

MASTER I SAMFUNNSØKONOMI

«Neste stasjon er Oslo S»

Påvirkes Boligpris av Endring i Reisetid med Jernbane?

Tonje Holmsen

Institutt for samfunnsøkonomi, NTNU

Desember 2013



FORORD

«Jeg har fullført masterstudiet i samfunnsøkonomi og er utrolig stolt!»

Nå som studietiden snart er ferdig føler jeg en enorm lettelse, frykt, takknemlighet og ikke minst lykke for at jeg har klart å gjennomføre. Som min farmor sier: «Nå har du gått på skole i 20 år, noe må du ha lært»

Masterstudiet har vært krevende, lærerikt, morsomt og veldig spennende. Uten deg Marianne Lundgaard hadde dette ikke vært resultatet. Du fortjener en stor takk for at du er den du er og for at du utgjør den ene halvdel av et ekstremt bra team!

Jeg vil også takke min dyktige veileder Jørn Rattsø for at du er lett å prate med, gir konstruktive tilbakemeldinger og alltid er tilgjengelig.

Takk til mamma og pappa for at dere støtter meg uansett hva jeg finner på. Takk til onkel Erik og mine gode venner Erlend Dancke Sandorf og Kim Andre Skogstrand for å ha hjulpet meg med lesing av korrektur, formatering og innspill til hva som kan gjøre oppgaven bedre. Det har vist seg at dere ikke bare er snille, men veldig flinke!

I de tyngste stundene har jeg erfart at det finnes engasjerte og dyktige fagfolk der ute som er villig til å bruke mange timer på en helt vanlig samfunnsøkonomistudent. Ulf Erik Bakke i NSB og Christer Tonheim i NAF er blant disse menneskene. Det er betryggende å vite for en som skal ut i arbeidslivet at det finnes så engasjerte og hyggelig mennesker som dere.

Jeg vil også takke deg Niclas Herlovsen for at du alltid stiller opp og har gitt mye for at jeg skulle realisere studiene.

Om ikke annet, skal det i hvert fall ikke være noe tvil om at alle feil er min egne.

Tonje Holmsen

Oslo, 1. Desember 2013

Innholdsfortegnelse

1. INNLEDNING	1
1.1 Problemstilling	2
1.2 Disposisjon	2
2. TEORETISK TILNÆRMING	3
2.1 Verdsetting av boligattributter ved hedonisk metode	3
2.2 Verdsetting av reisetid ved hedonisk metode.....	5
2.2.1 Husholdningenes budfunksjon for reisetid og tilgjengelighet.....	7
3. EMPIRISK MODELLERING	10
3.1 Anvendt modell	13
4. ØKONOMETRISKE UTFORDRINGER.....	15
4.1 Multikollinearitet.....	16
4.2 Endogenitet i modellen.....	17
4.2.1 Utelatt variabelproblem.....	17
4.2.2 Målefeil	19
4.2.3 Simultanitet	21
4.3 Difference in difference	21
4.3.1 Utbygging av Romeriksporten	22
4.3.2 Svakheter ved difference in difference.....	24
5. DATABESKRIVELSE	26
5.1 Presentasjon av data	26
5.1.1 Den avhengige variabelen	26
5.1.2 De uavhengige variablene	27
6. RESULTAT	33
6.1 Referansemodell.....	34
6.1.1 Effekten av reisetid, tilgjengelighet og inntekt på boligpriser	35
6.1.2 Effekten av kontroll variable på boligpriser.....	36
6.2 Robusthetstester	38
6.2.1 Effekten av MKM med postnummerdummyer	38
6.2.2 Effekten av MKM med kommunedummyer	39
6.2.3 Effekten av MKM uten geografiske variabler.....	40
6.2.4 Oppsummering	41
6.3 Estimeringsresultater ved difference-in-difference	42

6.4 Robusthetstester	43
6.4.1 Effekten av en endring i reisetid	43
6.4.2 Forventningseffekt	44
6.4.3 Oppsummering	44
7. KONKLUSJON	45
8. REFERANSER	47
APPENDIKS	51

TABELLER

1. Estimeringsresultater for MKM.....	34
2. Estimeringsresultater for difference in difference.....	42
A.1 Variabelbeskrivelse.....	51
A.2 Deskriptiv statistikk.....	52
A.3 Korrelasjonsmatrise for ikke-geografiske variabler.....	53
A.4 Korrelasjonsmatrise for geografiske variabler.....	54
A.5 Estimeringsresultater for MKM.....	55
A.6 Estimeringsresultater for difference in difference.....	57

FIGURER

1. Husholdningens budfunksjon.....	9
2. Reisetid med jernbane til CBD.....	28
3. Reisetid med bil til CBD.....	30
A.7 Antall nybygde boliger.....	58
A.8 Toglinjer NSB.....	59

KART

1. Studieområdet oppdelt i kommuner.....	13
--	----

1. INNLEDNING

Jernbaneinvesteringer gjort i Norge de siste 20 årene har hatt som mål å forbedre jernbanetransporten. En slik forbedring vil være med å øke tilgjengeligheten ved redusert reisetid i bestemte områder. Det er naturlig å tenke seg at slike transportinvesteringer har konsekvenser for de nærliggende områdene som vil gi fordeler som reflekteres i boligprisene.

I en rapport gjort av transportøkonomisk institutt som kartlegger reisemønster ved langpendling innenfor intercitytriangelet¹ viser at transportmiddelvalget avhenger av reisetiden. Til tross for at bil fremdeles er hoved transportmiddelet ved pendling til og fra arbeidsplassen anses togtilbudet som en nødvendig forutsetning for å kunne pendle. Rapporten fastsetter at når reisetiden med tog konkurrerer med reisetiden med bil vil flere velge å reise med tog (Tøi rapport 1201/2012). Et område med et velutviklet togtilbud og kort reisetid til bykjernen vil være en attraktiv kvalitet for husholdninger i en boligkjøpsituasjon.

Det er ikke bare kortere reisetid som er en attraktiv kvalitet ved en bolig. Husholdningene prefererer en rekke kvaliteter som spiller inn ved fastsettelse av boligpris. Denne analysen ønsker å finne husholdningenes marginale betalingsvilje² for redusert reisetid ved å bruke hedonisk metode. Denne markedsbaserte verdsettingsmetoden danner det teoretiske fundamentet for vår multippel lineære regresjon som anvendes for å kunne estimere den kausale sammenhengen mellom boligpris og reisetid med jernbane.

Vår analyse tar utgangspunkt i et bestemt antall kommuner i Stor-Oslo. Stor-Oslo inkluderer vanligvis Oslo kommune, hele Akershus fylke og flere kommuner i fylkene Buskerud, Vestfold og Østfold. Utgangspunktet for denne analysen er 11 tilfeldig utvalgte kommuner som har en jernbanestasjon. Kommunene som observeres er Opepgård, Moss, Rygge, Eidsberg, Skedsmo, Ullensaker, Bærum, Drammen, Holmestrand, Nittedal og

¹ Intercity-triangelet er vanlig betegnelse på jernbanelinjene mellom Skien, Lillehammer og Halden gjennom Oslo, og omfatter Vestfoldbanen, Østfoldbanen, Gardermobanen og Dovrebanen til Lillehammer (JBV.no).

² Hvilken pris husholdningen er villig til å betale for en enhet til av godet.

Eidsvoll. Studiet har en monosentrisk tilnærming der vi antar at kommunenes befolkning er avhengig av å pendle inn til Oslo på grunn av arbeid. Med dette som utgangspunkt ønsker vi å finne den kausale effekten av reisetid på boligpriser.

1.1 Problemstilling

Det er gjort en rekke studier på sammenhengen mellom avstand til bykjernen og boligpriser internasjonalt og til dels nasjonalt. Flesteparten av disse studiene har konkludert med at det finnes en signifikant sammenheng mellom avstand og boligpris. Det forekommer færre studier som har som hensikt å finne den kausale sammenhengen mellom reisetid til bykjernen og boligpriser. På grunn av stor satsing gjennom nasjonal transportplan vil det bli enklere å pendle fra enkelte kommuner utenfor bykjernen, slik at reisetid vil ha en stor betydning for pendlere i årene som kommer. Utbygging av dobbeltspor og jernbanetunnel på strekningen fra Oslo til Ski vil medføre kortere reisetid fra ulike steder i Akershus til Oslo S enn fra flere steder innad i Oslo (jernbaneverket.no). Det er derfor interessant og relevant å studere om det er en kausal sammenheng mellom reisetid og boligpriser. Dette gir grunnlag for undersøkelsens problemstilling:

«Påvirkes boligpriser av endring i reisetid med jernbane?».

1.2 Disposisjon

I kapittel 2 utledes teorigrunnlaget for analysen gjennom en hedonisk modell for verdsetting av reisetid til bykjernen og avstand til kollektivtilbud. I kapittel 3 presenteres vår empiriske modellspesifikasjon. I kapittel 4 diskuterer de økonometriske utfordringene ved estimering med minste kvadraters metode (MKM). Deretter drøftes difference in difference som en alternativ metode for å håndtere problemet med endogenitet. I kapittel 5 presenterer estimeringsresultatene ved MKM og difference in difference samtidig som at vi tester for robustheten til resultatene. Til slutt vil vi oppsummere og konkludere i kapittel 7.

2. TEORETISK TILNÆRMING

For å undersøke analysens problemstilling anvendes teori fra Rosen om hedonisk metode for estimering av attributtpriser presentert av Liv Osland (2001). Denne markedsbaserte metoden er utgangspunktet for min empiriske analyse og utleder husholdningens maksimale betalingsvillighet for reisetid og avstand til kollektivtilbud. Ved å ta utgangspunkt nyklassisk teori grunnlagt av Alonso (1964) skal vi ved hedonisk metode verdsette reisetid til bykjernen og tilgjengelighet til kollektivtilbudet. Metoden gir oss mulighet til å utvikle en boligprismodell som forklarer hvilke boligattributter som husholdningene verdsetter og hvor mye man verdsetter de ulike attributtene.

2.1 Verdsetting av boligattributter ved hedonisk metode

Den hedoniske metoden baserer seg på teorien om at prisen på en vare i et gitt marked reflekterer kvaliteten på en rekke attributter ved varen, og at hvert attributt har en implisitt pris. De fleste hedoniske studier baserer seg på boligmarkedet. Det var Lancaster (1966) som først introduserte denne måten å tenke på et gode på. Han hevdet at det ikke er godet i seg selv som gir konsumenten nytte, men en rekke attributter eller kvaliteter knyttet til godet. Rosen (1974) baserte seg på dette da han utviklet det teoretiske grunnlaget for hedonisk metode.

Det er naturlig å tenke at prisen på en bolig kan bestemmes av blant annet størrelsen, antall soverom, alder på boligen, boligtype og som denne oppgaven ønsker å belyse plassering av boligen gitt ved reisetid inn til byens sentrum (CBD)³ og avstand til kollektivtilbudet. Grunntanken er at dersom to boliger er identiske på alle måter, men at den ene boligen har mye kortere reisetid til CBD og bedre tilgjengelighet, vil denne boligen trolig ha en høyere pris. Prisforskjellen mellom de to boligene vil da gi oss den implisitte prisen som også reflekterer konsumentenes marginale betalingsvillighet for godet (Rosen 1974).

Bolig viser seg å være et komplekst gode som er vanskelig å konkretisere i en enkel modell. Ved å ta i bruk en boligmarkedsmodell må vi foreta enkelte forutsetninger. I den forbindelse er det sentralt og ta utgangspunkt i en modell som har sitt utspring i nyklassisk

³ CBD (central business district) forklares som sentrum i byen. Dette området identifiseres ved at det er preget av butikker og kontorlokaler. Uttrykket vil i oppgaven bli referert til som Oslo sentralstasjon.

lokaliseringsteori for byområder. Gjennombruddet for denne teorien forbindes med «Access-Space-Trade-Off» modellen (Alonso, 1964). I sin mest grunnleggende form tar den utgangspunkt i et homogent landskap hvor det finnes ett bysenter, der all sysselsetting foregår. Den klassiske modellen til Alonso-Muth-Mills forklarer hvordan boligprisene varierer innenfor et storbyområde når man kun ser på beliggenheten til boligen. Kjernen i modellen er at plassering av bolig nær CBD reflekteres i høyere boligpriser. Det er naturlig å tenke seg at en reduksjon i reisetiden vil få samme effekt på boligprisene (Baum-Snow, 2006).

Rosens modell er statisk og tar utgangspunkt i at et gode kan betraktes som en vektor bestående av n objektivt målte attributter $z = (z_1, \dots, z_n)$. For godet «bolig» kan attributtene deles opp i to hovedgrupper:

- ✓ Attributter knyttet til selve boligen (z_{bj}) som for eksempel boligareal.
- ✓ Attributter knyttet til boligens geografiske plassering (z_{lj}), som for eksempel avstandsvariabler og sosiale faktorer.

Vi antar at boligpris er en funksjon av ulike boligattributter z , både av geografisk (z_{lj}) og ikke-geografisk (z_{bj}) karakter. Denne modellen ønsker å finne husholdningens (h) maksimale betalingsvilje for redusert reisetid til bykjernen (z_{tr}) og avstand til kollektivtransporttilbudet (z_{φ}). Dette gir oss den implisitte prisfunksjonen for boliger gitt ved:

$$P_h(z) = P_h(z_{bj}, z_{sj}) = P_h(z_{tr}, z_{\varphi}, z_{kj}) \quad (1)$$

$$z_{kj}, j = 1, \dots, n \quad \forall j \neq z_{tr}, z_{\varphi} \quad k \in [b, l]$$

2.2 Verdsetting av reisetid ved hedonisk metode

Modellen antar at hver husholdning (h) har en nyttefunksjon og tilpasser seg slik at nytten maksimeres.⁴

$$U_h(c, z) \tag{2}$$
$$c > 0, z > 0$$

Husholdningens nyttefunksjon (U_h) blir gitt av konsum av ulike attributter knyttet til boligen (z) og konsum av alle andre goder med unntak av godet bolig (c).

Nyttefunksjonen antas å være strengt konkav som betyr at begge godene er normale.

Nytten maksimeres under budsjettbetingelsen

$$w_h = c + P(z_{tr}, z_\varphi, z_{kj}) + w_h A(z_{tr}, z_\varphi) \tag{3}$$

der

$$A(z_{tr}, z_\varphi) = \min[z_{tr}, z_\varphi]$$

$$A(z_{tr}, z_\varphi) \leq 0$$

$A(z_{tr}, z_\varphi)$ er et mål på boligens plassering der z_φ er attributtet for tilgjengelighet og angir lengden fra boligen til det nærmeste kollektivknutepunkt, og z_{tr} er boligattributtet for reisetid til CBD. Reisetiden øker med økt avstand (r) til CBD og økt t , der $t = \frac{1}{v}$ og representerer den inverse av farten (v) (Baum-Snow, 2006).

