiske analyser.....	95

Figurligste

Figur 2.1: Et utdrag av de 20 største hotellkommunene i Norge (SSB, 2023).	3
Figur 3.1: Konsumtilpasningen i den hedoniske prismodellen (basert på Rosen, 1974).....	12
Figur 3.2: Produksjonstilpasningen i en hedonisk prismodell (basert på Rosen, 1974).....	14
Figur 3.3: Markedslikevekt under den hedoniske prismodellen (basert på Rosen, 1974).....	15
Figur 4.1: Skjerm bilde av Booking.com (Booking.com, 2023a).....	27
Figur 5.1: Segmentert rompris-histogram med observasjoner og tetthet.....	37
Figur 5.2: Residualplott av hotellene i By.	39
Figur 5.3: Residualplott av hotellene i Distrikts-Norge	39
Figur 6.1: Geografisk plassering av hotellene i utvalget.....	43
Figur 6.2: Segmentandelen i utvalget.	44
Figur 6.3: Prisvariasjon blant hotellene innenfor hvert segment.	45
Figur 6.4: Observerte bærekraftsertifiseringer i hvert segment.	46
Figur 6.5: Gjennomsnittlige kundevurderinger av hotellattributtene i hvert segment.....	47
Figur A5.1: Sammenhengen mellom pris og de uavhengige variablene.....	89
Figur A6.1: De ti største hotellkonsentrasjonene i utvalget.....	90
Figur A6.2: Antall tilbakemeldinger per hotell i hvert segment.....	91
Figur A7.1: Boksplokk av residualene for By og DN.....	95
Figur A7.2: Boksplokk av residualene for hvert segment i By.	95
Figur A7.3: Boksplokk av residualene for hvert segment i DN.	96

Tabelliste

Tabell 7.1: Korrelasjonsmatrise for de inkluderte variablene.....	50
Tabell 7.2: Regresjonstabell for Modell 1; By og DN.....	52
Tabell 7.3: Regresjonstabell for Modell 2.	64
Tabell A6.1: Deskriptiv statistikk for hotellromprisene i Norge på segmentnivå.	91
Tabell A6.2: Deskriptiv statistikk for de inkluderte attributtene.	92
Tabell A6.3: Deskriptiv statistikk for gjennomsnittlige vurderinger på segmentnivå.....	92
Tabell A7.1: Regresjonstabell for alle hoteller i det norske hotellmarkedet.....	93
Tabell A7.2: Regresjonstabell for Oslo-markedet.	94
Tabell A7.3: Residualer for By og DN i Modell 1.	97
Tabell A7.4: Residualer for By-segmentene i Modell 2.	97
Tabell A7.5: Residualer for DN-segmentene i Modell 2.....	97
Tabell A7.6: Durbin-Watson test statistikk.	98

Tabell A7.7: VIF-matrise for By og DN.....	98
Tabell A7.8: VIF-matrise med elimineringsmetode.	98
Tabell A7.9: VIF-matrise for By på segmentnivå.	99
Tabell A7.10: VIF-matrise for DN på segmentnivå.	99
Tabell A7.11: Standardavvik for variablene i Modell 2.	99

1 Innledning

Med bruk av nettbaserte bestillingsplattformer kan kunder ikke bare bestille hotellrom, men også gi offentlige tilbakemeldinger om sine opplevelser. Dette har ført til i et betydelig volum av tilgjengelige data for hotellene, som gir økt innsikt i gjestenes opplevelser og preferanser. Dette gjør det mulig for hotellene å tilpasse sine tilbud i samsvar med denne innsikten (Inversini & Masiero, 2014).

I litteraturen anvendes en prisingsteori kalt hedonisk prising, for å undersøke betalingsviljen til konsumenter, basert på ulike attributter. En informativ studie av Zhang et al. (2011) viser hvordan kundetilbakemeldinger kan anvendes til å analysere prissettingen av dobbeltrom i hotellmarkedet i New York ved bruk av hedonisk prising. Vi har latt oss inspirere av denne studien i forbindelse med oppgavens tema.

Et annet relevant aspekt for hoteller er deres miljømessige engasjement. Booking.com, en av verdens største nettbaserte bestillingsplattformer (Inversini & Masiero, 2014), rangerer hotellene på en skala fra null til tre, basert på deres prioritering av bærekraft (Booking.com, 2023). Til tross for denne rangeringen, eksisterer det tilsynelatende begrenset forskning på hvordan ulike hotellattributter og bærekraft påvirker hotellromprisene i Norge.

Derfor vil vi i denne oppgaven undersøke om norske hotellrompriser blir påvirket av kundevurderinger og hotellenes bærekraftsertifisering. Vi vil samle inn data fra Booking.com, en anerkjent nettbasert bestillingsplattform (Inversini & Masiero, 2014).

1.1 Problemstilling og avgrensning

Formålet med oppgaven er få en grundig forståelse av hvordan norske hotellrompriser kan påvirkes av kundevurderinger gitt på Booking.com og hotellenes eventuelle bærekraftsertifiseringer. Dette leder oss til følgende problemstilling:

Hvordan påvirkes hotellromprisene i det norske markedet av kundevurderinger og hotellenes bærekraftsertifiseringer?

Oppgaven har to hovedbegrensninger: For det første er vårt fokus begrenset til det norske markedet, og derfor kan vi ikke uten videre generalisere resultatene til andre land. Vi vil kun undersøke hotellrompriser og faktorene som påvirker dem, basert på data fra Booking.com. Vi

tar dermed ikke hensyn til tilsvarende informasjon fra andre plattformer som Expedia eller Airbnb.

Til tross for disse begrensningene vil oppgaven gi en omfattende innsikt i betydningen av kundevurderinger for hotellrompriser, samt hvordan bærekraftsertifiseringer kan påvirke prissettingen. Ved å grundig utforske denne problemstillingen, bidrar oppgaven til å utvide kunnskapen om hotellmarkedet i Norge.

1.2 Oversikt av oppgaven

I dette avsnittet gir vi en kort presentasjon av oppgavens struktur. Kapittel to gir en generell oversikt over bransjeinformasjonen som er relevant for temaet vi undersøker. Dette vil inkludere informasjon om markedet, relevante aktører og tredjepartskanaler. I kapittel tre vil vi gå mer i dybden på det teoretiske rammeverket rundt oppgaven. Vi vil presentere hedonisk pristeori og gjengi tidligere studier som benytter attributtprising. Deretter, i kapittel fire, vil vi presentere oppgavens utvalg og forklare hvordan vi har samlet inn data. I kapittel fem vil vi se nærmere på metoden vi har benyttet oss av. I kapittel seks vil vi presentere detaljer om den konkrete datamengden vi har benyttet. I kapittel syv vil vi presentere resultatene fra våre modeller, og diskutere hva funnene våre betyr i lys av det teoretiske rammeverket som presenteres i kapittel tre. Vi vil også diskutere mulige implikasjoner av funnene. Til slutt vil vi konkludere oppgaven og komme med forslag til fremtidig forskning i kapittel åtte.

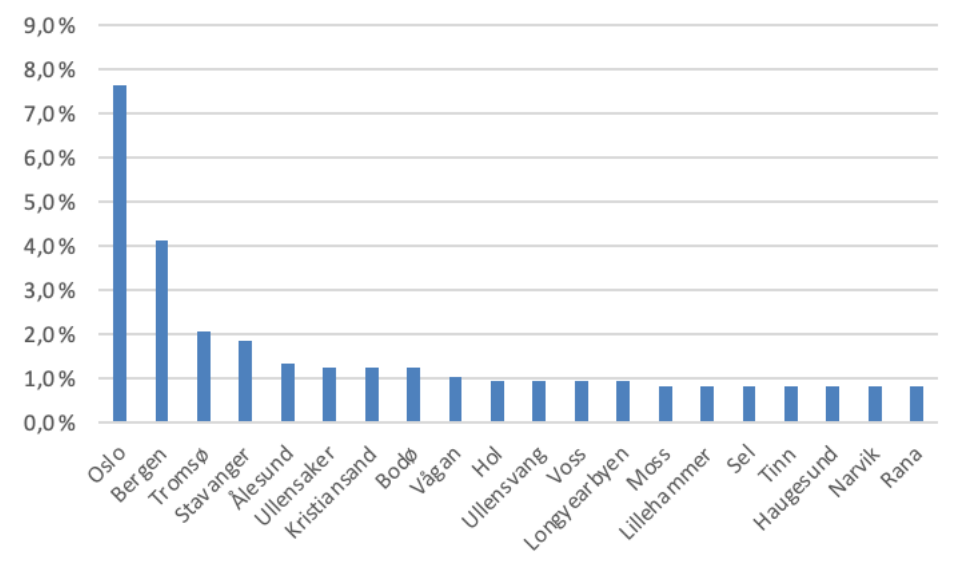
2 Generell bransjeinformasjon

I denne delen av oppgaven vil vi gi en beskrivelse av det norske hotellmarkedet. Vi vil utforske ulike aspekter som er sentrale for å forstå bransjen og de utfordringene vi står overfor i analysene. Først vil vi presentere en oversikt over hotellaktørene i Norge, deres roller og bidrag til markedet. Deretter vil vi dykke ned i prisvariasjonene i hotellromprisene og identifisere faktorer som påvirker disse forskjellene. Til slutt vil vi beskrive klassifiseringssystemene, herunder segmentrangering og bærekraftsertifisering, før vi ser på betydningen av tredjepartskanaler i hotellindustrien. En solid forståelse av dette bakgrunnsbildet vil danne grunnlaget for videre analyser og konklusjoner i oppgaven.

2.1 Det norske hotellmarkedet

Hotellkonsentrasjonen i Norge varierer betydelig mellom ulike kommuner. De største konsentrasjonene av hoteller finner vi i kommunene som huser de største byene, mens flere mindre kommuner har kun ett hotell (SSB, 2023). SSB (2023) definerer en hotellvirksomhet som en aktør som tilbyr rom for overnatting for én eller flere netter, og som tilbyr daglig rengjøringstjeneste, selv om servicenivået kan variere mellom hotellene.

I Figur 2.1, som illustrerer andelen hotellvirksomheter for de 20 største hotellkommunene, kan vi se at Oslo, Bergen og Tromsø er de tre mest hoteltette kommunene. Oslo har 74 hoteller (7,6 %), etterfulgt av Bergen med 39 hoteller (4,1 %) og Tromsø med 20 hoteller (2,1 %), av de totalt 968 registrerte hotellvirksomhetene i landet (SSB, 2023).



Figur 2.1: Et utdrag av de 20 største hotellkommunene i Norge (SSB, 2023).

Selv om hotellkonsentrasjonen er størst i Oslo og Bergen, er det også flere mindre byer og tettsteder som har et relativt høyt antall hoteller i forhold til befolkningstall og størrelse. Eksempelvis Oslo Lufthavn som tiltrekker seg gjester til Ullensaker kommune (Hotelia AS, 2023), som er den syvende største hotellkommunen i Norge (SSB, 2023).

2.1.1 Største hotellkjeder

Det norske hotellmarkedet er preget av at flere hotellvirksomheter tilhører samme merkenavn, som defineres som hotellkjeder (Ivanova & Ivanov, 2015). De fem største hotellkjedene i markedet er Strawberry, Scandic, Thon Hotels, Radisson og Classic Norway. Til sammen har aktørene over 50 % av alle hotellvirksomheter i Norge (Hotelia AS, 2023). Disse selskapene har et bredt spekter av hotellsegmenter som dekker ulike kundebehov, fra lavprishoteller til luksushoteller (Classic Norway Hotels AS, 2023; Scandic Hotels Group AB, 2023; Strawberry, 2023; Thon Hotels, 2023).

2.2 Prisvariasjonene i markedet

Hotellenes rompriser må gjenspeile kundenes totalopplevelse. Denne inkluderer vurderinger av faktorer som eksempelvis romkvalitet, tjenester og omgivelser. Kundene forventer god verdi for prisen de betaler, og hotellprisene bør derfor være i tråd med den kvaliteten og verdien hotellene tilbyr. I den sammenheng er en vanlig praksis i hotellbransjen å benytte seg av dynamisk prising for å bedre møte etterspørsel, hotellattributter og kundesegmenter (Lewis & Shoemaker, 1997).

En konkret tilpasning som hotellene kan gjøre, er å ta hensyn til variasjoner i etterspørselen og deretter justere prisene i tråd med dette. Ved høy etterspørsel er kundene ofte villig til å betale mer, og hotellene kan dermed øke sine priser. Motsatt kan hotellene senke prisene i lavsesonger for å tiltrekke seg gjester (Lewis & Shoemaker, 1997). For vår oppgave, som skal undersøke hvordan kundevurderinger og hotellenes bærekraftsertifiseringer påvirker rompriser, er det nødvendig å ta hensyn til at enkelte priser kan være unormalt høye grunnet stor etterspørsel i den undersøkte perioden. Dette hensynet vil vi diskutere ytterligere i kapittel 5.1.4 om reliabilitet.

2.3 Klassifisering av hotellene

Tidligere ønsket hotellene gjerne å ha en *one-size-fits-all*-tilnærming. Denne tankegangen har bransjen etter hvert forlatt, og en stigende trend i hotellmarkedet er å bygge hoteller som passer en bestemt målgruppe (Khan et al., 2018). Kundekategorisering, spesielt i fritids- og forretningsreisende, blir stadig viktigere på grunn av økt konkurranse (Khare & Bhagwat, 2019). Ved å kategorisere kundene i ulike grupper med forskjellige preferanser og behov, kan hotellene tilpasse sine produkter disse gruppene. Dette for å oppfylle de reisendes unike krav (Khan et al., 2018). Hotellets segmentrangering (stjernerangering) er et eksempel på en klassifisering som differensierer kundene (Mohsina et al., 2019). En annen klassifisering er bærekraftsertifisering (Cavero-Rubio & Amorós-Martínez, 2008).

2.3.1 Segmentrangering av hotellene

En vanlig praksis i hotellbransjen er å rangere hotellene etter segmentskala (stjerneskala). Skalaen gir kundene en oversikt over hotellkvaliteten, fra en enkel og billig overnatting på hoteller i segment én, til den høyeste kvaliteten hos hoteller i segment fem (Andersson, 2000). Målet er i hovedsak å gi en nivåindikasjon på romkvaliteten, servicenivået og fasilitetene som tilbys (Narangajavana & Hu, 2008). I vår undersøkelse vil vi segmentere hotellene fra én til fem, og velger å definere stjernerangeringen som segment.

Et problem ved å klassifisere hotellene fra én til fem er at flere aktører har forsøkt å manipulere seg opp et segment (Israeli, 2002). Et annet problem er at enkelte land ikke har en offisiell segmentrangering fordi myndighetene ikke fastslår en slik rangering. Norge er et slikt land (Norsk Hotellklassifisering, 2023). Flere private aktører har likevel sine egne segmentvurderinger, eksempelvis Booking.com og Expedia. Derimot så er ikke alltid segmentrangeringen mellom plattformene nødvendigvis konsistente, noe som fremkommer av en undersøkelse av hotellmarkedet i Hong Kong. Der hadde 75 % av hotellene ulik rangering på tvers av plattformene (Guillet & Law, 2010). For vår studie, som vil ta utgangspunkt Booking.com, betyr dette at undersøkelsen ikke nødvendigvis kan generaliseres til andre plattformer.

Segmenteringen kan også være misvisende med hensyn på kvalitet. Dette skyldes for eksempel utdaterte kriterier for segmentrangeringer. Til tross for dette, så kan segmentrangering bidra til å redusere effektene av asymmetrisk informasjon i hotellmarkedet (Martin-Fuentes, 2016).

2.3.2 Bærekraftsertifisering

Stadig flere er opptatt av å redusere sitt karbonavtrykk når de reiser. For å møte denne etterspørselen, har flere overnattingsleverandører begynt å tilby miljøvennlige alternativer. For eksempel gjennom grønne hoteller (Kuminoff et al., 2010).

Bærekraftsertifiseringer er en måte å differensiere hotellene på (Cavero-Rubio & Amorós-Martínez, 2008). Fra et bærekraftsperspektiv er bærekraftsertifiseringer den vanligste metoden for hotellene å differensiere seg på (Segarra-Oña et al., 2012). Bærekraftseringsprogrammer gir i så måte hotellgjestene mer nøyaktig informasjon om hotellets miljøytelse. Programmene inneholder ofte en logo og et hierarkisk system som gir kundene en øyeblikkelig bekreftelse på at hotellet er forpliktet til å redusere sin miljøpåvirkning. Graden av bærekraft vil på den måten bidra til den nevnte differensieringen. Å ta del i et slikt program kan også bidra til å motvirke at kundene beskylder hotellene for grønnvasking (Bernard & Nicolau, 2022). Kundene kan på sin side benytte Booking.com til å se hotellets bærekraftsertifiseringsnivå, hvor Booking.com graderer hotellenes bærekraftsertifiseringsnivå fra null til tre, der tre er det beste nivået (Booking.com, 2023c).

Siden mange reisende er opptatt av sin egen påvirkning på miljøet, kan informasjon om hotellets bærekraftsertifisering være verdifull under bestillingsprosessen. Dette kan også føre til at hoteller med slike sertifiseringer får et konkurransefortrinn over de som ikke har en slik sertifisering (Bernard & Nicolau, 2022). Jungelen av sertifiseringer kan også medføre at konsumentene ikke gjenkjenner sertifiseringene, som igjen kan bidra til økt skepsis, og dermed ingen prispåvirkning (Chen & Chang, 2013).

En studie av Geerts (2014) viser at andelen bærekraftsertifiseringer økte avhengig av hvilket segmentnivå hotellet befant seg i. I samme studie finner forfatteren det vanskelig å konkludere med om det faktisk hjalp hotellene å være bærekraftsertifiserte. Denne utfordringen var særlig knyttet til at hotellene ikke markedsførte seg godt nok.

2.4 Tredjepartskanaler (OTA)

Informasjon- og kommunikasjonsteknologi (IKT) har preget hotellnæringen sin utvikling. Særlig tilknytningen til internettbaserte kundevurderinger på tredjeparts distribusjonskanaler har satt sitt fotfeste. Et eksempel på en slik kanal er TripAdvisor, som blir definert som en *nettbasert tilbakemeldingstjeneste (OR, Online rating)* (Inversini & Masiero, 2014). En annen

type IKT-tjeneste er kanalene som både selger hotellrom og tilbyr en tilbakemeldingstjeneste. Slike nettsider blir definert som *Internettbaserte reiselivsaktører (OTA, online travel agents)*. Eksempler på slike tjenester i hotellnæringen er Booking.com, Expedia og Hotels.com (Inversini & Masiero, 2014).

Booking.com er den tjenesten konsumentene benytter mest (Inversini & Masiero, 2014). Et utvalg hotellsjefer i London, forklarer til Inversini og Masiero (2014) at kanalen genererer høyest antall nettsalg av hotellrom, sammenlignet med tilsvarende konkurrenter. Hele 51,2 % av hotellsjefene rangerer denne kanalen som viktigst. På en annen side er det kun 60 % av alle hotellvirksomheter som benytter seg av alle tilgjengelige distribusjonskanaler (Guillet & Law, 2010). Med bakgrunn i disse funnene vil vi benytte oss av Booking.com sin plattform til å innhente data for denne undersøkelsen.

For kundene gir tredjepartskanalene tilgang til et bredere utvalg av hotellalternativer, samt muligheten til å sammenligne priser og finne gode tilbud. Plattformene gir også kundene muligheten til å lese tilbakemeldinger fra tidligere gjester, noe som kan gi dem et bedre beslutningsgrunnlag ved valg av hotell (Webb, 2016).

Håndteringen av tilbakemeldinger er også ulik mellom plattformene. Noen aksepterer tilbakemeldinger fra alle, uavhengig om de har bodd på hotellet, mens andre plattformer kun aksepterer tilbakemeldinger fra tidligere gjester ved det respektive hotellet. På Booking.com kan kun tidligere gjester (innen seks måneder) vurdere sitt opphold, mens på TripAdvisor kan hvem som helst dele sin opplevelse (Dina & Chevalier, 2014).

For tilbudssiden har tredjepartskanaler vært en mulighet for økt synlighet og tilgang til en ekstra salgskanal (Inversini & Masiero, 2014). Kanalene gir også hotellene økt kunnskap om kundenes kjøpsatferd (Brown et al., 2007). Med dette kan de tilpasse tjenestene og dermed øke kundetilfredsheten (Felix & Clever, 2014). Hvis en tilbakemelding indikerer at en bestemt egenskap ved hotellet er spesielt verdifull, kan hotellene vurdere å øke prisen på produktet eller tilby flere alternativer som inkluderer denne egenskapen (Nangpiire et al., 2022).

For kundene, så muliggjør også kanalene en enkel enveis kommunikasjon med andre kunder. Et fenomen i denne sammenhengen er at hvis en kunde har hatt en dårlig opplevelse med et hotell og annonserer dette gjennom tilbakemeldingskanalen, så vil dette ha større effekt på

andre gjesters oppfatning av hotellet enn positive tilbakemeldinger (Chatterjee P. , 2006). Kundene oppfatter det også som mindre risikabelt å bestille hotellrom hos hoteller med tilbakemeldinger fra tidligere gjester (Gretzel & Yoo, 2008).

Det er imidlertid også flere utfordringer knyttet til bruk av OTAer. Et eksempel er hvis hotellaktører manipulerer vurderingene i sin favør. Mayzlin et al. (2014) finner for eksempel at en av tilbyderne ble rangert med 3,92 av 5,00 hos Expedia, mens samme tilbyder var i snitt rangert med 3,52 av 5,00 hos TripAdvisor. En forklaring kan være at hotellene manipulerer konkurrentenes OTA-vurdering, samtidig som de rangerer seg selv høyt (Dina & Chevalier, 2014).

En siste utfordring er kostnadene tilknyttet tredjepartskanalanene. For Booking.com er denne kostnaden i snitt 10,1 % av salgssummen, og kan gjøre at flere hotellaktører velger å ikke benytte seg av tredjepartskanalanene (Inversini & Masiero, 2014).

2.4.1 Kundevurderinger på Booking.com

Tredjepartskanalanene har ulike kvantifiserbare kriterier som kundene kan gi tilbakemeldinger på. Hos Booking.com er det syv slike vurderingskriterier: *beliggenhet, fasiliteter, komfort, personale, renhold, WiFi (gratis)* og *verdi for pengene* (Booking.com, 2023a). Poengsummen for hvert kriterium varierer fra 1,0 til 10,0. En presis beskrivelse av hver variabel er vedlagt i Appendiks A1.

2.5 Oppsummering av generell bransjeinformasjon

I dette kapitlet har vi presentert den generelle bransjeinformasjonen, som vi mener er nødvendig for å forstå hvor komplekst det norske hotellmarkedet er, der hotellkonsentrasjonen er ulik mellom regionene og segmentrangeringen også er ulik. Til slutt gir bransjeinformasjonen innsikt i de utfordringene tredjepartskanalanene fordrer i vår undersøkelse.

3 Teoretisk rammeverk

Formålet til oppgaven er å se på hvordan kundevurderinger og hotellenes bærekraftsertifiseringer påvirker romprisene, gjennom hedonisk pristeori. I delkapittel 3.1 vil vi gjengi den hedoniske pristeorien og dens anvendelse på hotellmarkeder. Deretter vil vi se på begrensninger ved den hedoniske pristeorien og utfordringene knyttet til den. I kapitelets nest siste del, vil vi beskrive relevante studier som har sett på hvordan ulike hotellattributter og hotellenes bærekraftsertifisering påvirker rompriser. I siste del av kapittelet vil vi presentere halo-effekten, et relevant psykologisk fenomen som kan påvirke vår analyse.

3.1 Den hedoniske prismodellen

3.1.1 Implementering av den hedoniske pristeorien

Den hedoniske pristeorien tar utgangspunkt i at prisen på et gode er bestemt av dets nyttebærende attributter, som igjen er avhengig av kundenes preferanser for disse attributtene. For eksempel kan en kunde være villig til å betale en høyere pris for et hotell med høy komfort eller en bedre beliggenhet, mens de kan akseptere en lavere pris for et hotell med lavere servicegrad og dårligere beliggenhet. Gjennom en økonometrisk tilnærming, spesifikt gjennom en hedonisk regresjonsanalyse, kan man analysere hvordan kundene verdsetter disse attributtene og dermed bestemme den optimale prisen for godet (Rosen, 1974).

Hotellbransjen er et eksempel på et marked hvor hedonisk prising ofte anvendes av forskere til å analysere hvordan ulike hotellattributter påvirker rompriser (Sánchez-Peres & Illescas-Manzano, 2019). Prisen på et hotellrom bestemmes da av ulike attributter tilknyttet rommet, som for eksempel størrelse, fasiliteter, beliggenhet og renslighet. Gjennom å forstå hvordan kundene verdsetter slike attributter, det vil si hvor mye de er villige til å betale for dem, kan tilbyderne bestemme den optimale prisen på godet. Dette gjøres gjennom en regresjonsanalyse, hvor prisfunksjonen kan være log-lineære former (Rosen, 1974). Ved å anvende den hedoniske pristeorien på data fra hotellbransjen, kan en dermed få innsikt i hvordan ulike attributter påvirker hotellromprisene og dermed gjøre mer informerte prissettingsbeslutninger (Sánchez-Peres & Illescas-Manzano, 2019).

3.1.2 Utledning av den hedoniske prismodellen

Prisen på en hotellovernatting avhenger, som tidligere forklart, av flere egenskaper. For å beregne rompriser i markeder bestående av ikke-homogene priser, vil vi anvende Rosens (1974) hedoniske prismodell. En detaljert utledning av denne vil bli presentert i dette underkapittelet.

Rosen (1974) deler den hedoniske prismodellen i to steg. Første steg innebærer å estimere marginalprisen på hvert attributt ved hjelp av en regresjon. Dette gir oss et mål på prisen. Neste steg er å finne den inverse etterspørselskurven for godet gjennom de implisitte prisene fra første steg. Den inverse etterspørselskurven representerer konsumentenes marginale betalingsvilje, som ifølge Rosen (1974) øker med inntekt.

3.1.2.1 Markedspris og likevekt under den hedoniske prismodellen

Den hedoniske prismodellen består av n attributter, som sammen utgjør følgende vektor i ligning 3.1. De resterende ligningene fra 3.1 til 3.16 er basert på Rosen (1974):

$$Z = (z_1, z_2, \dots, z_n) \quad (3.1)$$

I en analyse av hotellrompriser er det naturlig å undersøke konsumentenes betalingsvilje for ulike attributter som beliggenhet, komfort, service, renhold og andre prispåvirkende attributter. Ved å gjøre en marginal endring i ett av attributtene og se hvordan dette påvirker sluttprisen på godet, blir det mulig å måle attributtens bidrag til sluttprisen. Prisen på en hotellovernatting blir dermed bestemt av følgende funksjon, hvor totalprisen $P(Z)$ består av hvert enkelt attributts implisitte pris (Rosen, 1974):

$$P(Z) = P(z_1, z_2, \dots, z_n) \quad (3.2)$$

Denne markedsprisen bestemmes også av forholdene for likevekt mellom nyttemaksimerende konsumenter og profittmaksimerende produsenter (Osland, 2001). Videre er prisfunksjonen $P(Z)$ definert ved hvert punkt langs aksene. Dette hjelper produsentene og konsumentene med å ta informerte valg i forbindelse med de respektive attributtbuntene (goder), som kjøpes eller selges. På den måten maksimerer markedsaktørene sine respektive nyttefunksjoner i det punktet på aksene som samsvarer med tilhørende likevekt mellom tilbud og etterspørsel. I dette punktet er det ikke mulig for aktørene å forbedre sine valg. Samtidig bestemmes markedsprisene, $p(z)$, ut fra konsumentenes preferanser og produsentenes kostnader (Rosen, 1974).

3.1.2.2 Etterspørselssiden

For konsumentene kan vi tenke oss at vi har et sett med attributter i vektoren fra ligning 3.1. Deres nyttefunksjon ved konsum av et produkt med markedsverdi $P(Z)$, er dermed gitt ved $U = (x, z_1, z_2, \dots, z_n)$, der x står for andre konsumerte goder. Denne faktoren inkluderes fordi den enkelte konsument i hotellmarkedet sin strengt konkave nyttefunksjon også vil påvirkes av andre konsumerte goder. Disse andre konsumerte godene betegnes hver for seg med x_j , og settes for enkelhets skyld lik 1. Videre settes den enkelte konsuments inntekt, y , i enheter av x_j . Dermed er $y = x + p(z)$ (Rosen, 1974). Nyttefunksjonen avhenger også av hotellets attributter, z_i . Dette fordi konsumentene ikke bare mottar nytte fra godet, men også fra attributtene det bærer (Lancaster, 1966).

Med hensyn på deres respektive budsjettbetingelse, vil konsumentene maksimere sin nytte ved å velge det optimale nivået for x og z_1, z_2, \dots, z_n :

Max:

$$U_j = U(X, Z) \text{ gitt } Y_j = X + P(Z) \quad (3.3)$$

hvor

X = Konsum av andre goder enn hotellattributtene

Y_j = Konsument j sin inntekt

$P(Z)$ = Hedonisk prisfunksjon

Lagrange for denne målfunksjonen, som maksimerer funksjonen gitt dens begrensninger, er

$$L = U(x, z_1, z_2, \dots, z_n) + \lambda(y - x - p(z)) \quad (3.4)$$

Hvis konsumenten skal maksimere sin nytte med hensyn til budsjettbetingelsen, så må de tilpasse seg slik at både x og z_1, z_2, \dots, z_n tilfredsstiller førsteordens betingelse samt deres budsjett. Dette betyr at marginal substitusjonsrate mellom to attributter må være lik det marginale prisforholdet mellom de to respektive attributtene (Rosen, 1974).

$$\frac{\frac{\partial p}{\partial z_i}}{\frac{\partial p}{\partial z_j}} = \frac{U_{z_i}}{U_{z_j}} \quad \forall i, j = 1, \dots, n, i \neq j \quad (3.5)$$

$$\frac{\partial p}{\partial z_i} = p_{z_i} = \frac{U_{z_i}}{U_x} \quad \forall i = 1, \dots, n \quad (3.6)$$

Konsumentens maksimale betalingsvilje for ulike kombinasjoner av z , for et gitt nyttenivå u med inntekt y , blir dermed verdien eller budfunksjonen til konsumenten (Rosen, 1974). Budfunksjonen defineres som

$$\theta(z_1, z_2, \dots, z_n, u, y) \quad (3.7)$$

På implisitt form blir denne

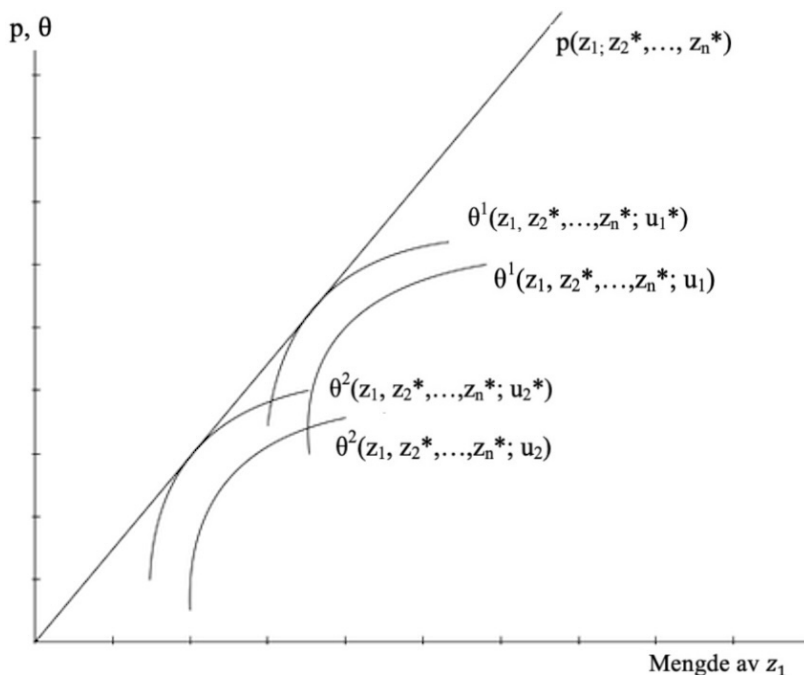
$$U(y - \theta, z_1, z_2, \dots, z_n) = u \quad (3.8)$$

Da er optimalt nyttenivå for den enkelte konsument, u^* , en funksjon av inntekt y og godets attributter (Rosen, 1974).