$A(z_{tr}, z_\varphi)$ forutsetter at husholdningene (h) velger å bosette seg der reisetiden (z_{tr}) og avstanden til kollektivtilbudet (z_φ) er minst. Intuisjonen bak $A(z_{tr}, z_\varphi) \leq 0$ er at redusert fart (v), økt avstand til CBD (r) og økt avstand til kollektivtilbudet (φ) fører til at $A(z_{tr}, z_\varphi)$ minker. Dette betyr at husholdningene preferer redusert reisetid til CBD og økt

⁴ En husholdning regnes som alle personer som bor i samme bolig (regjeringen.no). For enkelthetsskyld antar vi at en husholdning består av en person i denne analysen.

tilgjengelighet til kollektivtransporttilbudet. Dersom $A(z_{tr}, z_{\varphi}) = 0$ er boligen plassert i CBD (Ibid).

Budsjettbetingelsen gitt av studiet til Baum-Snow (2006) antar at reisetid og inntekt er eksogent gitt i modellen. Vi ser at husholdningen bruker all sin tid på enten å pendle eller arbeide, slik at økt reisetid og avstand til kollektivtilbudet utgjør en kostnad for husholdningen. Desto mer tid man bruker på å pendle, desto mindre tid bruker man på arbeid. Dette reflekteres i husholdningens inntekt (w_h) ved at den reduseres ved økt reisetid og avstand (Ibid). Husholdningen bruker hele sin inntekt på konsum av boligattributter, andre goder enn bolig og pendling. Teorien bygger på at prisfunksjonen $P(z_{tr}, z_{\varphi}, z_{kj})$ har en form som sikrer en entydig indre løsning på nyttemaksimeringsproblemet. Prisen på annet konsum (c) er konstant og lik 1 (Osland, 2001).

I optimum vil den marginale substitusjonsraten mellom z og c være lik den partiell deriverte av prisfunksjonen med hensyn på de respektive boligattributter. For å finne den marginale nytten husholdningen har av godet reisetid og tilgjengelighet relativt til andre goder holdes de andre boligattributtene (z_{kj}) konstant.

$$\frac{\partial U / \partial z}{\partial U / \partial c} = P(z_{tr}, z_{\varphi}, \bar{z}_{kj}) \quad (4)$$

Høyresiden i likning (4) svarer til de hedoniske prisene på boligattributtene reisetid til CBD og avstanden til kollektivtilbudet. Den angir helningen til prisfunksjonen i punkter for optimal mengde av redusert reisetid til CBD og redusert avstand til kollektivtilbudet ($A(z_{tr}, z_{\varphi})$).

Ved å totaldifferensiere likning (4) med hensyn på $A(z_{tr}, z_{\varphi})$ får vi:

$$\frac{\partial P(z_{tr}, z_{\varphi}, \bar{z}_{kj})}{\partial A(z_{tr}, z_{\varphi})} > 0 \quad (5)$$

Likning (5) forteller oss at redusert reisetid til CBD og redusert tilgjengelighet til kollektivtransporttilbudet øker boligprisen (Brueckner, 1986) Den intuitive forklaringen bak den geografiske oppførselen til $P(z_{tr}, z_{\varphi}, \bar{z}_{kj})$ er at husholdninger med lang reisetid til CBD og redusert tilgjengelighet til kollektivtransporttilbudet må få kompensasjon for lang og kostbar pendling⁵. Denne kompensasjonen tar form i reduserte boligpriser (Ibid). En økning i $A(z_{tr}, z_{\varphi})$ viser også at pendling fører med seg en kostnad i form av redusert inntekt. Desto mer tid man bruker på å reise desto mindre tid bruker man på arbeid.

Videre undersøkes effekten av en økning i inntekt på boligpris. Ved å totaldifferensiere likning (4) med hensyn på w_h får vi:

$$\frac{\partial P(z_{tr}, z_{\varphi}, \bar{z}_{kj})}{\partial w_h} > 0 \quad (6)$$

Likning (6) viser at en økning i inntekt (w_h) fører til en økning i boligpris. Enkel mikroøkonomi for normale goder presiserer at en økning i inntekt fører til økt etterspørsel som igjen øker prisen på godet. Med utgangspunkt i at husholdningene verdsetter plassering av bolig, vil en inntektsøkning føre til at husholdningen ønsker å flytte til et område med kortere reisetid til CBD. Når boligprisen presses opp vil den positive inntektseffekten spises opp av økte boligprisene slik at husholdningene bosetter seg i et område med lenger reisetid fra CBD. Dette flater ut effekten av økt inntekt og totaleffekten vil avhenge av hvor mye inntekten øker i forhold til boligpris. Vi ser at boligpris og inntekt vokser i samme retning.

2.2.1 Husholdningenes budfunksjon for reisetid og tilgjengelighet

Husholdningenes budfunksjon defineres som maksimal betalingsvillighet for $A(z_{tr}, z_{\varphi})$ når nyttenivå og inntekt holdes konstant. Budfunksjonen er en indifferenskurve som gjør

⁵ Med kostbar pendling refererer man direkte til kostnaden ved reisetid, som fremgår av budsjettbetingelsen.

det mulig å studere boligprisen ved alternative kombinasjoner av boligens plassering i relasjon til subjektive priser og markedspriser, heller enn til et annet gode.

$$\theta_h = \theta \left(A(z_{tr}, z_\varphi)_h, \bar{z}_{kj}, w_h, U_h \right) \quad (7)$$

Husholdningenes budfunksjon for plassering av boligen, er det husholdningen er villig til å betale for plassering av boligen gitt deres nytte.

Budfunksjonen kan utledes ved å ta utgangspunkt i de optimale verdiene for konsum av boligattributter z^* og for konsum av andre goder c^* , slik at

$$c^* = w_h \left(1 - A(z_{tr}^*, z_\varphi^*) - P(z_{tr}^*, z_\varphi^*, \bar{z}_{kj}) \right)$$

Innsatt i nyttefunksjonen gir dette

$$\begin{aligned} U_h &= \left(z^*, w_h \left(1 - A(z_{tr}^*, z_\varphi^*) \right) - P(z_{tr}^*, z_\varphi^*, \bar{z}_{kj}) \right) = U_h^* \\ &= U \left(z, 1 - A(z_{tr}^*, z_\varphi^*) \right) - \theta_h \end{aligned} \quad (8)$$

Dersom nyttenivået er konstant lik U^* og inntekt er gitt er det rimelig å forutsette at betalingsvilligheten θ_h , er den prisen man faktisk betaler, $P(z_{tr}^*, z_\varphi^*, \bar{z}_{kj})$. Uttrykket (8) definerer implisitt relasjonen for maksimal betalingsvillighet for andre plasseringer av boligen enn den optimale.

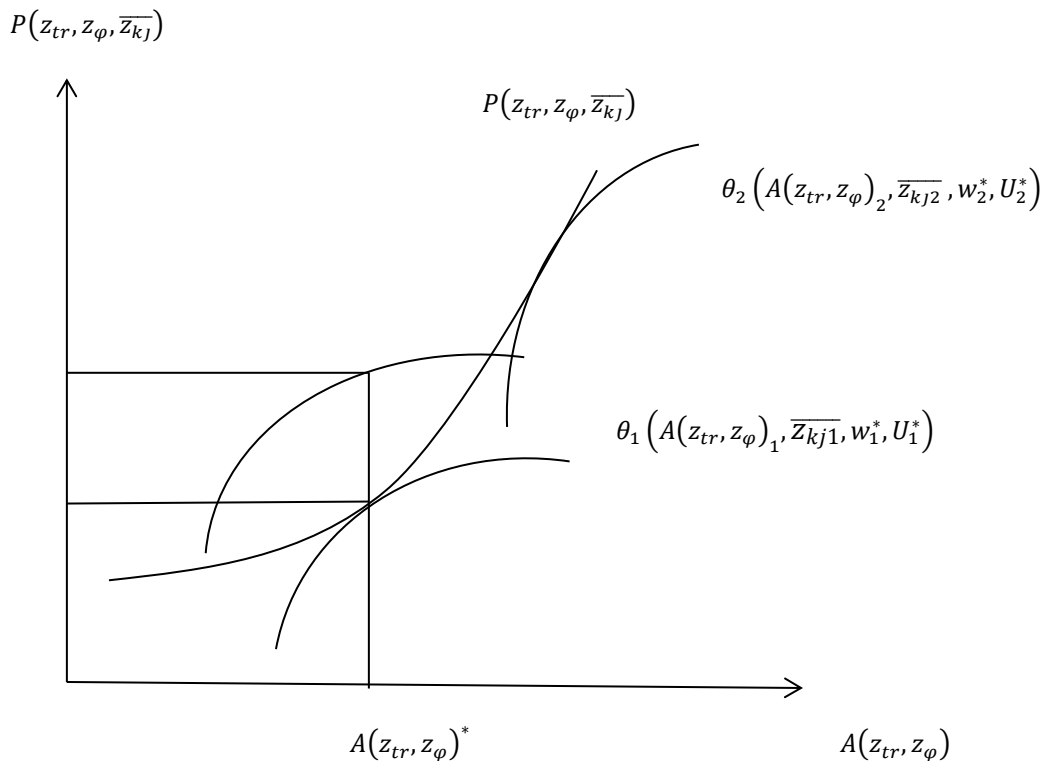
Budfunksjonen (7) vil variere med valgt inntekt og nyttenivå. Ved implisitt derivasjon av (7) gir følgende resultat

$$\frac{\partial \theta_h}{\partial z} = \frac{\frac{\partial U_h}{\partial z}}{\frac{\partial U_h}{\partial c}} > 0 \quad (9)$$

Likning (9) tolkes som maksimal betalingsvillighet for en partiell økning i $A(z_{tr}, z_{\varphi})$ da vi antar at øvrige boligattributter holdes konstant (\bar{z}_{kj}). Siden nyttefunksjonen er strengt konkav er det mulig å vise at $\partial^2 \theta_h / \partial z^2 < 0$ (Rothenberg et al. 1991). Dette betyr at betalingsvilligheten er positiv men avtakende for en partiell økning i $A(z_{tr}, z_{\varphi})$.

Grafisk gir budfunksjonen et sett av indifferenskurver til hvert nyttenivå. Den vertikale aksene måler betalingsvilligheten og den horisontale aksene måler mengden av attributtet reisetid og avstand til kollektivtilbudet. Figuren illustrerer hvordan to husholdninger med ulik marginal betalingsvilje for reisetid til CBD og tilgjengelighet til kollektivtilbudet tilpasser seg gitt den implisitte prisen for godene.

Figur 1: Husholdningenes budfunksjon



Det optimale for hver husholdning er å tilpasse seg slik at deres marginale betalingsvilje er lik den marginale prisen for en bedre plassering. Figuren viser at den marginale

betalingsviljen for attributtene reisetid til CBD og avstand til kollektivtilbudet er avtagende. Når reisetid og avstand reduseres, vil $A(z_{tr}, z_{\varphi})$ øke og boligprisen reduseres.

I figuren kan man se at for en gitt pris maksimerer husholdningene sin nytte ved å velge det alternativet som gir best plassering av boligen. Budfunksjon θ_1 representerer en husholdning som prefererer kortere reisetid til CBD og kortere avstand til kollektivtilbudet enn husholdningen med budfunksjon θ_2 . Ved å se på husholdningenes atferd gjennom enkel mikroøkonomi ser vi at husholdningene velger den plasseringen som de har råd til og som maksimerer deres nytte (Brueckner, 1986).

3. EMPIRISK MODELLERING

Det er gjort en rekke studier på kapitalisering av arbeidsplassstilgjengelighet gjennom boligpriser. Allerede i 1964 presenterte Alonso sin grunnleggende tanke om at boligprisene ville falle ved økt avstand til CBD. Med utgangspunkt i «Access-Space-Trade-Off» modellen forsøkte Liv Osland og Inge Thorsen (2005) gjennom en hedonisk tilnærming å teste hypotesen om at pendlere prefererer plassering av boligen nærmest mulig arbeidsplassen. Datagrunnlaget for deres analyse besto av 2788 enkelt observasjoner for salg av bolig i den sørvestlige delen av Norge. Regionen består av 13 kommuner er delt i 98 postnumre med Stavanger som CBD. Analysen baserte seg på antagelsene om at all sysselsetting er plassert i bykjernen. Osland og Thorsen (2005) har hovedsakelig fokusert på effekten av generelle geografiske kvaliteter snarere enn å forklare boligprisene i en bestemt region. De bruker avstand og reisetid til CBD som mål for å forklare den geografiske variasjonen i deres data.

I likhet med analysen til Osland og Thorsen skal vi ta for oss forholdet mellom bolig- og arbeidsmarkedet i et monosentrisk studie. Dette gjøres ved å estimere den implisitte prisfunksjonen i likning (1) ved å bruke hedonisk metode. På denne måten finner vi den estimerte betalingsvilligheten for ulike attributter i de ulike kommunene som studeres. Vårt bidrag er å se på reisetid til CBD som et mål på arbeidsplassstilgjengelighet og som forklarer en del av den geografiske variasjonen i boligprisene. Undersøkelsen til Osland og Thorsen (2005) ser bort i fra andre geografiske boligattributter som bidrar med å

forklare regional variasjon i boligprisene. De antar at effekten av disse attributtene forklares av variablene for plassering av boligen, som reisetid og avstand til CBD. Dette skiller seg fra vår analyse som har flere variabler som inkluderer informasjon om det geografiske området som studeres.

Analysen vår tar utgangspunkt i 11 norske kommuner i Stor-Oslo regionen som har minst en jernbanestasjon plassert i kommunen. Kommunene som observeres er Oppegård, Moss, Rygge, Eidsberg, Skedsmo, Ullensaker, Bærum, Drammen, Holmestrand, Nittedal og Eidsvoll. Kommunene som er valgt ligger innenfor en region der det er naturlig å anta at kommunens bosatte pendler til Oslo, hovedsakelig på grunn av arbeid. Undersøkelsen baserer seg på at all reise foretatt av husholdningen skjer til og fra arbeidsplassen.

I en rapport gjort av transportøkonomisk institutt for langpendling innenfor intercity triangelet viser at bilbruken økte sterkt på midten av 80-tallet på bekostning av alle andre type transportmidler. Denne utviklingen fortsatte utover 90-årene og bil er fremdeles hoved transportmidlet for beboere i Oslo og Akershus. Variasjonen i transportmiddel er størst i CBD og avtar i områdene rundt (rapport 1201/2012). Samme rapport viste også at husholdningene velger det transportmiddelet som krever minst reisetid. Rapporten bekrefter at flere reiser med tog dersom de har kort avstand hjemmefra til jernbanestasjonen. Andelen som reiser med tog synker gradvis med økende distanse mellom bosted og jernbanestasjon, mens bilandelen øker. Dette er i tråd med teorien presentert ovenfor om at husholdningene preferer kort reisetid til CBD samt god tilgjengelighet til jernbanestasjonen. Dette betyr at reisetid er viktig for pendlere.

For å estimere om betalingsvilligheten av reisetid til CBD med jernbane dannes en hedonisk regresjon. Ved å bruke Minste kvadraters metode (MKM) skal vi teste analysens hovedhypotese om at redusert reisetid med jernbane øker boligprisene. Mer spesifikt skal vi estimere den generelle modellen gitt ved likning (1). For å estimere (1) med MKM må den defineres additivt. I tillegg blir det et spørsmål om hvilke boligattributter som skal inkluderes i analysen og hvilken funksjonsform som skal benyttes. Alle disse valgene gir forutsetninger for den empiriske modellen, som kan være mer eller mindre realistiske.