Hvis vi så differensierer budfunksjonen med hensyn på z_i , får vi førsteordensbetingelsen, som er konsumentens implisitte marginalverdi av attributt z_i , gitt deres inntekts- og nyttenivå (Osland, 2001).

$$\theta_{z_i} = \frac{U_{z_i}}{U_x} > 0 \quad \forall i = 1, \dots, n. \quad (3.9)$$

Når marginalbudet er lik marginalprisen for et attributt, vil nyttenivået være optimalt. Da er $\theta_{z_i} = p_{z_i}$ (Rosen, 1974). Vi kan illustrere dette med Figur 3.1, som viser budfunksjonene til to konsumenter. Her antas en konstant marginalpris.



Figur 3.1: Konsumtilpasningen i den hedoniske prismodellen (basert på Rosen, 1974).

Konsumentenes nyttefunksjoner kommer frem av indifferenskurvene i Figur 3.1. Disse kurvene er konkave og dermed avtakende når prisen øker. Budfunksjonene til hver enkelt konsument er også forskjellige, ettersom deres nyttefunksjoner er ulike. En konsument med høy inntekt eller med preferanse for mer kostbare attributter, vil tilpasse seg høyere i diagrammet langs prisfunksjonen (θ^1), sammenlignet med en konsument med lavere inntekt eller mindre kostbare preferanser (θ^2). En logisk slutning er at nyttenivået stiger desto lengre ned en beveger seg i diagrammet. Dette fordi konsumenten kjøper et gode som tilbyr mer av et gitt attributt z_i (Rosen, 1974).

3.1.2.3 Tilbudssiden

Produsentene må på sin side avgjøre hvor mange enheter av hotellrom de skal selge, samt hvilken kombinasjon av attributter disse hotellrommene skal ha. Ifølge Rosen (1974) bestemmes profitten deres dermed av:

$$\pi = M \cdot p(z) - C(M, z_1, \dots, z_n, \beta) \quad (3.10)$$

hvor

$p(z)$ er den implisitte prisfunksjonen for et hotellroms attributter

M er antall hotellrom med attributter z

β er en skiftparameter for kostnadene

Førsteordens betingelser for produsentene blir da

$$\frac{\partial \pi}{\partial z_i} = \pi_{z_i} = M \cdot P_{z_i} - C_{z_i} = 0 \Rightarrow p_{z_i} = \frac{C_{z_i}}{M}, \forall i = 1, \dots, n \quad (3.11)$$

$$\frac{\partial \pi}{\partial M} = \pi_M = p(z) - C_M = 0 \Rightarrow p(z) = C_M, \forall i = 1, \dots, n \quad (3.12)$$

Ifølge ligning 3.12 bør det respektive hotellet kun selge hotellrom – med et bestemt sett av attributter – hvor marginalinntekten for hvert attributt er lik marginalkostnaden av å produsere hotellrommet. Videre viser ligning 3.12 at antall hotellrom som tilbys må være slik at prisen på hotellrommet er lik marginalkostnaden av å selge et ekstra hotellrom (Rosen, 1974).

På lik linje med at konsumentene har en budfunksjon, så har produsentene en offerfunksjon. Denne defineres som den enhetsprisen produsenten er villig til å godta for et hotellrom med

attributter z , så lenge dette gir en fortjeneste π , gitt skiftparameter β (Rosen, 1974). Offerfunksjonen defineres som

$$\phi(z_1, \dots, z_n, \pi, \beta) \quad (3.13)$$

Vi kan dermed finne den enhetsprisen produsentene er villige til å akseptere ved å løse for ϕ fra følgende funksjon

$$\pi = M \cdot \phi - C(M, z_1, \dots, z_n) \quad (3.14)$$

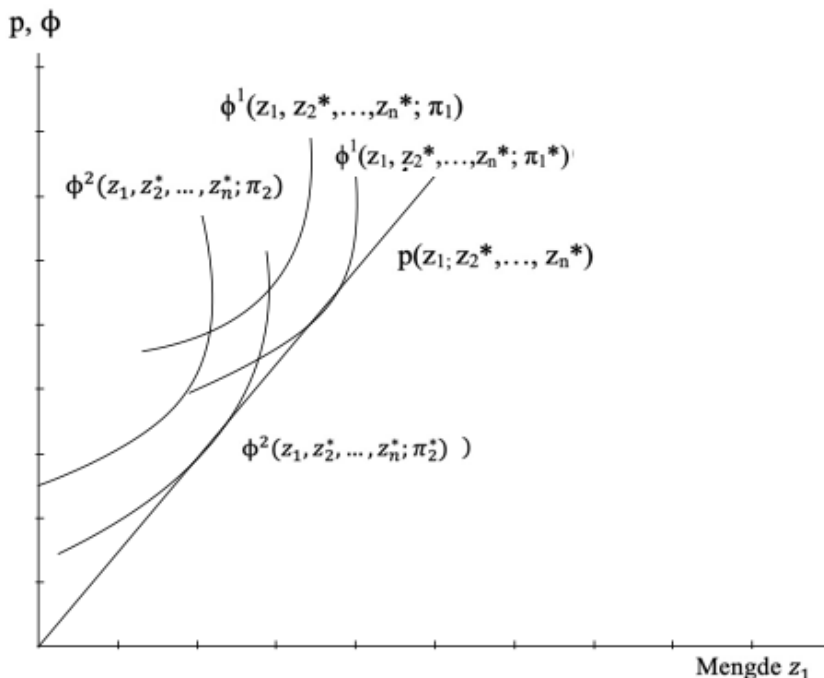
Dette gir

$$\phi = \frac{\pi - C(M, z_1, \dots, z_n)}{M} \quad (3.15)$$

Deretter kan vi finne førsteordens betingelse for offerfunksjonen ved å differensiere ligning 3.15 med hensyn på z_i

$$\phi_{z_i} = \frac{C_{z_i}}{M} > 0 \quad \forall i = 1, \dots, n \quad (3.16)$$

Med dette kan vi finne det optimale nivået hvor marginaltilbudet av attributt i er lik dens marginalkostnad (Rosen, 1974). Dette er hvor $p(z)$ tangerer $\phi(z, \pi^*, \beta)$ i Figur 3.2 nedenfor.



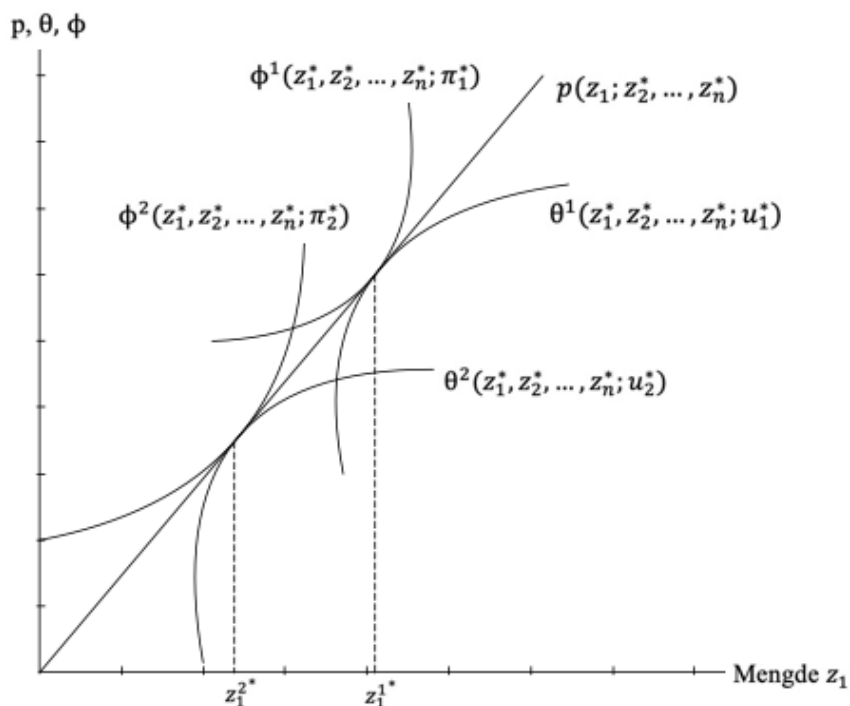
Figur 3.2: Produksjonstilpasningen i en hedonisk prismodell (basert på Rosen, 1974).

Figur 3.2 viser tilpasningen til to forskjellige produsenter, gitt deres offerfunksjoner ϕ^1 og ϕ^2 . Kurvene for produsentenes offerfunksjoner ligger i et z_1 - ϕ plan, som skjærer gjennom

indifferensoverflaten ved de optimale verdiene for andre attributter enn z_1 . Tilpasningen er slik at optimalt produksjonsnivå ligger der offerfunksjonene ϕ^1 og ϕ^2 tangerer den lineære prisfunksjonen $p(z)$. Videre vil produsenten med offerfunksjon ϕ^2 produsere færre enheter av z_1 , sammenlignet med produsenten med offerfunksjon ϕ^1 . Dette fordi sistnevnte har et komparativt fortrinn i forbindelse med produksjonen og kostnader, slik at denne produsenten kan produsere flere enheter av z_1 (Rosen, 1974).

3.1.2.4 Markedslikevekt

Nå som vi har innsikt i hvordan etterspørsels- og tilbudssiden av markedet tilpasser seg, kan vi finne markedslikevekten, som gir oss markedsprisen. Dette oppstår ved det punktet hvor konsumentenes budfunksjoner og produsentenes offerfunksjoner møtes og treffer prisfunksjonen, og det er denne likevekten som bestemmer markedsprisen. (Rosen, 1974). Likevekten reflekterer konsumentenes ønske om å betale en lavest mulig budpris for maksimal nytte, samt produsentenes ønske om en høyest mulig offerpris for maksimal profitt (Osland, 2001). Markedslikevekten er vist i Figur 3.3 nedenfor.



Figur 3.3: Markedslikevekt under den hedoniske prismodellen (basert på Rosen, 1974).

I Figur 3.3 ser vi markedslikevekten mellom konsumentenes budfunksjoner og produsentenes offerfunksjoner. For to konsumenter og produsenter, vil tilpasningen ligge langs den lineære prisfunksjonen. Optimalt konsum og produksjon av attributt z_1 blir da z_1^{1*} og z_1^{2*} , respektivt. Dette til en gitt pris, som er bestemt av hvor budfunksjonene og offerfunksjonene tangerer hverandre i prisfunksjonen. Likevekten representerer dermed den hedonistiske prisfunksjonen (Rosen, 1974).

3.1.3 Begrensninger ved hedonisk pristeori

I vår oppgave, hvor vi anvender den hedoniske prismodellen til å analysere hotellrompriser, er det viktig å være oppmerksom på begrensningene ved modellen. Chin og Chau (2003) påpeker at modellen krever mye data som kanskje ikke er tilgjengelig og at estimatene kan være forutinntatte grunnet utelatte variabler. Dette kan føre til upålitelige resultater, spesielt hvis informasjonen om prisene til et gode og dets relevante attributter ikke er av tilfredsstillende kvalitet. Videre kan en stor mengde uavhengige variabler skape estimeringsproblemer og utfordringer knyttet til multikollinearitet (Sirmans et al., 2005; Stock & Watson, 2020). Disse begrensningene vil vi beskrive i nærmere detalj i kapittel 5. Vi må også ta hensyn til de når vi tolker resultatene fra våre hedoniske regresjonsmodeller i delkapittel 7.2 og 7.3.

3.1.4 Oppsummering av hedonisk pristeori

I dette delkapittelet har vi gjengitt den hedoniske pristeorien og påpekt dens relevans i vår analyse. Vi betrakter hotellrom som differensierte goder med ulike attributter, for eksempel beliggenhet og fasiliteter. Prisen på et hotellrom bestemmes derfor av de implisitte prisene på hotellrommets attributter.

Hotellenes produksjonskostnader bidrar til å opprettholde markedslikevekten på tilbudssiden, mens konsumentenes betalingsvilje, basert på deres preferanser og budsjettbetingelser, bidrar fra etterspørselssiden. Markedslikevekten viser i så måte hvordan prisene på hotellrom og deres attributter danner hedoniske prisfunksjoner, som kan brukes til å måle hvordan attributtkvalitet og implisitte priser påvirker sluttprisen på hotellrommet.

Vi har også diskutert utfordringene ved å estimere attributtene effekt, noe som vil være avgjørende i vår analyse. Spesielt når tilgjengelige data om attributtene er begrenset eller hvis relevante variabler blir utelatt. Vi må også være oppmerksomme på eventuell multikollinearitet

når vi inkluderer mange variabler for å estimere prisfunksjonen, da dette kan påvirke estimatene og tolkningen av resultatene i vår analyse.

3.2 Hedoniske prisstudier av hotellmarkeder

I dette avsluttende delkapittelet vil vi vise til studier som har anvendt hedonisk pristeori på variablene *beliggenhet, fasiliteter, komfort, personale, renhold, WiFi* og *bærekraftsertifisering*, som vi skal undersøke i denne oppgaven. Vi vil se på hvordan disse hotellattributtene påvirker romprisene i disse studienes respektive markeder.

Delkapittelet starter med å utforske tidligere studier som har analysert om *segmentnivå* påvirker hotellromprisen. Deretter ser vi på hvordan variablene *beliggenhet, fasiliteter, komfort, personale, renhold, WiFi* og *bærekraftsertifiseringsnivå* påvirker romprisene.

3.2.1 Segmentnivå

Segmentrangering, bedre kjent som stjernerangering i hotellbransjen, indikerer hotellets kvalitet fra én til fem og skiller hotellene fra hverandre. Vi har valgt å benevne det som «segmentnivå» i denne oppgaven, da det ikke finnes et offisielt stjernerangeringssystem i Norge. Flere studier undersøker segmentnivåets betydning for rompriser i ulike markeder, som vi nå skal gjengi.

Zhang et al. (2011) identifiserte at hotellromprisene i New York ble påvirket av segmentnivå, og en økning i segmentnivå førte til en prisøkning på 35,8 %. Segmentnivået var også den variabelen som hadde størst betydning for romprisene i deres studie. I en studie utført av Espinet et al. (2003) i Costa Brava, finner forfatterne at forskjellen i rompris fra segment én og to var minimal, mens forskjellen fra segment tre og fire kunne resultere i en prisøkning på opptil 50 %. Dette tyder på at et hotell i segment tre kan øke inntektene ved å oppgradere til segment fire. Imidlertid kunne ingen konklusjoner trekkes for hotellene i segment fem, da Espinet et al. (2003) ikke inkluderte disse hotellene i analysen.

En annen studie utført av Israeli (2002), viser at segmentnivå har en signifikant positiv effekt på hotellromprisene i byene Eilat, Tel Aviv og Jerusalem. Eilat med en effekt på 35,15 %, Tel Aviv på 47,13 % og Jerusalem på 56,39 %. De påpeker at segmentrangeringen er den viktigste prediktoren for hotellromprisene i de undersøkte regionene.

Castro & Ferreira (2018) finner også at segmentnivå har en signifikant effekt på romprisene i Lisboa. Deres utvalg inkluderer hoteller i segment tre, fire og fem. Ved bruk av dummyvariabler for segmentene, kunne de evaluere segmentnivåenes effekt på pris per økning i segment, relativt til segment tre. Hoteller i segment fire hadde 7,1 % høyere predikert pris, sammenlignet med segment tre. På samme måte hadde hoteller i segment fem en 39,2 % høyere predikert pris. Dermed var segmentnivåets effekt på rompris størst i segment fem. Castro & Ferreira (2018) konkluderer derfor med at segmentnivå er en viktig faktor for prisbestemmelsen, samt et signal om kvalitet.

De gjennomgåtte studiene indikerer at det er forskjell mellom de geografiske markedene i hvordan segmentnivå påvirker hotellromprisen, men samtlige finner at en økning i segmentnivå gir en høyere pris (Castro & Ferreira, 2018; Espinet et al., 2003; Israeli, 2002; Zhang et al., 2011).

3.2.2 Beliggenhet

Beliggenhet refererer til plasseringen av et hotell. En god beliggenhet kan innebære nærhet til turistattraksjoner og sentrum (Zhang et al., 2011). For den generelle faktoren *beliggenhet*, så finner Castro og Ferreira (2018) i hotellmarkedet i Lisboa at et ekstra poeng i tilbakemeldingene (fra én til ti), vil gi en prisøkning på 18,6 %. Zhang et al. (2011) finner på sin side at et ekstra poeng (fra én til fem) gir en prisøkning på 12,4 % i hotellmarkedet i New York. Ingen av disse forfatterne diskuterer faktoren *beliggenhet* ytterligere.

Flere studier undersøker imidlertid betydningen av *beliggenhet*, særlig knyttet opp mot avstanden til sentrum. I Bull (1994) sin studie fra Australia, påpekes det at beliggenhet er den viktigste faktoren for å oppnå en høy rompris. Forfatteren finner at romprisen faller gjennomsnittlig med 8,5 % for hver ekstra engelsk mil (ca.1,6 km.) unna sentrum. Thrane (2007) undersøker på sin side betydningen av beliggenhet for romprisene i Oslo, men finner ingen signifikant sammenheng mellom nærhet til sentrum og rompris. I Espinet et al. (2003) sin studie av strandhoteller i Spania, finner de at hotellene som ligger i førstelinje mot stranden oppnår i snitt 19,4 % høyere rompris enn hotellene i andre linje. De finner imidlertid ikke at nærhet til sentrum har betydning for hotellromprisene.

Thrane (2007) finner heller ingen sammenheng mellom avstanden til sentrum og hotellromprisene i sin studie av hotellmarkedet i Oslo. White & Mulligan (2010) finner i sin studie av Four Corners-regionen at støynivået i området er en vel så viktig faktor som urbane omgivelser. Dette kan indikere hvorfor Thrane (2007) ikke finner en prisøkning ved nærhet til sentrum, i og med at støynivået antageligvis er høyere i Oslo sentrum. Thrane (2007) sitt funn kan også sees i sammenheng med Arbel & Pizam (1977), som finner at hotellgjestene gjerne kan reise opp til tjue minutter, uten at det går ut over betalingsviljen.

Oppsummert ser vi at beliggenhet kan være en viktig faktor for hotellromprisene (Arbel & Pizam, 1977; Bull, 1994; White & Mulligan, 2010).

3.2.3 Fasiliteter

Fasiliteter refererer til de tjenestene og tilbudene som er tilgjengelige for hotellets gjester. Eksempelvis parkering, frokost og treningsfasiliteter (Thrane, 2007). Flere studier har sett på den overordnede variabelen *fasiliteter*, altså en felles variabel for alle typer hotellfasiliteter. Forfatterne av disse studiene finner at fasilitetene har en signifikant effekt på romprisen (Andersson, 2008; Castro & Ferreira, 2018; Saez et al., 2003).

For den overordnede variabelen *fasiliteter*, finner Espinet et al. (2003) en effekt på 8,5 % i Spania, mens Andersson (2008) finner en effekt på 4,6 % i Singapore og Castro og Ferreira (2018) finner en effekt på 30,8 % i Lisboa. Det er imidlertid interessant å se nærmere på hvilke fasiliteter som påvirker romprisene. Bull (1994) finner for eksempel at hoteller med restaurant kan øke romprisene med 20 %. Thrane (2007) finner i Oslo at gratis parkering har en effekt på 18,6 %, minibar en effekt på 26,6 % og hårføner en effekt på 38,7 %. Thrane (2007) påpeker at de høye koeffisientene skyldes at de få hotellene som ikke tilbyr minibar og hårføner er lavprishoteller, og derav blir prisforskjellene store.

De tidligere studiene indikerer at fasiliteter kan påvirke romprisene på forskjellige måter, avhengig av markedet og undersøkelsesmetoden (Andersson, 2008; Bull, 1994; Castro & Ferreira, 2018; Chen & Rothschild, 2010; Saez et al., 2003).

3.2.4 Komfort

Komfort refererer til gjestenes opplevelse av bekvemmelighet og velvære under hotelloppholdet. Dette kan inkludere en god seng (White & Mulligan, 2010). Flere studier undersøker sammenhengen mellom oppfattet komfort og hotellrompriser (Castro & Ferreira, 2018; Stringam et al., 2010).

Stringam et al. (2010) finner i sin studie at komfort er sterkt korrelert med total tilfredshet i de 100 største byene i USA. De rangerer faktisk komfort og service (personale) som de aller viktigste faktorene for å oppnå høye gjennomsnittlige vurderinger, selv for andre attributter. Imidlertid konkluderer de med at det er vanskelig å gi konkrete anbefalinger for å forbedre komfortvurderingene, da faktorer som er viktige for ett individ kanskje ikke har samme betydning for et annet. Dette understreker kompleksiteten og subjektiviteten i gjestenes oppfatning av komforten på hotellrom.

For Lisboa-markedet, indikerer studien til Castro og Ferreira (2018) at det ikke er en sammenheng mellom komfort og pris, uten å underbygge dette funnet ytterligere. Dette gir likevel en indikasjon på hvordan kundene verdsetter denne variabelen i en storby.

I sum viser disse studiene at det er en kompleks sammenheng mellom hotellrompriser og oppfattet komfort (Castro & Ferreira, 2018; Stringam et al., 2010).

3.2.5 Personale

Personale refererer til gjestenes evalueringer og vurderinger av hotellets ansatte og servicekvaliteten de leverer. Flere studier har sett på hotellpersonalets effekt på hotellrompriser. Variabelen benevnes imidlertid ulikt i forskjellige studier, der «Service» (Zhang et al., 2011), «Room Service» (Schamel, 2012) og «Staff» (Castro & Ferreira, 2018) er noen av eksemplene. Vi likestiller begrepene som presentert ovenfor med variabelen *personale* hos Booking.com.

Flere ulike regioner har blitt undersøkt for korrelasjon mellom servicenivå og hotellrompriser, med sprikende resultater. Castro og Ferreira (2018) og Zhang et al. (2011) finner ingen signifikant sammenheng mellom personale og rompriser i henholdsvis Lisboa og New York når de samler alle segmenter i en modell. I segmentanalysen til Zhang et al. (2011) finner de derimot en signifikant effekt innenfor luksussegmentet (Segment 4-, 4,5- og 5), der personale har en positiv effekt på romprisen med 40 %.

For Oslo finner Thrane (2007) at personale har en negativ effekt på rompriser med -12,3 %. Schamel (2012) finner på sin side en negativ effekt i en liten, ikke-navngitt by, i intervallet [-0,169 , -0,072]. Disse funnene er overraskende, da vi antar at et godt personale har en positiv effekt på romprisene. Thrane (2007) sin antagelse er at hoteller som tilbyr ekstra gode personaltjenester, eksempelvis gjennom romservice, har en direkte inntekt som kompenseres med lavere hotellrompriser.

Totalt sett viser disse studiene sprikende resultater i sammenhengen mellom hotellrompris og personale (Castro & Ferreira, 2018; Schamel, 2012; Thrane, 2007; Zhang et al., 2011).

3.2.6 Renhold

Renhold er en av tilbakemeldingsvariablene som gjestene kan vurdere på Booking.com etter sitt hotellopphold. Dette refererer eksempelvis til standarden på rengjøringen (Booking.com, 2023a). Zhang et al. (2011) finner ingen signifikant effekt av renhold på romprisene, verken for hotellmarkedet som helhet eller når hotellene inndeles i ulike segment. Forfatterne diskuterer imidlertid den begrensede variasjonen i renholds nivået mellom hotellene, og reflekterer over at dette kanskje ikke er et kriterium som skiller hotellene fra hverandre når kundene velger hotellrom.

Zhang et al. (2011) får delvis støtte fra Castro og Ferreira (2018). Studiene enes om at det ikke er grunnlag for å konkludere med at renslighet påvirker hotellromprisene, gitt at hotellmarkedet ikke deles inn i ulike segmenter. Imidlertid finner Castro og Ferreira (2018) en sammenheng mellom renhold og priser i segment fem, med en prisøkning på 68,8 % ved en gjennomsnittlig økning i tilbakemeldingene på én. Dette funnet var overraskende høyt, med de drøfter ikke dette ytterligere.

En annen relevant studie som kan bidra til å belyse dette temaet, er en hedonisk prisanalyse av hostelbransjen (Santos, 2016). Dette er relevant fordi hostelbransjen representerer et lavere kvalitetsnivå, men kan likevel gi en indikasjon på hvordan renhold prioriteres i overnattingssektoren. I Santos' (2016) studie, som inkluderer over 8000 hosteller i mer enn 2000 byer i 134 land, finner forfatteren at renhold er den viktigste variabelen for kundene når de velger hostel, med en prisøkning på 31,7 %. Dette indikerer at ved lavere prisnivå kan renholdet være svekket (Santos, 2016).

Borges et al. (2015) finner på sin side en negativ koeffisient på -0,21 % i en samlet modell for alle tilgjengelige variabler og segmenter. De velger imidlertid å ekskludere denne variabelen fra sine videre undersøkelser, med argument om høy korrelasjon til andre attributter.

Vi kan oppsummere med at studiene viser at renhold kan ha ulik innvirkning på hotellromprisen, avhengig av markedet (Borges et al., 2015; Castro og Ferreira, 2018; Santos, 2016; Zhang et al., 2011).

3.2.7 WiFi

WiFi refererer til tilgjengeligheten og kvaliteten på trådløst Internett på rommet og fellesområdene. Det var utfordrende å finne studier som ser på hvordan WiFi påvirker kundens betalingsvilje for hotellrom. Vi har imidlertid identifisert tre relevante studier. I disse studiene ser forfatterne på hvordan WiFi påvirker romprisene i Algarve, Taipei og Karibia (Soler et al., 2019; Chen & Rothschild, 2010; Lorde et al., 2019).

I Algarve finner Soler et al. (2019) ingen signifikant sammenheng mellom tilgang på WiFi og hotellrompriser, men drøfter ikke dette funnet i stor grad. Chen og Rothschild (2010) finner derimot at WiFi har en positiv effekt på romprisen på 22,3 % i Taipei i ukedager, hva gjelder kundesegmentet forretningsreisende. For fritidssegmentet, som primært reiser i helgene, finner de ingen signifikant effekt. Chen og Rothschild (2010) forklarer at forretningsreisende tilbringer mer tid på hotellrommet, og vektlegger derfor WiFi i større grad enn fritidssegment, som tilbringer mer tid utenfor hotellrommet.

For Airbnb-markedet i Karibia, finner Lorde et al. (2019) at tilgang til Internett har en signifikant effekt på prisene, men med en negativ effekt på -15,5 %. Forskerne forklarer dette med at siden de fleste tilbyderne i det lokale Airbnb-markedet inkluderer WiFi (96,9 %), så reduserer de prisen for å skille seg fra konkurrentene.

I sum tyder resultatene fra disse studiene på at WiFi påvirker prisene ulikt, avhengig av marked og kundegruppe (Chen & Rothschild, 2019; Lorde et al., 2019; Soler et al., 2019).

3.2.8 Bærekraftsertifisering

Dersom et hotell prioriterer bærekraft, er dette tydelig merket på Booking.com. Denne markeringen refererer til hotellenes spesifikke bærekraftsertifiseringsnivå, som reflekterer bærekraftstandarden til hotellene. Skalaen går fra null til tre (Booking.com, 2023c). Det er imidlertid ikke en enighet om hvorvidt bærekraftsertifiseringer påvirker rompriser, noe varierende resultater fra litteraturen viser (Cavero-Rubio & Amorós-Martínez, 2008; Geerts, 2014; Soler et al., 2019).

Geerts (2014) finner at pris, beliggenhet og servicekvalitet er mer kritiske utvalgsriterier for en hotellovernatting i London enn bærekraftsertifisering, og dermed var ikke en slik sertifisering lønnsom for hotellene.

Kuminoff et al. (2010) undersøker om kundene er villig til å betale mer for et miljøvennlig hotellrom i Virginia, og finner at de er villige til å betale 9,9 % mer for slike rom. Sánchez-Ollero et al. (2014) finner også at hoteller som implementerer miljøtiltak i Andalusia, Sør-Spania, har høyere hotellrompriser med en gjennomsnittlig økning på 10,35 %. Sánchez-Ollero et al. (2014) forklarer dette med at kundene ser en nytte i at hotellet har utført miljøvennlige tiltak, selv om de ser større nytte i andre attributter som badebasseng og spaanlegg.

Bernard og Nicolau (2022) finner at bærekraftsertifiseringer kan ha en positiv effekt på bedriftens omdømme. For eksempel kan gjenkjennelige tredjepartssertifiseringer bidra til å forbedre bedriftens bærekraftomdømme, og dermed redusere mistanker om grønnvasking (Bernard & Nicolau, 2022). Bernard & Nicolau (2022) finner også at et hotells bærekraftsertifiseringer kan resultere i en kortvarig økning i markedsverdien. De identifiserer imidlertid ingen langsiktige og varige effekter. Dette antyder at økningen i markedsverdien er midlertidig og oppstår kun rundt tidspunktet for kunngjøringen av sertifiseringen (Bernard & Nicolau, 2022). Siden studiet både ser på markedsverdien og omdømmet til hotellene, og dermed ikke spesifikt på rompris, så forklarer ikke resultatene hvordan romprisen påvirkes av bærekraftsertifiseringer. Det er likevel et nevneverdig funn som forteller oss noe om hvorfor hotellene kan velge å sertifisere seg.

Vi kan oppsummere med at tidligere studier finner ulike resultater for effekten av bærekraft på rompriser. (Claver-Cortés et al., 2007; Geerts, 2014; Kuminoff et al., 2010; Sánchez-Ollero et al., 2014).

3.2.9 Oppsummering av variablene

Gjennom vår gjennomgang av relevant litteratur har det blitt tydelig at segmentvariabelen spiller en avgjørende rolle i bestemmelsen av hotellrompriser. Dette gjelder på tvers av ulike markeder og geografiske områder, og indikerer at hotellenes kvalitet i stor grad blir signalisert gjennom deres tilhørighet til et spesifikt segment.

Vi har også observert at variablene beliggenhet, fasiliteter, komfort, personale, renhold, WiFi og hotellenes bærekraftsertifisering har en variabel innvirkning på hotellromprisene. Resultatene fra ulike regioner og forskningsmetoder viser en betydelig variasjon og umuliggjør en definitiv konklusjon. Den spesifikke konteksten og de unike forholdene for hvert enkelt marked må vurderes når man ser på effekten av tilbakemeldinger og bærekraftsertifisering på hotellromprisene.

3.3 Halo-effekten i tilbakemeldingene

Etter å ha sett nærmere på hvordan ulike variabler påvirker hotellrompriser i litteraturen, vil vi nå se på et fenomen som tilsynelatende påvirker kundenes tilbakemeldinger. Dette fenomenet kalles halo-effekten. Basert på hva vi har sett i litteraturen, synes det relevant å forstå denne effekten når vi undersøker tidligere gjesters subjektive tilbakemeldinger. Halo-effekten handler om hvordan en person tenderer mot å generalisere positive eller negative inntrykk av en ting, person eller tjeneste, på grunnlag av en enkelt positiv eller negativ egenskap eller handling (Nicolaua et al., 2020). En person som har en positiv oppfatning av noe på grunn av en spesifikk egenskap, kan med andre ord anta at alt annet ved dette også er bra.

Hvis vi knytter dette til kundeopplevelser og tilbakemeldinger, kan en halo-effekt oppstå ved at en kunde som er svært fornøyd med et bestemt attributt ved et hotell, også kan være tilbøyelig til å gi positive tilbakemeldinger på alt annet ved hotellopplevelsen. Dette selv om andre deler av opplevelsen ikke var like bra (Nicolaua et al., 2020). For eksempel, hvis en gjest har hatt en fantastisk opplevelse på hotellrommet, så kan det føre til at de gir positive tilbakemeldinger på alt fra personale til beliggenheten, til tross for at disse aspektene ikke nødvendigvis var så imponerende.

Hvis et hotell ønsker å øke kundetilfredsheten, kan de fokusere på å optimalisere et enkelt attributt som har stor innvirkning på kundeopplevelsen. Nicolaua et al. (2020) finner at når kundene vurderer sin hotellopplevelse, vil 50 % av det hotellattributtet som skal vurderes (eksempelvis *personale*) bli påvirket av noe som ikke omhandler det spesifikke attributtet. Vi vil derfor ta hensyn til halo-effekten i våre analyser i delkapittel 7.2 og 7.3, der vi vil ta en korrelasjonsanalyse. Dette for å undersøke om variablene kan påvirke hverandre.

3.4 Konklusjon av teori og formulering av forskningsspørsmål

I vår gjennomgang av relevant litteratur, har vi fått en grundig forståelse av hedonisk prising og hvordan det kan anvendes på hotellrompriser. Vi har gjengitt flere studier som har anvendt hedonisk pristeori, og påvist dens effektivitet i å analysere hotellrompriser. Deretter har vi sett nærmere på variablene, som gjestene har mulighet til å gi tilbakemeldinger på Booking.com, og observert at resultatene varierer avhengig av marked og undersøkelsesmetode.

I tillegg har vi undersøkt hvordan hotellenes bærekraftsertifiseringer kan påvirke hotellromprisene, og vi har sett at dette også varierer avhengig av marked og hotelltype. Den store variasjonen i resultatene fra tidligere studier gjør det spennende å sammenligne og kontrastere disse funnene med våre egne resultater for det norske markedet. Derfor har vi valgt å gjennomføre en hedonisk prisanalyse for å besvare følgende forskningsspørsmål:

Hvordan påvirkes hotellromprisene i det norske markedet av kundevurderinger og hotellenes bærekraftsertifiseringer?

Vi ønsker at denne studien skal bidra til økt forståelse av hva som driver hotellromprisene i Norge, og gi innsikt i hvilken påvirkning kundevurderinger og bærekraftsertifiseringer har for hotellromprisene. I den kontekst vil vi også kunne undersøke om det er andre faktorer, som segmenttilhørighet, som påvirker hotellromprisene. Oss bekjent, er det få studier som har fokusert på dette området i det norske hotellmarkedet.

Vår undersøkelse vil ha en dimensjon som skiller tettsteder med høyest hotellkonsentrasjon fra steder med lav hotellkonsentrasjon. På denne måten skiller vi to ulike markeder fra hverandre, i et forsøk på å danne mer homogene markeder for undersøkelsen. Denne metoden har vi ikke observert i tidligere studier. Vi vil forsvare denne distinksjonen nærmere i delkapittel 7.1.

4 Data og datainnsamling

I dette kapitlet vil vi redegjøre for utvalget og innsamlingen av data. Formålet er å forklare hvordan vi sikrer et representativt utvalg, som vil gi pålitelige og gyldige resultater. I dette tilfellet består utvalget av alle hoteller med tilstrekkelig data i det norske hotellmarkedet. Dataene ble samlet inn fra Booking.com sin database, hvor vi har benyttet web-skraping til å hente ut denne informasjonen gjennom deres nettsider. Avslutningsvis vil vi kort diskutere noen etiske problemstillinger ved web-skraping.

4.1 Utvalg

For å sikre at studien er gyldig og pålitelig er det viktig å ha et godt utvalg. Ved seleksjon av utvalg må det være representativt for den populasjonen vi ønsker å studere, som for oss er det norske hotellmarkedet. Et lite representativt utvalg vil gjøre det umulig å stole på resultatene og generalisere funnene til hele markedet, da de bare vil gjelde for det spesifikke og begrensede utvalget (Johannessen et al., 2011).

For å finne et representativt utvalg, bestemte vi oss for å ta utgangspunkt i det norske hotellmarkedet, som både har en geografisk dimensjon og en produkt-dimensjon. Nærmere bestemt alle hoteller i det norske hotellmarkedet med tilstrekkelig datagrunnlag for samtlige inkluderte variabler. For oss var det avgjørende med et stort utvalg, da dette kan gi en større grad av presisjon i resultatene, samt øke påliteligheten av våre funn.

Vi har innhentet data fra Booking.com sin database for norske hoteller, da denne OTAen er rangert av hotellaktørene som viktigste distribusjonssted for salg av hotellovernattinger (Inversini & Masiero, 2014). I tråd med White og Mulligan (2010), valgte vi å innhente data fra en tilfeldig midtukes dato (11.04.2023). Til forskjell fra White og Mulligan (2010) hentet vi kun data fra én dato – et øyeblikksbilde (tverrsnitt) – mens de innhentet midtukes data for årets fire sesonger (paneldata). De fant imidlertid små forskjeller mellom sesongene. Basert på dette, valgte vi å begrense studien til et øyeblikksbilde innenfor én sesong. Dette gjorde også datainnhenting og -behandlingen mer overkommelig, gitt den tidsrammen oppgaven har. Vi kan underbygge dette valget med en studie av Schamel (2012), hvor resultatene var tilnærmet like, uavhengig av ukedag og sesong.

I kapittel 2, Generell bransjeinformasjon, presenterte vi Guillet & Law (2010) sitt funn om at ikke alle hotellaktører benytter seg av alle tilgjengelige distribusjonskanaler. Dette kan

betraktes som en begrensning ved vår undersøkelse, da vi vil gå glipp av hotellene som ikke er tilgjengelige på Booking.com. Vi anser imidlertid antallet innsamlede hoteller, som er 720 av 968 hoteller (SSB, 2023), som stort nok til å regnes som et representativt utvalg.

4.2 Innhenting av data

Booking.com samler inn kundevurderinger for hotellene ved å oppfordre kunder til å gi en score på en skala fra 1 til 10 for ulike attributter, der 1 er dårligst og 10 er best. Gjennomsnittet av kundevurderingene for hvert attributt presenteres i hotellinformasjonen på Booking.com, som vist i Figur 4.1 nedenfor. Informasjonen om hvorvidt hotellet er et bærekraftig valg vises i det grønne feltet til høyre for stjernene, øverst i Figur 4.1. Dette indikerer bærekraftsertifiseringsnivået [0–3], der ett blad representerer nivå 1, og tre blader representerer toppnivået (Booking.com, 2023c). Booking.com viser også antall kundetilbakemeldinger og en gjennomsnittlig totalscore for overnattingsstedet.



Figur 4.1: Skjerm bilde av Booking.com (Booking.com, 2023a).

Vi benyttet web-skraping for å trekke ut data fra alle tilgjengelige hoteller i det norske hotellmarkedet på Booking.com. Web-skrapingen var en effektiv metode, siden det tok i underkant av to timer å innhente all informasjon. Dette står i kontrast til Soler et al. (2019), som tastet inn sin informasjon i løpet av en 17-dagers periode, noe som ikke er ideelt da hotellpriser sjelden er faste (Lopes et al., 2017). I forbindelse med web-skrapingen benyttet vi oss av programmeringsspråket Python og nettlesertjenesten Google Chrome. Vi startet med å importere lenkene til hvert hotells nettside på Booking.com. Deretter opprettet vi en ordbok

som inneholder informasjon om hotellene vi ønsket å skrape. Vi brukte forskjellige teknikker for å finne informasjon om hotellenes navn, by, pris, segmentnivå, antall vurderinger, overordnet poengsum, personale, fasiliteter, renhold, komfort, valuta for pengene, beliggenhet, WiFi og bærekraftgrad. Vi brukte en rekke prøve-og-feile-setninger for å fange opp feil som kunne oppstå da vi hentet ut informasjon fra nettsidene, slik som når ulike faktorer ikke eksisterer på hotellenes nettsider. En annen utfordring var at enkelte hoteller ikke var tilgjengelige på Booking.com. Det er imidlertid verdt å merke seg at vi kun brukte én nettside, noe som kan ha påvirket resultatene. Med denne prosessen fikk vi et initialt utvalg på 720 hoteller av totalt 968 på landsbasis (SSB, 2023).

Til slutt eksporterte vi datasettet til en Excel-fil. Med dette kunne vi lagre all informasjonen som ble hentet ut fra Booking.com på en enkel og strukturert måte. Det ble dermed lettere å behandle dataene, samt å utføre videre analyser. I Excel-arket har vi kvalitetssikret dataene ved å besøke lenkene til hvert hotell. Dette innebar å undersøke og validere verdiene for segmentnivå, bærekraftsertifiseringsnivå og kundevurderinger.

Det komplette datasettet kunne så behandles i det integrerte utviklingsmiljøet RStudio, hvor vi først rensket datasettet ved å fjerne rader med manglende verdier (*Not Availables*, NA). Dette var nødvendig fordi vi ikke kan utføre multiple regresjoner på datasett med manglende verdier for de uavhengige variablene (Studenmund, 2014). Noen av hotellene manglet eksempelvis data for variabelen *WiFi*. Vi fjernet også hoteller med ti eller færre tilbakemeldinger. Dette fordi vi mener at disse ikke har tilstrekkelig antall kundevurderinger til å være representativt, noe som vil senke kvaliteten på resultatene og begrense vår evne til å generalisere funnene til det norske hotellmarkedet. Ifølge Studenmund (2014) kan vi akseptere å miste noen observasjoner gjennom slik databehandling. Datasettet inneholdt også variablene *overall* (helhetsvurdering) og *verdi for pengene*, som vi velger å ekskludere fra våre kommende regresjonsmodeller. Dette fordi *overall* er gjennomsnittet av alle tilbakemeldingsvariablene, mens *verdi for pengene* er lik den avhengige variabelen (hotellrompris) delt på *overall*. Dermed er disse variablene ikke uavhengige. Etter datarensingen anser vi det endelige utvalget på 680 hoteller som akseptabelt. Tabell A6.1 til Tabell A6.3 viser deskriptiv statistikk for datasettet.

4.2.1 Etikk ved Web-skraping

Web-skraping reiser også noen spørsmål rundt lovlighet og etikk. Nettsider kan for eksempel inneholde sensitiv informasjon som persondata (Mitchell, 2018). I den forbindelse var vi nøye

med å forsikre oss om at nettsidene vi skrapet på Booking.com ikke inneholdt navn, telefonnumre, epost-adresser eller annen personlig informasjon som kunne identifisert kunder eller andre aktører. Ved å beskytte ulike parters anonymitet, mener vi at det er forsvarlig å web-skrape nettstedet. Vi har heller ikke web-skrapet Booking.com mer enn én gang og vi har holdt oss til ett øyeblikksbilde. I enkelte tilfeller kan ekstensiv web-skraping over lengre perioder, hvor en henter ut større datasett enn det vi har gjort, medføre problemer som redusert kvalitet på nettstedets tjenester. Nettstedet kan for eksempel slutte å fungere eller få redusert hastighet under web-skraping, som igjen kan skade eieren eller brukerne av nettstedet i form av effektivitetstap. Videre har nettstedets eiere også kostnader tilknyttet drift og vedlikehold av servere, og disse kostnadene kan øke når web-skrapingen krever stor serverkapasitet.

Det er også verdt å merke seg at informasjonen vi ønsket å innhente hverken er av litterært eller kunstnerisk arbeid, men heller statistikk og fakta. Dermed er denne informasjonen ikke beskyttet av opphavsrettslover. Statistiske data og fakta er heller ikke beskyttet av patenter, som er skal beskytte eierskapet til oppfinnelser. Vi har heller ikke skrapet informasjon som kan bryte med lover om beskyttelse av varemerker (Mitchell, 2018).

Oppgaven bidrar også til økt kunnskap om hotellrompriser og resultatene vil samtidig være offentlig tilgjengelig. Dette vil i fra et samfunnsperspektiv kunne bidra til å veie opp for eventuelle uetiske aspekter ved web-skraping.

5 Metode

I den følgende delen av oppgaven vil vi forklare metoden vi skal benytte oss av til å generere resultatene i kapittel 7. Dette kapittelet vil derfor fokusere på vår forskningstilnærming, valg av studiedesign, hvordan modellene vi skal benytte oss av fungerer, samt hvilke problemer vi kan støte på i forbindelse med dette. Vi ønsker å sikre at undersøkelsens validitet og reliabilitet svarer til forskningsspørsmålet (Skilbrei, 2019): *Hvordan påvirkes hotellromprisene i det norske markedet av kundevurderinger og hotellenes bærekraftsertifiseringer?*

For å adressere forskningsspørsmålet, vil vi anvende hedonisk pristeori og multiple regresjonsanalyser. I delkapittel 5.1 vil vi først beskrive valg av forskningstilnærming og studiedesign, for så å forklare hvordan vi vil sikre reliable og valide resultater. Deretter vil vi, i delkapittel 5.2, introdusere og forklare den multiple regresjonsmodellen. Der vil vi også løfte relevante problemer som vi må ta hensyn til i vår bruk av denne metoden. Vi vil også løfte frem styrker og svakheter med multipel regresjonsanalyse i vår kontekst.

5.1 Forskningstilnærming, studiedesign og kvalitetssikring

5.1.1 Forskningstilnærming

I litteraturen skilles det mellom induktiv og deduktiv metode. En induktiv tilnærming kan være aktuell når et tema i liten eller ingen grad er forsket på tidligere. Dermed kan en danne en teori basert på forskningen. En deduktiv metode søker derimot å bekrefte eller styrke eksisterende teori (Johannessen et al., 2011). I denne oppgaven vil vi komme med nytt bidrag til eksisterende teori, der vår oppgave vil ha en by-land-dimensjon, som skiller vår oppgave fra tidligere studier. Derfor vil vi anvende deduktiv metode til å besvare forskningsspørsmålet.

5.1.2 Studiedesign

Det er vanlig å skille mellom eksplorativt, kausalt og deskriptivt studiedesign, eller en kombinasjon av disse tilnærmingene. Ved en eksplorativ tilnærming kunne vi generert idéer og hypoteser om vår problemstilling, der vi eksempelvis kunne ha samlet inn data gjennom intervjuer eller observasjoner (Creswell & David, 2018). En kausal forskning konkluderer om en endring i én variabel fører til endring i en annen variabel. En deskriptiv tilnærming kvantifiserer forekomsten av ulike fenomene (Johannessen et al., 2011).

På grunn av oppgavens tidsramme, har vi valgt å kun benytte oss av et kausalt design for å undersøke om det er en kausal sammenheng mellom bærekraftsertifiseringer, kundevurderinger og hotellrompriser. Et annet studiedesign vil være egnet for videre studier, noe vi vil diskutere i kapittel 8.

5.1.3 Kvantitativ metode

Siden vi har valgt en deskriptiv tilnærming, som innebærer å kvantifisere sammenhenger mellom variabler, vil vi benytte oss av kvantitativ metode. Kvantitativ metode er en forskningstilnærming som fokuserer på å samle inn og analysere numeriske data, og det er en hyppig brukt metode på tvers av ulike fagområder (Johannessen et al., 2011).

En fordel med kvantitativ metode er at den gir rom for å samle inn store mengder data, noe som kan gi et representativt bilde av virkeligheten (Johannessen et al., 2011). Dette kan bidra til å styrke konklusjonene vi trekker i oppgaven. Imidlertid har kvantitativ metode også noen begrensninger og ulemper, spesielt når det gjelder datainnsamlingsprosessen. Det kan for eksempel være vanskelig å fange opp nyanser i kundenes opplevelser og erfaringer, samtidig som det kan være utfordrende å velge hvilke variabler som skal inkluderes i studien (Johannessen et al., 2011). Vi vil derfor diskutere hvorvidt oppgaven preges av disse utfordringene, og reflektere over hvordan de har påvirket våre funn og konklusjoner.

5.1.4 Reliabilitet

For å sikre at våre resultater gir et representativt bilde av virkeligheten, så må de være reliable. Reliabilitet refererer til graden av pålitelighet i resultatene av en studie (Saunders et al., 2007). Vårt utvalg dekker store deler av det norske hotellmarkedet, noe vi anser som representativt for våre geografiske områder, og dermed som reliabelt. Når det gjelder prisvariabelen, som endrer seg kontinuerlig basert på etterspørsel, så utfordres reliabiliteten fordi vi har valgt å studere et dagsbilde. For eksempel kan et stort arrangement i Arendal, som Arendalsuka, føre til høyere etterspørsel og dermed påvirke romprisene på det aktuelle tidspunktet. Dette vil sannsynligvis påvirke reliabiliteten av resultatene. Resultatene for denne uken kan dermed ikke overføres til de resterende ukene i året. Vi har imidlertid ikke funnet noen store arrangementer som fant sted i Norge den 11. april 2023. Det er også verdt å merke seg at selv om resultatene i en studie er reliable, så betyr ikke det nødvendigvis at de er valide eller overførbare (Johannessen et al., 2011). Derfor må vi også vurdere studiens validitet.

5.1.4 Validitet

Validitet handler om å evaluere om forskningsdesignet og de anvendte måleinstrumentene faktisk måler det vi ønsker å undersøke (Creswell & David, 2018). Studien er gyldig (valid) hvis resultatene kan generaliseres til populasjonen den ønsker å representere (Krumsvik, 2014).

For å undersøke om hotellprisene blir påvirket av kundevurderinger og bærekraftsertifiseringer, må vi forsikre oss om at studien faktisk måler dette. Vi har tatt hensyn til andre faktorer som kan påvirke hotellromprisene, slik som tidspunkt på året og ekstreme verdier (uteliggere). Vi har også gått gjennom hvert hotell i utvalget for å sikre at de er representative for det norske hotellmarkedet, og da først og fremst ved å avgrense utvalget til hoteller med norsk adresse. En mulig risiko er at vi har gått glipp av en betydelig andel reisende, gitt at de ikke anvender OTAer. Jobbreisende kan tenkes å utgjøre en større andel av slike reisende, siden mange bedrifter har avtaler med ordinære reisebyråer som Berg-Hansen.

Ved å sikre at resultatene av studien er både reliable og valide, kan vi være trygge på at de gir en pålitelig og gyldig beskrivelse av det vi undersøker – altså hvordan kundevurderinger og bærekraftsertifiseringer påvirker prissettingen for hoteller i det norske hotellmarkedet.

5.1.4.1 Ekstern validitet

Ekstern validitet handler om i hvilken grad resultatene av studien kan generaliseres til andre populasjoner, situasjoner eller tidsperioder (Krumsvik, 2014). I vårt tilfelle kan økt ekstern validitet oppnås ved å vurdere om funnene er overførbare til utenlandske hotellmarkeder. Basert på de sprikende resultatene fra tidligere studier, er det rimelig å anta at de nasjonale hotellmarkedene varierer fra land til land, noe som vil begrense overførbarheten av våre funn. Vi vil derfor drøfte resultatene opp mot tidligere studier, for å se om det er noen sammenhenger mellom andre regioner og vårt område.

For å øke generaliserbarheten av funnene, har vi inkludert ulike typer hoteller i utvalget. Vi har for eksempel inkludert både store og mindre hoteller, samt hoteller med ulikt bærekraftnivå. Vi har også tatt hensyn til om resultatene kan generaliseres til andre tidspunkter. Siden hotellromprisene kan ha sesongvariasjoner, kan det være viktig å undersøke om funnene gjelder for hele året, eller om det er forskjeller i prissetting avhengig av tidspunktet for studien.

Til slutt er det viktig å merke seg at studien er basert på Booking.com sine data. Resultatene av studien kan dermed være vanskelig å overføre til lignende aktører som TripAdvisor og Expedia. Dette er også et mål på ekstern validitet, hvor lignende resultater basert på data fra disse aktørene vil medføre økt ekstern validitet for våre resultater.

5.2 Multippel regresjon

Vi vil som forklart undersøke hvordan kundetilbakemeldinger og hotellenes bærekraftsertifiseringer (uavhengige variabler) påvirker hotellrompriser (avhengig variabel) i det norske hotellmarkedet. Derfor benytter vi oss av multiple regresjonsmodeller, som lar oss estimere hver uavhengig variablers effekt på den avhengige variabelen, mens vi holder de andre uavhengige variablene konstante. Målet er å forstå hvordan forklaringsvariablene, X_{ji} , påvirker utfallet av responsvariabelen, Y_i (Stock & Watson, 2020). Altså hvordan hotellrompriser påvirkes av *segment, beliggenhet, fasiliteter, komfort, personale, renhold, WiFi* og *bærekraft*.

Modellen antar en lineær sammenheng mellom responsvariabelen og forklaringsvariablene. Den antar også at forklaringsvariablene ikke er perfekt korrelerte, at det ikke eksisterer ekstreme observasjoner (uteliggere) i utvalget, og at målingene ikke inneholder systematiske feil (Stock & Watson, 2020). Vi vil i den multiple regresjonsmodellen bruke minste kvadraters metode (OLS) til å estimere koeffisientene til $\beta_0, \beta_1, \dots, \beta_k$.

Den multiple regresjonsmodellen kan ifølge Stock og Watson (2020) spesifiseres med:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i, i = 1, \dots, n, \quad (5.1)$$

hvor

Y :	Predikert verdi av den avhengige variabelen
β_0 :	Konstantledd (y-skjæringspunkt når alle andre parametere er lik 0)
β_k :	Koeffisient
X :	Uavhengig variabel
i :	Antall observasjoner av hver variabel, $i = 1, 2, \dots, n$
k :	Antall uavhengige variabler
u_i :	Modellens feilledd (variasjonen i estimatet av Y)

Vi vil spesifisere modellen ytterligere med valg av funksjonsform. Dette avhenger av hvilke variabler vi inkluderer i modellen og hva vi estimerer. Rosen (1974) anbefaler semi-log (eller log-lineær) spesifisering av prisfunksjonen. De uavhengige variablene er som regel da ikke logaritmiske, med unntak av noen tilfeller hvor dette er logisk. Castro og Ferreira (2018) anvender for eksempel logaritmisk form for antall kommentarer. Vi har testet ulike log-modeller for å finne den best egnete funksjonsformen, og landet til slutt på en log-lineær form med den naturlige logaritmen av den avhengige variabelen ($\ln(Pris)_i$). Dette vil ifølge Rosen (1974) forbedre regresjonsmodellens forklaringskraft. En log-lineær form er også anvendt av Zhang et al. (2011), White & Mulligan (2010) og Castro & Ferreira (2018) i deres hedoniske studier av hotellrompriser.

5.2.1 Minste kvadraters forutsetninger for modellen

Multiple regresjonsmodeller følger fire forutsetninger når minste kvadraters metode brukes, og er (Stock & Watson, 2020):

- LS.1: Betinget gjennomsnitt lik null
- LS.2: $(X_{1i}, X_{2i}, \dots, X_{ki}, Y_i), i = 1, \dots, n$ er identisk og uavhengig distribuert (i.i.d.)
- LS.3: Store uteliggere er uvanlig eller ikke tilstedeværende
- LS.4: Fravær av perfekt multikollinearitet

5.2.1.1 Betinget gjennomsnitt lik null

Den betingede fordelingen til feilleddet u_i , gitt $X_{1i}, X_{2i}, \dots, X_{ki}$, skal ha et gjennomsnitt lik null for å innfri LS.1. Dette er forventningsverdien til feilleddet i modellen. Denne forutsetningen er brutt når andre variabler som tilhører feilleddet i en multippel regresjonsmodell (utelatte variabler), påvirker Y_i . Brudd på LS.1 skyldes vanligvis seleksjonsprosessen, utelatte variabler, simultan kausalitet og målefeil i de uavhengige variablene (Stock & Watson, 2020). Vi vil nå se nærmere på disse.

Seleksjonsprosessen har vi adressert ved å sikre en grundig seleksjon av relevante data, i samsvar med litteraturen, jamfør diskusjonen om intern validitet i avsnitt 5.1.4. Vi har heller ikke identifisert mulige utelatte variabler som kan påvirke resultatene. De inkluderte variablene tilfredsstillende fenomenet vi ønsker å undersøke, i samsvar med forskningsspørsmålet. Disse omfatter den avhengige variabelen *pris* og de uavhengige variablene som er underlagt kundevurderinger og bærekraftsertifisering. Vi har ikke identifisert simultan kausalitet mellom variablene i utvalget, noe som også tyder på at LS.1-forutsetningen er oppfylt. For å redusere

målefeil har vi sikret oss nøyaktige målinger av variablene gjennom web-skraping av Booking.com. Dette bidrar til å minimere eventuelle unøyaktigheter i analysen. Basert på våre tiltak mener vi at vi har tatt hensyn til og adressert potensielle brudd på LS.1-forutsetningen i vår analyse.

5.2.1.2 Identisk og uavhengig distribusjon (i.i.d.)

Den andre forutsetningen er at $(X_{1i}, X_{2i}, \dots, X_{ki}, Y_i), i = 1, \dots, n$ er uavhengig og identisk distribuerte tilfeldige variabler. LS.1 gjelder for utvalg som tilfeldig trekkes fra en populasjon, hvor det er like sannsynlig å inkludere hvert medlem av populasjonen. Denne forutsetningen vurderes ut fra hvilke type data en studerer og hvordan en anvender dem (Stock & Watson, 2020). Dermed er LS.2 ikke en strengt nødvendig forutsetning for alle dataformater eller studiedesign. I vårt tilfelle er utvalget ikke en tilfeldig andel av populasjonen vi studerer. Dette fordi vi har innhentet informasjon om alle tilgjengelige hoteller på Booking.com, i stedet for å trekke et tilfeldig utvalg av de tilgjengelige hotellene. Vår studie er dermed et eksempel på en studie hvor LS.2-forutsetningen ikke holder. Dette fordi den ikke passer med våre data, som er et tverrsnitt av alle tilgjengelige hoteller på Booking.com. Utvalgsprosessen vår er beskrevet mer detaljert i avsnitt 4.1.

5.2.1.3 Fravær av store uteliggere

Tredje forutsetning er at store uteliggere skal være fraværende eller sjeldne. Dette har vi tatt hensyn til da vi undersøkte utvalget i kapittel 4. Våre ekstreme observasjoner er høyprisede hoteller i segment fire og fem. Siden OLS-estimatoren til koeffisientene kan være sensitive for slike uteliggere (Stock & Watson, 2020), valgte vi å undersøke disse nærmere. Vi valgte å beholde dem fordi prisen var høy, men ikke feil. Slike uteliggere skal ikke elimineres uten videre, men undersøkes slik vi har gjort (Studenmund, 2014).

5.2.1.4 Fravær av perfekt multikollinearitet

For å oppfylle den fjerde forutsetningen, må vi unngå perfekt multikollinearitet, der en av regressorene er en perfekt lineær kombinasjon av de andre. Uperfekt multikollinearitet derimot, betyr at regressorene er høyt, men ikke perfekt korrelert. Perfekt multikollinearitet hindrer beregning av OLS-estimatoren, mens uperfekt multikollinearitet kan føre til upresise estimater. I vår analyse har vi valgt våre variabler basert på vårt forskningsspørsmål og inkludert dem i modellen. Uperfekt multikollinearitet vil ikke hindre oss i å estimere regresjonslinjene, men det

kan påvirke presisjonen av estimatene. Det er også viktig å merke seg at graden av multikolaritet kan variere mellom ulike øyeblikksbilder eller sesonger. (Stock & Watson, 2020; Studenmund, 2014). Vi vil evaluere graden av multikolaritet for våre modeller i kapittel 7.

5.2.2 Andre forutsetninger

I tillegg til minste kvadraters forutsetninger for multiple regresjonsmodeller, så må vi ta hensyn til et par andre forutsetninger (Stock & Watson, 2020):

- Linearitet
- Normalitet
- Homoskedastisitet

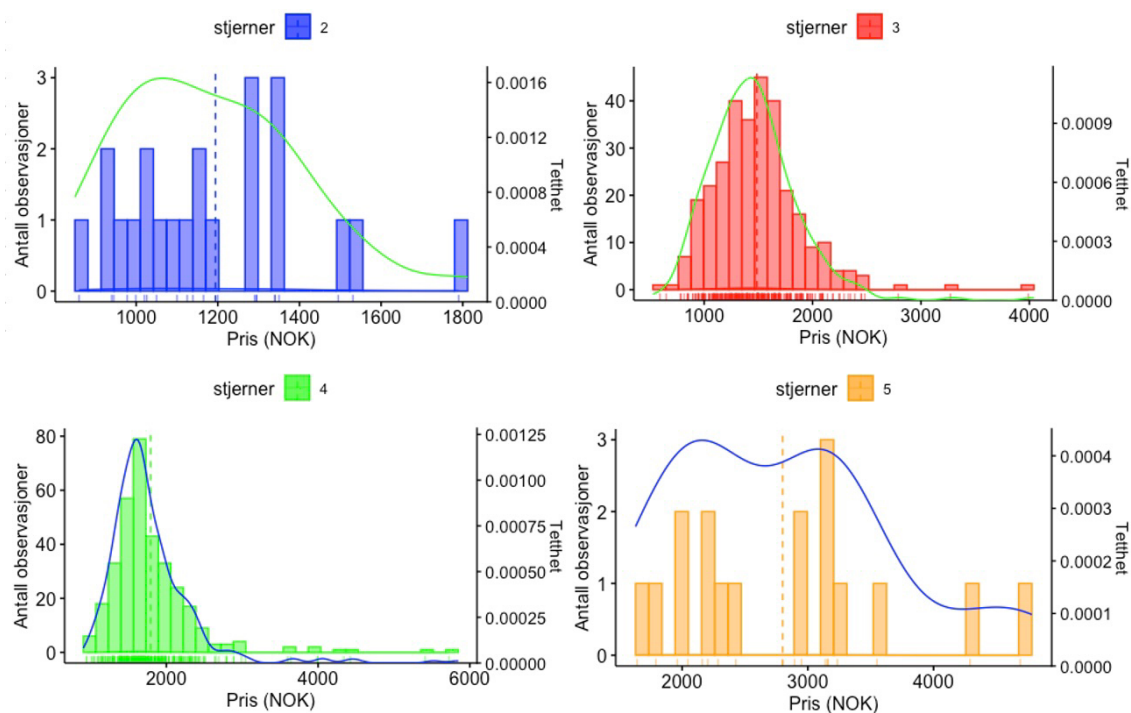
Vi vil nå beskrive disse forutsetningene og vurdere om regresjonsmodellen avviker fra dem.

5.2.2.1 Linearitet

Linearitet vil si at modellen har en lineær sammenheng mellom den avhengige variabelen Y_i og de uavhengige variablene X_{ji} . For å undersøke disse sammenhengene for hvert sett av avhengig og uavhengig variabel – *pris* og *beliggenhet*, *pris* og *fasiliteter*, osv. – konstruerte vi et plott i Figur A5.1. Som viser lineære sammenhenger mellom responsvariabelen og de predikerende variablene, hvor *beliggenhet*, *fasiliteter*, *komfort*, *personale*, *renhold* og *WiFi* har en konveks og positiv sammenheng med *pris*. Dette taler også for bruk av en semi-logaritmisk funksjonsform, slik vi har argumentert for i delkapittel 5.2.

5.2.2.2 Normalitet

Forutsetningen om normalitet innebærer at residualene i regresjonsmodellen må være normalfordelte (Studenmund, 2014). Et histogram av fordelingen i datasettet illustrerer hvorvidt forutsetningen om normalitet er innfridd Figur 5.1 nedenfor viser fordelingen for segment to, tre, fire og fem.



Figur 5.1: Segmentert rompris-histogram med observasjoner og tetthet (RStudio).

Fordelingene i Figur 5.1 viser at deler av segmentene i utvalget ligner en normalfordeling, med noe overvekt mot venstre i histogrammene. Antall observasjoner av pris innenfor hvert segment er illustrert med søyler, hvor hvert segment har sin unike farge. De vertikale, stiplede linjene viser gjennomsnittlig pris i hvert segment, med tilhørende farge. Fordelingene i segment tre og fire har lengre haler mot høyre i diagrammene grunnet høyprisede hoteller. Vi kunne eliminert disse hotellene for å oppnå et mer normalfordelt utvalg, men da ville vi utelatt en vesentlig del av markedet og dermed begrenset vår analyse. Derfor vurderer vi disse fordelingene som nært normalfordelt og egnet for våre regresjonsmodeller. Observasjonene i segment to og fem er ikke normalfordelte. Dette skyldes et begrenset utvalg på 21 observasjoner i segment to og 16 observasjoner i segment fem. Dermed er segment to og fem mindre egnet til regresjonsanalyser enn segment tre og fire. Med utgangspunkt i fordelingene i Figur 5.1, kan vi se at segment tre og fire danner majoriteten av datagrunnlaget for de kommende regresjonsanalysene.