Som forklart i den hedoniske pristeorien er den marginale betalingsvilligheten for redusert reisetid avtakende. Teorien brukte reisetid som en generell notasjon for all transport. Den anvendte modellen tar utgangspunkt i at husholdningen har to transportmiddel valg slik at arbeidsplassstilgjengeligheten er gitt ved reisetid med tog og reisetid med bil.

I motsetning til Osland og Thoresen (2005) inkluderer vår undersøkelse geografiske variabler som forklarer den regionale variasjonen i boligprisene i tillegg til reisetid. Mer spesifikt er dette variabler som omfatter sosiale faktorer i de ulike kommunene. Disse geografiske variablene sier noe om hvor attraktiv kommunen er og antas å ha effekt på boligprisen. Alle variablene i modellen, både geografiske og ikke geografiske utgjør de ulike modellspesifikasjonene som skal estimeres med MKM. Variablene som inkluderes presenteres i kapitel 6.

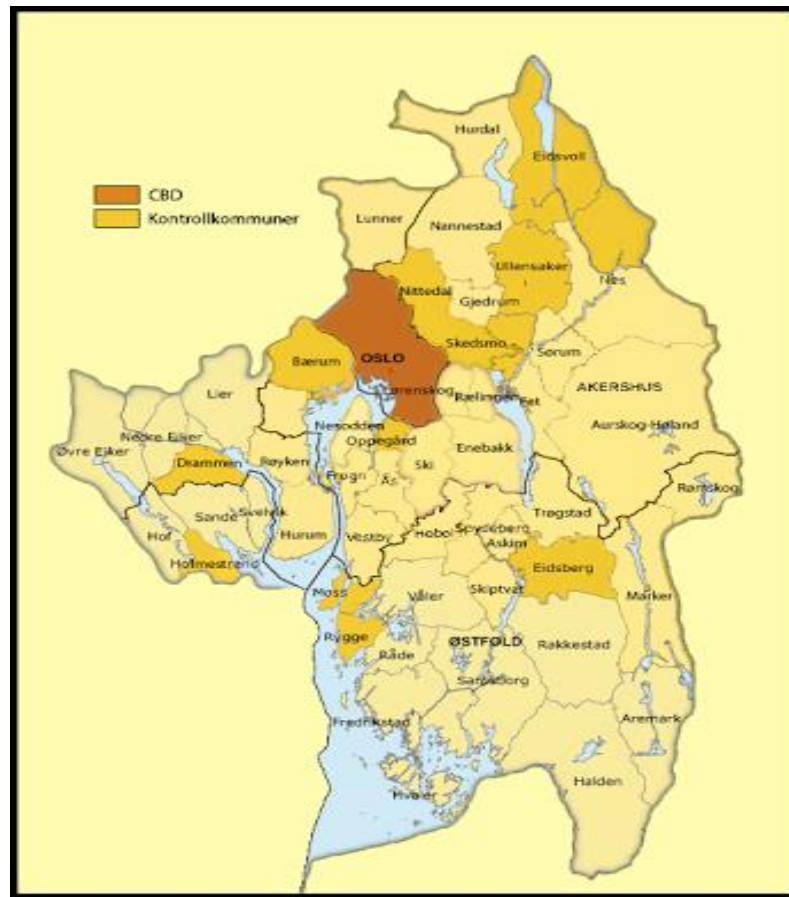
M1: Modell som forklarer geografisk variasjon med retningsdummyer og sosiale faktorer

M2: Modell uten geografiske variable, med unntak av reisetid og tilgjengelighet.

M3: Modell som forklarer geografisk variasjon med kommunedummyer og sosiale faktorer.

M4: Modell som forklarer geografisk variasjon med postnummerdummyer og sosiale faktorer.

Kart 1: Analysens studieområde, oppdelt i kommuner



3.1 Anvendt modell

Den anvendte modellen bygger på den implisitte prisfunksjonen i likning (1) som er gitt ved redusert form. Ved å legge til en rekke geografiske og ikke geografiske attributter vil vi danne en multippel lineær regresjonsmodell som anvendes for å kunne estimere den kausale sammenhengen mellom boligpris og reisetid med jernbane. Målet med analysen er å estimere en empirisk modell vi tror ligger så nært som mulig den sanne populasjonen og teste om redusert reisetid med jernbane gir en økt signifikant effekt på boligprisene. Den anvendte modellen følger den hedoniske metoden for attributtpriser. Den hedoniske regresjonen er formulert på følgende måte:

$$\ln P(z_{tr}, z_{\varphi}, z_{kj})_i = \beta_0 + \beta_1 \ln z_{tr(tog)_i} + \beta_2 \ln z_{tr(bil)_i} + \beta_3 \ln w_i + D_f \sum_{f=1}^4 z_{\varphi_i} + \beta_f \sum_{f=1}^{13} z_{kj} + u_i \quad (10)$$

Denne modellen skal forsøke å estimere den implisitte prisen for ulike boligattributter og fanger opp de sammenhengene som forklarer boligprisen. Fotskrift i referer til kommunen boligen er plassert (se A.1. for variabelbeskrivelse). De ulike boligattributtene som inkluderes i modellen er nærmere forklart i kapittel 6.

Det benyttes en dobbellogaritmisk funksjonsform ved estimering av modellen med unntak av variablene for barnehagedekning (z_{bhd}_i), anmeldte lovbrudd (z_{anm}_i) og arbeidsledighet (z_{arbled}_i), da disse måles i prosent av befolkningen. Vi skal se nærmere på disse variablene i kapittel 5. De sistnevnte variablene kombinerer log-transformerte og lineære ledd og kan tolkes som semi-elastisiteter. De andre koeffisientene tolkes som elastisiteter. Estimering av denne modellen gir oss grunnlag for å forkaste eller beholde analysens hovedhypotese.

H_0 : Redusert reisetid med jernbane har ikke positiv effekt på boligprisene ($\beta_1 \geq 0$)

H_1 : Redusert reisetid med jernbane øker boligprisene ($\beta_1 < 0$)

4. ØKONOMETRISKE UTFORDRINGER

I dette kapitlet diskuteres de økonometriske utfordringene ved denne empiriske undersøkelsen. Samtidig presenteres de ulike metodevalgene for estimering.

MKM vil bli brukt som utgangspunkt for metodediskusjonen. MKM er en metode for å estimere sammenhengen mellom den avhengige variabelen og forklaringsvariablene i en multipl linear regresjons modell for å minimere summen av kvadrerte avvik (Wooldridge, 2009). Denne metoden er utgangspunktet ved estimering av vår modell. Det er en rekke forutsetninger som må være oppfylt for at MKM skal gi konsistente og forventningsrette estimat.

- (1) $P(z_{tr}, z_{\varphi}, z_{kj})_i = \beta_0 + \beta_1 z_{tr(tog)_i} + \beta_2 z_{tr(bil)_i} + \dots + u_i$
- (2) $E(u_i) = 0$
- (3) $var(u_i, z_i) = \sigma^2$
- (4) $cov(u_j, u_i) = 0$ for $j \neq i$
- (5) $u_i \sim N(0, \sigma^2)$

Vi kan se av (1) at det forutsettes at populasjonsmodellen er lineær i parameterne. (2) sier at forventningen til restleddet er lik null og at estimatorene for parameterne må være konsistente. Estimatorene er konsistente når $P(\widehat{z_{tr}, z_{\varphi}, z_{kj}})$ konvergerer mot den sanne verdien til boligpris når antall observasjoner går mot uendelig. At estimatorene er konsistente ansees ofte som et minstekrav. (2) indikerer også at estimatorene er forventningsrette, det vil si at forventningen til estimatoren er lik estimatoren $E(\beta_i) = \beta_i$. (3) er forutsetningen om konstant varians. Denne impliserer at for alle verdier av de eksogene variablene er verdiene til $P(\widehat{z_{tr}, z_{\varphi}, z_{kj}})$ fordelt rundt forventningsverdien. Dersom denne antagelsen ikke holder er dataene heteroskedastiske. Forutsetning (4) sier at kovariansen mellom restleddene skal være null. Dersom vi tilfeldig plukker boligprisen i kommune i , skal denne være uavhengig av boligprisen i kommune j . Dersom denne forutsetningen brytes har vi autokorrelasjon i restleddene. (5) viser at restleddet og

dermed også boligpris skal være normalfordelt rundt sin forventningsverdi. Dersom forutsetningene (1) - (5) holder vil estimatene være BLUE (Gauss Markov teoremet). Det vil si at estimatorene har minste varians, er en lineær funksjon av observert data og forventningsrette. Gauss Markov teoremet indikerer også at estimatorene ikke må forveksles med de sanne populasjonsverdiene (Woolridge, 2009).

4.1 Multikollinearitet

Et økonometrisk eksperiment som dette kan være preget av multikollinearitet. Det vil si at det oppstår korrelasjon mellom to eller flere forklaringsvariabler i modellen. Dersom flere av disse forklaringsvariablene beveger seg i et systematisk mønster er dette å betrakte som multikollinearitet (Woolridge, 2009).

Dersom vi har et tilfelle med perfekt multikollinearitet er det eksakt lineær sammenhengen mellom to eller flere forklaringsvariabler i modellen, som ofte oppstår ved konstruksjon av data. Et eksempel på tilfelle med perfekt multikollinearitet er når vi introduserer like mange dummyvariabler som vi har kategorier (dummy variabel fellen). Perfekt multikollinearitet bryter med forutsetningene til MKM.

Vår analyse ser ut til å være preget av multikollinearitet da teorien for verdsetting av reisetid forutsetter at husholdningen bruker all sin tid på pendling og arbeid.

Husholdningene som er plassert i et område med lang reisetid til CBD substituerer bort noe av tiden som går med på reise til mindre tid på arbeidsplassen. Dette fører til en reduksjon i husholdningenes inntekt. En inntektsøkning for husholdningene fører til at de ønsker å flytte til et område med kortere reisetid til CBD. Denne intuisjonen viser at det er en sterk sammenheng mellom inntekt og reisetid til CBD, og styrker tanken om en negativ lineær sammenheng mellom disse variablene.

Av korrelasjonsmatrisen A.4 ser vi av korrelasjonskoeffisientene at inntekt og reisetid er relativt høye sammenliknet med flere av de andre forklaringsvariablene. En korrelasjonskoeffisient er et mål på i hvor stor grad variablene varierer «i takt». Disse koeffisientene har verdier fra -1 til 1, der -1 viser perfekt negativ korrelasjon og 1 viser perfekt positiv korrelasjon (Woolridge, 2009). Mer spesifikt ser vi at inntekt korrelerer

negativt med reisetid med tog (-0,53) og reisetid med bil (-0,47). Disse verdiene viser en relativt høy negativ korrelasjon sammenliknet med de andre variablene i modellen. Og disse negative verdiene bekrefter antagelsen om lineær sammenheng mellom inntekt og reisetid.

Multikollinearitet er vanskelig å definere og ved å tolke korrelasjonsmatrisene er det viktig å poengtere at multikollinearitet ikke bryter noen av antagelsene for MKM, da variablene ikke lider av perfekt multikollinearitet. utfordringen blir å finne om sammenhengene er slik at de vil påvirke kvaliteten på estimatet. Når estimatene er upresise med høy varians vil dette kunne skape "gale" fortegn og estimat som er følsomme for små endringer i dataene. Det kan også skape problemer ved at vi ikke får signifikante verdier samtidig som R^2 er høy uten at det er høy signifikans på parameterne til de eksogene variablene (Woolridge, 2009).

4.2 Endogenitet i modellen

Bolig er et svært komplekst gode når det kommer til prissetting. Det er vanskelig å lage en modell som klarer å fange opp alle de ulike forholdene som er med på å påvirke prisen på boligen. I tillegg til en god modellspesifikasjon kreves et godt datasett uten for mange feil og mangler. Dette gir modellformuleringen og datamaterialet en rekke utfordringer ved å estimere med MKM. Brudd på eksogenitetsforutsetningen er en stor utfordring når det kommer til hedonisk analyse (Graves et al. 1988). Dersom modellen er preget av endogenitet vil vi få forventningsskjev estimat. Dette problemet oppstår når forklaringsvariablene i modellen korrelerer med restleddet. Vi har tre kilder til endogenitet; utelatt variabelproblem, målefeil og simultanitet (Woolridge, 2009).

4.2.1 Utelatt variabelproblem

Siden bolig er et heterogent gode og alle husholdninger har ulik nytte og preferanser knyttet til de ulike kvalitetene ved en bolig, vil det være komplisert å identifisere og implementere alle relevante forklaringsvariabler som forklarer en del av variasjonen i boligprisene. Denne analysen prøver derfor å trekke frem de viktigste attributtene ved en bolig som gjør den attraktiv og forklarer boligprisen. Det er riktignok uklart hvilke

attributter som forklarer boligprisene på en mest nøyaktig måte og det er fare for å utelate relevante forklaringsvariabler i modellen. Vi skal se at modellen har utelatt flere variabler som trolig fanges opp av modellens restledd.

Datasettet fra eiendomsverdi AS er befasst med stor grad av manglende observasjoner på flere interessante variabler, som er med på å påvirke boligprisen. Dette er ikke nødvendigvis et problem hvis de manglende observasjonene er tilfeldig. Konsekvensen av dette er at estimatene blir mindre presise fordi at størrelsen på utvalget reduseres, men estimatene vil fremdeles være forventningsrette. Dersom manglene ikke er tilfeldig vil det føre til at utvalget som brukes i regresjonen ikke gjenspeiler populasjonen. Dette fører til brudd på forutsetning (4) om tilfeldig utvalg (Woolridge, 2009). Et problem i vårt datasett er at flere sentrale forklaringsvariabler er preget av manglende observasjoner i så stor grad at de utelates fra modellene. Vi velger å kutte variabler for tomteareal, antall etasjer, renovering av boligen, antall rom og eierform som optimalt ville vært representert i vår hedoniske modellspesifikasjon.

Vår modellspesifikasjon går utfra at husholdningene pendler med bil og tog. Det vil si at reisetidvariabler som omfatter transportmidler som buss, sykkel, moped og andre fremkomstmidler blir utelatt fra modellen. Dette er variabler som sier noe om tilgjengeligheten til CBD som trolig vil fanges opp av restleddet. I vår modell blir reisetid sett på som en kostnad for husholdningene ved at de må velge mellom å bruke tiden på å reise eller å jobbe. Lenger reisetid til CBD utgjør en større kostnad for husholdningene ved redusert arbeidstid og mindre inntekt. Pendling utgjør ofte en større kostnad enn kun reisetid i praksis ved å inkludere drivstoffutgifter, kapitalkostnader og andre avgifter på bil samt billettpris på kollektivtransport. Denne kostnaden ved å pendle ser analysen bort ifra.

Som nevnt ovenfor vil vår analyse inkluderer geografiske variabler i tillegg til reisetid til CBD. Disse geografiske variablene er med på å forklare den regionale variasjonen i boligprisene. Variablene sier noe om de attraktive forholdene i de ulike kommunene som for eksempel sosiale faktorer i kommunen og fasiliteter knyttet til boligens plassering. Det er riktignok vanskelig å inkludere alle sosiale faktorer og fasiliteter som husholdningene preferer, både fordi de er vanskelig å måle og fordi preferansene er svært ulike. Av en

rekke geografiske attributter kan det tenkes at tilgang til friluftsområder er et av mange eksempler som er med på å påvirke boligprisen. Dette er en attraktiv fasilitet for mange husholdninger som utelates fra modellen. Dette betyr at vår modellspesifikasjon utelater viktige variabler som fanges opp av modellens restledd. Dette kan føre til forventningsskjevne estimat som betyr at forutsetning (2) ved MKM ikke holder.

Utelatt variabelproblem oppstår når en relevant forklaringsvariabel er utelatt fra den modellen som estimeres. Mer konkret betyr det at forklaringsvariabelen forklarer en del av variasjonen i boligprisen og er korrelert med en eller flere av forklaringsvariablene slik at det oppstår skjevhet i estimatene (Wooldridge, 2009).

4.2.2 Målefeil

I de fleste økonomiske analyser vil variablene ofte i større eller mindre grad være beheftet med målefeil, og det er ikke usannsynlig at dette preger våre data både i den avhengige og i de uavhengige variablene. Målefeil betyr at observert data avviker fra den sanne verdien på dataene. Datasettet fra Eiendomsverdi er preget av målefeil da vi kan observere flere feilregistreringer i forklaringsvariablene som representerer kvaliteter ved boligen. Den deskriptive statistikken A.2 indikerer at vi har målefeil i forklaringsvariablene. Vi ser blant annet av minimums- og maksimumsverdien at vi har veldig lave og høye verdier for boligareal i datasettet. Det er vanskelig å tenke seg at et salg av en bolig med et boligareal på $1 m^2$ er et reelt salg. Vi ser også at alder på boligen i det aktuelle salgstidspunktet er oppført med negative verdier. Siden datasettet tar for seg boliger som allerede er bygd, indikerer de negative verdiene målefeil i datasettet.

Det er også en rekke svakheter med forklaringsvariablene knyttet til lokalisering av boligen som indikerer målefeil. Variabelen for endring i reisetid til CBD med bil innehar svakheter for strekningene Eidsvoll, Ullensaker, Drammen, Holmestrand, Moss og Rygge ved at de er oppgitt med teoretisk reisetid inn til de ulike knutepunktene ved bygrensen⁶. Det er rimelig å anta at tilgjengeligheten for reise med bil er mye større enn for reise med

⁶ Ringnes i Ås (E6 sør), Asker (E18 Vest) og Skedsmovollen (E6 Nord).

tog, da pendlerne kan benytte bilen direkte fra boligen til arbeidsplassen. Dataene for reisetid med bil har en svakhet ved at den er målt fra et bestemt punkt på E6/E18 i den respektive kommunen. Dette betyr at forsinkelser og kødannelser innad i kommunen ikke blir tatt hensyn til i datasettet. Variabelen for reisetid med bil sier derfor ikke noe om den faktiske tiden det tar fra selve boligen til Bispelokket i Oslo. Disse strekningene vil trolig ha omfattende målefeil i reisetiden. Til tross for utfordringene oppgitt ovenfor vil disse dataene være det beste målet på endringen i reisetid med bil, da det er vanskelig å finne et eksakt mål på reisetiden.

Dummyvariablene som representerer avstand fra boligen til jernbanestasjon i den respektive kommunen, er regnet ut manuelt. Dataene for boligen er oppgitt med postnummer som lokaliseringinformasjon. Med denne informasjonen er det tatt et tilfeldig utvalg av boligadresser i postnummerregionen og regnet ut en gjennomsnittsavstand fra postnummerets adresser til jernbanestasjonen. Svakheten med disse dummyvariablene er at de ikke er presise. Postnummerregioner er områder med variert størrelse som avhenger fra kommune til kommune. Vi kan for eksempel se at en liten kommune med relativt få postnumre vil være utstyrt med større målefeil enn store kommuner med mange postnumre. Det er også mange postnummerregioner som er preget av stor spredning i boligområder slik at et gjennomsnitt for avstanden til jernbanestasjonen vil kunne avvike fra den sanne avstanden. Videre har vi forutsatt at kommunene kun er representert med en jernbanestasjon. Modellspeifikasjonen har ikke tatt høyde for at det finnes flere jernbanestasjoner i kommunen eller at husholdningene i kommunen benytter seg av jernbanestasjoner i nabokommunene. Analysen tar utgangspunkt i at avstandsdummyene allikevel sier noe om hvor boligen er plassert og vil gi et signifikant og riktig bilde på hva avstand til jernbanestasjonen har å si for boligprisene. Dette indikerer at postnummerdummyene kan fange opp mye av de effektene som allerede er gitt uttrykk for i andre variabler.

Målefeil i den avhengige variabelen, hvis målefeilen er ukorrelert med forklaringsvariablene kan føre til høyere varians og standardavvik for MKM-estimatene. Målefeil i forklaringsvariablene av mer fundamental karakter kan gi alvorlige konsekvenser dersom de er ukorrelert med de uobserverbare variablene. Dette kan føre til

inkonsistente og skjeve estimat (Woolridge, 2009). Graves et al. (1988) peker på utelatte forklaringsvariabler, feil funksjonsform og målefeil i data som noen av de vanligste problemene ved hedonisk analyse og at disse utfordringene kan medføre brudd på forutsetningene for minste kvadraters metode (MKM).

4.2.3 Simultanitet

Problemet med simultanitet oppstår hvis en eller flere av de inkluderte forklaringsvariablene blir bestemt endogent i modellen og dermed avhenger av venstresidevariabelen. Forklaringsvariablene korrelerer med restleddet på grunn av en toveiskausalitet (Woolridge, 2009).

I de fleste økonometriske analyser er målet å kunne antyde at en variabel har effekt på en annen variabel. Denne effekten kalles kausaleffekt. Å estimere den kausale effekten av ulike boligattributter på boligpriser kan være utfordrende da en rekke av disse er formulert slik at de bestemmes simultant med boligprisene. I vår analyse har vi med utgangspunkt i den hedoniske teorien gjort antagelser om at inntekt og reisetid er eksogent gitt i modellen sammen med de andre boligattributtene. Vi tar derfor utgangspunkt i at vår modell er preget av enveis kausalitet som vil si at de inkluderte forklaringsvariablene forklarer vestresidevariabelen (Ibid).

4.3 Difference in difference

Endogenitet kompliserer studiet av effekten av reisetid på boligpriser med MKM, da dette kan føre til forventningsskjeve estimat. For å redusere problemet med endogenitet skal vi ta utgangspunkt i difference in difference metode for estimering.

Denne analysen benyttes for å kartlegge endring i reisetid på boligpriser. Regresjoner av endringsvariabler gir en indikasjon på bevegelser i faktorer over tid. Metoden Difference in difference estimerer forskjeller i endring mellom kommunene fra ulike andeler av utvalget ved å se på en kontroll- og en behandlingsgruppe (Woolridge, 2009).

Difference in difference analyse kan bidra til å finne løsninger på våre økonometriske utfordringer. Kilden til skjevheter i MKM estimatene grunnet utelatte variabler som korrelerer med restleddet elimineres når konstante variabler differensieres bort. Videre kan enkelteffekter av faktorer som endres simultant med boligpris isoleres ved å inkludere endring i andre variabler i regresjonen. Dersom disse varierer over tidsrommet vi studerer vil de ikke differensieres bort, og effekten av en endring i reisetid vil kunne avdekkes. Dette er de store styrkene ved å bruke difference in difference estimering i analysen (Ibid).

4.3.1 Utbygging av Romeriksporten

Strategien er å undersøke om en endring i reisetid i Skedsmo som følge av utbyggingen av Romeriksporten har hatt en effekt på boligprisene i kommunen. Datasettet som brukes er det samme som ble brukt ved estimering av MKM.

I 1992 bestemte Stortinget at Gardemoen skulle fungere som ny hovedflyplass og det ble samtidig vedtatt at tog skulle være det viktigste transportmiddelet. I sammenheng med dette ble den nye jernbanestrekningen til Gardemoen bygd med en dobbeltsporet høyhastighetsbane som skulle gå igjennom Romeriksporten. Jernbanetunnelen stod ferdig i 1999 og reduserte den generaliserte reisetiden med tog fra Lillestrøm til Oslo sentralstasjon med 11 minutter, se figur 1. Dette danner grunnlag difference in difference hypotesen:

H_0 : Romeriksporten har ikke positiv effekt på boligprisene i Skedsmo

H_1 : Romeriksporten fører til økte boligpriser i Skedsmo

Difference in difference benyttes for å kartlegge effekten av redusert reisetid som direkte effekt av utbyggingen av Romeriksporten på boligprisen. Regresjoner av endringsvariabler gir en indikasjon på bevegelser i faktorer over tid. Difference in difference er en metode som estimerer forskjeller i endring mellom kontroll- og behandlingsgruppe. Strategien er å studere om det har skjedd en endring i boligprisene i Skedsmo kommune sammenliknet med analysens kontrollkommuner etter utbyggingen av

jernbanetunnelen. Denne metoden er et godt supplement til MKM da den kan bidra til å korrigere for endogenitet i modellen.

Formålet er å identifisere en endring i boligprisene i Skedsmo kommune sammenliknet med kontrollkommunene etter utbyggingen av tunnelen. Dette gjøres ved å betrakte en behandlingsgruppe og en kontroll gruppe. Behandlingsgruppen (Skedsmo) blir direkte påvirket av utbyggingen ved redusert reisetid til CBD, mens kontrollgruppen ikke reduserer sin reisetid på grunn av utbyggingen. For å kontrollere for systematiske forskjeller trenger vi to grupper for før og etter utbyggingen.

$$Kommune = \left\{ \begin{array}{l} 0 \text{ kontrollgruppe} \\ 1 \text{ behandlingsgruppe (Skedsmo kommune)} \end{array} \right\}$$

$$\text{år} = \left\{ \begin{array}{l} 0 \text{ perioden før utbyggingen} \\ 1 \text{ perioden etter utbyggingen} \end{array} \right\}$$

Vi ser at dersom kommunedummyen tar verdien 1 indikerer dette kommunen som blir påvirket av utbyggingen og 0 for alle andre. Dersom dummyen for år tar verdien 1 indikerer dette perioden etter utbyggingen. Modellspesifikasjonen som skal estimeres ved bruk av difference in difference tar formen

$$\ln P(z_{tr}, z_{\varphi}, z_{kj})_i = \beta_0 + \alpha_0 \text{år} + \beta_1 \text{kommune} + \alpha_1 \text{år} * \text{skedsmo} + \text{andre variable} + u_i$$

$\widehat{\alpha}_1$ er difference in difference estimatoren som uttrykker den estimerte effekten av utbyggingen på boligprisen i Skedsmo sammenliknet med de kontrollerende kommunene.

$$\widehat{\alpha}_1 = \frac{\overline{\ln P(z_{tr}, z_{\varphi}, z_{kj})_{\text{etter 1999, skedsmo}}} - \overline{\ln P(z_{tr}, z_{\varphi}, z_{kj})_{\text{etter 1999, ikke skedsmo}}}}{\overline{\ln P(z_{tr}, z_{\varphi}, z_{kj})_{\text{før 1999, skedsmo}}} - \overline{\ln P(z_{tr}, z_{\varphi}, z_{kj})_{\text{før 1999, ikke skedsmo}}}}$$

β_0 betraktes som modellens konstantledd og viser gjennomsnittlig boligpriser i Stor-Oslo inkludert Skedsmo før utbyggingen. α_0 fanger opp effekten på boligprisene i perioden etter utbyggingen i alle kommunene i modellen, og β_1 gir effekten av boligprisene i Skedsmo kommune.

Vi tester følgende hypotese:

$$H_0: \alpha_1 \leq 0 \text{ (ingen eller redusert effekt)}$$

$$H_1: \alpha_1 > 0 \text{ (positiv effekt)}$$

Dersom H_0 forkastes antar vi at redusert reisetid som en direkte effekt av Romeriksporten har en signifikant positiv effekt på boligprisene i Skedsmo.

Forutsetningene for at diff-in-diff estimatoren skal gi riktig effekt er

1. At modellen er riktig spesifisert.
2. $E(u_i) = 0$ uttrykker at estimatorene er forventningsrette, altså at forventningen til estimatoren er lik estimatoren $E(\widehat{\alpha}_1) = \alpha_1$.
3. At restleddet u_i er ukorrelert med de andre forklaringsvariablene.

$$\text{cov}(u_i, \text{år}) = 0$$

$$\text{cov}(u_i, \text{kommune}) = 0$$

$$\text{cov}(u_i, \text{år} * \text{kommune}) = 0$$

4.3.2 Svakheter ved difference in difference

Difference in difference modellen inkluderer både boligattributter knyttet til selve boligen og de geografisk avhengige boligattributtene som er knyttet til lokalisering av boligen i tillegg til reisetid. I og med at datasettet består av kombinerte tverrsnitt vil boligene som er solgt etter åpningen av Romeriksporten trolig være systematisk forskjellige fra de som er solgt før utbyggingen. Det er derfor viktig å kontrollere for karakteristikk ved boligen som kan ha vært forskjellig. Selv om karakteristikkene er like for periodene før og etter

åpningen av jernbanetunnelen vil standardfeilen til difference in difference estimatet $\hat{\alpha}_1$ reduseres.

Difference in difference metoden krever at gruppene som undersøkes er like i forkant av et naturlig eksperiment. Dette kan by på utfordringer når vi de observerte gruppene er av geografisk karakter. Geografiske områder som kommuner er svært sammensatte og har ulike kvaliteter. Det er tilnærmet umulig å finne to kommuner som er identiske uansett tidspunkt. Det er derfor problematisk å sammenlikne Skedsmo med kommuner som er helt like før utbyggingen, da de kontrollerende kommunene er systematisk forskjellig. Dersom difference in difference estimatet indikerer en økning i boligprisene på grunn av utbyggingen er det vanskelig å vite om hele eller deler av denne effekten også var til stede i utvalget før utbyggingen inntraff.

En annen svakhet ved å bruke difference in difference er at metoden ser bort i fra tendenser mellom kontroll- og behandlingsgruppen som skjer samtidig. Hvis det er andre faktorer som påvirker forskjellen i trender mellom de to gruppene vil estimatene bli forventningsskjevne. Det kan tenkes at andre forhold en endring i reisetid påvirket boligprisene i Skedsmo kommune i 1999 som modellen ikke fanger opp. Slik at effekten av Romeriksporten vil bli under- eller overvurdert. Denne svakheten kan også påvirkes i motsatt retning, ved at ulike forhold i de kontrollerende kommunene påvirker boligprisene så sterkt at effekten av Romeriksporten frafaller.