Figur 5.1 illustrerer også tettheten i hvert segment, representert med tetthetskurver. Den svingende formen på tetthetskurven i segment fem støtter argumentet om at observasjonene i dette segmentet er mindre normalfordelte, sammenlignet med de andre segmentene. Tetthetskurven til segment to viser likevel at observasjonene er mer normalfordelte enn i segment fem, i og med at formen har mindre kurvaturer. Observasjonene i segment to er dog

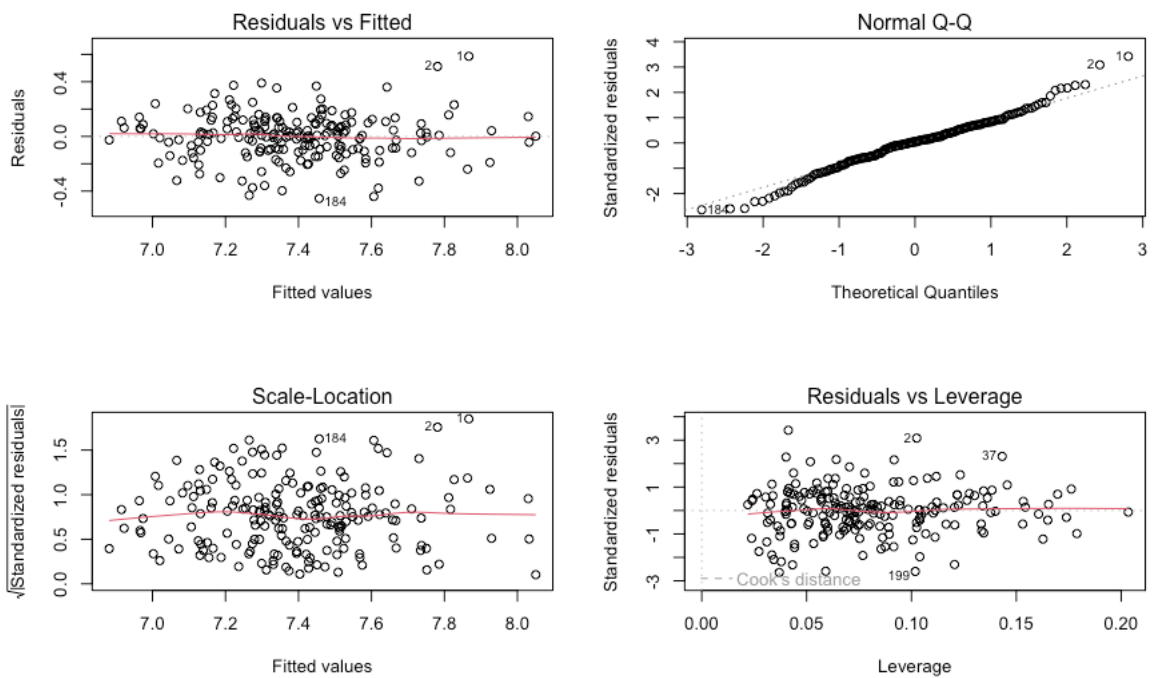
ikke tilfredsstillende normalfordelt. Tetthetskurvene i segment tre og fire støtter at observasjonene i disse segmentene er nært sagt normalfordelte. Disse tetthetskurvene er smale, særlig i segment fire, hvilket betyr at færre av observasjonene befinner seg utenfor et standardavvik på begge sider av medianen i segmentene.

5.2.2.3 Homoskedastisitet

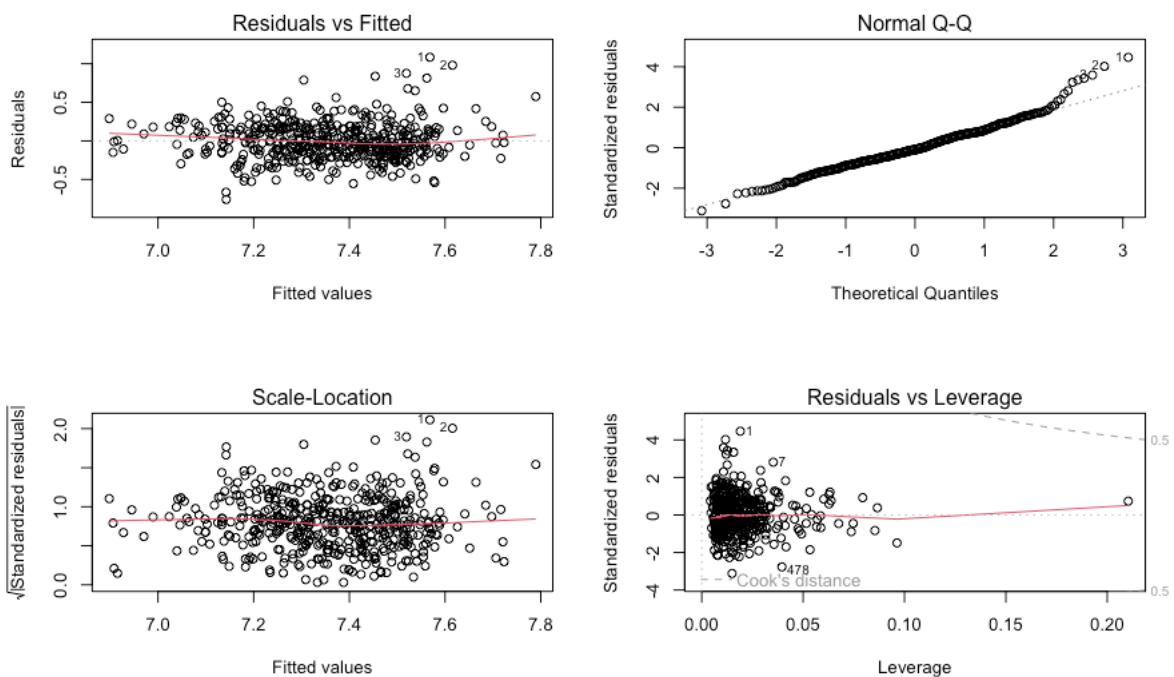
Til slutt har vi forutsetningen om homoskedastisitet, hvilket betyr at variansen i feilleddet skal være konstant for alle observasjoner. Hvis variansen varierer med forklaringsvariablene, har vi derimot heteroskedastisitet. Heteroskedastisitet kan påvirke modellens koeffisienter og gjøre standardfeilene upresise. Homoskedastisitet sikrer derimot konsistente og effektive estimater i analysen (Stock & Watson, 2020; Studenmund, 2014). Vi har testet for homoskedastisitet med en Breusch-Pagan test. En p -verdi på under 5 % betyr at vi har heteroskedastisitet i modellen, og homoskedastisitet hvis over 5 % (Curto & Pinto, 2011; Studenmund, 2014). Testen gav en p -verdi på 11,21 %, hvilket betyr at forutsetningen om homoskedastisitet er innfridd.

5.2.3 En nærmere undersøkelse av residualene

Et plott av residualene vil kunne bekrefte og synliggjøre om de multiple regresjonsmodellene faktisk følger forutsetningen om homoskedastisitet. Vi har allerede evidens for at datasettet ikke har heteroskedastisitet, jamfør tidligere Breusch-Pagan test, men vi ønsker å visualisere dette. Vi presenterer derfor residualplott for residualene til dimensjonene; By (Figur 5.2) og Distrikts-Norge (By) (Figur 5.3) nedenfor. Bakgrunnen for inndeling i dimensjoner er forklart i kapittel 7.1.



Figur 5.2: Residualplott av hotellene i By (Kilde: RStudio).



Figur 5.3: Residualplott av hotellene i Distrikts-Norge (Kilde: RStudio).

De røde linjene representerer gjennomsnittet til residualene i våre observasjoner, mens de grå stiplede linjene representerer de estimerte regresjonslinjene. Vi kan se at de røde linjene er nært horisontale og følger de grå stiplede linjene, noe som tyder på en god tilpasning for

regresjonsmodellene. Øverst til venstre i Figur 5.2 og Figur 5.3 er eksempelvis feilleddenes gjennomsnitt tilnærmet horisontalt og sentrert rundt null. Det er verdt å merke seg at hvis vi fjerner de høyest prisede hotellene i segment fire og fem, så vil linjen trolig være nær perfekt horisontal. Dette fordi disse få ekte uteliggerne, som vi mener er relevante for analysen, bidrar til å gjøre de røde linjene lettere konkave. Basert på plottene i Figur 5.2 og Figur 5.3, så er vi fornøyde med residualene. Dette fordi de indikerer at det ikke forekommer noen seriøse tendenser i datasettet, som kan forstyrre regresjonsmodellenes validitet.

Normal Q-Q viser at flertallet av de ekte residualene ligger langs linjen for de teoretiske residualene til en perfekt modell, altså den lineære stiplede linjen. Vi ser også at de identifiserte sanne uteliggerne befinner seg et stykke unna de teoretiske residualene i hver ende av den stiplede linjen. Dette betyr at utvalgene ikke følger en perfekt normalfordeling, slik vi allerede har identifisert i avsnitt 5.2.2.2.

Videre viser plottene nede til høyre i Figur 5.2 og Figur 5.3 at ingen av residualene ligger utenfor Cooks avstand, som er et mål på hvor mye hver observasjon påvirker regresjonsmodellen. Mer spesifikt er Cooks avstand et mål på påvirkning, som baserer seg på differansen mellom de estimerte regresjonsparameterne $\hat{\beta}$ og hva de endres til når vi eliminerer en observasjon (Chatterjee & Hadi, 2015). En observasjon med høy Cooks avstand vil ha stor innflytelse på den estimerte regresjonslinjen, slik at den kan endre seg vesentlig hvis observasjonen elimineres. Dermed vil observasjonen kunne være et problem for reliabiliteten til modellen. Grensen for Cooks avstand er representert med tykke stiplede linjer. Denne er lik 50 % av F -fordelingen til utvalgene med $p + 1$ og $n - p - 1$ frihetsgrader som innflytelsesrike punkt (Chatterjee & Hadi, 2015). I Figur 5.2 ligger denne et sted oppe til høyre, men utenfor plottet. I Figur 5.3 ser vi grensen oppe i høyre hjørne av plottet, hvor ingen av observasjonene ligger utenfor denne. Dermed har ingen av observasjonene i By eller DN en uvanlig stor innflytelse på regresjonsmodellen.

Selv om det fremkommer av Figur 5.2 og Figur 5.3 at residualene ikke følger en perfekt teoretisk modell, så kan vi med utgangspunkt i disse figurene underbygge konklusjonen fra Breusch-Pagan-testen, altså at modellene innfrir forutsetningen om homoskedastisitet.

5.2.4 Korrelasjon

Et annet viktig hensyn er i hvilken grad variablene i regresjonsmodellene korrelerer med hverandre. Det vil si styrken på den lineære sammenhengen mellom to variabler, X og Y , for et utvalg med n observasjoner. Korrelasjonskoeffisienten (r_{XY}) mellom to variabler ligger mellom -1 og 1, og er lik 1 hvis $X_i = Y_i$ og -1 hvis $X_i = -Y_i$ (Stock & Watson, 2020). Ved sterke sammenhenger kan regresjonsmodellene bli upresise, slik at vi kan få problemer med å estimere reliable og valide regresjonskoeffisienter (Studenmund, 2014). Vi vil derfor utføre en korrelasjonsanalyse i delkapittel 7.2, for å avdekke eventuelle korrelasjoner.

5.2.5 Autokorrelasjon

Modellenes residualer kan også være preget av autokorrelasjon, som vil si at feilleddene i modellen korrelerer med hverandre. Dette forekommer som regel i tidsserier (paneldata), men kan også forekomme blant observasjonene i et øyeblikksbilde. I et øyeblikksbilde, kan for eksempel spesielle hendelser i et geografisk område påvirke de respektive markedsforholdene, slik at tilbud og etterspørsel påvirkes i nærliggende geografiske områder (Studenmund, 2014). Vi antar da at romprisene i de nærliggende geografiske områdene vil endres. De spesielle hendelsene kan for eksempel være sesongbaserte, slik som påskeferier, eller forbundet med større arrangementer som idrettsmesterskap. Dermed vil en test for autokorrelasjon kunne bidra til å indikere om utvalget påvirkes av spesielle hendelser i våre geografiske områder. Vi vil derfor benytte en Durbin-Watson test for å teste for autokorrelasjon i residualene. I den kontekst, så påpeker White & Mulligan (2010) at hotellrompriser forventes å ha en viss grad av autokorrelasjon. Det kan dermed hende at vi avdekker autokorrelasjon i residualene.

5.2.6 Tilpasning av regresjonsmodellene

For å tilpasse regresjonsmodellene til dataene i utvalget, anvendte vi programmene RStudio og Stata. Vi importerte datasettet som vi innhentet gjennom web-skraping i disse programmene, før vi programmerte regresjonsmodellene. For å implementere vår metode, spesifiserte vi modellene med den log-lineære funksjonsformen vi argumenterer for i avsnitt 5.2.2. Resultatene vi fikk ut fra disse programmene ble deretter formatert som tabeller og figurer gjennom ulike funksjoner i programmene.

5.2.7 Modellens forklaringsgrad

Etter å ha estimert regresjonskoeffisientene i våre modeller, må vi vurdere deres prestasjon og validitet gjennom ulike statistikker, tester og diagrammer. Disse vil beskrive hvor godt

modellene passer til å besvare problemstillingen. Vi vil drøfte potensielle begrensninger ved modellene og hvordan vi kan adressere disse i lys av forklaringsgrad. To vanlige oppsummerende statistikker er regresjonens standardfeil (SER) og R-kvadrat (R^2). Disse måler hvor godt OLS-estimatet til de multiple regresjonslinjene beskriver dataene (Stock & Watson, 2020). I tillegg til disse vil vi evaluere og kommentere F -statistikk og konfidensintervaller i kapittel 7.

5.3 Oppsummering av metode

I kapittel 5 har vi beskrevet og begrunnet vårt valg av metodisk tilnærming og undersøkt hvorvidt våre observasjoner egner seg til vår undersøkelse. Vi kan konkludere med at dataene tilfredsstillende forutsetningene til multiple regresjonsmodeller, og vil derfor fortsette oppgaven med å presentere hvordan dataene vil fordele seg i modellene.

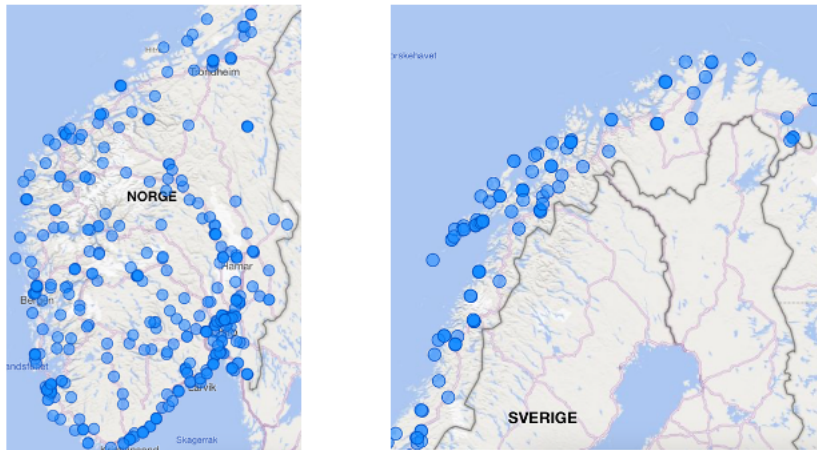
6 Presentasjon av data

I dette kapittelet vil vi presentere de innhentede dataene om hotellenes geografiske plassering, segmentnivåer, bærekraftsertifisering og kundevurderinger. Dette vil gi en oversikt over våre observasjoner.

6.1 Geografisk plassering av hotellene

Norge er et langstrakt land med hoteller spredt fra nord til sør, og våre observasjoner dekker store deler av landet, som man kan se i

Figur 6.1 under. Kartet til venstre viser våre observasjoner sør for tettstedet Hell i Trøndelag, mens kartet til høyre viser observasjonene nord for Hell.



Figur 6.1: Geografisk plassering av hotellene i utvalget (PowerBI).

Vi observerer at hotellkonsentrasjonen i Norge varierer og er ulik fra tettsted til tettsted. Den geografiske sorteringen av hotellene er basert på tettstedenes plassering, som baserer seg på hotellenes oppgitte adresser på Booking.com. Den største konsentrasjonen observerer vi i hovedstaden Oslo, med 79 hoteller. I Bergen er det 36, mens Stavanger kommer på tredje plass med 20 hoteller, etterfulgt av henholdsvis Trondheim, Tromsø, Gardermoen, Kristiansand, Svolvær, Bodø og Harstad med henholdsvis 19, 16, 12, 11, 9, 8 og 7 hoteller. Denne variasjonen gjør det utfordrende å få et helhetlig bilde av hotellmarkedet i Norge. Vi vil diskutere hvordan vi har håndtert denne utfordringen i delkapittel 7.1.

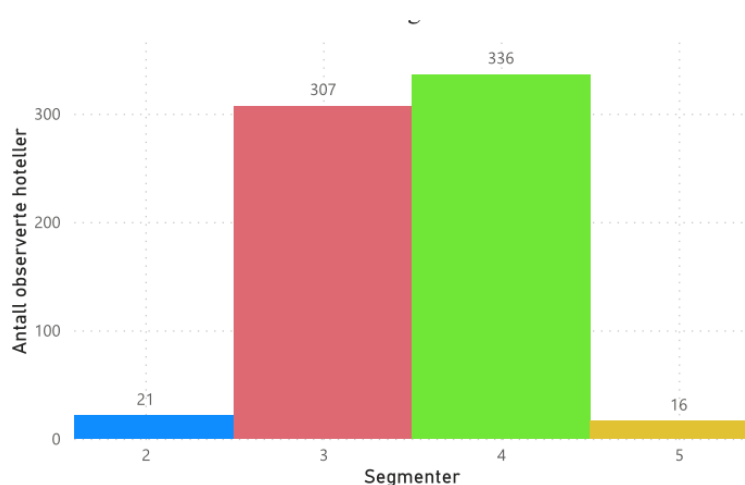
6.2 Fordeling på segmentnivå

For å få bedre innsikt i variasjonene mellom segmentene, vil vi presentere våre observasjoner fra hvert segment. Dette vil gi en dypere forståelse av segmentnivåene, og vil være nødvendig for å forstå valgene vi har gjort i regresjonsmodellene. Som tidligere forklart, har vi valgt å segmentere og benevne hotellene i et segmentrangeringssystem, på en skala fra én til fem (Guillet & Law, 2010). Dette er en egen vurdering av Booking.com, da vi ikke har et offisielt segmenteringssystem i Norge (Norsk Hotellklassifisering, 2023). Videre er segment én ekskludert fra oppgaven på grunn av mangel på observasjoner.

I dette delkapittelet vil vi først presentere andelen hoteller i hvert segment, før vi ser på prisvariasjoner, bærekraftsertifiseringer og til slutt kundevurderingene innad i hvert segment.

6.2.1 Observerte segmentnivåer

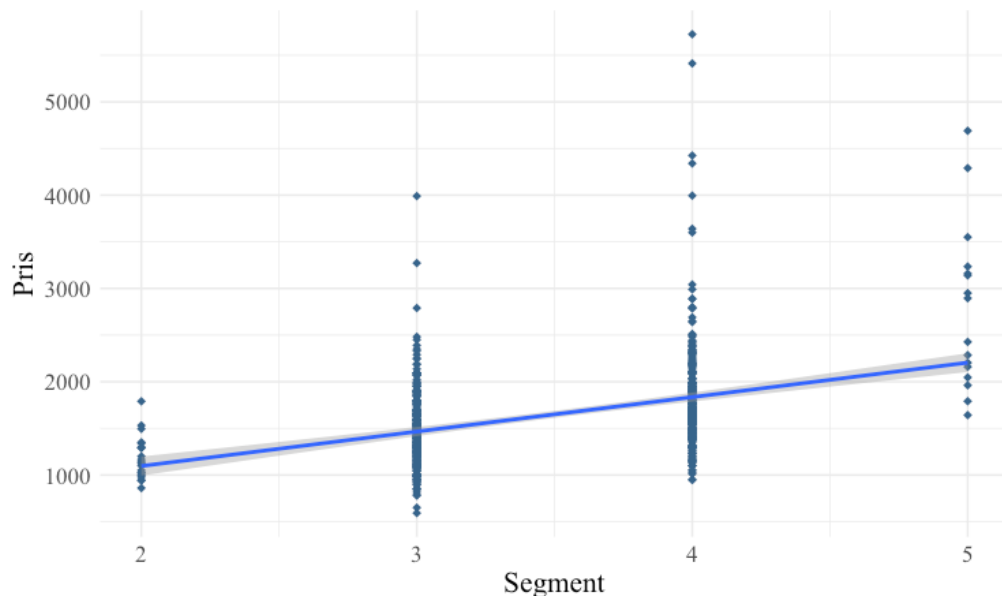
Segmentstørrelsene er ulike og varierer fra nivå 2 til 5. De to største segmentene er tre og fire. Segment tre utgjør 45 % ($n = 307$) av observasjonene i utvalget, mens segment fire utgjør 49 % ($n = 336$). Den minste andelen finner vi i segment fem, som utgjør 2 % ($n = 16$), og den nest minste andelen finner vi i segment to, som utgjør 3 % ($n = 21$) av utvalget. Segment én er ekskludert fra utvalget på grunn av for få observasjoner (2) og tilbakemeldinger (<10) til at undersøkelsene blir representative. Vi har grundig innhentet data for alle segmentene i vår studie. Likevel domineres studien av hoteller i segment tre og fire, da disse var tilgjengelige på Booking.com på den aktuelle dagen. En visualisering av fordelingen presenteres nedenfor i Figur 6.2, der segmentnivået er på x-aksen og andelen observerte hoteller er på y-aksen.



Figur 6.2: Segmentandelene i utvalget.

6.2.2 Observerte priser på segmentnivå

Nedenfor har vi konstruert et punktdiagram som viser prisvariasjonene for våre observasjoner (Figur 6.3), der segmentnivå er på x-aksen og prisnivået på y-aksen. For en mer detaljert beskrivelse, se Tabell A6.1.



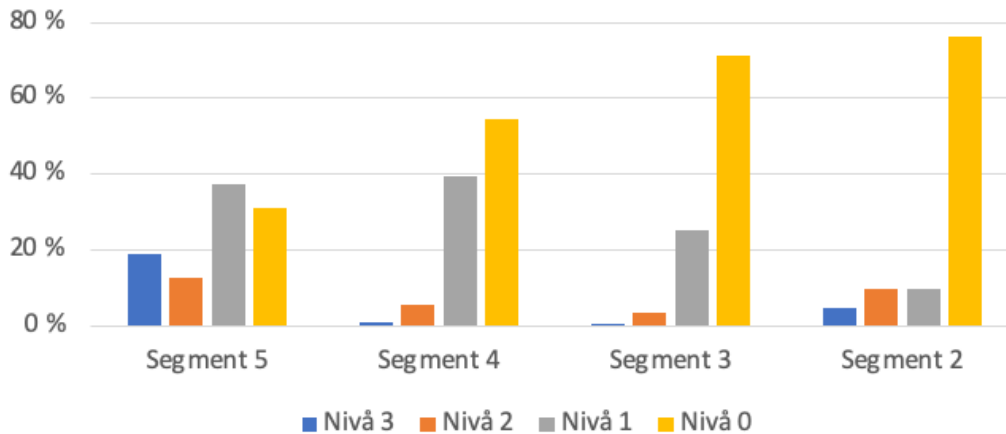
Figur 6.3: Prisvariasjon blant hotellene innenfor hvert segment.

Vi observerer flere trekk ved våre observasjoner, som er visualisert i Figur 6.3 og deskriptivt forklart Tabell A6.1. For det første varierer prisene betydelig mellom de ulike segmentene. Gjennomsnittsprisene øker fra segment to til fem, noe den positive helningen på den blå linjen indikerer. Den høyeste prisen er representert av et hotell i segment fire, med en pris på 5 725,- per natt. Dette hotellet ligger ved en attraktiv vinterdestinasjon. Ved innsamlingsdatoen kan det ha vært høy etterspørsel i området, noe som kan bidra til å forklare den høye prisen. På den andre enden av skalaen finner vi et hotell i segment tre, som har den laveste prisen på 592,- per natt.

Den samlede statistikken i Tabell A6.1 viser at observasjonene har et høyt standardavvik (fra 229,48 til 878,78), som indikerer at det er stor forskjell på hvor mye kundene kan forvente å betale innenfor hvert segment. Standardavviket øker i takt med segmentnivået. Vi undersøkte de høye verdiene nærmere, men fant ingen feil blant observasjonene. Dette skyldes at alle hotellrom i utvalget har like forutsetninger.

6.2.3 Bærekraftsertifiseringsnivå

Når det gjelder bærekraftsertifiseringer, ønsker vi å undersøke om kundenes betalingsvilje endres ved hoteller som er rangert som bærekraftige på en skala fra 0 til 3. For å gjøre dette har vi samlet inn informasjon om hotellenes bærekraftsertifiseringer fra Booking.com. I Figur 6.4 nedenfor har vi oppsummert andelen hoteller med bærekraftsertifisering 0–3 på y-aksen, oppdelt etter segmenttilhørighet på x-aksen.



Figur 6.4: Observerte bærekraftsertifiseringer i hvert segment.

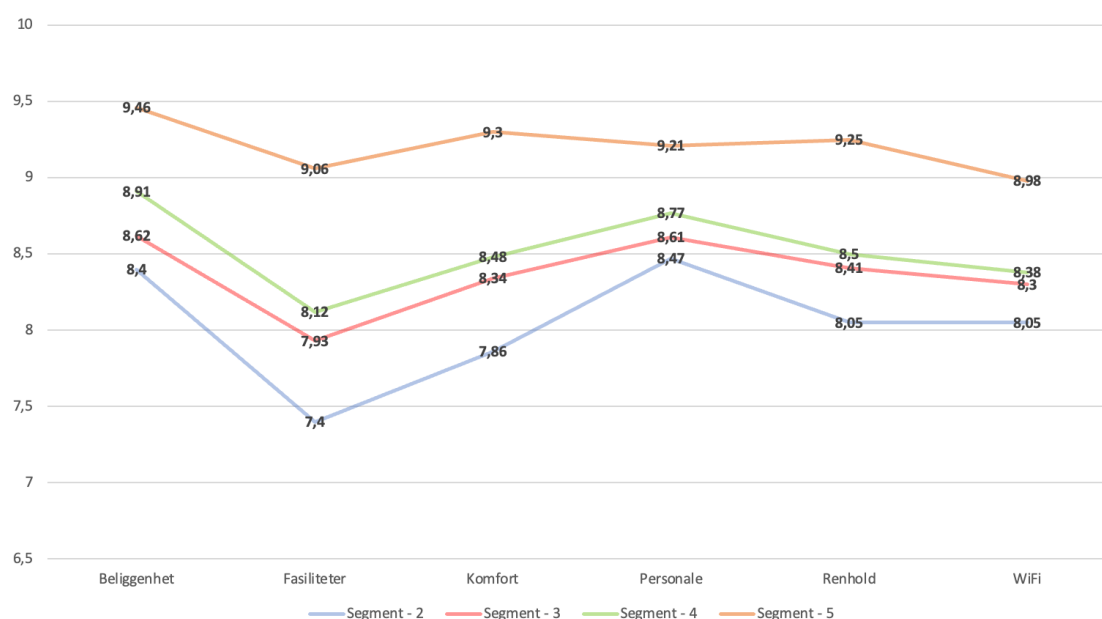
Våre observasjoner viser at prioriteringen av bærekraft hos hotellene varierer innen hvert segment og mellom segmentene. I sum har 62,2 % av hotellene bærekraftsertifiseringsnivå 0, mens 37,8 % befinner seg på nivå 1 til 3.

Kun 8 av totalt 680 hoteller (1,2 %) i utvalget er sertifisert på nivå 3, hvorav 3 tilhører segment fem, 3 tilhører segment fire, ett tilhører segment tre, og ett tilhører segment to. I tillegg er 32 hoteller (4,7 %) sertifisert på nivå 2, hvorav 2 tilhører segment fem, 18 tilhører segment fire, 10 tilhører segment tre, og 2 tilhører segment to. 217 hoteller (31,9 %) har sertifisering på nivå 1, hvorav 6 tilhører segment fem, 132 tilhører segment fire, 77 tilhører segment tre, og 2 tilhører segment to. Til slutt er det 423 hoteller (62,2 %) som har en sertifisering på nivå 0, hvorav 5 tilhører segment fem, 183 tilhører segment fire, 219 tilhører segment tre, og 16 tilhører segment to.

Våre observasjoner indikerer at hoteller i høyere segmenter oftere prioriterer bærekraft. Dette funnet kan sees i sammenheng med Geerts (2014), som også finner flest bærekraftige hoteller i de mest luksuriøse segmentene.

6.2.4 Innhentende kundevurderinger

For å forstå betydningen av kundevurderingene og hvordan de vil påvirke våre regresjoner, er det hensiktsmessig å se hvordan de ulike segmentene scorer i forhold til hverandre. I Figur 6.5 nedenfor, vises hvordan kundene i gjennomsnitt har vurdert sitt hotellopphold på Booking.com (y-aksen) for våre observasjoner, fordelt på tilbakemeldingsvariablene (x-aksen). Den lysegrønne linjen viser det samlede gjennomsnittet for alle segmentene, mens de resterende fire linjene er gjennomsnittet innenfor hvert segment [2–5].



Figur 6.5: Gjennomsnittlige kundevurderinger av hotellattributtene i hvert segment.

Når vi ser på segmentene hver for seg, kan vi observere at segment fem har de høyeste gjennomsnittlige vurderingene for de inkluderte variablene, med verdier mellom 8,98 og 9,46. Segment fire har også høye vurderinger, med gjennomsnitt på mellom 8,12 og 8,91. Segment tre har gjennomsnittlige vurderinger på mellom 7,93 og 8,62, mens hotellene i segment to har de laveste gjennomsnittlige vurderingene på mellom 7,4 og 8,47.

For en bedre forståelse av kundevurderingene, har vi inkludert deskriptiv statistikk i Tabell A6.3. Denne viser tydelige variasjoner mellom hotellsegmentene, både for standardavvik og minimum- og maksimumsverdier. Samlet sett gir kundevurderingene god innsikt i hotellgjestenes subjektive vurderinger, og vi ser at vurderingene øker i takt med segmentnivået, der gjennomsnittsverdiene er lavest i segment to og høyest i segment fem.

7 Analyse og diskusjon

I dette kapitlet presenterer vi resultatene fra våre regresjonsmodeller og diskuterer hvordan kundevurderinger og hotellenes bærekraftsertifiseringer kan påvirke prisen på et hotellrom, og knytter dette opp mot tidligere studier.

Vi vil analysere to ulike modeller i denne oppgaven, hvor vi kontrollerer for segment på to ulike måter. Den første modellen, som vi har kalt Modell 1, er en samlet modell der vi har slått sammen alle segmentene. En utfordring ved denne modellen er at vi har mange observasjoner i segment tre og fire, noe som gjør at disse segmentene vil prege modellen. Denne store segmentkonsentrasjonen gjør at vi har konstruert en Modell 2, der segmentene holdes adskilt.

7.1 Inndeling av markedet

Som belyst i delkapittel 6.1, er våre observasjoner spredt rundt om i Norge, og hotellkonsentrasjonene varierer betydelig. Det var derfor nødvendig å undersøke om det var betydelige forskjeller mellom ulike regioner i Norge. Ved å utføre en samlet regresjonsmodell for hele Norge (Tabell A7.1) observerte vi at det var betydelige forskjeller i hvilke faktorer som er viktige for henholdsvis by og land, som gjorde at vi ikke kunne se hele Norge som ett marked. Derfor deles datasettet i både Modell 1 og Modell 2 opp i By og Distrikts-Norge (heretter kalt DN), som vi nå vil forklare nærmere.

Ved å lage regresjonsmodeller for hver av de åtte største hotellkonsentrasjonene, identifiserte vi at markedene for Oslo, Bergen, Trondheim, Stavanger, Gardermoen, Tromsø, Kristiansand og Bodø hadde små forskjeller mellom både koeffisientene og signifikansnivået til variablene. Alle disse stedene har høy konsentrasjon av hoteller, og danner derfor dimensjonen By. Vi kunne også valgt andre måter å definere dimensjonen By på. For eksempel kunne vi utført regresjoner med de 10, 25 eller 30 største byene. Da vi testet disse, ble resultatene mer sprikende, og det var tydelig at det ikke var greit å lage slike inndelinger. Konsistensen i sammenhengen mellom de uavhengige variablene og hotellromprisene ble da mindre, slik at tilliten til modellen ble lavere. En mulig forklaring på hvorfor resultatene ble mer sprikende med flere byer, er at flere av de mindre byene har lavere hotellkonsentrasjon. Det kan også hende at når vi inkluderer flere byer i modellen, så blir modellen mer heterogen. Da øker variabiliteten mellom observasjonene i modellen, slik at resultatene blir mindre reliable og

valide. Vi konkluderte derfor med at vår By-dimensjon, med de åtte største hotellkonsentrasjonene, samlet de regionene som var likest.

Vi har tatt med Gardermoen som en del av By, til tross for at det ikke er definert som én by. Dette valget ble gjort fordi vi mener det er relevant å sammenligne det med de andre byene i undersøkelsen. Gardermoen har et høyt antall hoteller innenfor et lite område, noe som gir kundene et bredt spekter av valgmuligheter. En annen grunn til å inkludere Gardermoen er at flyplasshotellene i Bergen også er tatt med i Bergen. Derfor ville det være feil og unøyaktig å ekskludere Gardermoen fra dimensjonen By. Flyplasshotellene i Bergen ble inkludert på grunn av deres nærhet til Bergen sentrum, og at det samtidig er et reelt alternativ for kunder som ønsker å være både i sentrum og nær flyplassen. Selv om flyplasshotellene ikke ligger direkte i sentrum, er de enkelt tilgjengelige med kort reisevei. Derfor ville det være unøyaktig å ekskludere dem fra By-dimensjonen.

På grunn av få hoteller enkelte steder, ble det nødvendig å samle disse i én kategori. Dette genererte dimensjonen DN. Observasjonene i DN var annerledes enn tettstedene med høy konsentrasjon, og derfor valgte vi å se på forskjellene mellom by og land og gjøre inndelingen deretter.

For å ta hensyn til betydelige prisvariasjoner mellom byene, har vi inkludert dummyvariabler for hver av de åtte byene i regresjonsmodellene. Dette bidrar til å korrigere for eventuelle store forskjeller i prisingen mellom de ulike byene og sikrer en mer nøyaktig analyse av resultatene.

For dimensjonen By har vi en konsentrasjon på 2, 60, 248 og 11 i henholdsvis segment to, tre, fire og fem. Dette gjør at undersøkelsen er dominert av hotellene i segment tre og fire og få observasjoner i segment to og fem. En konsekvens av dette er at estimatene av regresjonslinjene i Modell 2 blir begrenset for segment to og fem. For DN har vi henholdsvis 19, 248, 208 og 5 observasjoner i segment to, tre, fire og fem. Også her får vi en stor konsentrasjon rundt segment tre og fire og få observasjoner i segment to og fem. Dermed begrenses også de estimerte regresjonslinjene for segment fem i DN.

Mangel på observasjoner i segment to og fem i By og fem i DN, gjør at vi får problemer med å estimere deres regresjonslinjer. Til tross for dette, velger vi å beholde dimensjonsinndelingene; By og DN, da dette gir den mest representative studien for 97% av utvalget i Modell 2.

7.2 Korrelasjonsanalyse

I delkapittel 3.4 forklarte vi Halo-effekten, hvor kunder som uttrykker misnøye med en spesifikk variabel også har en tendens til å vurdere andre variabler negativt. Gitt denne effekten er det sannsynlig at det kan oppstå korrelasjoner mellom variablene vi undersøker. Vi anser det derfor nødvendig å grundig undersøke om det er en slik sammenhengen i våre observasjoner. Derfor har vi gjennomført en korrelasjonsanalyse basert på våre observasjoner, som presenteres i Tabell 7.1.

Tabell 7.1: Korrelasjonsmatrise for de inkluderte variablene.

	Ln(Pris)	Segment	Beliggenhet	Fasiliteter	Komfort	Personale	Renhold	WiFi	Bærekraft
Ln(Pris)	1,00								
Segment	0,45	1,00							
Beliggenhet	0,39	0,30	1,00						
Fasiliteter	0,46	0,39	0,52	1,00					
Komfort	0,46	0,35	0,51	0,95	1,00				
Personale	0,39	0,20	0,46	0,67	0,68	1,00			
Renhold	0,41	0,28	0,46	0,93	0,97	0,71	1,00		
WiFi	0,18	0,16	0,25	0,46	0,47	0,36	0,46	1,00	
Bærekraft	0,14	0,20	0,18	0,14	0,12	0,08	0,09	0,13	1,00
<i>Note:</i>	Foruten Bærekraft og Ln(Pris), er alle signifikant på $p < 1\%$.								

Tabell 7.1 viser at flere av variablene er høyt korrelert med hverandre. De høyeste verdiene finner vi mellom *renhold* og *fasiliteter* (0,93), *komfort* og *fasiliteter* (0,95) og *renhold* og *komfort* (0,97). Dette indikerer en sterk positiv sammenheng mellom variablene, som kan gjøre regresjonsmodellen upresis, slik at det blir vanskelig å isolere effekten av hver variabel. Dette kan begrense vår evne til å trekke konklusjoner om sammenhenger i utvalget (Studenmund, 2014).

At flere variabler korrelerer med hverandre er likevel ikke overraskende. Det er ikke uvanlig at service- og kvalitetsvariabler er høyt korrelerte (Lai et al., 2018). Abratea et al. (2011), Borges et al. (2015), Castro og Ferreira (2018) og Stringam et al. (2010) fant høye korrelasjonsverdier i sine hedoniske prisingstudier av hotellmarkeder. De har valgt å behandle variablene ulikt, enten gjennom eliminering eller inklusjon. Eksempelvis valgte Abratea et al. (2011) å beholde

samtligte variabler, selv med høye korrelasjonsverdier. Dette fordi en slik modell var mer anvendelig og styrket deres resultater. Stringham et al. (2010) mener at deres høye korrelasjonsverdier kan skyldes at kunder med et godt helhetsinntrykk av hotellet, også gir en høy score på alle tilbakemeldingsvariablene, og motsatt ved et dårlig helhetsinntrykk. Dette er med andre ord et tegn på en Halo-effekt, som også kan være tilfelle for våre data. En annen årsak til de høye korrelasjonsverdiene kan være hotellets iboende kvalitet, som i så måte medfører høye eller lave kundevurderinger over hele linjen (Stringham et al., 2010). Vi kan heller ikke utelukke at våre korrelasjonsverdier reflekterer dette fenomenet.

Vi kan imidlertid argumentere for å beholde de høyt korrelerte variablene, siden eliminering kan svekke nøyaktigheten til regresjonsmodellene (White & Mulligan, 2010). Med bakgrunn i våre undersøkelser av alternative regresjonsmodeller samt ulike tilnærminger i litteraturen, velger vi å beholde samtlige variabler. Dette fordi vi ønsker å undersøke sammenhengen mellom disse og hotellrompriser, som gir oss et bedre grunnlag til å besvare forskningsspørsmålet.

7.2 Modell 1

Modell 1 er en modell der alle hotellene er samlet i én regresjon, fordelt på dimensjonene By og DN. Dette vil danne grunnlaget for vår analyse av hvordan kundevurderinger og bærekraftsertifiseringer påvirker prissettingen av dobbeltrom i to de dimensjonene. Samtidig er det, oss bekjent, få studier som undersøker hvordan kundevurderinger og bærekraftsertifiseringer påvirker hotellromprisen i det norske hotellmarkedet. Derfor anser vi det som viktig å skape en helhetlig forståelse av hotellmarkedet og identifisere hvilke variabler som påvirker prissettingen, slik vi gjør i Modell 1.

En utfordring ved Modell 1 er at hovedtyngden av observasjonene våre utgjøres av noen få segmenter. For dimensjonen By har vi kun 2 observasjoner i segment to, 60 observasjoner i segment tre, 128 observasjoner i segment fire og 11 observasjoner i segment fem. For DN har vi 19 observasjoner i segment to, 247 observasjoner i segment tre, 208 observasjoner i segment fire og kun 5 observasjoner i segment fem. Totalt utgjør observasjonene i segment tre og fire hele 94 % av utvalget. Disse segmentene vil dermed ha størst innflytelse på de estimerte regresjonslinjene i våre modeller. Følgende Tabell 7.2 presenterer resultatene fra Modell 1.

Tabell 7.2: Regresjonstabell for Modell 1; By og DN.

	Avhengig variabel:			
	Ln(Pris)			
	By		DN	
	Koeffisient	SD	Koeffisient	SD
Segment	0,172***	0,026	0,119***	0,021
Beliggenhet	0,116***	0,028	0,067***	0,024
Fasiliteter	0,062	0,093	0,061	0,054
Komfort	0,323**	0,151	0,154*	0,079
Personale	0,089**	0,041	0,089***	0,031
Renhold	(0,236)*	0,136	(0,164)**	0,077
WiFi	(0,008)	0,039	(0,014)	0,017
Bærekraft	0,028	0,020	0,021	0,019
Oslo	0,233***	0,066		
Bergen	0,068	0,070		
Trondheim	0,205***	0,077		
Stavanger	0,330***	0,075		
Gardermoen	0,149*	0,083		
Tromsø	0,185**	0,078		
Kristiansand	0,094	0,082		
Konstant	3,557***	0,329	5,325***	0,246
Observasjoner	201		479	
R²	0,640		0,283	
Residual SE	0,175		0,245	
	(df = 185)		(df = 470)	
F-statistikk	21,932*** (df = 15; 185)		23,171*** (df = 8; 470)	
Note:	*p<0,1; **p<0,05; ***p<0,01; negative koeff. i parentes			

Basert på resultatene i Tabell 7.2 har vi satt opp regresjonsligninger for By og DN. $\ln(\text{Pris})$ er den naturlige logaritmen av prisfunksjonen, og de uavhengige variablene representerer de inkluderte hotellattributtene. For By har vi også dummyvariabler. Feilleddet i regresjonsligningene er representert med ε_i .

By:

$$\begin{aligned} \ln(\text{Pris}) = & 3,557 + 0,172 \text{ Stjerner}_i + 0,116 \text{ Beliggenhet}_i + 0,062 \text{ Fasiliteter}_i + 0,323 \text{ Komfort}_i \\ & + 0,089 \text{ Personale}_i - 0,236 \text{ Renhold}_i - 0,008 \text{ WiFi}_i + 0,028 \text{ Bærekraft}_i + 0,233 \text{ Oslo}_i \\ & + 0,068 \text{ Bergen}_i + 0,205 \text{ Trondheim}_i + 0,330 \text{ Stavanger}_i + 0,149 \text{ Gardermoen}_i \\ & + 0,185 \text{ Tromsø}_i + 0,094 \text{ Kristiansand}_i + \varepsilon_i \end{aligned}$$

DN:

$$\begin{aligned} \ln(\text{Pris}) = & 5,325 + 0,119 \text{ Stjerner}_i + 0,067 \text{ Beliggenhet}_i + 0,061 \text{ Fasiliteter}_i + 0,154 \text{ Komfort}_i \\ & + 0,089 \text{ Personale}_i - 0,164 \text{ Renhold}_i - 0,014 \text{ WiFi}_i + 0,021 \text{ Bærekraft}_i + \varepsilon_i \end{aligned}$$

7.2.1 Residualer

Etter å ha presentert resultatene fra Modell 1, vil vi nå vurdere modellens presisjon. Vi vet fra kapittel 5 at dataene innfrir forutsetningene til multiple regresjonsmodeller, men vi må også vurdere om resultatene fra regresjonsmodellene kan forklare prisvariasjonen i By og DN. Modell 1 har uforklarte avvik mellom de observerte verdiene av den avhengige variabelen, $\ln(\text{Pris})_i$, og dens predikerte verdi, $\ln(\widehat{\text{Pris}})_i$. Dette avviket er beregnet med forskjellen mellom de estimerte regresjonskoeffisientene og de faktiske verdiene til de uavhengige variablene. For å vurdere regresjonsmodellens nøyaktighet vil vi analysere residualene i modellen. Residualene kan gi oss et mål på spredningen eller variabiliteten til feilleddene i modellen (Studenmund, 2014). Dette vil i så måte hjelpe oss med å vurdere hvorvidt det uforklarte avviket i Modell 1 er et problem eller ikke. Vi har inkludert boksplokk for residualene i Modell 1 i Figur A7.1, og statistikk over residualene i Tabell A7.9.

7.2.1.1 Residualer for By

R^2 , som forklarer hvor stor andel av variasjonen i $\ln(\text{Pris})$ som blir forklart av attributtene, er 61,1 % for By – en verdi som anses som god. Vi ser ut fra Figur A7.1 og Tabell A7.3 at residualene har et minimum på -0,453, et første kvartil på -0,097, en median på 0,008, et tredje kvartil på 0,103, og et maksimum på 0,587. Dermed ligger halvparten av residualene mellom -0,097 og 0,103. Samtidig er majoriteten av residualene små og sentrert rundt null, som styrker modellens presisjon. Videre viser residualenes standardfeil på 0,175 med 185 frihetsgrader, at

residualene er noe spredt. Dette har en negativ påvirkning på modellens presisjon. På den andre side styrkes modellen av at F -statistikken er lik 21,93 med 15 og 185 frihetsgrader og en p -verdi mindre enn 0,01. Dette betyr at modellen har en bedre tilpasning enn en nullmodell, hvor alle uavhengige variabler settes lik null. Vi kan dermed oppsummere med at Modell 1 har en god tilpasning til dataene i By.

7.2.1.2 Residualer for DN

For dimensjonen DN forklarer attributtene 28,3 % (R^2) av variasjonen i $\ln(\text{Pris})$. Modellen forklarer en del av variasjonen, men det er fortsatt mye uforklart variasjon i den avhengige variabelen, forklaringskraften anses derfor som moderat. Residualene har et minimum på -0,759, et første kvartil på -0,153, en median på -0,028, et tredje kvartil på 0,154, og et maksimum på 1,084. For DN betyr det at halvparten av residualene befinner seg mellom -0,152 og 0,154. Vi ser, på samme måte som for By, at residualene er små og sentrert rundt null. Dette styrker modellens presisjon. Videre er residualenes standardfeil på 0,245 med 470 frihetsgrader. Denne spredningen er med på å svekke modellens presisjon. Til slutt er modellens F -statistikk på 23,17 med 8 og 470 frihetsgrader, med en p -verdi på mindre enn 0,01. Dette betyr at regresjonslinjen for DN også er bedre tilpasset enn en nullmodell. Modell 1 har dermed også en god tilpasning til DN-dataene. De uavhengige variablene kan dermed forklare store deler av prisvariasjonen i DN.

7.2.1.3 Oppsummering av residualene til Modell 1

Vi kan oppsummere med at residualene viser at Modell 1 har moderat til god forklaringskraft. Majoriteten av residualene er små og sentrert rundt null for både By og DN, hvilket tyder på at modellen er nøyaktig. På den andre side viser residualenes standardfeil at residualene har en viss spredning rundt null for både By og DN. Dette kan svekke modellens nøyaktighet. Vi vet imidlertid også at modellen er bedre enn en nullmodell for både By og DN. I sum tyder dermed residualene på at modellen har en god tilpasning, med en noe bedre tilpasning for DN sammenlignet med By.

7.2.1.4 Durbin-Watson test for autokorrelasjon

Videre har vi testet for autokorrelasjon i residualene med en Durbin-Watson test. DW-statistikk for By og DN fremkommer i Tabell A7.6. Testen sammenligner forskjellen mellom hver påfølgende residual med gjennomsnittet av residualene. Autokorrelasjon testes oftest for i

tidsserier, men det kan likevel forekomme autokorrelasjon i tverrsnittstudier, gitt at et fenomen påvirker det vi studerer i det aktuelle øyeblikksbildet (Studenmund, 2014). Resultatene fra testen har en p -verdi på over 5 % for By og under 5 % for DN. Dette indikerer fravær av autokorrelasjon i residualene til By og autokorrelasjon i residualene til DN (Chatterjee & Hadi, 2015; Curto & Pinto, 2011). Autokorrelasjonen i DN er også positiv siden Durbin-Watson-statistikken til DN-residualene er på 0,523. Dette betyr at residualene ikke er uavhengige av hverandre. Det har imidlertid ikke vært mulig å eliminere effekten av disse autokorrelasjonene med alternative regresjonsmodeller. I avsnitt 5.2.4 om autokorrelasjon, presenterer vi sannsynlige årsaker til autokorrelasjon i residualene. Det er mulig at spesielle hendelser i ulike geografiske områder har påvirket markedsforholdene slik at hotellromprisen påvirkes i nærliggende områder. Samtidig påpeker White og Mulligan (2010) at hotellrompriser generelt forventes å ha en viss grad av autokorrelasjon. Vi har også observert i tidligere studier, for eksempel Soler et al. (2019), at det er vanlig praksis å anta at hotellmarkeder er relativt homogene, slik at man kan kontrollere for autokorrelasjon. Gjennom å anta at DN består av relativt homogene markeder, kan vi dermed kontrollere for effektene av autokorrelasjon, gitt at de har betydning for resultatene.

7.2.2 VIF-analyse

Vi har også gjennomført en VIF-analyse av Modell 1 for å vurdere multikollinearitet mellom variablene i modellen. Resultatene for VIF-analysen ligger i Tabell A7.7 og Tabell A7.9. Basert på tidligere studier (Castro & Ferreira, 2018; Zhang et al., 2011), kan vi velge å bruke en tommelfingerregel om at VIF-verdier på under 10 er akseptable nivåer av multikollinearitet. Høyere verdier kan anses som problematiske fordi de er tegn på sterk multikollinearitet (Curto & Pinto, 2011).

Resultatene viser at variablene *segment* og *beliggenhet* har lave VIF-verdier på henholdsvis 1,45 og 1,31 for By, og 1,23 og 1,6 for DN. Variabelen *bærekraft* har også lave VIF-verdier på 1,06 for By og 1,04 for DN. På den andre side har variablene *fasiliteter*, *komfort* og *renhold* høyere VIF-verdier. *Fasiliteter* har en VIF-verdi på 16,71 for By og 11,47 for DN, mens *komfort* har en VIF-verdi på 37,74 for By og 17,67 for DN. *Renhold* har en VIF-verdi på 27,48 for By og 14,93 for DN. Dette indikerer en betydelig grad av multikollinearitet mellom disse variablene i både By og DN.

Med utgangspunkt i resultatene over, kan vi argumentere for å ekskludere samtlige eller deler av variablene med høye VIF-verdier fra Modell 1. Eksempelvis er *renhold* og *komfort* to variabler med høye VIF-verdier i Modell 1, noe som kan redusere nøyaktigheten til modellen. I Tabell A7.8 undersøker vi om variablenes VIF-verdier reduseres ved å eliminere variabler med høy VIF. Vi ser at VIF-verdiene reduseres til under 10 for alle gjenværende variabler når vi eliminerer *komfort* og *renhold*. Dette kan tenkes å være en bedre modell, fordi den fjerner eventuelle problemer med multikollinearitet. Vi kan underbygge dette med Borges et al. (2015) og Castro og Ferreira (2018), som valgte å ekskludere variabler med høye VIF-verdier. Castro & Ferreira (2018) ekskluderte kun *renhold* fordi denne variabelen korrelerte med flest variabler. På den annen side kan *komfort* og *renhold* teoretisk sett ha betydning for den avhengige variabelen $\ln(\text{Pris})$, og dermed styrke gyldigheten og påliteligheten til resultatene (Stock & Watson, 2020). Derfor velger vi å beholde disse variablene, selv med høy grad av multikollinearitet. Vi kan også se til Zhang et al. (2011) som hevder at godt renhold er en forventning blant kundene. Dermed kan et lavt renholds nivå påvirke kundenes oppfatning av produktet, og dermed deres tilbakemeldinger. Videre kan Halo-effekten være en bakgrunn for de høye VIF-verdiene (Nicolau et al., 2020).

7.2.3 Attributtenes effekter

Vi har nå vurdert residualene til Modell 1 samt i hvilken grad modellens regresjonslinjer preges av multikollinearitet og autokorrelasjon. I det følgende avsnittet vil vi diskutere resultatene fra Modell 1 og se dette i lys av tidligere studier av andre markeder. Innledningsvis gjentar vi at en utfordring ved Modell 1 er at store deler av observasjonene preges av segment tre og fire, som derfor også vil prege den følgende analysen.

7.2.3.1 Segment

I Modell 1 ser vi en betydelig positiv sammenheng mellom *segment* og prisene for en hotellovernatting i både By og DN. Resultatene viser at et hotell med høyere segmentrangering kan oppnå høyere priser, med en estimert økning på 17,2 % for By ($p < 0,01$) og 11,9 % for DN ($p < 0,01$). Dette er en generell oppfatning fra tidligere studier, der Israeli (2002), Zhang et al. (2011), Castro & Ferreira (2018) og Soler et al. (2019) finner at segmentnivå har en signifikant effekt på pris i henholdsvis Israel, New York, Lisboa og Algarve. Soler et al. (2019) finner den minste effekten på 9,8 % i Algarve, mens Israeli (2002) finner den største effekten i Jerusalem på 65 %. Zhang et al. (2011) finner et slags midtpunkt mellom de to førstnevnte studiene, hvor

en ekstra stjerne indikerer en økning i rompris på 35,8 % i New York. En forklaring på at vår studie har så ulike koeffisienter, kan være at vi har et geografisk sammensatt datasett. Altså at vi har med By og DN, istedenfor én by eller ett mindre geografisk marked.

Våre funn indikerer en tendens der segmenter spiller en mer avgjørende rolle i områder med høy hotellkonsentrasjon, og dette reflekteres også i høyere koeffisienter for byområder sammenlignet med distriktene. Flere faktorer kan bidra til dette mønsteret. Større byer tiltrekker seg typisk flere turister og har et bredere utvalg av hotellalternativer. Dette kan føre til økt konkurranse mellom hotellene i byene, noe som gjør at hoteller i høyere segment kan sette høyere priser basert på deres bedre vurderinger av kvalitet. På en annen side kan hotellmarkedet i distriktene ha færre hotellalternativer, og dermed være mindre konkurranseutsatt. Dette kan resultere i at også hoteller i lavere segment kan sette relativt høyere priser. Modell 1 støtter denne antakelsen, og antyder at det er en sammenheng mellom segmentrangering og høyere priser.

7.2.3.2 Beliggenhet

I Modell 1 finner vi en signifikant sammenheng mellom den rangerte beliggenheten til hotellene og prisene. Resultatene viser at hoteller med en bedre beliggenhet har høyere priser, med en estimert økning på 11,6 % for By ($p < 0,01$) og 6,7 % for DN ($p < 0,01$). Resultatene indikerer dermed at hoteller med en attraktiv beliggenhet – et tenkt eksempel kan være nært turistattraksjoner eller i sentrale områder – kan ha en prisfordel sammenlignet med hoteller med mindre gunstig beliggenhet.

Koeffisienten til DN (6,7 %) er lavere enn for By (11,6 %), noe som tilsier at hotellets beliggenhet påvirkes av antall hotellalternativer gjestene har. I DN har gjestene ofte få eller ingen alternativer å velge mellom. Et eksempel på dette er Thon Hotel Hallingdal. Hotellet ligger avsidesliggende til mellom Gol og Geilo, langs en av hovedveiene mellom Oslo og Bergen. Det er også det eneste hotellet i bygden (Booking.com, 2023b), som betyr at kundene har få eller ingen valgmuligheter. For hotellene som ligger i en by, så kan kundene ha mange valgmuligheter. I Oslo har kundene eksempelvis over 70 valgmuligheter (SSB, 2023). Dette kan bidra til å forklare hvorfor koeffisienten er lavere i DN.

Espinet et al. (2003) finner at nærhet til strender er viktig i distriktene i Spania, og gir en økning i rompris. Dette kan oppfattes som en attraksjonsverdi for enkelte kundesegmenter. Et

tilsvarende eksempel kan være Spidsbergseter hotell, som ligger avsides til, uten nærliggende hovedveier. Beliggenheten til dette hotellet og andre hoteller i DN viser seg likevel å ha en signifikant og positiv effekt på hotellrompriser. Dette indikerer at selv om beliggenheten kan være avsides eller mindre sentral, kan andre faktorer som utsikt, natur eller unike opplevelser i området fortsatt påvirke prisene positivt. Dette viser at oppfatningen av en god beliggenhet er relativ og avhengig av kontekst og marked. Samtidig er ikke beliggenhet like viktig i små tettsteder (DN) som i store byer, hvor forskjellene er mer markante. Dette fordi tettstedene ofte har kort vei til alt, lite trafikkstøy, osv.

Hvis vi ser på By i Modell 1, kan en god beliggenhet bety nærhet til sentrum. På lik linje med hva Bull (1994) finner, kan den observerte effekten av beliggenhet indikere at avstanden til sentrum har en betydelig effekt for By. På en annen side kan effekten av beliggenhet også forklares på andre måter: White og Mulligan (2010) påpekte at mengden støy var et like viktig utvalgs-kriterium som hotellets fysiske beliggenhet. Arbeil og Pizzam (1977) forklarer på sin side at betalingsviljen ikke nødvendigvis reduseres utenfor sentrum, hvis gjestene har tilgang til nærliggende kollektivtransport. Dette kan bety at en god beliggenhet for dimensjonen By inkluderer andre faktorer enn nærhet til sentrum, slik som lite støy, tilgjengelighet til kollektivtransport, parker, strender, skiløyper og lignende.

7.2.3.3 Fasiliteter

Tilbakemeldingsvariabelen *fasiliteter* omfatter tjenestene som gjestene kan nyte under oppholdet, slik som parkering, treningssenter, svømmebasseng, restaurant og spa. Estimatene i Modell 1 viser en liten prisøkning på 6,2 % for By og 6,1 % for DN, men disse estimatene er ikke signifikante ($p > 0,1$). Vi kan derfor ikke fastslå en sammenheng mellom fasiliteter og rompriser i hotellbransjen i det norske markedet, når vi betrakter segmentene samlet.

Det finnes flere mulige årsaker til at fasiliteter ikke har en statistisk signifikant effekt på romprisene. En mulig årsak kan være begrenset variasjon i de fasilitetene som tilbys av hotellene i datasettet, noe som kan ha begrenset vår evne til å finne en signifikant effekt. En annen mulig årsak kan være at hotellene i utvalget primært tilhører segment tre (45 %) og fire (49 %). Fasilitetene innad i disse segmentene kan være relativt like på tvers av hotellene. Dermed får vi i liten grad fanget opp effekten i de andre segmentene.

Det er oss bekjent få tidligere studier som støtter våre funn fra det norske markedet. Det er tilsynelatende en utbredt oppfatning at ulike fasiliteter har en signifikant effekt på romprisene (Andersson, 2010; Bull, 1994; Espinet et al., 2003; Thrane, 2007). Selv om vår analyse ser på den generelle variabelen *fasiliteter*, og ikke spesifikke attributter som for eksempel hårføner, som er en del av fasilitetene noen hoteller tilbyr, vil det være hensiktsmessig å drøfte et tidligere studier i lys av våre funn.

Bull (1994) forklarer i sin studie fra Australia at et hoteller med restauranter påvirker romprisene med 20 %. Våre observasjoner befinner seg primært i segment tre og fire, og hvor hotellrestaurantene kan være relativt like. Det gjør at kundene i det norske markedet ikke vektlegger en slik fasilitet ved bestilling av hotellrom, og dermed påvirker den heller ikke romprisene.

7.2.3.4 Komfort

Komfort refererer til gjestenes opplevelse av bekvemmelighet og velvære under oppholdet. For både By og DN finner vi en positiv og statistisk signifikant sammenheng mellom komforten på hotellene og romprisene, med en estimert økning på 32,3 % ($p < 0.05$) og 15,4 % ($p < 0.1$). Dette tyder på at hoteller med høyere komfortnivå tenderer mot høyere priser i både By og DN.

Vi kan trekke paralleller mellom våre signifikante og positive funn og Stringam et al. (2010), som finner at komfort er den faktoren som har størst innflytelse på generell tilfredshet blant kundene. Dette kan bidra til å forklare våre høye og signifikante koeffisienter.

Det er bemerkelsesverdig at våre funn avviker fra Castro og Ferreira (2018), som ikke finner noen signifikant effekt av *komfort* på romprisene i Lisboa. Forfatterne diskuterer imidlertid ikke denne variabelen ytterligere. Det ser derfor ut til å være regionale forskjeller i hotellmarkedet, da våre variabler skiller seg markant fra Lisboa (Castro & Ferreira, 2018), og det ser ut som dimensjonene By og DN kan ha unike kjennetegn ved seg. Dette kan skyldes at variasjonen i opplevd komfort på hotellene er ulik mellom hotellene i Norge.

7.2.3.5 Personale

Som en tilbakemeldingsvariabel hos Booking.com, refererer *personale* til gjestenes opplevelse av service og gjestfrihet under oppholdet. I Modell 1 ser vi at *personale* har en signifikant

positiv effekt på romprisene både i By og DN. Koeffisienten er lik 8,9 % for begge markedene, men signifikansverdien varierer mellom By ($p < 0,05$) og DN ($p < 0,01$). Vi ser dermed at effekten av personale har en sterkere signifikans i DN sammenlignet med By.

Thrane (2007) finner i sin Oslo-studie at *romservice* – en del av *personale* – har en negativ og signifikant effekt (-12,3 %) på romprisene. Funnet indikerer at en tilstedeværelse av et dyktig og effektivt personale, har en negativ effekt. Thrane (2007) sin antagelse er at hoteller som tilbyr romservice har en direkte inntekt fra denne ekstra tjenesten, og dette kompenseres med lavere hotellrompriser. Thrane (2007) sitt resultat skiller seg med 21,3 % fra vårt resultat i By, hvilket tyder på at effekten av *personale* på romprisene kan variere mellom ulike regioner i Norge. Resultatene for en Oslo-modell kommer frem i Tabell A7.2. Et viktig poeng i den sammenheng, er at våre resultater for en Oslo-modell (9,8 %) er forskjellig fra Thrane (2007) sine funn. Våre resultater samsvarer imidlertid med våre forventninger, da de stemmer overens med et økonomisk rasjonale om at bedre personale kan bidra til høyere hotellrompriser.

Vi kan også vise til studier av Castro og Ferreira (2018) og Zhang et al. (2011), som har undersøkt effekten av *personale* på romprisene i henholdsvis Lisboa og New York. Disse finner ingen signifikant effekt av *personale* på romprisene i de nevnte byene. Det er imidlertid verdt å merke seg at hotellkonsentrasjonen er større i Lisboa og New York, sammenlignet med By og DN. Resultatene fra Modell 1 viser en signifikant effekt av *personale* i By ($p < 0,05$) og DN ($p < 0,01$), mens de tidligere nevnte studiene ikke fant en slik signifikant effekt. Dette indikerer at når hotellkonsentrasjonen øker i et geografisk område, vil det ha en mindre betydelig og signifikant effekt på betalingsviljen knyttet til personalet.

7.2.3.6 Renhold

Renhold, som inkluderer renslighet av rom og fellesområder, viser interessant nok en negativ og signifikant effekt på hotellromprisen for både By og DN. En økning på 1,0 for *renhold* har en negativ effekt på -23,6 % ($p < 0,1$) i By, og en negativ effekt på -16,4 % i DN ($p < 0,05$). Dette er et uventet signifikant fortegn for *renhold*, som kan svekke tilliten til modellen. Resultatene kan imidlertid ses i sammenheng med Borges et al. (2015), som også finner at *renhold* påvirker hotellromprisen med -0,21 %. Borges et al. (2015) valgte å ekskludere *renhold* som en variabel i sin undersøkelse, mens vi valgte å inkludere den selv om variabelen korrelerer med flere av de andre variablene. Eksklusjon ville redusert VIF-verdien til de andre variablene, som vist i

Tabell A7.7. Vi mener likevel at *renhold* er en viktig variabel for å besvare vårt forskningsspørsmål, hvilket taler for å inkludere den i modellen.