Ved å bruke difference in difference som en metode for å kartlegge effekten av utbyggingen av Romeriksporten på boligpris må vi ta høyde for en forventningseffekt. Som nevnt tidligere ble det annonsert at Romeriksporten skulle bygges allerede i 1992. Det kan tenkes at markedsmekanismene rundt denne annonseringen førte til en forventning om økte boligpriser i Skedsmo kommune som gjorde at flere kjøpte bolig i Skedsmo allerede før Romeriksporten åpnet. Det vil si at effekten av utbyggingen på boligprisen utspilte seg allerede ved annonseringen. Denne effekten fanges ikke opp når vi tar høyde for tiden etter åpningen av tunnelen, slik at vi må korrigere for dette i en ny modellspesifikasjon.

Disse svakhetene gir grunnlag for å estimere tre ulike modeller, for å teste for robustheten ved estimering med difference in difference. Vi skal derfor se på resultatene ved tre ulike modellspesifikasjoner for å korrigere for noe av modellens svakheter.

M5: Modell med alle kontrollkommunene som sammenlikningsgrunnlag.

M6: Modell som utelater de største kontrollkommunene Moss, Drammen, Bærum og Eidsvoll.

M7: Modell som tester for forventningseffekt.

5. DATABESKRIVELSE

I dette kapitlet beskrives datamaterialet som er brukt i analysen. Vi skal foreta en systematisk forklaring av modellens variabler og deres antatte effekt på boligprisen.

5.1 Presentasjon av data

Analysen tar utgangspunkt i 59 362 enkelt observasjoner for salg av boliger med tilhørende geografiske og ikke geografiske boligattributter i perioden 1995 - 2006. Boligene er spredt over 11 kommuner tilhørende Stor-Osloregionen. Data for de inkluderte variablene er i hovedsak hentet fra Eiendomsverdi AS, Statistisk Sentralbyrå (SSB), Norsk Samfunnsvitenskapelig Datatjeneste (NSD), Triona og Norges Statsbaner (NSB).

5.1.1 Den avhengige variabelen

Den avhengige variabelen, boligpris, er den variabelen analysen ønsker å forklare.

Hovedkilden til boligprisdatabasen er Eiendomsverdi AS, som er et selskap som overvåker og registrerer aktivitet og utvikling i det norske eiendomsmarkedet. Tilgang til databasen er ikke offentlig. Variabelen for boligpriser består av 59 362 enkeltobservasjoner for salg av bolig målt i norske kroner i et bestemt år, kommune og postnummer.

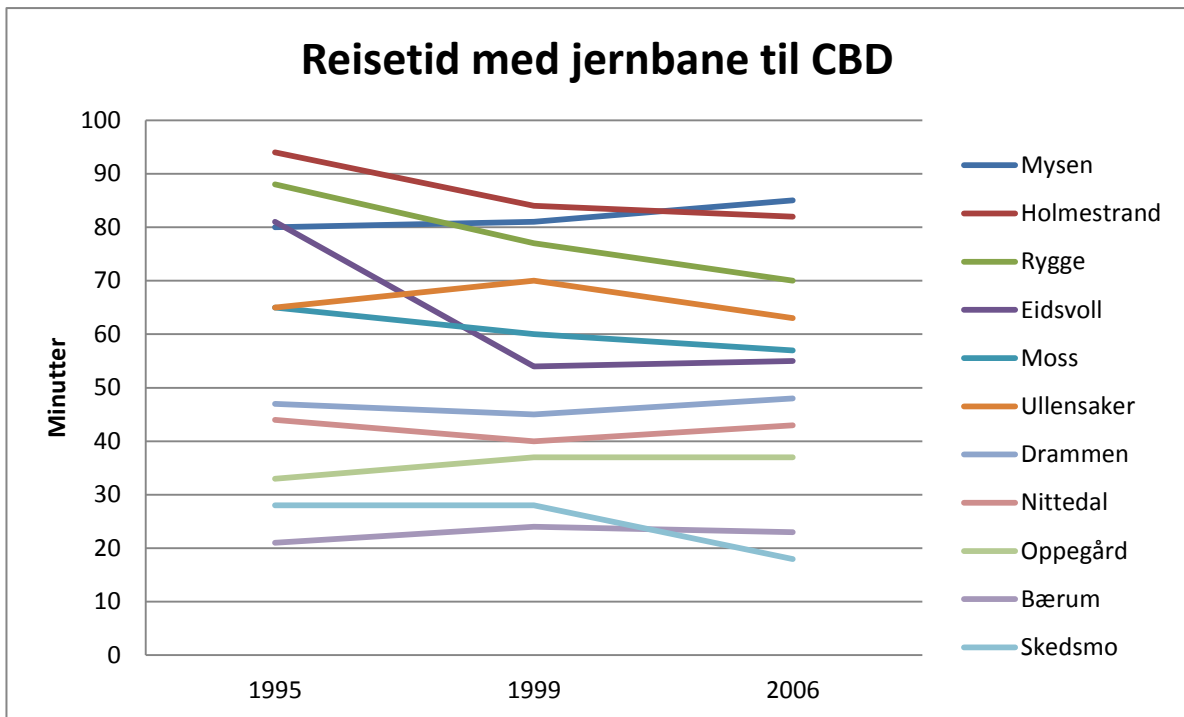
5.1.2 De uavhengige variablene

De uavhengige variablene er de variablene som forklarer boligprisen. Analysens interessevariabel er reisetid med jernbane, men for å kunne si noe om dette må vi ta høyde for øvrige variabler som vil påvirke boligprisen. Med utgangspunkt i teorien om den hedoniske prisfunksjonen presentert ovenfor, det empiriske studiet til Osland og Thorsen (2005) og de dataene som er tilgjengelige inkluderes følgende variabler i analysen:

- ✓ Variabler som er direkte knyttet til boligen som boligareal, bruksareal, antall soverom, boligens alder og boligtype.
- ✓ Seks variabler knyttet til lokalisering av eiendommen. Endring i reisetiden med jernbane fra kommunens jernbanestasjon til Oslo S og endring i reisetid med bil fra kommunens jernbanestasjon til Bispelokket i Oslo. I tillegg har vi fire dummyvariabler som måler kilometer avstand til jernbanestasjonen fra boligen, 133 postnummerdummyer og 11 kommunedummyer som viser hvilken postnummerregion og kommune boligen er plassert. Det er også lagt inn dummyvariabler for øst-, vest- og nordlig retning ut av CBD.
- ✓ Variabler som beskriver sosiale faktorer som omfatter egenskaper ved befolkningen og kommunen boligen er plassert. Disse geografiske variablene omfatter gjennomsnittlig inntekt per innbygger, arbeidsledighet, barnehagedekning, anmeldte lovbrudd, nettoinnflytting og boligbygging i kommunen.
- ✓ 12 dummyvariabler for boligens omsetnings år fra 1995-2006.

Reisetid med jernbane inn til Oslo Sentral stasjon ($z_{tr(tog)}$) er modellens interessevariabel og viser den generaliserte reisetiden målt i minutter fra kommunens jernbanestasjon til Oslo S. Den generaliserte reisetiden utgjør den gjennomsnittlige reisetiden med tog pluss halvparten av frekvensen. Frekvensen er den tiden du må vente på neste tog. Den generaliserte reisetiden sier derfor noe om den totale tiden det tar å reise til Oslo S inkludert den tiden man må vente på toget. Denne variabelen sammen med endring i reisetid med bil presenterer et geografisk skille i modellen og fanger opp de infrastrukturinvesteringene som er gjort på de ulike strekningene i perioden 1995 til 2006. dataene for denne variabelen er utregnet manuelt fra NSB sine rutebøker.

Figur 2: Reisetid med jernbane til CBD



Figuren viser endringen i reisetid med jernbane for de ulike kommunene i perioden 1995-2006 målt i minutter. Dersom vi tar en nærmere titt på enkelt kommunene skal vi se at endringen i reisetiden med jernbane kommer som følge av de infrastrukturinvesteringene som er gjort. Disse investeringene har to grupperinger hvor den mest sentrale er de investeringene som er gjort ved å lage dobbeltspor og bygge jernbanetunnel. I tillegg vil reisetiden endres ved å foreta en økning eller reduksjon i antall avganger. Denne operasjonen vil også falle inn under investeringer som er gjort på jernbane og følgelig ha en effekt på den generaliserte reisetiden.

Mest oppsiktsvekkende er endringene i Eidsvoll der reisetiden har redusert med 27 minutter i perioden 1995 til 1999. Reduksjonen er et resultat av at NSB innførte flere avganger fra Eidsvoll til Oslo S i den følgende perioden. Dette er en direkte konsekvens av at nettoinnflytningen i kommunen hadde en sterk vekst i samme periode (tabell 06913, ssb.no). Flere mennesker i kommunen fører til flere reisende med tog. En annen effekt verdt å merke seg er endringen i reisetid ved strekningen fra Moss til Oslo S. I 1996 ble

det bygget dobbeltspor fra Ski til Sandbukta ved Moss som fikk en effekt på den generaliserte reisetiden både på strekningen fra Rygge og fra Moss. Skedsmo kommune i nordlig retning fikk redusert reisetid med 10 minutter fra 1999 – 2006 som følge av Romeriksporten.

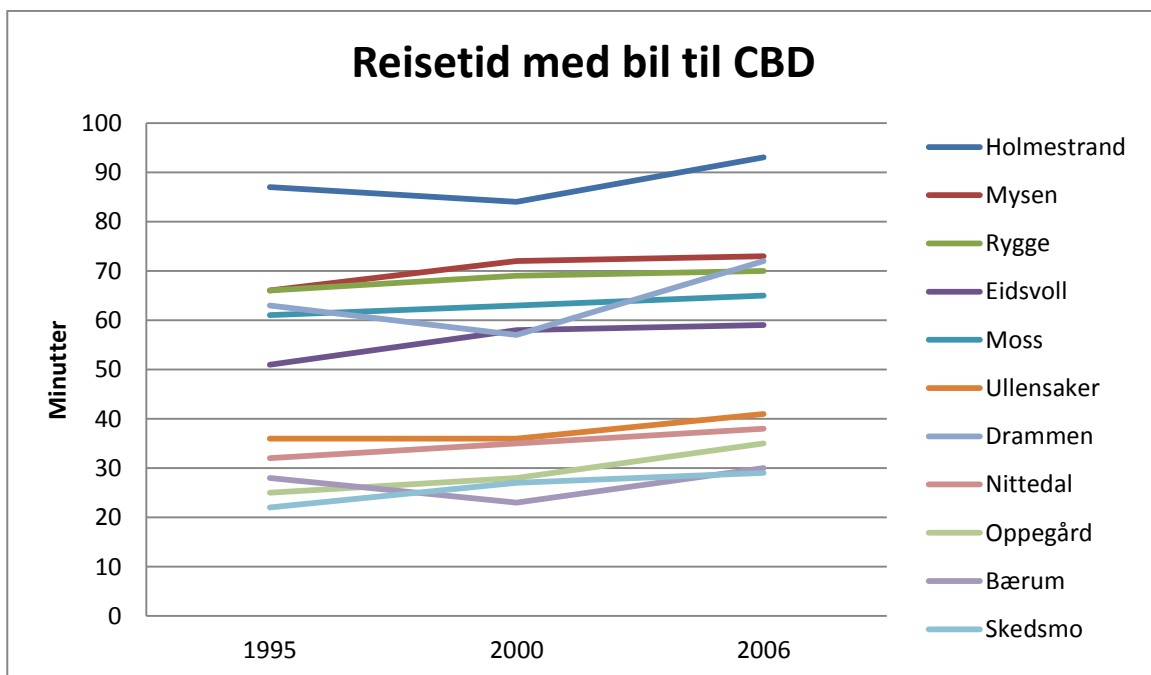
Reisetid med jernbane til Oslo S forventes å ha en negativ effekt på boligprisene. Denne antagelsen følger analysens teoretiske utgangspunkt og hovedhypotese om at redusert reisetid øker boligprisene.

Avstanden fra boligen til jernbanestasjonen (z_{ϕ_i}) er et annet attributt knyttet til lokalisering som antas å ha en effekt på boligpriser. Avstanden måles i gjennomsnittlig antall kilometer kjøreavstand fra boligen til kommunens jernbanestasjon. Avstanden regnes ut ved å ta et tilfeldig utvalg av adresser i det respektive postnummeret der boligen er lokalisert og finne postnummerregionens gjennomsnittsavstand til jernbanestasjonen. Variabelen er representert ved fire dummyvariabler som tar verdiene 0 - 2 km, $2\text{km} < \text{avstand} \leq 4\text{km}$, $4\text{ km} < \text{avstand} \leq 6\text{ km}$ og avstand over 6 km. Dersom verdien er 1 indikerer dette at det er den respektive avstanden til jernbanestasjonen, og 0 dersom avstanden representeres ved en av de andre dummyene. Variabelen beskriver stasjonstilgjengeligheten som forventes å ha en positiv effekt på boligprisene. Det antas at det er mer attraktivt å bo nærmest mulig jernbanestasjonen. Koeffisientene for dummyvariablene vil derfor være positive med høyest verdi for den korteste distansen. Dummyvariabelen for $4\text{ km} \leq 6\text{ km}$ utgår fra modellen for å hindre perfekt multikollinearitet.

Reisetid med bil ($z_{tr(bil)_i}$) er en variabel inkludert i modellen som forklarer reisetiden med bil fra kommunens jernbanestasjon til Bispelokket i Oslo målt i antall minutter. Denne variabelen måler den teoretiske reisetiden inn til punktene Ringnes i Ås (E6 sør), Asker (E18 Vest) og Skedsmovollen (E6 Nord) fra ulike punkt ved utkjøring på motorvei i den respektive kommunen. Fra de nevnte punktene inn til Bispelokket måles middels tidsbruk i morgenrushet da vi antar flest mennesker pendler til arbeid. Denne registrerte reisetiden tar høyde for forsinkelser inn til Bispelokket som ivaretar effekten av kø og andre forsinkelser. Den teoretiske reisetiden for de utvalgte årene er gitt av Triona, og

reisetiden i morgenrushet gitt ved middels tidsbruk er hentet fra rapporter gjort av Prosam. Bruk av bil til arbeidsplass er sentral ved pendling og utgjør en stor del av arbeidsreisene. Det investeres mye i infrastruktur på vei som reduserer reisetid, som vil ha en naturlig effekt på hvor husholdningene ønsker å bosette seg. Denne variabelen antas å ha samme negative effekt på boligprisene som reisetid med jernbane.

Figur 3: Reisetid med bil til CBD



Figur 3 viser endringen i reisetid med bil for perioden 1995 til 2006. Vi ser at flere av kommunene har hatt en økende tendens i reisetiden med bil til CBD sammenliknet med reisetid med jernbane (figur 2) i samme periode. Og det ser ut til at reisetid med jernbane har hatt en større variasjon i det aktuelle tidsrommet enn reisetid med bil. Måling av reisetid med biltransport er noe mer komplisert enn reisetid med jernbane, da reise med bil i større grad avhenger av trafikk.

Modellen inkluderer 133 *postnummerdummyer* ($post_i$) som representerer den postnummerregionen boligen er lokalisert. Disse dummyene ønsker å forklare forskjellene

mellom de ulike postnummerregionene. Det antas at noen postnummerregioner er mer attraktive å bo i enn andre og vil derfor ha forskjellig koeffisientverdier. De ulike postnummerregionene kan for eksempel ha ulike kvaliteter som gjør det mer attraktivt å bo i den spesifikke regionen.