Videre finner Zhang et al. (2011) og Castro & Ferreira (2018) ingen signifikant sammenheng mellom *renhold* og hotellromspriser, i henholdsvis New York og Lisboa. Zhang et al. (2011) forklarer dette med at renholdet er forholdsvis standardisert, og at de derfor ikke finner signifikante resultater. Dette gjør funnene fra det norske markedet overraskende.

En mulig årsak til de negative koeffisientene kan være at andre faktorer som korrelerer med renhold, har en større påvirkning på hotellromsprisene. For eksempel kan beliggenhet, fasiliteter og komfort være sterke drivere for romprisene, og renholds nivået kan ha en indirekte påvirkning gjennom disse faktorene.

7.2.3.7 WiFi

Kundevurderingene for tilgjengeligheten og kvaliteten på trådløst internett (*WiFi*) finner vi ikke signifikant for rompriser, verken for By eller DN. Samtidig er standardavviket lavt i forhold til de andre testede variablene. Dette tyder på at tilgang til WiFi har blitt en standardfunksjon på hoteller, og derfor ikke påvirker romprisene.

Selv om vår analyse ikke viser en signifikant effekt av *WiFi* på romprisene i det norske hotellmarkedet, har tidligere studier funnet betydelige effekter i andre markeder. For eksempel finner Lorde et al. (2019) og Chen og Rotschild (2010) signifikante sammenhenger mellom *WiFi* og hotellrompriser, i henholdsvis Airbnb-markedet i Karibia og hotellmarkedet i Taipei. Dette indikerer at det norske hotellmarkedet skiller seg fra disse markedene. Tilgjengeligheten av trådløst Internett kan potensielt ha en signifikant effekt på romprisene i spesifikke regioner eller kundesegmenter i Norge. Vi har imidlertid ikke funnet en slik effekt i de to dimensjonene vi har testet i Modell 1. En alternativ forklaring på forskjellene mellom Norge, Karibia og Taipei kan være tilgjengeligheten av mobildata, som kan redusere behovet for WiFi i Norge.

7.2.3.8 Bærekraft

Vi finner ingen signifikant effekt av bærekraftsertifiseringsnivå på hotellromprisene i det norske markedet, verken i By eller i DN. Dette er i tråd med lignende studier utført i Spania av Cavero-Rubio og Amorós-Martínez (2008) og i London av Geerts (2014). Ingen av dem

observerer prisvariasjoner mellom bærekraftsertifiserte hoteller og ikke-sertifiserte hoteller. Geerts (2014) forklarer dette med at bærekraftsertifiserte hoteller ikke markedsfører seg godt nok som bærekraftige. Samtidig var pris, beliggenhet og servicekvalitet viktigere kriterier enn bærekraft for valg av hotell. Vi kan trekke paralleller fra Geerts (2014) sine funn til det norske markedet, der kundene heller ikke virker interessert i hotellenes bærekraftsertifisering.

Våre funn fra det norske markedet skiller seg imidlertid fra det sørspanske markedet. Sánchez-Ollero et al. (2014) finner at implementering av bærekraftige tiltak kan føre til økte rompriser (5,15 %). Denne studien er imidlertid begrenset til to dimensjoner; By og DN, noe som kan påvirke overførbarheten av resultatene. Det kan derfor hende at betydningen av bærekraft for hotellrompriser kan variere i enkelte regioner i Norge også, men vi finner ingen slike signifikante effekter i verken By eller DN i Modell 1.

7.2.4 Oppsummering av Modell 1

Vi vil nå kort oppsummere funnene fra Modell 1 før vi ser på Modell 2, hvor vi deler markedet inn etter segmenter. Hovedproblemet med Modell 1 er at 94 % av observasjonene tilhører segment tre (45 %) og fire (49 %), noe som fører til en skjevhet i resultatene mot disse to segmentene. Til tross for dette gir Modell 1 flere nevneverdige funn.

Vi finner en positiv og signifikant sammenheng mellom *segmentnivå* og rompriser, med en koeffisient på 17,2 % i By og 11,9 % i DN. Dette er i tråd med tidligere forskning (Israeli, 2002; Soler et al., 2019; Zhang et al., 2011). Videre observerer vi en signifikant sammenheng mellom *beliggenhet* og hotellrompriser, med koeffisienter på 11,6 % i By og 6,7 % i DN. Dette er også i tråd med tidligere forskning (Arbel & Pizam, 1977; Bull, 1994; White & Mulligan, 2010).

Våre funn skiller seg imidlertid fra tidligere studier (Andersson, 2010; Bull, 1994; Espinet et al., 2003; Thrane, 2007) ved at vi ikke observerer en signifikant sammenheng mellom *fasiliteter* og hotellrompriser i det norske markedet. Dette skyldes antagelig begrenset variasjon blant hotellfasilitetene i vårt datasett.

Vi finner også at *komfort* påvirker hotellromprisene med 32,3 % i By og 15,4 % i DN. Dette er i tråd med tidligere studier (Castro & Ferreira, 2018; Stringam et al., 2010). Videre observerer vi en positiv effekt av *personale* på hotellromprisene, med en økning på 8,9 % i begge markeder.

Dette skiller seg noe fra tidligere studier av større byer (Castro & Ferreira, 2018; Thrane, 2007; Zhang et al., 2011).

Overraskende nok, finner vi en negativ effekt av *renhold* på hotellromprisene, med en nedgang på henholdsvis 23,6 % i By og 16,4 % i DN. Dette avviker fra tidligere studier (Castro & Ferreira, 2018; Zhang et al., 2011).

Videre viser våre funn ingen signifikant effekt av *WiFi* på hotellromprisene, noe som er i tråd med våre forventninger. Til slutt finner vi ingen betalingsvillighet for *bærekraft* i det norske hotellmarkedet, noe som samsvarer med funn fra London (Geerts, 2014) og skiller seg fra funn i det sørs spanske markedet (Cavero-Rubio & Amorós-Martínez, 2008). Disse funnene bidrar til vår forståelse av bærekraftens betydning for hotellromprisene i det norske markedet.

7.3 Modell 2

Et av hovedproblemene ved Modell 1 er dens skjevhet mot segment tre og fire, der 94 % av våre observasjoner faller inn under disse segmentene. Samtidig finner vi forskjellige koeffisienter mellom segmentene når vi kjører regresjoner for hvert segment. Vi har derfor konstruert Modell 2, som bidrar til å belyse hvordan kundevurderinger og bærekraftsertifiseringer påvirker hotellromprisene innad i hvert segment. Dette vil bidra til forskningen på segmentforskjeller. En slik modell har blitt benyttet i flere tidligere studier, som Borges et al. (2015) Castro og Ferreira (2018) og Zhang et al. (2011), men ikke i Norge. Vår studie skiller seg imidlertid fra disse ved at vi inkluderer bærekraftvariabelen og en by/land-dimensjon.

Skjevheten i segmentene er også en utfordring med Modell 2. Vi har få observasjoner og dermed frihetsgrader i noen av segmentene, noe som hindrer oss i å estimere en fullstendig regresjonslinje for disse segmentene. Dette gjelder segment to og fem i By, samt segment fem i DN. Til tross for dette anser vi det som hensiktsmessig å benytte både den geografiske og den segmentale inndelingen, da dette gir det mest representative bildet for 97 % av utvalget, selv om vi ikke klarer å estimere regresjonslinjene for segment to og fem i By, samt segment fem i DN.

Som tidligere forklart i Modell 1, er observasjonene preget av store prisforskjeller mellom byene. Derfor inkluderer vi dummyvariabler for hver av byene i By. I den sammenheng er Bodø

referansebyen, slik at den er lik null og dermed ikke inkludert i regresjonstabellen for Modell 2. Resultatene fra Modell 2 fremkommer av Tabell 7.3 nedenfor. Vi har også vedlagt en tabell med variablenes standardavvik i Tabell A7.11.

Tabell 7.3: Regresjonstabell for Modell 2.

Segmentnivå	Avhengig variabel: Ln(Pris)							
	By				DN			
	2	3	4	5	2	3	4	5
Beliggenhet	(0,35)	0,033	0,146***	(280,369)	(0,081)	0,091***	0,032	(5,539)
Fasiliteter	-	(0,017)	0,113	(86,322)	(0,130)	(0,032)	0,225**	(173,20)
Komfort	-	0,778**	0,073	591,683	0,690**	0,259**	0,014	204,532
Personale	-	0,120*	0,083	358,359	0,184**	0,094**	0,066	14,164
Renhold	-	(0,561)*	(0,047)	(607,870)	(0,575)**	(0,205)*	(0,097)	-
WiFi	-	0,016	(0,014)	(158,766)	(0,153)***	(0,009)	(0,019)	-
Bærekraft	-	0,020	0,036	14,499	(0,030)	(0,0001)	0,047*	-
Oslo	-	0,190	0,259***	(69,056)				
Bergen	-	(0,040)	0,115	18,569				
Trondheim	-	0,227	0,189**	46,696				
Stavanger	-	0,165	0,375***					
Gardermoen	-	0,145	0,165*					
Tromsø	-	0,133	0,222**					
Kristiansand	-	0,008	0,090					
Konstant	10,13	4,003***	4,156***	1.730,495	7,584***	5,612***	5,616***	(402,55)
Observasjoner	2	60	128	11	19	247	208	5
R²	1,00	0,470	0,490	1,000	0,698	0,190	0,188	1,000
Residual SE		0,186 (df = 45)	0,165 (df = 113)		0,136 (df = 11)	0,238 (df = 239)	0,251 (df = 200)	
F-statistikk		2,851*** (df=14;45)	7,743*** (df=14;113)		3,625** (df=7;11)	8,007*** (df=7;239)	6,594*** (df=7;200)	
Note:	*p<0,1; **p<0,05; ***p<0,01; negative verdier i parentes.							

7.3.1 Residualer

I Modell 2 ser vi flere utfordringer knyttet til få observasjoner for noen segmenter (manglende frihetsgrader), samt at segment tre og fire dominerer modellen. Dette gjør at vi følgelig vil undersøke og diskutere residualene for hvert segment i By og DN, før vi presenterer og analyserer resultatene. I Appendikset har vi inkludert boksplokk i Figur A7.2 og Figur A7.3, som viser residualene for henholdsvis By og DN. Boksplokkene bidrar til å visualisere denne diskusjonen. Vi har også inkludert tabeller med statistikk over residualene i hvert segment i Tabell A7.4 og Tabell A7.5.

7.3.1.1 Residualer for By-segmentene

Vi vil først se på residualene til By-segmentene, som presentert i Figur A7.2 og Tabell A7.4. Segment to og fem har for få observasjoner til å analysere deres residualer. For segment tre og fire er residualene jevnt fordelt rundt null, med unntak av en ekstremt negativ avvikende verdi i segment tre, og en ekstremt positiv avvikende verdi i segment fire.

Residualene for segment tre har en median på 0,046, og halvparten av residualene befinner seg mellom -0,092 og 0,119. De er relativt små og sentrert rundt null, hvilket styrker modellen. Residualenes standardfeil på 18,59 %, svekker modellen grunnet noe spredning. Modellen kan likevel forklare 81,41 % av prisvariasjonen. F -statistikken på 2,851 ($p < 0,01$) styrker på sin side modellen, siden den er bedre enn en nullmodell. Variasjonen (R^2) som blir forklart av attributtene er på 47 % for segment tre i denne dimensjonen, noe vi anser som generelt god.

I segment fire har residualene en median på -0,007, og halvparten av residualene ligger mellom -0,084 og 0,093. De er dermed små og sentrert rundt null. Dette tyder på at modellen gir presise estimater for de fleste observasjonene. Residualstandardfeilen er på 16,49 %, som betyr at residualene er litt mindre spredt i segment tre. Modellen styrkes videre av en F -statistikk på 7,743 ($p < 0,01$), hvilket er en bedre tilpasning enn en nullmodell. R^2 er 49 % for segment fire i By, som vi mener forklarer en betydelig del av prisvariasjon.

Resultatene viser dermed at Modell 2 er godt tilpasset dataene i segment tre og fire i By. Variablene i modellen kan forklare majoriteten av prisvariasjonen i segment tre og fire, og disse segmentene har også størst betydning i dimensjonen By.

7.3.1.2 Residualer for DN-segmentene

For DN-segmentene (se Figur A7.3 og Tabell A7.5), har vi for få observasjoner til å analysere residualene for segment fem. Vi vil derfor fokusere på residualene for segment to, tre og fire. Vi ser også fra Figur A7.3 at segment tre og fire har noen ekstreme observasjoner. Disse kan svekke modellens presisjon.

Residualene i segment to har en median på 0,00044. Residualene er små og konsentrert rundt null, siden halvparten ligger mellom -0,080 og 0,055. Vi får dermed presise estimater for majoriteten av observasjonene. Residualenes standardfeil er 13,55 % med 11 frihetsgrader. Dette er relativt få frihetsgrader, hvilket svekker modellen. Standardfeilen viser imidlertid at 86,45 % av prisvariasjonen kan forklares av koeffisientene. Modellen styrkes av en signifikant F -statistikk på 3,625 ($p < 0,05$). Prisvariasjonen som kan forklares av attributtene (R^2) er 70 % for segment to. Modellen forklarer dermed en stor andel av prisvariasjonen.

For segment tre er medianen av residualene -0,004. Halvparten av residualene befinner seg mellom -0,140 og 0,155, og er dermed små og konsentrert rundt null. Modellen gir dermed presise estimater. Standardfeilen er 23,79 %, slik at 76,21 % av variasjonen i $\ln(\text{pris})$ kan forklares med koeffisientene. Den signifikante F -statistikken på 8,007 ($p < 0,01$) er også med på å styrke modellen, som dermed er bedre enn en nullmodell. Vi ser at R^2 er 19 % for segment tre. Attributtene har derfor en begrenset forklaringskraft for variasjonen i den avhengige variabelen.

I segment fire har residualene en median på -0,026, hvor halvparten av residualene ligger mellom -0,175 og 0,116. Residualene er også små i segment fire, hvilket styrker modellens presisjon. Residualenes standardfeil er 25,11 %, som betyr at koeffisientene kan forklare 84,89 % av prisvariasjonen. Modellen styrkes også av en signifikant F -statistikk på 6,594 ($p < 0,01$). Vi ser igjen at modellen har en bedre tilpasning enn en nullmodell. Imidlertid er R^2 på 19 %, slik at de predikerende variablene også her har en begrenset forklaringskraft for prisvariasjonen.

Vi kan oppsummere med at Modell 2 er godt tilpasset dataene i segment to, tre og fire i DN, med forbehold om at segment to har få frihetsgrader. Dette må vi ta hensyn til i analysen på segmentnivå. Segment tre og fire har også noen ekstreme observasjoner som kan svekke modellens presisjon og validitet. Vi ser imidlertid at segment tre og fire vil ha størst betydning

for DN i Modell 2. Dette fordi residualene tyder på at de vil gi presise estimater for de fleste observasjonene, og fordi de har en bedre tilpasning enn en nullmodell.

7.3.2 VIF-analyse

VIF-verdiene for Modell 2 kommer frem i Tabell A7.9 og Tabell A7.10. Det var ikke mulig å estimere VIF-verdier for deler av attributtene til segment to i By og segment fem i DN, grunnet få observasjoner.

I By er VIF-verdiene for *beliggenhet*, *fasiliteter*, og *komfort* høye, hvor de fleste verdier i segment tre, fire og fem befinner seg mellom 10 og 30. *Fasiliteter* i segment fem har den høyeste verdien på 99,5. Videre har *personale* verdiene 1,89 og 1,83 i segment tre og fire, mens den er på 15,88 i segment fem. *Renhold*, *WiFi* og *bærekraft* har lave verdier i samtlige segment. *Renhold* har en verdi på henholdsvis 1,83, 1,62 og 3,25 i segment tre, fire og fem. *WiFi* ligger på 1,28, 1,28 og 1,71 for de samme segmentene, og *bærekraft* ligger på 1,10, 1,06 og 1,63. Vi ser dermed at VIF-verdiene for By på segmentnivå følger samme mønster som By i Modell 1. Dette tyder på at vi også har multikollinearitet innenfor hvert segment i By. Særlig for *beliggenhet*, *fasiliteter* og *komfort*.

I DN følger VIF-verdiene en lignende trend som i By. *Beliggenhet*, *fasiliteter* og *komfort* har høye VIF-verdier i segment to, tre og fire. *Beliggenhet* på 27,40, 22,64 og 10,35; *fasiliteter* på 21,73, 17,57 og 10,18; *komfort* på 17,36, 11,99 og 7,42 i segment to, tre og fire. Videre er VIF-verdiene for disse segmentene lave for *personale*, *renhold*, *WiFi* og *bærekraft*; 2,14, 2,22 og 2,35 for *personale*; 1,73, 1,56 og 1,38 for *renhold*; 1,34, 1,26 og 1,28 for *WiFi*; 1,18, 1,03 og 1,05 for *bærekraft*. I segment fem mangler vi VIF for *beliggenhet*, *komfort* og *renhold*, grunnet få observasjoner. I dette segmentet er *fasiliteter* på 6,57, *personale* på 6,23, *WiFi* på 1,39 og *bærekraft* på 1,00.

I sum, så er VIF-verdiene for Modell 2 høye for flere av de uavhengige variablene. Hvorvidt disse høye VIF-verdiene er et problem for nøyaktigheten til koeffisientene, samt stabiliteten til resultatene i våre regresjonsmodeller, har vi diskutert under Modell 1 i avsnitt 7.2.2. Vår konklusjon om at de ikke legger en sterk begrensning på nøyaktigheten til koeffisientene og stabiliteten til resultatene, holder fortsatt. Dette fordi signifikante regresjonskoeffisienter fortsatt kan forklare hvordan attributtene påvirker rompriser, selv med betydelig grad av multikollinearitet.

7.3.3 Attributtenes effekter

Når vi skiller segmentene, ser vi at det norske hotellmarkedet har sprikende koeffisienter avhengig av segment og region. Dette aspektet gjør det norske hotellmarkedet enda mer komplekst. Det er imidlertid ikke overraskende at koeffisientene avviker mellom segmentene og dimensjonene. Zhang et al. (2011) finner betydelige forskjeller mellom koeffisientene i deres segmentanalyse, mens Israeli (2002) finner betydelige forskjeller mellom de ulike regionene han undersøkte.

I denne delen av oppgaven vil vi undersøke i hvilken grad variablene påvirker prisvariabelen innen hvert segment. Herunder vil vi igjen drøfte våre resultater opp mot tidligere studier, der vi påpeker likheter og forskjeller. Dette vil bidra til en dypere forståelse av hvordan konsumentenes betalingsvilje varierer mellom by og distrikt, samt mellom ulike segmenter. En utfordring med Modell 2 er at vi har et begrenset antall observasjoner i segment to i By og segment fem i DN, noe som medfører en mangel på frihetsgrader. Vi klarer dermed ikke å estimere en fullstendig regresjonslinje for disse segmentene, dog får vi en fullgod regresjon for 97% av vårt utvalg. Videre vil vi påpeke at det er forventet å observere sprikende funn mellom segmentene. Dette er i tråd med Zhang et al. (2011), som også finner relativt sprikende funn mellom segmentene.

7.3.3.1 Beliggenhet

I byområdene har *beliggenhet* kun en signifikant effekt på betalingsviljen i segment fire, der en økning på 1,0 i tilbakemeldingene på beliggenhet fører til en økning i romprisen på 14,6 % ($p < 0,01$). I segment tre og fem har *beliggenhet* ingen signifikant effekt på pris, noe som antyder at kundene i disse segmentene verdsetter beliggenheten i mindre grad.

I distriktsområdene har *beliggenhet* kun signifikant effekt i segment tre, der romprisen øker med 8,5 % ($p < 0,01$) når tilbakemeldingene øker med 1,0. I likhet med byområdene, ser vi at *beliggenhet* kun har en signifikant effekt på romprisen i ett segment. Beliggenhet har dermed liten betydning i segment to, fire og fem i DN, mens for hoteller i segment tre virker en god beliggenhet viktig for å påvirke kundenes betalingsvilje. Dette kan skyldes at kundene har et begrenset hotellutvalg eller ingen valgalternativer i distriktsområdene. Eksempelvis når man reiser fra Oslo til Bergen og trenger en overnatting, så kan det kun være ett reelt hotellalternativ. Derfor er antagelig beliggenhet av liten betydning for gjestene i DN, foruten om segment tre. At vi heller ikke finner *beliggenhet* signifikant i segment fem for DN, kan skyldes at hotellene

i dette segmentet er i seg selv en destinasjon. Dermed er beliggenhet av mindre betydning. Vi kan for eksempel se på Refsnes Gods (Hotell Refsnes Gods, 2023) som et slikt eksempel. Funnet er også i tråd med Castro og Ferreira (2018), som ikke finner *beliggenhet* signifikant i segment fem.

I likhet med våre funn, finner Zhang et al. (2011) at beliggenhet har størst effekt på kundenes betalingsvilje i mellomsegmentet, med en påvirkning på 31,5 %. Zhang et al. (2011) finner heller ingen signifikant effekt av beliggenhet i lavprissegmentet, som kan sammenlignes med segment to i våre modeller. Dette tyder på at beliggenhet er en viktig del av hotellromprisen for hotellene i midtre segmentnivå, og at når hotellkonsentrasjonen øker – New York har større hotellkonsentrasjon enn våre dimensjoner – så påvirkes hotellromprisen i større grad av beliggenhet.

7.3.3.2 Fasiliteter

Fasiliteter påvirker prisene ulikt, avhengig av marked og segmentering. For byhotellene finner vi ingen signifikant sammenheng mellom *fasiliteter* og rompris, uavhengig av segment. Derimot finner vi en signifikant og positiv effekt i segment fire i DN, med en koeffisient på 22,5 % ($p < 0,05$). Ingen av de andre segmentene i DN viser en signifikant effekt.

Vårt funn om at *fasiliteter* har en signifikant og positiv effekt i segment fire i DN, er i tråd med studiene til Thrane (2007), Bull (1994) og Castro og Ferreira (2018). Alle disse finner en positiv og signifikant sammenheng mellom fasiliteter og rompris. Det er imidlertid kun Castro og Ferreira (2018) som har undersøkt denne sammenhengen på segmentnivå. De finner en signifikant effekt av fasiliteter i segment fire (33,7 %) og segment tre (9,2 %) i Lisboa. Vårt funn for segment fire i DN kan derfor sammenlignes med deres funn. Et eksempel på dette er Son Spa, som etter våre funn kan ta en høyere rompris på grunn av deres ekstra spa-fasiliteter. For kundene kan dette være en egenskap som tiltrekker dem til hotellet og dermed øker deres betalingsvilje, noe som gjør at denne variabelen da blir viktigere enn den fysiske beliggenheten til hotellet.

At fasiliteter ikke har en signifikant effekt på romprisene i By, utfordrer mye av den tidligere presenterte litteraturen (Castro & Ferreira, 2018; Thrane, 2007). Dette tyder på at steder med høy hotellkonsentrasjon i Norge skiller seg fra lignende markeder i andre land. Thrane (2007) sin studie av Oslo-markedet viser at parkeringsfasiliteter påvirker romprisene. Dermed kan det

norske markedet ha endret seg de siste femten årene, da våre resultater ikke viser en signifikant effekt, og heller ikke i Oslo (se Tabell A7.2 for en regresjonsmodell for regionen Oslo). Antagelig er dette fordi kollektivtilbudet i Oslo har blitt bedre siden 2007, og det har vært en villet politikk å begrense bilbruken i byen (Oslo kommune, 2023). Dette kan ha hatt en smitteeffekt også på hotellgjesters bilbruk, når det legges mindre til rette for privatbilisme og mer til rette for sykling og kollektivtransport. Dette funnet kan vi generalisere til resten av byene og kan være et argument for at *fasiliteter* ikke har en betydning for romprisene.

7.3.3.3 Komfort

For *komfort* finner vi en signifikant positiv effekt ($p < 0,05$) på pris i segment tre i bydimensjonen, med en koeffisient på 77,8 %. På den annen side finner vi ingen signifikante sammenhenger mellom *komfort* og rompris for hotellene i de andre segmentene i By.

For DN finner vi signifikante sammenhenger for hotellene i segment to og tre, men med betydelige forskjeller i påvirkningsgrad. En økning i score med 1,0 vil påvirke romprisen med 69 % ($p < 0,05$) i segment to og 25,9 % ($p < 0,05$) i segment tre. I segment fire finner vi ingen signifikant sammenheng. Standardavvikene for *komfort* er også relativt høye i segment to (19,2 %) og segment tre (11,5 %) i DN, noe som indikerer at også innenfor samme segment opplever kundene komforten ulikt. De høye koeffisientene våre kan skyldes at komfortvariabelen fanger opp ulike underkategorier, eksempelvis sengestørrelse, støy og atmosfære.

Våre signifikante funn er ikke overraskende fra et teoretisk og empirisk perspektiv. Komfort fremstår som en viktig egenskap ved valg av hotellovernatting, både intuitivt og i litteraturen. Våre høye koeffisienter underbygges av Stringam et al. (2010). I deres studie av de 100 største byene i USA, finner de at komfort er den faktoren med størst innflytelse på generell tilfredshet blant kundene, noe som kan forklare våre høye koeffisienter.

At vi finner flere ikke-signifikante verdier støttes av Castro og Ferreira (2018), som ikke finner noen signifikant sammenheng mellom komfort og pris i Lisboa-markedet. De tester imidlertid ikke variabelen på segmentnivå. For hotellene som ikke har signifikant effekt, ser det altså ut til at komforten oppleves som lik innenfor samme segment.

7.3.3.4 Personale

Gjestenes vurderinger av *personale* har en positiv effekt på 12,0 % ($p < 0,1$) i segment tre i byene. For de øvrige segmentene i By kan vi ikke fastslå en sammenheng, grunnet ikke-signifikante resultater.

I DN har *personale* en positiv og signifikant effekt på rompris i segment to, med en koeffisient på 18,4 %, og i segment tre med en koeffisient på 9,4 %. Standardavvikene for *personale* er lave i segment to (8,1 %) og tre (3,7 %) i DN, hvilket betyr at variasjonen i tilbakemeldingene på *personale* er liten i disse segmentene. Dermed virker det som om kundene i disse segmentene har en relativt lik opplevelse av personalet.

Funnet vårt om at *personale* har en positiv effekt på rompris i segment to og tre i DN, er informativt. Dette fordi vi ikke har identifisert lignende effekter i sammenlignbare studier fra andre regioner (Castro & Ferreira, 2018; Zhang et al., 2011). Fra et teoretisk perspektiv er det imidlertid ikke overraskende, da personalet normalt sett forventes å ha betydning for kundenes hotellopplevelse.

På sin side finner Zhang et al. (2011) at *personale* har en signifikant effekt (40 %) på romprisen i luksussegmentet (segment 4, 4,5 og 5) i New York. Vi finner imidlertid ingen signifikante verdier i disse segmentene, verken i By eller DN. At vi ikke finner signifikante verdier i de øvre segmentene, kan tyde på at variasjonen i *personale* er større i New York sammenlignet med Norge. Dette kan være årsaken til at de finner signifikante verdier i disse segmentene. I Norge kan personalets ferdigheter være så gode at personalet i liten grad skiller hotellene i de øvre segmentene.

7.3.3.5 Renhold

Variabelen *renhold* finner vi kun signifikant for segment tre i By, med en koeffisient på -56,1 %. For DN finner vi at en økning i score med 1,0, påvirker romprisen med -57,5 % ($p < 0,01$) i segment to og -20,5 % ($p < 0,1$) i segment tre. At resultatene har et negativt signifikant fortegn, bidrar til å svekke tilliten til modellen. Ut fra et økonomisk rasjonale skulle en økning i renhold ha en positiv effekt på romprisene, men våre funn viser det motsatte.

Vi kan imidlertid knytte flere argumenter til at hoteller i noen segmenter, som prioriterer høy renholdskvalitet, har lavere rompris. Det kan for eksempel være en strategisk beslutning for å

bygge et omdømme for renslighet, og dermed tiltrekke seg kunder som verdsetter dette aspektet sterkt. Vi diskuterte også i avsnitt 7.2.4.6 (Renhold – Modell 1) at det er mulig at andre variabler som korrelerer med renhold har større effekt på romspriser, og derfor får vi negative koeffisienter for *renhold*.

Våre ikke-signifikante verdier i de resterende segmentene støttes av Zhang et al. (2011) og Castro og Ferreira (2018), som heller ikke finner signifikante verdier i noen av segmentene. Zhang et al. (2011) forklarer dette med at renholds nivået er så høyt i markedet i New York at variasjonen mellom hotellene i liten grad påvirker romprisene. Vi kan argumentere for at det samme gjelder for segmentene med ikke-signifikante verdier i By og DN.

Som nevnt under i Modell 1 (avsnitt 7.2.3.5 Renhold) finner Borges et al. (2015) variabelen negativ og signifikant, og valgte å ekskludere den fra analysen. Imidlertid har vi valgt å beholde denne variabelen. Vi finner det mistenkelig at hoteller som prioriterer renhold får lavere pris for en overnatting. Renhold anses som et grunnleggende egenskap ved et hotell, og derfor sår vi også tvil om disse resultatene.

7.3.3.6 WiFi

Når vi ser på variabelen *WiFi* på segmentnivå, så finner vi at *WiFi* har en negativ og signifikant effekt på romprisen i segment to i DN (-15,3 %, $p < 0,01$). I de andre segmentene finner vi ingen signifikant effekt av *WiFi*, verken i By eller DN. Det negative fortegnet på koeffisienten i segment to i DN er derimot et overraskende funn, og kan bidra til å svekke tilliten til modellen. Dette fordi en intuitivt ville forvente at kundene var villig til å betale mer for god WiFi.

Våre funn kan forklares med at kundene i By og DN sannsynligvis i stor grad benytter seg av mobildata, og dermed ikke vektlegger behovet for rask og stabil trådløs Internett-tilkobling. Det kan også skyldes at tilgangen til Internett er standardisert blant hotellene, slik at det i liten grad differensierer dem. Funnet støttes av Soler et al. (2019), som finner at WiFi ikke har en signifikant effekt på romprisene hos hotellene langs Algarvekysten. Dette tyder på at kundene i dette markedet ikke vektlegger tilgang til WiFi ved bestilling av hotellrom, uten at resultatet blir nærmere begrunnet. Dette indikerer at våre funn kan sammenlignes med andre markeder.

Vårt funn om at *WiFi* har en negativ og signifikant effekt på -15,3 % ($p < 0,01$) i segment to i DN, indikerer at kundegruppen i dette segmentet er mer sensitiv for *WiFi* enn kundegruppene i

de resterende segmentene. Dette funnet er delvis i tråd med studiene til Schamel (2012) og Lorde et al. (2019), som begge finner at tilgang til Internett har en negativ effekt på kundenes betalingsvilje. Vår studie er imidlertid mer nyansert, da vi deler hotellene inn etter segment, mens Schamel (2012) og Lorde et al. (2019) aggregerer alle segment i en modell. Resultatene deres er likevel signifikante. Lorde et al. (2019) forklarer funnet sitt med at mange aktører inkluderer WiFi (96,9 %), og derfor reduserer prisen for å differensiere seg fra konkurrentene. Dette kan også gjelde for de norske distriktshotellene i segment to.

7.3.3.7 Bærekraft

I Modell 2 viser resultatene at hotellenes bærekraftsertifiseringer ikke har en signifikant effekt på romprisen i noen av segmentene i By. I DN finner vi en svak, men signifikant positiv effekt på 4,7 % ($p < 0,1$) i segment fire, men ingen signifikant effekt i de andre segmentene. Dette antyder at mer bemidlede kunder i distriktene verdsetter hotellenes bærekraftsertifisering, sammenlignet med gjestene i lavere segment.