Modellen inkluderer også *11 kommunedummyer (kommune)* som representerer de ulike kommunene som observeres. Disse dummyvariablene vil ha samme funksjon som postnummerdummyene og vil trolig ha ulike koeffisientverdier da noen kommuner er mer attraktive å bo i enn andre. For å unngå problemet med perfekt multikollinearitet utelater modellene dummyvariabel for Holmestrand.

I tillegg til dummyvariable for kommuner og postnummerregion inkluderer modellen tre *retningsdummyer* som representerer retningene nord, øst og vest fra CBD. Vi antar at disse dummyene samlet vil forklare mer av den geografiske variasjonen i boligprisen og si noe om boligveksten i de ulike retningene. Dummyvariabelen for nordlig retning ut av CBD er utelatt fra modellen for å hindre problemet med perfekt multikollinearitet.

Bolig- (z_{boa}) og bruksareal (z_{bra}) er variabler som uttrykker størrelsen på boligen målt i antall kvadratmeter. Boligareal er innvendig bruksareal med unntak av bod, loft, kjeller, veranda, kott, terrasser, balkonger, osv. Bruksareal er summen av det arealet i en bolig som ligger innenfor ytterveggene (snl.no). Bolig- og bruksareal er et attributt knyttet til selve boligen som forventes å ha en sterk positiv effekt på boligprisene. Desto større boligen er, desto høyere blir prisen på boligen.

Antall soverom (z_{sov}) er også et boligattributt som forventes å ha positiv koeffisient. Det er naturlig å anta at dersom antall soverom øker, øker størrelsen på boligen. Dette indikerer at boligprisen øker med antall soverom.

Alder på boligen (z_{alder}) er en variabel som beskriver hvor gammel boligen er på salgstidspunktet i antall år og forventer en negativ effekt på boligprisene. Jo eldre boligen er, jo høyere vil antatte vedlikeholdskostnader være. Standarden for en ny bolig vil normalt være bedre enn for en eldre bolig. Det er derfor rimelig å anta at desto eldre boligen er desto mindre attraktiv er den for kjøperen.

De fire boligtypene ($z_{boligtype}$) i datasettet betegnes som fire dummyvariabler.

Boligtypene som representeres er enebolig, leilighet, rekkehus og tomannsbolig. Disse betegnes ved dummyvariabler der verdien på dummyen er 1 dersom boligen er den respektive boligtypen, og 0 ellers. Det er naturlig å tenke seg at koeffisientverdiene er positive, og at enebolig har en høyere koeffisientverdi enn estimatet for leilighet. Dette bunner ut i at eneboliger oftest har større areal enn leilighet. Dummyvariabelen for rekkehus er utelatt fra modellen for å hindre perfekt multikollinearitet.

Barnehagedekning (z_{bhd}_i) angir antall barn i alderen 1-5 år som går i barnehage i kommunen målt i prosent av befolkningen. Variabelen måler hvor mange barn som får plass i kommunens barnehager i forhold til hvor mange som har søkt. Denne variabelen brukes som et mål på hvor god barnehagetjenestene er i den kontrollerende kommunen. De mellom 0-1 år er ikke inkludert da vi antar at de er hjemme på grunn av foreldrepermisjon. Denne variabelen antas å ha en positiv effekt på boligprisen.

Anmeldte lovbrudd (z_{anm}_i) angir andelen brudd på loven som er anmeldt i kommunen per år i prosent. Dette er en sosioøkonomisk faktor som kan tenkes å påvirke hvor husholdningene bosetter seg. Det er rimelig å anta at en økning i anmeldte lovbrudd vil føre til reduserte boligpriser da svært få ønsker å bo i et kriminelt belastet område.

Inntekt per capita (w_i) viser gjennomsnittlig inntekt per innbygger i kommunen. Dette er en variabel som vi venter vil ha en sterk og positiv effekt på boligprisen. Det er rasjonelt og tenke at dersom gjennomsnittlig inntekt for en privatperson i en kommune er høy relativt til andre kommuner, så har husholdningen i den respektive kommunen større kjøpekraft. Etterspørselen etter bolig vil dermed være høy i kommuner hvor innbyggerne har en høy gjennomsnittsinntekt. Disse dataene har en tidsserie fra 1995-2006 og hentet fra SSB.

Arbeidsledighet (z_{arbled}_i) viser prosentandelen av befolkningen i kommunen som er arbeidsledige. Denne forventes å ha en negativ effekt på boligprisene. Dersom mange av kommunens innbyggere er uten arbeid forventes det redusert etterspørsel etter boliger i kommunen. I kommuner med høy arbeidsledighet forventer vi at arbeiderne står mindre sterkt i lønnsforhandlinger samtidig som at det kan hende at flere velger å leie bolig enn å

kjøpe. I tillegg henger arbeidsledighet nøye sammen med de økonomiske konjunktorene og gir et uttrykk for hvordan husholdningene ser på fremtiden. Ved høy ledighet dempes investeringslysten som igjen fører til lav etterspørsel etter boliger og prisene på bolig presses ned.

Netto innflytting (z_{M_i}) viser totalt antall inn- og utflyttere til kommunen. Det er rimelig å anta at en økning i kommunens befolkning vil få en økning i boligprisene, da etterspørselen etter bolig øker når befolkningen blir større.

Boligbygging (z_{BB_i}) forventes å ha en negativ effekt på boligprisene. En kommune med sterk vekst i boligbyggingen vil få et økt tilbud av bolig som fører til at etterspørselen møtes. Dette vil normalt presse boligprisen ned. Variabelen for boligbygging er målt i antall nye boliger som bygges i kommunen.

Omsetningsår fremstilles som 12 dummyvariabler, en variabel for hvert år. Tidsintervallet strekker seg fra 1995 til 2006. Siden det ble fri prisfastsettelse på boliger på midten av 1980-tallet, har prisene gått oppover. Vi antar derfor at koeffisientene for omsetningsår øker hvert år. Dummy variabelen for 2003 er utelatt for å hindre perfekt multikollinearitet.

6. RESULTAT

I dette kapittelet presenteres resultatene av de estimerte modellene med MKM og difference in difference, M1-M7.

I modellene M1-M4 estimeres den hedoniske prisfunksjonen med MKM på grunnlag av datamaterialet som gir informasjon om pris og tilhørende boligattributter. MKM-resultatene av de ulike modellene med alle kontrollvariabler er beskrevet i A.5. Vi har estimert fire forklaringsmodeller der modell M1 er den modellen med MKM estimering som forklarer variasjonen i boligprisene best utfra antagelsene som er gjort, denne modellen vil bli referert til som referansemodellen.

Tabellfremstillingen (Tabell 1) av MKM resultatene legger vekt på effekten av reisetid, avstand til jernbanestasjonen og inntekt på boligpris, mens kontrollvariablene dokumenteres i appendiks. Effekten av kontrollvariablene på testresultatene kommenteres

bare i referansemodellen. Ved tolkning av resultatene er det fire punkter som er verdt å legge merke til; fortegnet til den estimerte koeffisienten, koeffisientens størrelse, signifikansnivå og modellens forklaringskraft. Med utgangspunkt i referansemodellen vil vi utføre en mengde robusthetstester ved å inkludere og ekskludere ulike geografiske variabler med unntak av reisetid, avstand til jernbanen og inntekt som forklares i alle modellene.

Som diskutert ovenfor preges MKM estimering av en rekke utfordringer med endogenitet. Som et verktøy for å kontrollere for noe av disse utfordringene skal vi gjøre en analyse med difference in difference ved å sjekke effekten før og etter et naturlig eksperiment ved å se på endringen i boligprisene etter utbyggingen av Romeriksporten. Estimaten for difference in difference kan derfor anses som mer troverdig.

6.1 Referansemodell

Tabell 1: Estimeringsresultater for MKM

$\ln P(z_{tr}, z_{\varphi}, z_{kj})$	M1	M2	M3	M4
<i>Observasjoner</i>	36890	40362	36890	19461
R^2	0,69	0,69	0,70	0,77
<i>Konstant</i>	0,104*** (0,569)	6,11*** (0,424)	10,92*** (1,06)	14,08*** (1,39)
$z_{tr(tog)_i}$	-0,096*** (,0076)	-0,11*** (0,005)	-0,0278* (0,01744)	-0,036* (0,02)
$z_{tr(bil)_i}$	-0,133*** (0,019)	-0,03*** (0,007)	0,0263 (0,0277)	0,14*** (0,031)
$z_{\varphi 1_i}$	0,201*** (0,005)	0,21*** (0,005)	0,2017*** (0,0051)	-0,276* (0,137)
$z_{\varphi 2_i}$	0,104*** (0,004)	0,10*** (0,004)	0,12022*** (0,0049)	0,029 (0,160)
$z_{\varphi 4_i}$	0,04*** (0,005)	0,048*** (0,005)	0,04942*** (0,0053)	0,54*** (0,153)
w	1,15*** (0,043)	0,94*** (0,015)	0,2268*** (0,0856)	-0,101 (0,1177)
<i>Omsetnings år</i>	Ja	Ja	Ja	Ja
<i>Boligkarakteristikker</i>	Ja	Ja	Ja	Ja
<i>Sosiale faktorer</i>	Ja	Nei	Ja	Ja
<i>Retningsdummy</i>	Ja	Nei	Nei	Nei
<i>Kommunedummyer</i>	Nei	Nei	Ja	Nei
<i>Postnummerdummyer</i>	Nei	Nei	Nei	Ja

Standardfeilene er representert i parentesene, *** $p < 0,01$, ** $p < 0,05$, * $p < 0,1$

6.1.1 Effekten av reisetid, tilgjengelighet og inntekt på boligpriser

MKM-resultatene for M1 i tabell 1 viser at dersom reisetiden med jernbane øker med 1 % vil boligprisene reduseres med 0,096 %. Samtidig ser vi at en økning i reisetiden med bil med 1 % fører til en reduksjon i boligprisene med 0,13 %. Begge effektene er signifikante på 1 % nivå. Dette er i tråd med antagelsene gjort ovenfor og indikerer at redusert reisetid til CBD øker boligprisene. Resultatene viser at reisetid med bil får en sterkere negativ effekt på boligprisen enn reisetid med jernbane. I og med at en større prosentandel pendlere benytter seg av bil fremfor tog utenfor Oslo vil dette være et naturlig å anta at reisetid med bil har en sterkere effekt på boligprisene enn reisetid med jernbane. I M1 er kommunene delt opp i vestlig, østlig og nordlig retning for Oslo. Det forventes ulik boligvekst i de ulike retningene. Av resultatene ser vi at det har vært høyest boligvekst i østlig retning og lavest i nord, som betyr at boligprisene har økt mer i Moss, Rygge og Oppegård enn for de andre retningene ut av Oslo.

Tilgjengeligheten til jernbanestasjonen representert ved avstandene 0-2 km, $2 \text{ km} \leq 4 \text{ km}$, og over 6 km. Dummyvariabelen for $4 \text{ km} \leq 6 \text{ km}$ utgår fra modellen for å hindre perfekt multikollinearitet. Estimaten er signifikante med høyest verdi for den korteste distansen som er i tråd med antagelsene gjort ovenfor. Resultatene tyder på at tilgjengeligheten til jernbanen forklarer variasjonen i boligprisedataene bedre enn reisetid til CBD da disse estimatene er relativt robuste med unntak av M4. Dette strider med Ahfeldts studie som undersøker verdien av tilgjengelighet til jernbanestasjon i et intercity området. Han konkluderer med at transporttilgangen har minimal effekt på boligprisene (Ahfeldt, 2010). Vår modellspesifikasjon tar ikke hensyn til at en boligplassering nær jernbanen kan føre med seg forurensningen og støy som betyr at en plassering «for» nær jernbanestasjonen kan virke negativt på resultatene for dummyvariabelen nærmest stasjonen. Disse utelatte variablene kan derfor føre til forventningsskjevne estimat for avstand. Det kan også tenkes at stasjonstilgjengeligheten hadde forklart mer av variasjonen i boligprisen dersom analysen hadde vært av en mer multisentrisk form⁷. Jernbanestasjonene i de observerte

⁷ Multisentrisk modell vil si en urbanmodell som antar at sysselsetting også foregår utenom cbd.

kommunene er plassert i kjernen av kommunen hvor det er naturlig å anta at sysselsetting hadde foregått dersom modellen hadde vært multisentrisk.

Estimeringsresultatet for gjennomsnittlig inntekt per innbygger i kommune i er positiv og er den variabelen med sterkest forklaringskraft i modellen. En økning med 1 % i inntekten vil føre til en 1,15 % økning i boligprisen. Inntekt er en variabel som forklarer mye av variasjonen i boligprisene og dette kommer til uttrykk i resultatene. Effekten er signifikant på et 1 % signifikansnivå. Dette er i tråd med teorien om at en inntektsøkning fører til at husholdningen har mer disponibel inntekt å bruke på boligattributter og konsum av andre goder. Andelen av inntekten som brukes til de forskjellige godene avhenger av preferansene til husholdningen. I teorien for verdsetting av reisetid og tilgjengelighet har vi antatt at en økning i inntekt vil føre til at husholdningene flytter til et område nærmere CBD. Resultatene kan tyde på at vi får en segregering etter boligstandard, da de med høy inntekt har høyere betalingsvilje for redusert reisetid og avstand til jernbanestasjonen. Denne intuisjonen indikerer at de med høyest inntekt vil plassere seg der reisetiden til Oslo og avstanden til jernbanestasjonen er kortest.

R^2 viser 69 % og er et mål på hvor godt modellen beskriver virkeligheten. R^2 vil ligge i intervallet mellom 0 og 1 der 1 er full forklaringseffekt og 0 er ingen forklaringseffekt. Det er vanskelig å anslå hva som er god forklaringseffekt, da dette vil variere for kompleksiteten i modellen (Woolridge, 2009). For boligmarkedet vil det være vanskelig å lage en modell med forklaringsgrad som er lik 1, men R^2 vil allikevel være en indikator på hvor godt variablene forklarer boligprisen. Resultatene gir en relativt høy R^2 som kan skyldes at mange variabler og få frihetsgrader kunstig presser forklaringskraften opp på et høyt nivå.

6.1.2 Effekten av kontroll variable på boligpriser

Ikke-geografiske variabler

Resultatet for referansemодellen M1 viser at dersom antall soverom øker med 1 % vil vi få en økning i boligprisene på 0,43 %. Denne effekten er signifikant på 1 % signifikansnivå. Estimeringsresultatet for boligens alder indikerer at boligprisen reduseres

med økt alder på boligen. Dette virker fornuftig ettersom jo eldre boligen er, desto høyere vil antatte vedlikeholdskostnader være. Dersom alderen på boligen øker med 1 % reduseres boligprisen med 0,078 %. Dette estimatet har også et signifikansnivå på 1 %.

Dummyvariablene for boligtype viser at boligprisen øker mest for enebolig og reduseres for leilighet. Det er naturlig at boligprisen er høyere for eneboliger enn leiligheter da eneboliger som oftest er større av areal. Allikevel strider den negative koeffisienten for leilighet med antagelsene som er gjort tidligere. Alle dummyvariablene for boligtype er signifikante på et 1 % nivå.

Geografiske variabler knyttet til sosiale forhold i kommunen

Resultatene for nettoinnflytting viser at en økning i nettoinnflytting i kommunen med 1 % reduserer boligprisene med 0,007 %. Effekten er signifikant på 1 % signifikansnivå. Dette strider mot antagelsen som er gjort tidligere. Det er naturlig å tenke seg at boligprisene øker når befolkningen i kommunen øker da flere vil etterspørre bolig. Også resultatet for arbeidsledighet, antall anmeldte lovbrudd og barnehagedekning strider mot antagelser gjort tidligere da estimatene indikerer den motsatte effekten enn tidligere antatt. Til tross for dette er alle kontrollvariablene signifikante på et 1 % nivå. Boligbygging i kommunen er den eneste kontrollvariabelen som beskriver sosiale forhold i referansemodellen som har forventet effekt på boligprisene i tråd med antagelsene. 1 % økning i boligbyggingen øker boligprisen med 0,01 %. Når veksten i boligbygging øker, øker tilbudet av boliger og etterspørselen møter tilbudet. Dette fører til en reduksjon i boligprisene. Dette estimatet er også signifikant på et 1 % nivå. Vi skal merke oss at boligprisene stiger jevnt i perioden 1995 til 2006.

6.2 Robusthetstester

For at vi skal være sikre på at resultatene i referansemodellen M1 er robuste presenteres robusthetstestene for modellen.

6.2.1 Effekten av MKM med postnummerdummyer

I M4 undersøker vi om de ulike postnumrene har betydning for resultatene. Dette er av relevans da vi antar at de ulike postnumrene forklarer forskjellene mellom de ulike postnummerregionene. Estimaten for reisetiden i denne modellen viser at en økning i reisetiden med tog til CBD med 1 % fører til en reduksjon i prisene på 0,036 %. Denne effekten er signifikant på et 10 % nivå. Effekten av reisetid med bil til CBD har derimot en positiv effekt på boligprisene med et signifikansnivå på 1 %. Reisetiden med bil til CBD samsvarer ikke med forventningene. Variablene som beskriver avstanden til jernbanestasjonen og gjennomsnittlig inntekt har også uventet effekt og lave t-verdier når vi kontrollerer for postnummerdummyene.

Antagelsene om at postnummerdummyer forklarer variasjon i boligpris i de ulike postnummerregionene fanges ikke opp i modellen da det oppstår høy korrelasjon mellom postnummerdummyene og avstandsdummyene. Vi ser også at det er multikollinearitet mellom postnummerdummyene som gjør at flere blir utelatt ved estimering i STATA. Ved å undersøke for robuste resultater fjernes postnummerdummyene for videre estimering. Det er allikevel viktig å poengtere at faren med å utelate postnummerdummyene kan være stor. Estimaten vil kunne bli påvirket av andre forskjeller mellom postnumre som ikke lenger vil bli fanget opp av dummyene. Dette kan være faktorer som forklarer attraktiviteten ved å bo i den enkelte postnummerregionen. Tilgjengeligheten til skoler, barnehager, friluftsområder er eksempler på dette. Ved å fjerne disse vil vi også fjerne de attraktive kvalitetene ved de ulike postnummerregionene.

M4 viser også en uventet verdi på estimaten for boligareal (z_{boa}) og bruksareal (z_{bra}). Disse variablene er preget av målefeil og mangelfulle observasjoner som gir utslag på estimatens signifikansnivå og fortegn. Estimatet for boligareal er negativt og indikerer at

en økning i boligareal gir en reduksjon i boligprisen. Bruksarealet derimot angir forventet effekt, men effekten er ikke signifikant. Av korrelasjonsmatrisen A.3 for de ikke-geografiske variablene ser vi at variablene for boligareal og bruksareal naturlig korrelerer med variabelen om antall soverom. Det er rimelig å anta at disse variablene korrelerer da de forklarer mye av det samme. På grunn av disse svakhetene utelates variablene for bolig- og bruksareal og vi bruker antall soverom (z_{sov}) for å forklare antagelsen om at en større bolig vil føre til en økning i boligprisene.

6.2.2 Effekten av MKM med kommunedummyer

Det undersøkes videre om resultatene er følsomme for ekskludering av postnumre ved å ta høyde for de ulike kommunene representert ved dummyvariabler i M3. I og med at reisetiden er gitt fra et punkt i kommunen vil dette kunne være med på å beskrive reisetiden til CBD bedre enn postnumrene. Resultatene viser at en økning i reisetid med jernbane til CBD med 1 % reduserer boligprisene med 0,027 % og en økning i reisetid med bil til CBD med 1 % øker boligprisene med 0,026 %. Reisetid med jernbane er signifikant på et 10 % nivå som indikerer at vi også i denne modellspeifiseringen må tolke estimatet med forsiktighet. Reisetid med bil er derimot ikke signifikant. Inntekt og avstandsvariablene er i tråd med de antagelsene gjort tidligere og signifikante på et 1 % nivå.

Sammenliknet med M4 blir estimatene for reisetid mindre i M3. Estimeringsresultatene i M3 er i likhet med M4 preget av upresise estimat med høy varians som skaper små koeffisientverdier og uventede fortegn. Det kan tyde på at forklaringsvariablene er preget av multikollinearitet. Korrelasjonskoeffisientene for reisetid korrelerer relativt sterk med flere av de geografiske variablene som indikerer en sterk sammenheng mellom flere av variablene. Ved estimering vil det bety at variablene for reisetid beskriver forholdene i kommunen som allerede fanges opp i de andre geografiske variablene.

6.2.3 Effekten av MKM uten geografiske variabler

I likhet med studiet til Osland og Thorsen (2005) skal vi estimere en modell uten geografiske variable, med unntak av reisetid og tilgjengelighet. Multikollinearitet mellom reisetid og de andre geografiske variablene kan bidra til høy varians som skaper små koeffisientverdier og «gale» fortegn. Vi skal teste for robustheten ved å bruke reisetid og tilgjengelighet som eneste forklaringsvariabel som beskriver den geografiske oppførselen til boligprisen.

Estimeringsresultatene for M2 viser at reisetid med tog til CBD får en sterkere negativ effekt på boligprisene enn ved estimering i de tidligere modellene. Boligprisene reduseres med 0,11 % dersom reisetiden med jernbane øker med 1 %. Reisetid med bil til CBD har også en negativ effekt på boligprisene som indikerer at 1 % økning i reisetiden reduserer boligprisen med 0,03 %. Begge effektene er signifikante på et 1 % nivå.

Forklaringskraften i M2 presentert ved R^2 er lik referansemodellen med 69 %. MKM-resultatene for Osland og Thorsen (2005) studie viser en R^2 verdi på 73,76 %, som er større enn for vår analyse. En av årsakene til dette kan være at de har informasjon om flere relevante forklaringsvariabler som forklarer boligattributter direkte knyttet til boligen som for eksempel renovering, antall toaletter og om boligen har garasje. Det kan også tenkes at variablene som er brukt for deres studie forklarer boligprisene i det aktuelle området bedre enn våre variable.

Denne modellen forklarer effekten av reisetid bedre enn modellen M3 og M4. Til tross for tilfredsstillende resultater for reisetid vil M2 trolig være preget av forventningsskjeve estimat da modellspesifikasjonen utelater en rekke geografiske variabler som er med på å forklare boligprisen. Det er vanskelig å si noe konkret om hvordan dette slår ut i estimatene, men det er grunnlag for å konkludere med at estimatene i større eller mindre grad er preget av utfordringen med endogenitet.

Det at vi ikke får signifikante verdier på flere av variablene i de ulike modellspesifikasjonene, samtidig som at R^2 er høy indikerer at estimeringsresultatene for MKM er lite troverdige.

6.2.4 Oppsummering

Robusthetstestene for MKM viser at M1 er ganske følsom for endringer i datautvalget som bør tas med når det gjøres tolkninger av resultatene. På den andre siden beholder koeffisientene for reisetid med jernbane sitt negative fortegn i alle modellene som er med på å styrke sannsynligheten for at økt reisetid med jernbane har en negativ kausal effekt på boligprisene.

Resultatene for MKM må tolkes med ytterst forsiktighet da det er tydelig at M1-M4 er preget av multikollinearitet mellom forklaringsvariablene, samtidig som at modellene preges av endogenitet. Til tross for at vi har identifisert disse utfordringene i modellen er det allikevel vanskelig å si noe nøyaktig om hvordan dette slår ut i estimatene. På grunnlag av resultatenes koeffisienter og signifikansnivå vil vi kunne forkaste nullhypotesen og konkludere med at økt reisetid med jernbane reduserer boligprisen ($\beta_1 < 0$). MKM resultatene vil ikke være troverdige da vi antar at modellen har utfordringer med endogenitet. For å bekrefte at økt reisetid reduserer boligprisene skal vi bruke difference in difference metode som et supplement til MKM.

6.3 Estimeringsresultater ved difference-in-difference

Dette avsnittet presenterer estimeringsresultatene fra difference in difference analysen. Metoden undersøker hvorvidt det har skjedd en endring i boligprisene etter utbyggingen av Romeriksporten. De ulike tilnærmingene ved estimering med difference in difference presenteres i M5-M7.

Tabell 2: Estimeringsresultater ved difference in difference

$\ln P(z_{tr}, z_{\varphi}, z_{kj})$	M5	M6	M7
<i>Observasjoner</i>	36890	14378	36890
R^2	0,67	0,67	0,65
<i>Konstant</i>	2,85*** (0,26)	-3,18*** (0,5)	-3,39*** (0,20)
≤ 1999	0,26*** (0,007)	-	0,15*** (0,012)
≤ 1996	-	0,07*** (0,013)	-
<i>Skedsmo</i>	0,05*** (0,014)	0,14*** (0,0192)	0,24*** (0,032)
Δ reisetid	-0,03** (0,014)	0,03* (0,015)	-0,08** (0,0318)

Standardfeilene er representert i parentesene, *** $p < 0,01$, ** $p < 0,05$, * $p < 0,1$

M5 viser om det har skjedd en endring i boligprisen i Skedsmo etter åpningen av Romeriksporten, når vi kontrollerer for de ti observerte kommunene i analysen. Det estimerte resultatet i modell M5 viser at boligprisene i Skedsmo har sunket med 0,033 % etter utbyggingen i 1999 sammenliknet med analysens kontroll kommuner. Den negative koeffisienten for endring i reisetid indikerer at boligprisene i Skedsmo er gjennomsnittlig lavere enn kontroll kommunene. Konstantleddet er positiv som betyr at kontrollkommunene har en positiv vekst i boligprisen i undersøkelsens tidsperiode. Denne regresjonen argumenterer for at utbyggingen har en negativ virkning på boligprisene i Skedsmo.

Som diskutert ovenfor kan det være flere årsaker til dette resultatet. Det kan tenkes at Skedsmo kommune la til rette for vekst i boligbyggingen etter annonsering av den nye jernbanetunnelen. Av den komplette tabellen i A.6 ser vi at Skedsmo kommune har hatt

en relativt sterk vekst i boligbyggingen siden 1998, med unntak av 2000-2001 som øker tilbudet av boliger i kommunen. Denne intuisjon fører til at tilbudseffekten i boligmarkedet er sterkere enn etterspørselseffekten, som igjen fører til at boligprisene presses ned. Dette fører til at den sanne effekten av Romeriksporten ikke kommer tydelig frem.

6.4 Robusthetstester

På samme måte som for resultatene med MKM skal det testes om difference in difference estimatene er robuste.

6.4.1 Effekten av en endring i reisetid

En annen årsak til den negative koeffisienten kan skyldes skjevt sammenlikningsgrunnlag i de kontrollerende kommunene. En forutsetning for at difference in difference skal være en effektiv metode er at kontroll og behandlingsgruppen skal være lik før det naturlige eksperimentet inntreffer. I denne analysen sammenlikner vi kommuner som vil være systematisk forskjellig fra hverandre på et hvert tidspunkt. De ulike kommunene vil ha forskjellige egenskaper som påvirker boligprisene slik at den sanne effekten av endringen i reisetid i Skedsmo ikke fanges opp.

For å jevne ut dette problemet utelates Drammen, Bærum, Eidsvoll og Moss fra settet med kontrollkommuner. Dette er analysens største kommuner som skiller seg mest fra Skedsmo. I og med at vår modellspesifikasjon er monosentrisk og antar at befolkningen i de ulike kommunene pendler til Oslo for å jobbe vil større kommuner trolig ha et høyere tilbud av arbeidsplasser som denne modellen ser bort ifra. Dette kan bety at boligprisene har en kraftigere vekst i de største kommunene. Det er også naturlig at disse kommunene vil ha størst nettoinnflytting som bidrar til økte boligpriser. Dette er et av mange eksempler på at sammenlikningsgrunnlaget er skjevt.

Estimeringsresultatene i M6 viser nå at boligprisene i Skedsmo har steget med 0,03 % etter utbyggingen i 1999 sammenliknet med gjennomsnittet i de inkluderte kontrollkommunene, Oppegård, Rygge, Eidsberg, Ullensaker, Holmestrand og Nittedal. I motsetning til M5 ser det ut til at utbyggingen hatt en positiv effekt på boligprisene i Skedsmo når vi sammenlikner med mer «like» kommuner. Selv om vi har utelatt de

største kommunene fra modellen for å korrigere for de største forskjellene vil vi fremdeles ha et problem med at kommunene som sammenliknes med Skedsmo ikke er like før utbyggingen. Det er også verdt å merke seg at konstantleddet i denne modellen er negativt som vil si at de kontrollerende kommunene har en negativ vekst i boligprisene.

6.4.2 Forventningseffekt

Tidligere studier om hvordan transportinvesteringer påvirker boligpriser viser at det knyttes en rekke forventninger til investeringene. Studiet til Chau og Ng (2008) hevder at forventningene reflekteres i boligprisene før investeringen er gjort. Det er derfor naturlig å tenke seg at det skjedde en forventningseffekt allerede ved annonsering av jernbanetunnelen. Som nevnt tidligere ble annonsering av utbyggingen allerede gjort i 1992. Vi skal derfor teste for en eventuell forventningseffekt som følge av markedsføringen her ved å bruke hele settet med kontrollkommuner. Svakheten med å teste for en forventningseffekt med våre data er at vi ikke fr målt den eventuelle forventningseffekten når informasjon som utbyggingen ble sluppet, da vi kun har data for perioden 1995-2006. Vi kan derfor ikke teste for forventningseffekt tidligere enn 1996.

Estimeringsresultatene i M7 viser at boligprisene i Skedsmo har sunket med 0,082 % sammenliknet med de kontrollerende kommunene etter 1996. Denne effekten er signifikant på et 10 % nivå. Dette resultatet tyder ikke på at det har vært en forventningseffekt som følge av Romeriksporten etter 1996.

6.4.3 Oppsummering

Estimeringsresultatene for difference in difference viser seg og ikke støtte analysens hypotese. Vi beholder nullhypotesen ($\alpha_1 \leq 0$) og konkluderer med at utbygging av Romeriksporten ikke har hatt økt effekt på boligprisene

Som diskutert ovenfor vil difference in difference metode ha en rekke svakheter. Mest avgjørende vil være modellens sammenlikningsgrunnlag når kontroll- og behandlingsgruppen måles som kommuner. Dette kan bidra til at grunnlaget for sammenlikningen er for svakt og de sanne effekten av en endring i reisetid ikke kommer tydelig nok frem. Vi vil derfor ikke kunne konkludere med at utbyggingen av Romeriksporten har ført til økte boligpriser i Skedsmo kommune.