Det er imidlertid ikke overraskende at vi ikke finner *bærekraft* signifikant for majoriteten av segmentene i By og DN. Lignende studier av Cavero-Rubio og Amorós-Martínez (2008) og Geerts (2014) finner ingen prisforskjeller mellom bærekraftsertifiserte og ikke-sertifiserte hoteller. De forklarer dette med at pris, beliggenhet og servicekvalitet er viktigere kriterier for kundene ved valg av overnattingsted. Vi kan derfor igjen trekke derfor paralleller til denne studien for de ikke-signifikante verdiene på segmentnivå, og støtter oss til at kundene prioriterer andre faktorer fremfor bærekraft.

At vi fant en betydning av bærekraftsertifisering i segment fire for DN, er imidlertid delvis i tråd med funnene til Kuminoff et al. (2010) og Sánchez-Ollero et al. (2014). De fant en positiv og signifikant betalingsvilje for «grønne» hotellrom i deres studier fra Virginia og Andalusia. Dette kan tyde på at kundene prioriterer bærekraftsertifiseringer i spesifikke regioner i Norge.

7.3.4 Oppsummering av Modell 2

Modell 2 gir oss innsikt i hvordan forbrukernes betalingsvilje for hvert attributt fordeler seg på segmentnivå, og avdekker klare forskjeller mellom By og DN, samt mellom ulike segmenter. Dette tyder på at det norske markedet ikke kan oppfattes som ett marked, men heller som flere ulike markeder.

En av hovedutfordringene i Modell 2 er estimering av noen segmenter, grunnet et begrenset antall observasjoner i disse segmentene, som dermed begrenser modellens frihetsgrader. Derfor har mye av diskusjonen vært knyttet til segment tre og fire.

På grunn av for få observasjoner er det ikke mulig å si noe fornuftig om segment to i By-dimensjonen. I segment tre finner vi overraskende negative koeffisienter for *renhold* (-56,1 %) og *komfort* (77,8 %). Resten av variablene hadde ingen signifikant påvirkning. I segment fire finner vi kun en sammenheng mellom rompris og *beliggenhet* på 14,6 %. I segment fem har vi heller ikke grunnlag for å trekke meningsfulle konklusjoner.

I DN finner vi i segment to signifikante koeffisienter for *komfort* (69 %), *personale* (18,4 %), *renhold* (-57,5 %) og *WiFi* (-15,3 %). I segment tre finner vi *beliggenhet* (9,1 %), *komfort* (25,9 %), *personale* (9,4 %) og *renhold* (-20,5 %) signifikante, mens i segment fire finner vi *fasiliteter* (22,5 %) og *bærekraft* (4,7 %) signifikante. I segment fem har vi for få observasjoner til å kunne trekke slutninger.

Modell 2 avdekker flere uventede funn. De negative koeffisientene for *renhold* i både By-segment tre (-56,1 %) og DN-segment to (-57,5 %) og tre (-20,5 %), er slike funn. Disse uventede resultatene strider med det etablerte økonomiske rasjonale om at bedre renhold vanligvis fører til høyere priser. Det avvikende mønsteret tvinger oss til å betvile modellen noe. På en annen side, finner for eksempel Borges et al. (2015) også en negativ koeffisient for renhold i sin undersøkelse.

Det er interessant å observere at bærekraft kun ser ut til å ha en betydelig effekt i segment fire i DN. Dette funnet antyder at godt bemidlede hotellgjester i distriktene er mer engasjerte i bærekraftige faktorer sammenlignet med andre segmenter, og at gjestene i de andre segmentene ikke vektlegger dette ved valg av hotell.

Et funn som stemmer overens med et økonomisk rasjonale, er at vi finner positive og signifikante verdier for *personale* i segment tre i DN, samt segment to og tre i By. Dette har vi ikke funnet i tidligere forskning (Castro & Ferreira, 2018; Thrane, 2007; Zhang et al., 2011).

Modell 2 gjør det tydelig at det norske hotellmarkedet skiller seg fra andre markeder. Imidlertid har vi ikke funnet studier som er direkte sammenlignbare, da de ikke inkluderer alle variablene

og den spesifikke dimensjonsinndelingen vi har benyttet (Castro & Ferreira, 2018; Thrane, 2007; White & Mulligan, 2010; Zhang et al., 2011).

Vi har også diskutert betydningen av ikke-signifikante koeffisienter. Vi betrakter disse som standardiserte innenfor hvert segment, og blir ansett som et grunnleggende egenskap av gjestene. Et eksempel på dette er *renhold*, som vi ikke finner signifikante koeffisienter for i flere av segmentene. Hotellene kan ikke velge å kutte ut renholdet, men nivået er såpass høyt at det i liten grad skiller hotellene fra hverandre.

Oppsummert viser resultatene fra Modell 2 at betydningen av de inkluderte attributtene varierer, avhengig av hotellsegment og dimensjonstilhørighet. Funnene våre støttes delvis av tidligere hedoniske prisingsanalyser av hotellrom, men vi observerer også klare forskjeller mellom markedene (Castro & Ferreira, 2018; Thrane, 2007; White & Mulligan, 2010; Zhang et al., 2011). Modell 2 styrker dermed påstanden om at det norske hotellmarkedet er preget av komplekse forhold mellom ulike faktorer som påvirker romprisen. Til slutt tolker vi modellen som at det er utfordrende å betrakte hele det norske hotellmarkedet som ett, da resultatene viser betydelige variasjoner mellom segmentene og de to dimensjonene.

8 Konklusjon og fremtidig arbeid

I dette avsluttende kapittelet vil vi se på de viktigste funnene og konkludere forskningsspørsmålet. Deretter vil vi diskutere oppgavens begrensninger og gyldighet, før vi kommer med forslag til videre arbeid.

8.1 Hovedfunn

Formålet med oppgaven har vært å undersøke «*Hvordan påvirkes hotellromprisene i det norske markedet av kundevurderinger og hotellenes bærekraftsertifiseringer?*». Vi har benyttet to modeller til å undersøke dette: Modell 1 er en samlet modell for segmentene, mens Modell 2 deler hotellene inn etter segmentnivå.

Ved å analysere dataene, fant vi betydelige forskjeller i faktorene som påvirker hotellrompriser i ulike regioner i Norge. Dermed kunne vi ikke betrakte det norske hotellmarkedet som ett marked, og valgte derfor å dele markedet inn i to dimensjoner: By og Distrikts-Norge (DN). Disse dimensjonene ble identifisert som de mest homogene inndelingene, basert på våre funn.

I Modell 1, som er en aggregert modell for alle segmentene, fant vi sammenfallende resultater mellom By og DN for koeffisientene, selv om det var en del forskjeller i størrelse. Modellen ble også sterkt preget av segment tre og fire for begge dimensjonene. Videre fant vi en tydelig indikasjon på at segmenttilhørighet har stor betydning for hotellromprisene i Norge. Variabelen har en større effekt enn kundevurderingene og bærekraftsertifiseringene. Dette funnet indikerer at mye av kvaliteten ligger innad i segmenttilhørigheten, og i mindre grad i de andre variablene. Dette funnet er innsiktsfullt, da det ikke finnes et offisielt segmentsystem i Norge. Et annet uventet funn, var at *renhold* hadde en negativ og signifikant effekt på hotellromprisen i både By (-23,6 %) og DN (-16,4 %). Dette avviker fra et økonomisk rasjonale. Vi fant også en positiv og signifikant koeffisient for *personale* på 8,9 % i begge dimensjonene. Dette har vi ikke observert i tidligere studier (Castro & Ferreira, 2018; Schamel, 2012; Zhang et al., 2011), men vi anser funnet som forventet, da det følger et økonomisk rasjonale.

I Modell 2, som deler hotellene etter segmentnivå, gjorde vi flere funn som bidrar til innsikt og til å besvare forskningsspørsmålet. Dessverre var datagrunnlaget for smalt for segment to og fem i By og segment to i DN, til at vi kunne trekke meningsfulle konklusjoner her. Et viktig funn er at variablene har betydelige forskjeller i koeffisienter og signifikansnivå, avhengig av

segment og dimensjon. Dette fremhever kompleksiteten i markedet. At resultatene er sprikende mellom segmentene er imidlertid ikke overraskende, da Zhang et al. (2011) også fant sprikende resultater mellom segmentene i deres studie fra New York.

Noen betydningsfulle funn inkluderer flere signifikante koeffisienter i segmentene, og illustrerer variasjonen i resultatene: I By-segment tre har *komfort* (77,8 %) en høy positiv effekt på romprisene, mens *renhold* (-56,1 %) har en betydelig negativ effekt. I By-segment fire påvirker *beliggenhet* (14,6 %) romprisene positivt. I DN-segment to er både *komfort* (69,0 %) og *personale* (18,4 %) viktige faktorer, mens *renhold* (-57,5 %) og *WiFi* (-15,3 %) har en negativ effekt. DN-segment tre viser at *beliggenhet* (9,1 %), *komfort* (25,9 %) og *personale* (9,4 %) har en positiv effekt på romprisene, mens *renhold* (-20,5 %) har en negativ effekt. I DN-segment fire har *fasiliteter* (22,5 %) en positiv effekt, og *bærekraft* (4,7 %) har en viss betydning for romprisene. Disse ulike funnene gir en tydelig indikasjon på variasjonen i kundenes preferanser.

Et annet viktig funn fra Modell 2 var at enkelte attributter tilsynelatende ikke påvirker hotellromprisene. Disse har altså en ikke-signifikant effekt. Dette indikerer at kundene ikke er villig til å betale ekstra for en forbedring av disse attributtene. De kan derfor anses som standardiserte forventninger til hotellene (Zhang et al., 2011). Et eksempel på dette er *WiFi*, som ikke hadde noen effekt på hotellromprisene, med unntak av en liten negativ effekt for segment to i DN. Hvis WiFi fjernes, vil det antagelig påvirke kundetilfredsheten negativt. Et annet eksempel er *personale*, som i flere segmenter ikke påvirket hotellromprisen. Variabelen har en positiv effekt på hotellromprisen i segment to i By, og i segment to og tre i DN. Dette tilsier imidlertid ikke at hotellene i segmentene med ikke-signifikante koeffisienter kan eliminere personalet. En ekstra innsats fra personalet vil likevel i liten grad differensiere hotellene med tanke på kundenes betalingsvilje.

Andre halvdel av forskningsspørsmålet dreier seg om konsumentenes betalingsvilje for hoteller som prioriterer bærekraft. Vi fant lite evidens for en slik betalingsvilje i det norske hotellmarkedet, med unntak av en svak effekt i segment fire i DN. Dette tyder på at segmenttilhørighet og andre egenskaper er viktigere kriterier for kundene ved valg av overnattingsted, enn i hvilken grad hotellet prioriterer bærekraft.

Vi finner dermed at *hotellromprisene i det norske markedet blir påvirket av kundevurderinger og hotellenes bærekraftsertifiseringer*, men kun i en viss grad. Markedet er komplekst, og det er vanskelig å generalisere i hvilken grad hvert attributt påvirker hotellromprisen. Dette avhenger av hotellenes segmenttilhørighet og hvilket marked hotellet opererer i.

8.2 Hva resultatene bidrar med

Først og fremst har oppgaven bidratt til forståelse av hvordan kundevurderinger og bærekraftsertifiseringer påvirker romprisen i det norske hotellmarkedet. Den belyser hva kundene er villige til å betale mer for, og hva de ikke er villige til å betale for.

Et unikt bidrag fra denne analysen, er inndelingen i By/DN-dimensjonen. Ved å dele det norske hotellmarkedet i to, de åtte største hotellkonsentrasjonene mot resten av landet, fikk vi frem forskjeller mellom by og land. Disse forskjellene viste seg dessuten å være betydelige. Dette bidrar til å belyse kompleksiteten i hotellmarkedet, og kan være med på å redusere tilliten til studier som har aggregert mange markeder i én modell.

Ved å fokusere på det norske hotellmarkedet og dets kompleksitet, gir oppgaven innsikt i unike trekk ved dette markedet og hvordan det skiller seg fra andre markeder. Vi har drøftet våre funn opp mot tidligere studier, og ikke funnet markeder som er helt sammenlignbare med det norske markedet. Vi har funnet flere forskjeller, eksempelvis at *personale* har en signifikant og positiv effekt på hotellromprisene i både By og DN for visse segmenter, noe som ikke er i tråd med tidligere studier av større byer. Dette bidrar til en mer helhetlig forståelse av hotellromprisene og attributtpåvirkningen på et nasjonalt nivå.

Vi benyttet oss av web-skraping for å trekke ut data fra alle tilgjengelige hoteller i det norske hotellmarkedet hos Booking.com. Denne effektive metoden brukte i underkant av to timer på å innhente vårt omfattende datagrunnlag. Vi fikk dermed et tverrsnitt av det norske hotellmarkedet, og ikke bare et semi-tverrsnitt. Dette er en klar styrke ved denne oppgaven, i motsetning til mange tidligere hedoniske prisstudier av hotellmarkeder. Soler et al. (2019) tastet eksempelvis inn sine data i løpet av en 17 dagers periode. Siden hotellromprisene sjelden holder seg stabile over en så lang periode (Lopes et al., 2017), kan dette utfordre studiens realibilitet og validitet. Web-skrapingen er derfor en styrke ved vår undersøkelse, og viser at det er mulig å innhente et stort datagrunnlag på relativt kort tid.

Studien utforsker også hotellenes bærekraftsertifiseringer og deres påvirkning på hotellrompriser. Ved å avdekke begrenset betalingsvilje for bærekraftsertifiseringer, bidrar oppgaven til forståelsen av kundenes holdninger og preferanser knyttet til bærekraft innenfor det norske hotellmarkedet.

Vi kan oppsummere med at oppgaven gir en dypere innsikt i hvordan kundevurderinger og bærekraftsertifiseringer påvirker hotellromprisen i det norske hotellmarkedet. Siden det ikke tidligere er gjort lignende studier i Norge, utgjør oppgaven et verdifullt bidrag til den eksisterende litteraturen på dette feltet.

8.3 Begrensninger og ekstern validitet

Vår analyse har visse begrensninger som kan ha påvirket resultatene og påliteligheten av studien. Først og fremst er oppgavens brede omfang en begrensning. Vi har undersøkt hotellattributter på et overordnet nivå og har ikke gjennomført grundige analyser av spesifikke faktorer. For eksempel kunne en mer detaljert analyse av fasiliteter gitt bedre innsikt i hvordan eksempelvis et svømmebasseng påvirker prisen. Derfor må våre konklusjoner betraktes som generelle og overordnede.

I utgangspunktet ønsket vi å teste samtlige variabler, noe som resulterte i høye VIF-verdier for flere av variablene. Disse verdiene indikerer en høy grad av kolinearitet mellom noen av de uavhengige variablene, og dermed multikollinearitet. Dette kan ha svekket reliabiliteten og validiteten til resultatene, da det blir vanskeligere å estimere effekten av de uavhengige variablene på den avhengige variabelen. Vi mener likevel at å redusere VIF-verdiene ved å eliminere uavhengige variabler, kan svekke reliabiliteten og validiteten til resultatene i større grad enn problemet med multikollinearitet. Dette fordi vi ønsket å teste samtlige av de uavhengige variablene, da de representerer tilbakemeldingsvariablene på Booking.com. De er dermed relevante for å besvare problemstillingen. Vi mener derfor at modellene er bedre egnet til å måle de fenomenene vi undersøker, men vi må likevel være oppmerksomme på at multikollinearitet kan ha påvirket resultatene og dermed begrenset den eksterne validiteten.

En annen begrensning er at vi kun har brukt Booking.com som datakilde. Selv om Booking.com er en av de mest brukte tredjepartskanalene (Inversini & Masiero, 2014), kunne inkludering av andre plattformer gitt et mer omfattende og representativt bilde av hotellmarkedet. Et vesentlig poeng i denne sammenhengen, er at vi hadde få observasjoner i segment to og fem i begge

dimensjoner. Det kan tenkes at vi ville hatt flere observasjoner i disse segmentene ved hjelp av alternative innsamlingsmetoder, samt inkludering av flere datakilder. Dette kunne sikret mer valide og representative regresjonslinjer for hvert segment.

Vi har også begrenset oppgaven til data fra én bestemt midtukes dag i april. Dette tverrsnittet av hotellmarkedet kan være påvirket av sesongvariasjoner og andre faktorer som er spesifikke for denne tiden av året. En mer langsiktig og omfattende studie som tar hensyn til slike variasjoner over flere sesonger, vil sannsynligvis gi en bedre forståelse av hotellmarkedets dynamikk.

Oppgaven er også begrenset av få observasjoner i to av segmentene. Dette ble tydelig i Modell 2, hvor vi undersøkte sammenhengene mellom de inkluderte variablene og hotellrompriser på segmentnivå. Segment to og fem i By hadde henholdsvis to og elleve observasjoner, mens segment fem i DN ble hadde fem observasjoner. Vi kunne dermed ikke estimere regresjonslinjen for disse segmentene. Dette skyldtes imidlertid en mangel på slike hoteller hos Booking.com, og ikke en begrensning i vår mulighet til å samle dem inn.

Når det gjelder ekstern validitet, må vi være forsiktige med å generalisere våre funn utenfor det norske hotellmarkedet. Siden dataene er basert på norske hoteller og priser fra Booking.com, kan det være variasjoner i hotellmarkedet og praksiser i andre land og regioner.

Vi vil også å påpeke at tidsrammen for oppgaven har begrenset vår evne til å inkludere mer detaljerte variabler og dybdeanalyser. Det er mulig vi har oversett eller utelatt viktige faktorer som kan ha påvirket resultatene.

8.4 Forslag til fremtidig arbeid

For å utvide og utforske problemstillingen i denne studien ytterligere, er det flere muligheter for fremtidig arbeid. Dette kan bidra til en mer omfattende forståelse av hvordan ulike attributter påvirker hotellromprisene.

Et forslag vil være å utføre en tilsvarende analyse i andre land, noe som vil gi et bredere sammenligningsgrunnlag og en dypere forståelse av hvordan hotellmarkedet varierer på tvers av ulike land. Dette kan bidra til å identifisere eventuelle likheter og forskjeller i hvilke attributter som påvirker romprisene i ulike hotellmarkeder.

Det er også mulig å utføre dybdestudier av spesifikke attributter. Eksempelvis kan man utforske enkeltattributter som spa-tilbud eller parkeringsplass mer grundig, og undersøke hvordan disse attributtene påvirker hotellromprisene. Dette vil bidra med en mer detaljert forståelse av hvilke faktorer som påvirker romprisene.

En annen interessant tilnærming vil være å analysere hvordan hotellromprisene endrer seg over tid som respons på endringer i kundevurderinger og hotellenes bærekraftsertifisering. Ved å undersøke hvordan hotellene reagerer på tilbakemeldinger fra kundene og bærekraftsertifiseringer, kan man få bedre innsikt i hvordan prisdannelsen påvirkes av disse faktorene og hvordan hotellene tilpasser seg markedsbehovene.

Videre kan det være verdifullt å anvende en kvalitativ metode til å forstå hotellaktørens perspektiver og praksiser knyttet til prissettingen. Dette kan omfatte intervjuer med sentrale hotellaktører for å få innsikt i deres beslutningsprosesser knyttet til prissetting. Kvalitative metoder kan i så måte gi en ny dimensjon for å forstå kompleksiteten bak prisdannelsen.

Denne oppgaven har også tatt for seg det samlede norske hotellmarkedet i to dimensjoner. En åpenbar mulighet for videre forskning vil være å utføre dypere undersøkelser av lokale forhold. Dette for å undersøke hvordan hotellromprisene og attributtpåvirkningen varierer mellom ulike geografiske områder.

De overnevnte forslagene gir flere muligheter for videreutvikling av vår studie eller for å bygge videre på den, og vil bidra til å utvide kunnskapen om hotellrompriser ytterligere.

Referanseliste

- Abratea, G., Capriello, A., & Fraquelli, G. (2011). When quality signals talk: Evidence from the Turin hotel industry. *Tourism Management*, ss. 912-921.
- Andersson, D. E. (2000). Hypothesis testing in hedonic price estimation - On the selection of independent variables. *The Annals of Regional Science*, ss. 293-304.
- Andersson, D. E. (2008). Hotel attributes and hedonic prices: an analysis of internet-based transactions in Singapore's market for hotel rooms . *The Annals of Regional Science volume*, ss. 229-240.
- Arbel, A., & Pizam, A. (1977). Some Determinants of Urban Hotel Location: The Tourists' Inclinations. *The Tourists' Inclinations. Journal of Travel Research*, , ss. 18-22.
- Bernard, S., & Nicolau, J. L. (2022). Environmental certification and hotel market value. *International Journal of Hospitality Management*, ss. 1-20.
- Booking.com. (2023a, mars 15). *Booking.com*. Hentet fra Fleischer's Hotel: <https://www.booking.com/hotel/no/fleichers.no.html?label=gen173nr-1BCAEoggI46AdIM1gEaKoBiAEBmAEduAEHyAEN2AEB6AEBiAIBqAIDuAKM08agBsACAdICJDIZN2Q1YTcyLWRiMDctNDA1Ni1iNTFhLTMzYjVkNDY0NjdiNtgCBeACAQ&sid=4c59d359de58354bad199dd448d763ec&aid=304142&ucfs=1&arphp>
- Booking.com. (2023b, mai 25). *Booking.com*. Hentet fra Thon Hotel Hallingdal: <https://www.booking.com/hotel/no/thon-hallingdal.no.html?aid=304142&label=gen173nr-1FCAEoggI46AdIM1gEaKoBiAEBmAEduAEHyAEN2AEB6AEB-AEMiAIBqAIDuAK9xMejBsACAdICJGJkODZmOWMzLWJkOWYtNGY2MS04NjZjLTY0NmQ2ZmIyNjJiZtgCBuACAQ&sid=73b9f8eb764885e88b7c57131f545e5a&de>
- Booking.com. (2023c, mai 25). *Partner.booking.com*. Hentet fra Becoming certified: <https://partner.booking.com/en-gb/learn-more/becoming-certified>
- Borges, I. R., Pereira, G. M., Matos, C. A., & Borchardt, M. (2015). Analysis of the relationship between the satisfaction of consumers and the prices offered on site booking.com. *Tourism & Management Studies*, ss. 64-70.
- Brown, J., Broderick, A. J., & Lee, N. (2007). word of mouth communication within online communities: conceptualizing the online social network. *Journal of Interactive Marketing*, ss. 2-20.

- Bull, A. O. (1994). Pricing a Motel's Location. *International Journal of Contemporary Hospitality Management*, ss. 10-15.
- Castro, C., & Ferreira, F. A. (2018). Online hotel ratings and its influence on hotel room rates: the case of Lisbon, Portugal. *Tourism & Management Studies*, ss. 63-72.
- Cavero-Rubio, J. A., & Amorós-Martínez, A. (2008). Environmental certification and Spanish hotels' performance in the 2008 financial crisis. *Journal of Sustainable Tourism*, ss. 771-796.
- Chatterjee, P. (2006). Online Reviews: Do Consumers Use Them? *Advances in consumer research. Association for Consumer Research*, ss. 129-133.
- Chatterjee, S., & Hadi, A. S. (2015). *Regression Analysis by Example (5th ed.)*. <https://www.perlego.com/book/2767432/regression-analysis-by-example-pdf>: Wiley Publications.
- Chen, C.-F., & Rothschild, R. (2010). An Application of Hedonic Pricing Analysis to the Case of Hotel Rooms in Taipei. *Tourism Economics*, ss. 685–694.
- Chen, Y.-S., & Chang, C.-H. (2013). Greenwash and Green Trust: The Mediation Effects of Green Consumer Confusion and Green Perceived Risk. *Journal of Business Ethics*, ss. 489-500.
- Chin, & Chau. (2003). A Critical Review of Literature on the Hedonic Price Model . *International Journal for Housing Science and Its Applications*, ss. 145-165.
- Classic Norway Hotels AS. (2023, mai 4). *Classic Norway Hotels*. Hentet fra Om Classic Norway Hotels: <https://classicnorway.no>
- Claver-Cortés, E., Molina-Azorin, J. F., & Pereira-Moline, J. (2007). The impact of strategic behaviours on hotel performance. *International Journal of Contemporary Hospitality Management*, ss. 6-20.
- Creswell, J. W., & David, J. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (5.th.ed)*. United States of America: SAGE Publications, Inc.
- Curto, J. D., & Pinto, J. C. (2011). The corrected VIF (CVIF). *Journal of Applied Statistics*, ss. 1499-1507.
- Dina, M. Y., & Chevalier, J. (2014). Promotional Reviews: An Empirical Investigation of Online Review Manipulation. *American Economic Association*, ss. 2421-2455.
- Espinet, J., Saez, M., Coenders, G., & Fluvià, M. (2003). Effect on Prices of the Attributes of Holiday Hotels: A Hedonic Prices Approach. *Tourism Economics*, ss. 165–177.

- Felix, C., & Clever, V. (2014). The Relationship between Hotel Rating and Customer Outcomes: Customer Perceived Service Quality and Customer Satisfaction. *Greener Journal of Business and Management Studies*, ss. 146-152.
- Geerts, W. (2014). Environmental certification schemes: Hotel managers' views and perceptions. *International Journal of Hospitality Management*, ss. 87-96.
- Gretzel, U., & Yoo, K.-H. (2008). Use and Impact of Online Travel Reviews. *Laboratory for Intelligent Systems in Tourism, Department of Recreation, Park and Tourism Sciences*, ss. 35-46.
- Guillet, B. D., & Law, R. (2010). Analyzing hotel star ratings on third-party distribution websites. *International Journal of Contemporary Hospitality Management*, ss. 797-813.
- Hotelia AS. (2023, mai 20). *Hotelia.no*. Hentet fra 4Q22: <https://www.hotelia.no/rapporter/4q22>
- Hotell Refsnes Gods. (2023, mai 28). *refsnesgods.no*. Hentet fra Things to do: <https://refsnesgods.no/en/things-to-do/>
- Inversini, A., & Masiero, L. (2014). Selling rooms online: the use of social media and online travel agents. *International Journal of Contemporary Hospitality Management*, ss. 272-292.
- Israeli, A. A. (2002). Star rating and corporate affiliation: their influence on room price and performance of hotels in Israel. *Hospitality Management* 21, ss. 405-424.
- Ivanova, M., & Ivanov, S. (2015). The Nature of Hotel Chains: An Integrative Framework. *International Journal of Hospitality & Tourism*, ss. 122-142.
- Johannessen, A., Tufte, P. A., & Christoffersen, L. (2011). *Introduksjon til samfunnsvitenskapelig metode (3.utg)*. Oslo: Abstrakt forlag AS.
- Khan, Y. H., Hakeem, A., & Naumov, N. (2018). The Use of Branding and Market Segmentation in Hotel Marketing: A Conceptual Review. *Journal of Tourism Intelligence and Smartness*, ss. 12-23.
- Khare, A., & Bhagwat, J. (2019). The Relevance and Utility of Market Segmentation for the International Hospitality Industry. *Atithya: A Journal of Hospitality*, ss. 4-7.
- Krumsvik, R. J. (2014). *Forskningsdesign og kvalitativ metode - ei innføring*. Bergen: Vigmonstad & Bjørke AS.
- Kuminoff, N. V., Zhang, C., & Rudi, J. (2010). Are Travelers Willing to Pay a Premium to Stay at a "Green" Hotel? Evidence from an Internal Meta-Analysis of Hedonic Price Premia. *Agricultural and Resource Economics Review*, ss. 468-484.

- Lai, I. K., Hitchcock, M., Lu, D., & Liu, Y. (2018). The Influence of Word of Mouth on Tourism Destination Choice: Tourist–Resident Relationship and Safety Perception among Mainland Chinese Tourists Visiting Macau. *Sustainability*, ss. 1-17.
- Lancaster, K. J. (1966). A New Approach to Consumer Theory. *Johns Hopkins University*, ss. 132-157.
- Lewis, R. C., & Shoemaker, S. (1997). Price-Sensitivity Measurement. *Cornell Hospitality Quarterly* 38, ss. 44-54.
- Lopes, I. C., Soares, F., & Silva, E. C. (2017). *Tourism Demand in the Algarve Region: Evolution and Forecast using SVARMA models*. Melville: AIP Publishing.
- Lorde, T., Jacob, J., & Weekes, Q. (2019). Price-Setting Behavior in a Tourism Sharing Economy Accommodation Market: A Hedonic Price Analysis of AirBnB Hosts in the Caribbean. *Tourism Management Perspectives*, ss. 251-261.
- Martin-Fuentes, E. (2016). Are guests of the same opinion as the hotel star-rate classification system? *Journal of Hospitality and Tourism Management*, ss. 126-134.
- Mitchell, R. (2018). *Web Scraping with Python: Collecting Data from the Modern Web. (2th ed)*. <http://cds.cern.ch/record/2041893>: O'Reilly Media, Inc.
- Mohsina, A., Rodriguesb, H., & Brochadob, A. (2019). Shine bright like a star: Hotel performance and guests' expectations based on star ratings. *International Journal of Hospitality Management* , ss. 103-114.
- Nangpiire, C., Silva, J., & Alves , H. (2022). Customer engagement and value co-creation/destruction: the internal fostering and hindering factors and actors in the tourist/hotel experience. *Journal of Research in Interactive Marketing*, ss. 173-188.
- Narangajavana, Y., & Hu, B. (2008). The Relationship Between the Hotel Rating System, Service Quality Improvement, and Hotel Performance Changes: A Canonical Analysis of Hotels in Thailand. *Journal of Quality Assurance in Hospitality & Tourism* , ss. 34-56.
- Nicolaua, J. L., Mellinasb, J. P., & Martín-Fuentesc, E. (2020). The halo effect: A longitudinal approach. *Annals of Tourism Research* 83, ss. 1-10.
- (2023, februar 11). *klassifisering.no*. Hentet fra OM OSS: <https://www.klassifisering.no/om-oss/>
- Osland, L. (2001). Den hedonistiske metoden og estimering av attributtpriser. *Norsk Økonomisk Tidsskrift* 115 , ss. 1-22.

- Oslo kommune. (2023, mai 28). *Klima Oslo*. Hentet fra Satsingsområde 3, transport: Sykkel, gange, kollektiv- og biltrafikk: <https://www.klimaoslo.no/article/3-satsingsomrader-ram-mot-2030-2/satsingsomrade-3/>
- Radisson Hotels Group. (2023, april 11). *www.radissonhotels.com*. Hentet fra Norge: <https://www.radissonhotels.com/no-no/destinasjon/norge>
- Rosen, S. (1974). Hedonic Pricing and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, 82(1), ss. 34-55.
- Santos, G. E. (2016). Worldwide hedonic prices of subjective characteristics of hostels. *Tourism Management*, ss. 451-454.
- Saunders, M., Lewis, P., & Thornhill, A. (2007). *Research Methods for Business Students (4th ed.)*. Harlow: Pearson Education Limited.
- Sánchez-Ollero, J. L., García-Pozo, A., & Marchante-Mera, A. (2014). How Does Respect for the Environment Affect Final Prices in the Hospitality Sector? A Hedonic Pricing Approach. *Sustainability and Finance*, ss. 31-39.
- Sánchez-Peres, M., & Illescas-Manzano, M. D. (2019). Modeling hotel room pricing: A multi-country analysis. *International Journal of Hospitality Management*, ss. 89-99.
- Scandic Hotels Group AB. (2023, april 11). *Scandichotels.no*. Hentet fra Nordens største hotellkjede: <https://www.scandichotels.no/landing-pages/scandic--rica-old>
- Schamel, G. (2012). Weekend vs. midweek stays: Modelling hotel room rates in a small market. *International Journal of Hospitality Management*, ss. 1113-1118.
- Segarra-Oña, M.-d.-V., Peiró-Signes, Á., Verma, R., & Miret-Pastor, L. (2012). Does Environmental Certification Help the Economic Performance of Hotels? Evidence from the Spanish Hotel Industry. *Cornell Hospitality Quarterly*, ss. 242-256.
- Sirmans, S. G., Macpherson, D. A., & Zietz, E. N. (2005). The Composition of Hedonic Pricing Models. *Journal of Real Estate Literature*, ss. 3-43.
- Skilbrei, M.-L. (2019). *Kvalitative metoder : Planlegging, gjennomføring og etisk refleksjon (1.utg)*. Bergen: Fagbokforlaget.
- Soler, I. P., Gemar, G., Correia, M. B., & Serra, F. (2019). Algarve hotel price determinants: A hedonic pricing model. *Tourism Management*, ss. 311-321.
- SSB. (2023, februar 21). *Statistisk sentralbyrå*. Hentet fra Kommersielle overnattingsverksemdet. Hotell, campingplassar, hyttegrender og vandrarheimar fordelt på fylke: <https://www.ssb.no/transport-og-reiseliv/reiseliv/statistikk/overnattingar>

- Stock, J., & Watson, M. (2020). *Introduction to econometrics (Fourth edition.; Global ed., The Pearson series in economics)*. Harlow: Pearson (Intl).
- Strawberry. (2023, mai 25). *www.strawberry.no*. Hentet fra Kart over alle våre hotell:
<https://www.strawberry.no/hotell/?resetDestination=true/>
- Stringam, B. B., Gerdes, J. H., & Vanleeuwen, D. M. (2010). Assessing the Importance and Relationships of Ratings on User-Generated Traveler Reviews. *Journal of Quality Assurance in Hospitality & Tourism*, ss. 73-92.
- Studenmund, A. H. (2014). *Using econometrics : A practical guide (6th ed.)* (Vol. 7). Harlow: Pearson Education Limited.
- Thon Hotel. (2023, april 11). *Thon Hotels*. Hentet fra Thonhotels.no:
<https://www.thonhotels.no/om-thon-hotels/om-olav-thon-gruppen/>
- Thrane, C. (2007). Examining the determinants of room rates for hotels in capital cities: The Oslo experience. *Journal of Revenue and Pricing Management volume*, ss. 313-323.
- Webb, T. (2016). From travel agents to OTAs: How the evolution of consumer booking behavior has affected revenue management. *Journal of Revenue and Pricing Management volume* , ss. 276–282.
- White, P. J., & Mulligan, G. F. (2010). Hedonic Estimates of Lodging Rates in the Four Corners Region. *The Professional Geographer*, ss. 533-543.
- Zhang, Z., Ye, Q., & Law, R. (2011). Determinants of hotel room price. An exploration of travelers' hierarchy of accommodation needs. *International Journal of Contemporary Hospitality Management*, ss. 972-981.

Appendiks

A1 Generell bransjeinformasjon

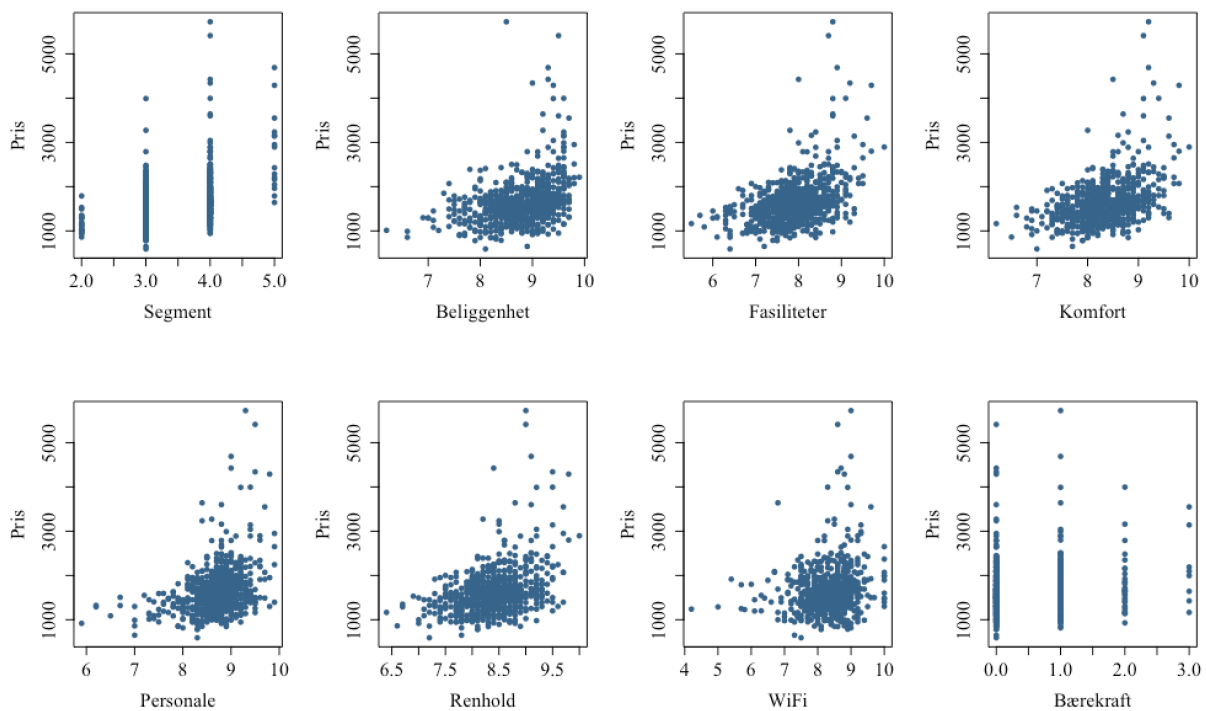
Tabell A1.1: Presis beskrivelse av variablene.

Variablene	Hva	Skala
Beliggenhet	Overnattingsstedet sin geografiske beliggenhet. En god beliggenhet kan eksempelvis være nærhet til attraksjoner og transportknutepunkter	1–10
Fasiliteter	Dette kan eksempelvis være parkering, treningssenter, basseng og spa	1–10
Komfort	Forteller noe om hvor behagelig oppholdet var. Dette kan eksempelvis gjelde senger, som kan påvirke gjestens komfort	1–10
Personale	Forteller noe om opplevd kundeservice.	1–10
Renhold	Forteller noe om hotellets renslighet.	1–10
WiFi	Forteller noe om kvaliteten og tilgjengeligheten av trådløs Internett-tilkobling ved hotellet.	1–10
Bærekraft	Forteller hvilken bærekraftsertifisering hotellet har hos Booking.com.	0–3

Tabell A1.1 gir en presis beskrivelse av variablene vi har inkludert i våre multiple regresjonsmodeller, samt hvilken skala de måles på.

A5 Metode

A5.1 Multipel regresjon

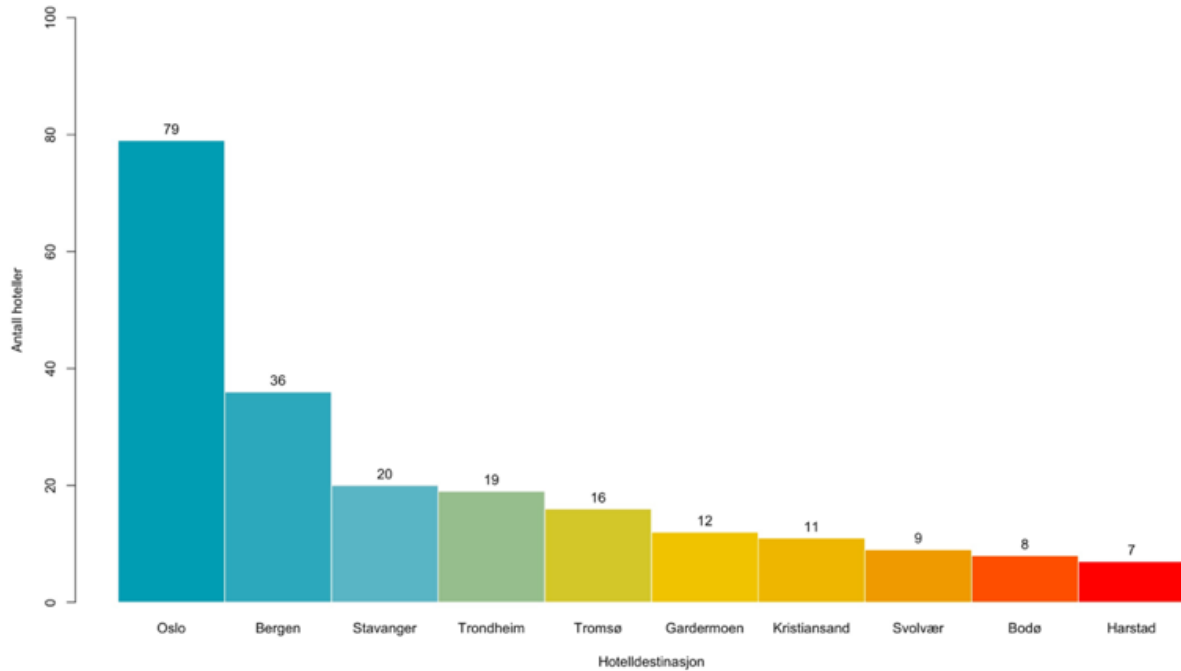


Figur A5.1: Sammenhengen mellom pris og de uavhengige variablene.

Figur A5.1 viser sammenhengen mellom responsvariabelen *Pris* og de predikerende variablene *Segment*, *Beliggenhet*, *Fasiliteter*, *Komfort*, *Personale*, *Renhold*, *WiFi* og *Bærekraft*. Vi ser at *segment* og *pris* har en positiv lineær sammenheng og at *bærekraft* og *pris* har en negativ lineær sammenheng. Videre har de resterende variablene *Beliggenhet*, *Fasiliteter*, *Komfort*, *Personale*, *Renhold* og *WiFi* også en positiv og lineær sammenheng med *Pris*, men med konvekset. Disse sammenhengene følger altså en konveks lineær trend. Denne trenden skyldes at punktene er konsentrert nede til høyre, med et fåtall ekte uteliggere oppe til høyre. Dette taler for en log-lineær funksjonsform, hvor vi tar den naturlige logaritmen av responsvariabelen *Pris*.

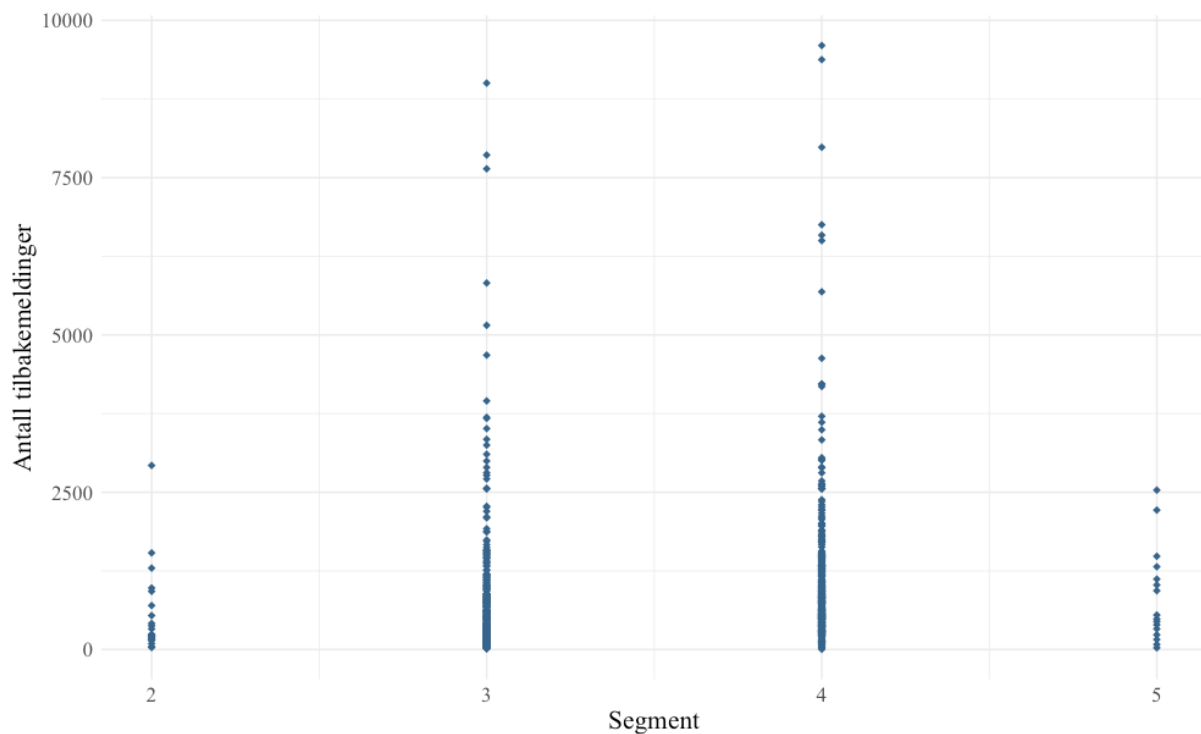
A6 Presentasjon av data

A6.1 Geografisk inndeling av data



Figur A6.1: De ti største hotellkonsentrasjonene i utvalget.

A6.2 Deskriptiv statistikk



Figur A6.2: Antall tilbakemeldinger per hotell i hvert segment.

Figur A6.2 viser hvor mange tilbakemeldinger hvert hotell i utvalget har mottatt på Booking.com, fordelt på segmentnivå. Hotellene i segment tre og fire har klart flest tilbakemeldinger, sammenlignet med hotellene i segment to og fem. Samtidig har flere av hotellene i disse segmentene mer enn 5 000 tilbakemeldinger, og enkelte har mellom 7 500 og 10 000 tilbakemeldinger. I segment to og fem har kun ett hotell mer enn 2 500 tilbakemeldinger. Figur A6.2 illustrerer dermed hvordan majoriteten av kundene i det norske hotellmarkedet benytter seg av hoteller i segment tre og fire.

Tabell A6.1: Deskriptiv statistikk for hotellromprisene i Norge på segmentnivå.

Segment	Mean	SD	Min	Maks
2,00	1 187,29	229,48	860,00	1 790,00
3,00	1 483,18	394,13	592,00	3 990,00
4,00	1 786,47	552,72	948,00	5 725,00
5,00	2 776,63	879,78	1642,00	4 690,00

Tabell A6.2: Deskriptiv statistikk for de inkluderte attributtene.

Variabel	Obs.	Mean	SD	Min.	Maks.
Beliggenhet	680	8,78	0,60	6,20	9,90
Fasiliteter	680	7,93	0,67	5,50	10,00
Komfort	680	8,34	0,58	6,20	10,00
Personale	680	8,70	0,51	5,90	9,90
Renhold	680	8,41	0,55	6,40	10,00
WiFi	680	8,34	0,68	4,20	10,00

Tabell A6.3: Deskriptiv statistikk for gjennomsnittlige vurderinger på segmentnivå.

	Segment 2				Segment 3			
	Min.	Maks.	Mean	SD	Min.	Maks.	Mean	SD
Beliggenhet	7,30	9,60	8,40	0,62	7,00	9,90	8,62	0,52
Fasiliteter	5,50	9,20	7,40	0,86	6,70	9,70	7,93	0,50
Komfort	6,20	9,00	7,86	0,76	7,20	9,80	8,34	0,45
Personale	6,40	9,20	8,47	0,72	6,70	9,90	8,61	0,39
Renhold	6,20	9,50	8,05	0,81	7,30	9,80	8,41	0,44
WiFi	5,70	9,20	8,05	0,91	5,40	10,00	8,30	0,54
	Segment 4				Segment 5			
	Min.	Maks.	Mean	SD	Min.	Maks.	Mean	SD
Beliggenhet	6,20	9,80	8,91	0,61	8,60	9,80	9,46	0,28
Fasiliteter	5,70	9,90	8,12	0,71	8,30	10,00	9,06	0,46
Komfort	6,30	9,90	8,48	0,61	8,60	10,00	9,30	0,38
Personale	5,90	9,90	8,77	0,59	8,40	9,90	9,21	0,39
Renhold	6,60	9,90	8,50	0,59	8,50	10,00	9,25	0,43
WiFi	4,20	10,00	8,38	0,77	8,30	9,60	8,98	0,36

A7 Analyse og diskusjon

A7.1 Alternative regresjonsmodeller

Tabell A7.1: Regresjonstabell for alle hoteller i det norske hotellmarkedet.

	<i>Avhengig variabel:</i> <i>Ln(Pris)</i>	
	Koeffisient	SD
Segment	0,137 ***	0,017
Beliggenhet	0,059 ***	0,019
Fasiliteter	0,012	0,046
Komfort	0,213 ***	0,070
Personale	0,097 ***	0,026
Renhold	(0,143) **	0,067
WiFi	(0,019)	0,015
Bærekraft	0,012	0,014
Konstant	5,017 ***	0,202
Observasjoner	680	
R²	0,335	
Justert R²	0,327	
Residual SE	0,234 (df = 671)	
F-statistikk	42,260 *** (df = 8; 671)	
Note:	*p<0,1; ** p<0,05; *** p< 0,01	

Regresjonslinjen til den generelle regresjonsmodellen for det norske hotellmarkedet blir:

$$\begin{aligned} \ln(\text{Pris}) = & 5,017 + 0,137 \text{ Stjerner}_i + 0,059 \text{ Beliggenhet}_i + 0,012 \text{ Fasiliteter}_i \\ & + 0,213 \text{ Komfort}_i + 0,097 \text{ Personale}_i - 0,143 \text{ Renhold}_i - 0,019 \text{ WiFi}_i \\ & + 0,012 \text{ Bærekraft}_i + \varepsilon_i \end{aligned}$$

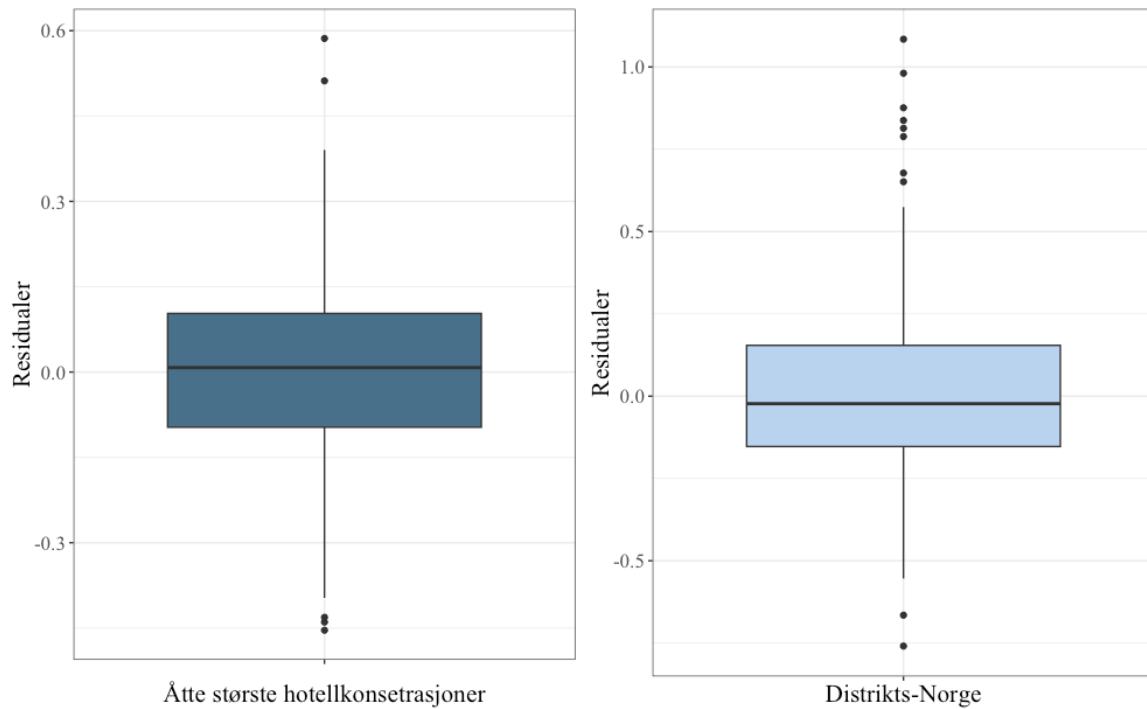
Tabell A7.2: Regresjonstabell for Oslo-markedet.

	<i>Avhengig variabel:</i>	
	Ln(Pris)	
	Koeffisient	SD
Segmentnivå	0,198***	0,051
Beliggenhet	0,172***	0,051
Fasiliteter	0,039	0,207
Komfort	0,495	0,301
Personale	0,053	0,075
Renhold	(0,407)	0,284
WiFi	0,0002	0,072
Bærekraft	0,050	0,042
Konstant	3,608***	0,564
Observasjoner	79	
R²	0.604	
Justert R²	0.559	
Residual SE	0.216 (df = 70)	
F-statistikk	13.344*** (df = 8; 70)	
Note:	*p<0.1; **p<0.05; ***p<0.01; negative koeff. i parentes	

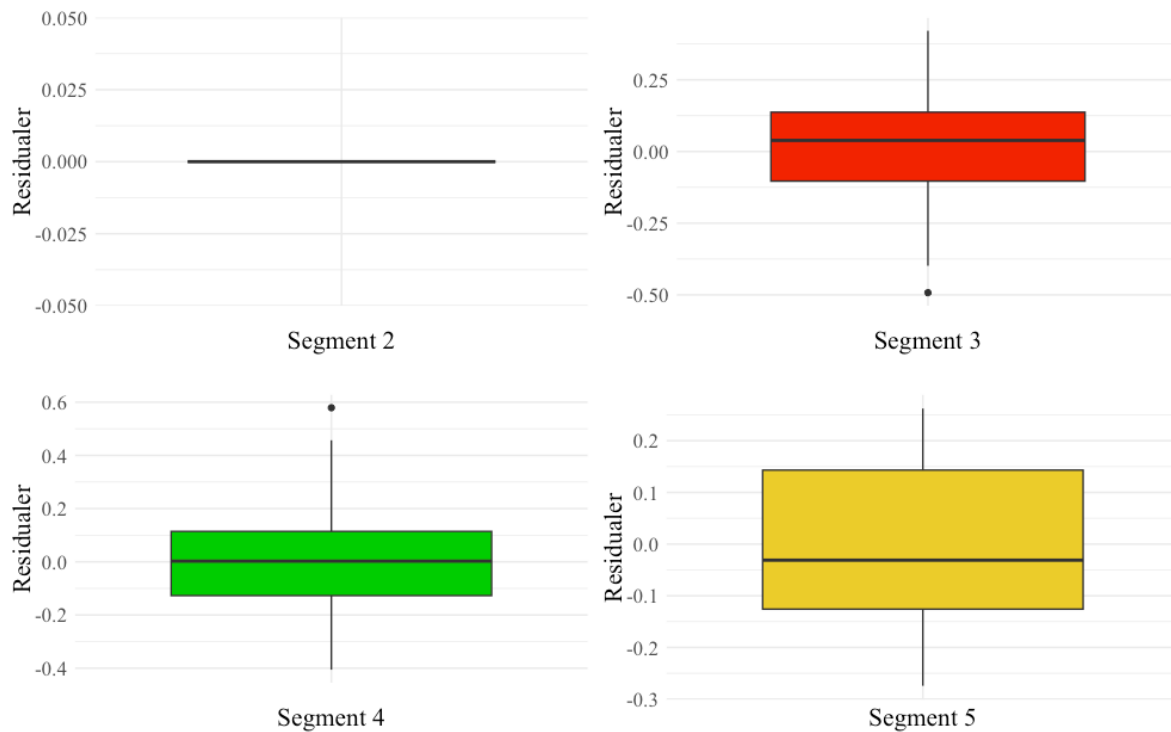
Regresjonslinjen til hotellmarkedet i Oslo blir da:

$$\begin{aligned}
 \text{Ln}(\text{Pris}) = & 3,608 + 0,198 \text{ Stjerner}_i + 0,172 \text{ Beliggenhet}_i + 0,039 \text{ Fasiliteter}_i \\
 & + 0,495 \text{ Komfort}_i + 0,053 \text{ Personale}_i - 0,407 \text{ Renhold}_i + 0,0002 \text{ WiFi}_i \\
 & + 0,050 \text{ Bærekraft}_i + \varepsilon_i
 \end{aligned}$$

A7.2 Statistiske analyser

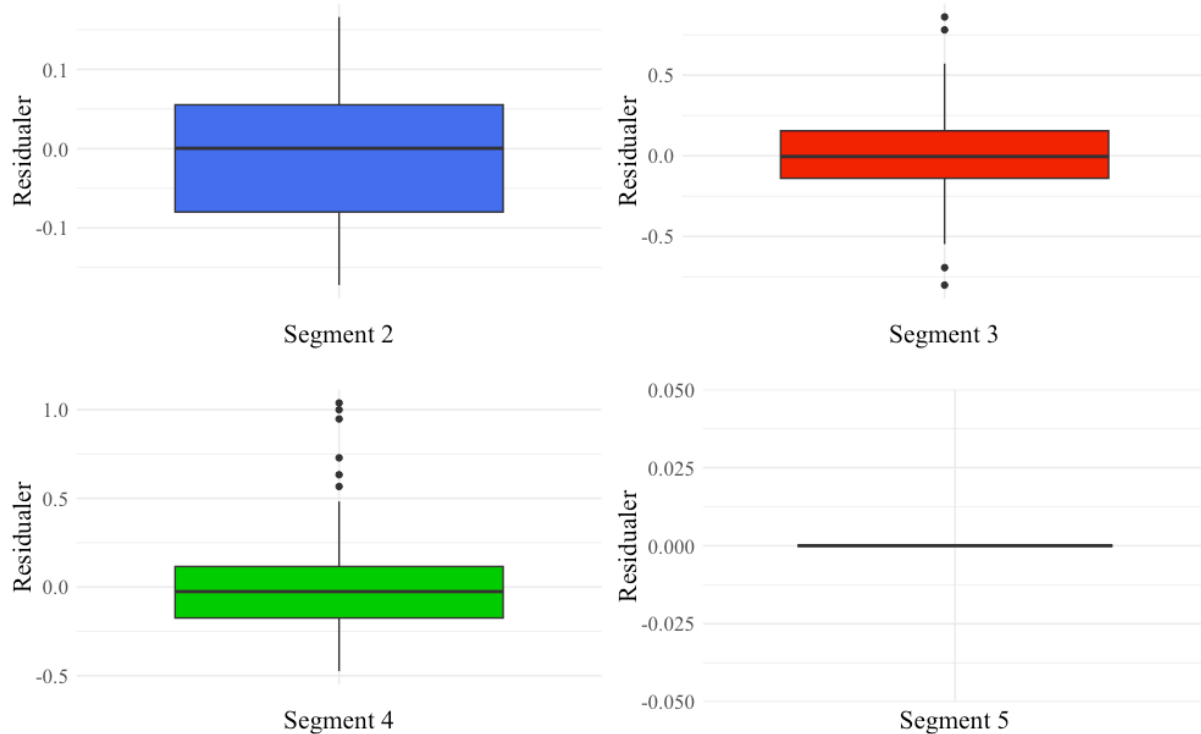


Figur A7.1: Bokplott av residualene for By og DN.



Figur A7.2: Bokplott av residualene for hvert segment By.

Figur A7.2 illustrerer hvordan segment tre og fire danner grunnlag for mer nøyaktige analyser. Disse har langt flere observasjoner sammenlignet med segment to og fem, uavhengig av region. Intervallene rundt gjennomsnittet (tykk horisontal linje) er normale og relativt mindre enn for segment fem. Vi ser også at det ligger en ekstrem observasjon på nesten -0,50 i segment tre og en på rundt 0,50 i segment fire.



Figur A7.3: Boksplokk av residualene for hvert segment i DN.

Det fremkommer av Figur A7.3 at residualene for DN følger et invertert mønster fra de åtte største. Segment to har nå et bredt intervall sammenlignet med segment tre og fire, mens dette gjaldt segment fem blant de åtte største. Nå er også boksplokket for segment fem manglende grunnet fraværende datagrunnlag, på samme måte som for segment to blant de åtte største. Vi mener at disse to boksplokkene illustrerer hvordan segment tre og fire dominerer våre analyser.

Tabell A7.3: Residualer for By og DN i Modell 1.

	Min.	K1	Median	K3	Maks.	Standardfeil	F-stat.
By	(0,454)	(0,097)	0,008	0,103	0,586	0,175	21,930*
DN	(0,756)	(0,153)	-0,023	0,154	1,084	0,245	23,170*
Note:	*p<0,01; negative verdier i parentes						

Tabell A7.4: Residualer for By-segmentene i Modell 2.

Segment	Min.	K1	Median	K3	Maks.	Standardfeil	F-stat.
2	-	-	-	-	-	-	-
3	(0,375)	(0,092)	0,046	0,119	0,320	0,186	2,851*
4	(0,469)	(0,084)	(0,007)	0,093	0,480	0,165	7,743*
5	-	-	-	-	-	-	-
Note:	*p<0,01; negative verdier i parentes						

Tabell A7.5: Residualer for DN-segmentene i Modell 2.

Segment	Min.	K1	Median	K3	Maks.	Standardfeil	F-stat.
2	(0,172)	(0,080)	0,00044	0,055	0,166	0,136	3,625*
3	(0,802)	(0,140)	(0,004)	0,155	0,861	0,238	8,007*
4	(0,474)	(0,175)	(0,026)	0,116	1,038	0,251	6,594*
5	-	-	-	-	-	-	-
Note:	*p<0,01; negative verdier i parentes						

Tabell A7.6: Durbin-Watson test statistikk.

Modell	Durbin-Watson	p-verdi
By	1,212	0,389
DN	0,523	<0,01

Tabell A7.7: VIF-matrise for By og DN.

	By	DN
Segment	1,45	1,23
Beliggenhet	1,31	1,60
Fasiliteter	16,71	11,47
Komfort	37,74	17,67
Personale	2,09	2,20
Renhold	27,48	14,93
Bærekraft	1,06	1,04
Gjennomsnitt	11,20	6,42

Tabell A7.8: VIF-matrise med elimineringsmetode.

	VIF 1	VIF 2	VIF 3	VIF 4	VIF 5	VIF 6	VIF 7	VIF 8
Segment	1,40	1,39	1,31	1,37	1,40	1,35	1,35	1,39
Beliggenhet	1,50	1,53	1,50	1,37	1,43	1,46	1,46	-
Fasiliteter	12,47	0,84	-	10,14	12,45	11,88	4,56	12,28
Komfort	20,22	15,35	20,00	-	20,12	11,39	-	10,84
Personale	2,21	2,28	2,10	2,20		2,10	2,07	2,11
Renhold	16,52	13,03	16,42	9,31	15,65	-	-	16,09
WiFi	1,40	-	1,30	1,30	1,30	1,30	1,29	1,30
Bærekraft	1,08	1,07	1,07	1,08	1,08	1,07	1,07	1,07

I Tabell A7.9 har vi kjørt regresjonen for Modell 1 og testet hva VIF-verdien blir ved å iterativt ekskludere hver variabel. I VIF 2 har vi eksempelvis fjernet den uavhengige variabelen *WiFi*.

Tabell A7.9: VIF-matrise for By på segmentnivå.

Segmentnivå	2	3	4	5
Beliggenhet	-	29,86	28,19	99,5
Fasiliteter	-	27,86	19,05	63,81
Komfort	-	12,08	13,19	20,69
Personale	-	1,89	1,83	15,88
Renhold	1,00	1,83	1,62	3,25
WiFi	-	1,28	1,28	1,71
Bærekraft	-	1,10	1,06	1,63

Tabell A7.10: VIF-matrise for DN på segmentnivå.

Segmentnivå	2	3	4	5
Beliggenhet	27,40	22,64	10,35	-
Fasiliteter	21,73	17,57	10,18	6,57
Komfort	17,36	11,99	7,42	-
Personale	2,14	2,22	2,35	6,23
Renhold	1,73	1,56	1,38	-
WiFi	1,34	1,26	1,28	1,39
Bærekraft	1,18	1,03	1,05	1,00

Tabell A7.11: Standardavvik for variablene i Modell 2.

Segment	2	3	4	5	2	3	4	5
Beliggenhet	-	0,075	0,032	-	0,068	0,031	0,042	-
Fasiliteter	-	0,162	0,127	-	(.147	0,071	0,089	-
Komfort	-	0,301	0,188	-	0,192	0,115	0,119	-
Personale	-	0,067	0,056	-	0,081	0,037	0,066	-
Renhold	-	0,293	0,163	-	0,214	0,105	0,125	-
WiFi	-	0,081	0,046	-	0,039	0,021	0,032	-
Bærekraft	-	0,050	0,026	-	0,038	0,030	0,027	-