7. KONKLUSJON

I denne undersøkelsen analyseres effekten av reisetid med jernbane på boligpriser.

Analysens hovedhypotese er at redusert reisetid med jernbane øker boligprisen. MKM gir oss signifikante resultater og forventet effekt på reisetid med jernbane og de negative estimatene ser ut til å være robuste for de ulike modellspesifikasjonene. Dette bekrefter analysens hovedhypotese om at redusert reisetid med jernbane øker boligprisen.

Estimeringsresultatene for reisetid med bil, avstand til jernbanestasjonen og inntekt er mer følsomme for små endringer dataene. En av årsakene til dette er at MKM resultatene kan være preget av multikollinearitet med upresise estimat med høy varians som skaper "gale" fortegn. Men også problemer av mer alvorlig karakter som endogenitet som fører til forventningsskjevne resultat. Når vi står ovenfor disse utfordringene vil MKM resultatene være lite troverdige selv om vi får signifikante verdier for reisetiden med jernbane.

For å styrke MKM-resultatene for reisetid med jernbane bruker vi difference in difference metode for estimering. Målet med denne metoden er å korrigere for endogenitet ved å differensiere bort de konstante variablene. De utelatte variablene som korrelerer med restleddet elimineres når konstante variabler differensieres bort. Estimeringsresultatene med difference in difference styrker ikke MKM resultatene da det ser ut til at boligprisene i Skedsmo har hatt en reduksjon etter 1999 sammenliknet med analysens kontrollkommuner. Også ved testing for en mulig forventningseffekt fra 1996 viser det seg at boligprisen er redusert. Det er allikevel indikasjon på at redusert reisetid som direkte effekt av Romeriksporten, har hatt en økt effekt på boligprisene når vi tester for kommuner med et bedre sammenlikningsgrunnlag. Metoden viser seg å ha en rekke utfordringer som gjør at vi ikke får frem den sanne effekten av utbyggingen. Dette betyr at vi ikke kan konkludere med at Romeriksporten har ført til økte boligpriser i Skedsmo.

Selv om denne analysen ikke gir oss tilfredsstillende resultat som bekrefter våre hypoteser, gir MKM resultatene allikevel intensiver for videre forskning på denne aktuelle problemstillingen. Ved bruk av difference in difference vil det være aktuelt å utvide datasettet ved å inkludere flere år. På denne måten vil vi kunne estimere forventningseffekten fra det punktet informasjonen om utbyggingen ble sluppet. Utviding av datasettet vil også være et hjelpemiddel for å korrigere for multikollinearitet.

Med utgangspunkt i samme datasett vil instrumentvariabelmetoden vært interessant ved videre analyse da dette er en mulig metode for å mildne problemet med utelatte variabler. En annen interessant metode for videre analyse ville vært å undersøke om boligprisen er påvirket av spatial avhengighet mellom kommunene i likhet med Osland og Thorsen (2005). Spatial avhengighet er tilfelle dersom boligprisene avhenger av geografiske områder eller egenskaper. Ved å bruke metoden for spatial autorkorrelasjon kan vi undersøke om det oppstår geografisk avhengighet mellom de utelatte variablene i kommune i og kommune j ($i \neq j$) (LeSage og Pace, 2009). Dette er en svært relevant metode for videre forskning da det er naturlig å tenke seg at boligpris i en kommune vil avhenge av boligprisen i en annen kommune.

8. REFERANSER

Adair, A. McGreal, S. Smyth, A. og Ryley, T. (2000): House prices and accessibility: The testing of Relationships within the Belfast urban area, Housing studies. Downloaded 21.01.13 11:01.

Ahlfeldt, M. Gabriel. (2010): The Train has left the station: Do Markets Value Intracity Access to Intercity Rail Connections? London School of Economics and Political Science, Departement of Geography and Environment 312-335.

Alonso, W (1964): Location and Land Use. Toward a General Theory of Land. Cambridge, MA: Harvard University Press.

Baum-Snow, Nathaniel (2006): Suburbanization and Transportation in the Monocentric Model. Brown University, Department of Economics, Box B.

Baum-Snow, Nathaniel og Kahn, Matthew E (2005): Effects of Urban Rail Transit Expansions: Evidence from Sixteen Cities, 1970–2000.

Brueckner, Jan K. (1986): The structure of urban equilibria: A unified treatment of the Muth-Mills model. University of Illinois at Urbana-Champaign.

Fejarang, R. A. (1994): Impact on Property Values: A study of the Los Angeles metro rail. Whashington, DC: Transportation Research Board (preprint, Transportation Research Board, 73rd Annual Meeting, January 9-13).

Fiva, Jon H og Kirkebøen, Lars J. (2011): Information Shocks and the Dynamics of the Housing Market. The Scandinavian Journal of Economics.

Landis J. D. (1994): The California Urban Futures Model: a new generation of metropolitan simulation models. Environment and Planning B: Planning and Design 21(4) 399 – 420

LeSage, J.P. (1999): The theory of spartial econometrics, Department of economics, University of Toledo

Osland, L. og Thorsen, I. (2005): Effects on housing prices of urban attraction and labor market accessibility. Working papers in economic Department of Economics University of Bergen.

Osland L, Thorsen I og Gitlesen J P. (2007): Housing Price Gradient in a Region with One Dominating Center. JRER. Volume 29. No.3.

Rosen, S. (1974): Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. The Journal of Political Economy, 82(1), 34-55.

Yinger, J. (1979): Estimating the relationship between location and the price of housing. Journal of Regional Science. Volume 19, Issue 3, pages 271–286, August 1979.

Yinger, J. (2005): Housing and Commuting: The theory of Urban Residential Structure. E-book, version 1.0 April 2005.

Yiu C.Y, Wong S.K (2005): The Effects of Expected Transport Improvements on Housing Prices. Urban Studies Journal Foundation.

Data referanser:

Rutebok for Norge 4/94, 25. september – 31. desember 1994

Rutebok for Norge 4/99, 26. september 1999 – 8. januar 2000

Rutebok for Norge 4/04, 16. august 2004 – 8. januar 2005

Rutebok for Norge 1/11, gyldig fra 1. mars

Internett adresser:

http://www.ssb.no/a/publikasjoner/pdf/notat_201210/notat_201210.pdf (ssb.no 13.05.2013)

<http://www.regjeringen.no/nb/dep/sd/dok/regpubl/stmeld/2008-2009/stmeld-nr-16-2008-2009-/1.html?id=548838> (regjeringen.no 16.05.2013)

<http://www.jernbaneverket.no/no/Prosjekter/Prosjekter/follobanen/Follobanen/> (jernbaneverket.no 24.05.13)

<http://snl.no/.search?query=bruksareal&search=>>

http://www.google.no/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=1&ved=0CCsQFjAA&url=http%3A%2F%2Fwww.akershus.no%2Ffile%2F7b21ce9f7b125484fee6eff6cf8eb6f55%2F12presentasjon_0237%2520Eidsvoll.pdf&ei=nmWjUffFAqOE4ASj2oHgBg&usg=AFQjCNE03XKlAeaKBkpBxLvAqrUPtWfk6Q&sig2=-noY5O5M3as-1IPosHH0Fg&bvm=bv.47008514,d.bGE (akershusfylkeskommune.no 27.05.13)

<http://www.regjeringen.no/nb/dep/krd/dok/regpubl/stmeld/2006-2007/stmeld-nr-31-2006-2007-/10/4.html?id=469340> (regjeringen.no 29.05.13)

<http://www.regjeringen.no/en/dep/krd/documents/white/Propositions/20022003/Stmeld-nr-31-2002-2003-/10.html?id=403087> (regjeringen.no)

<http://www.ssb.no/priser-og-prisindekser/artikler-og-publikasjoner/hva-driver-utviklingen-i-boligprisene>

<http://www.abcnyheter.no/bolig/faktainnhold/090105/hva-bestemmer-boligprisene>

<http://www.regjeringen.no/nb/dep/ud/dok/nou-er/2012/nou-2012-2/27/2/2.html?id=669841> (regjeringen.no 27.08.2013)

<http://www.regjeringen.no/nb/dep/bld/dok/nouer/1996/nou-1996-13/3/2/8.html?id=340890> (regjeringen.no 25.11.2013)

<https://www.toi.no/reisevaner-og-mobilitet/langpendlerne-foretrekker-tog-article30999-213.html> (27.08.2013)

<https://www.toi.no/getfile.php/Publikasjoner/T%C3%98I%20rapporter/2012/1201-2012/1201-hele%20rapporten%20nett.pdf> (rapport 1201/2012)

Tidsskrifter:

Nytt dobbeltspor Oslo-Ski, Samfunnsgeografiske perspektiver”, Samfunns- og næringslivsforskning AS (SNF) 30.5 2008

Osland L. (2001) Den hedoniske metoden og estimering av attributtpriser. Norsk Økonomisk Tidsskrift Nr 1, Årg 115.

Rapporter:

Øystein Engebretsen, Liva Vågane, Inge Brechan, Anne Gjerdåker (2012) Langpendling innenfor intercitytriangelet. Tøi rapport 1201/2012.

Prosam.org (april 1998). Fremkommelighetsundersøkelser for bil i Oslo og Akershus 1995,1996 og 1997. Rapport 55.

Prosam.org (mars 2003). Fremkommelighetsundersøkelser for bil i Oslo og Akershus 2000, 2001 og 2002. Rapport 101

Prosam.org (januar 2007). Fremkommelighetsundersøkelser for bil i Oslo og Akershus 2005 og 2006. Rapport 146.

Prosam.org (mars 2013). Reisevaner i Oslo og Akershus. Rapport 202.

APPENDIKS

A1: Variabelbeskrivelse av ulike boligattributter

	Variabelnavn
$\ln P(z_{tr}, z_{\varphi}, z_{kj})_i$	Boligpris i kommune i
$z_{tr(tog)}_i$	reisetid med jernbane fra kommune i til CBD
$z_{tr(bil)}_i$	reisetid med bil fra kommune i til CBD
w_i	gjennomsnittlig inntekt per innbygger i kommune i
z_{φ}_i	avstand til jernbanestasjon i kommune i
z_{boa}	boligareal
z_{bra}	boligens bruksareal
z_{sov}	antall soverom
z_{alder}	boligens alder
z_{arbled}_i	arbeidsledighet i kommune i
z_{M}_i	netto innflytting i kommune i
z_{anm}_i	antall anmeldte lovbrudd i kommune i
z_{bhd}_i	barnehagedekningen i kommune i
z_{BB}_i	boligbygging i kommune i
z_{ene}	enebolig
z_{leil}	leilighet
z_{rekke}	rekkehus
z_{to}	tomannsbolig
$post_i$	postnummer i kommune i
$Skedsmo$	Kommune
≤ 1999	År
$\Delta reisetid$	år * $skedsmo$

A2: Deskriptiv statistikk

VARIABEL	OBS	MEAN	STD.DEV	MIN	MAX
$P(z_{tr}, z_{\varphi}, z_{kj})$	59362	1511968	798626,3	5000	13 000000
$z_{tr(tog)}$	59362	43,88	18,83	18	94
$z_{tr(bil)}$	59362	43,89	18,65	22	93
W	59362	215819,8	48201	13800	361100
z_{boa}	35861	217,30	171,14	1	421
z_{bra}	42243	406,3	308,17	1	853
z_{sov}	43222	2,71	1,79	0	254
z_{alder}	59362	28,62	27,70	-17	406
z_{arbled}	59362	138,28	74,93	7	261
z_M	59362	351,09	320,98	-757	1134
z_{anm}	59362	135,27	58,22	1	216
z_{bhd}	59362	140,44	48,50	11	237
z_{BB}	59362	272,62	202,11	4	894

A3: Korrelasjonsmatrise for de ikke-geografiske variablene

	Z_{boa}	Z_{bra}	Z_{sov}	Z_{alder}	Z_{ene}	Z_{leil}	Z_{rekke}	Z_{to}
Z_{boa}	1							
Z_{bra}	0,58	1						
Z_{sov}	-0,39	-0,22	1					
Z_{alder}	-0,12	-0,04	0,13	1				
Z_{ene}	-0,21	-0,03	0,42	0,17	1			
Z_{leil}	0,47	0,26	-0,56	-0,24	-0,54	1		
Z_{rekke}	-0,20	-0,19	0,05	0,1	-0,39	-0,33	1	
Z_{to}	-0,11	-0,11	0,1	-0,03	-0,26	-0,22	-0,16	1

A.4: Korrelasjonsmatrise for de geografiske variablene

	$z_{tr(tog)}$	$z_{tr(bil)}$	w	z_{arbled}	z_M	z_{anm}	z_{bhd}	z_{BB}	Moss	Eidsberg	R	Oppegård	B	Skedsmo	Nittedal	Ullensaker	Eidsvoll	D	HS
$z_{tr(tog)}$	1																		
$z_{tr(bil)}$	0,77	1																	
w	-0,53	-0,47	1																
z_{arbled}	0,31	0,53	-0,30	1															
z_M	-0,27	-0,15	0,14	0,14	1														
z_{anm}	0,24	0,29	-0,41	0,38	0,48	1													
z_{bhd}	-0,45	-0,25	0,76	-0,01	0,08	-0,31	1												
z_{BB}	-0,35	-0,24	0,40	-0,03	0,57	0,19	0,32	1											
Moss	0,28	0,34	-0,24	0,29	-0,06	0,16	-0,11	0,2	1										
Eidsberg	0,22	0,18	-0,11	-0,04	-0,18	-0,09	-0,03	-0,1	-0,05	1									
Rygge (R)	0,23	0,18	-0,14	-0,04	-0,13	-0,31	-0,16	-0,2	-0,05	-0,02	1								
Oppegård	-0,11	-0,23	0,16	-0,26	-0,44	-0,45	0,3	-0,2	-0,09	-0,03	-0,04	1							
Bærum (B)	-0,57	0,49	0,53	-0,47	0,17	-0,24	0,27	0,3	-0,15	-0,06	-0,07	-0,11	1						
Skedsmo	-0,53	-0,47	0,03	0,03	0,3	0,13	0,08	0,05	-0,14	-0,06	-0,07	-0,11	-0,2	1					
Nittedal	0,02	-0,10	0,01	-0,16	-0,32	-0,44	-0,06	-0,2	-0,08	-0,03	-0,03	-0,06	-0,1	-0,10	1				
Ullensaker	0,37	-0,09	-0,01	-0,04	0,22	0,37	-0,16	0,2	-0,11	-0,04	-0,05	-0,09	-0,15	-0,15	-0,08	1			
Eidsvoll	0,17	0,16	-0,16	-0,07	-0,06	-0,01	-0,26	-0,07	-0,07	-0,03	-0,03	-0,05	-0,1	-0,09	-0,05	-0,07	1		
Drammen (D)	0,19	0,53	-0,20	0,52	0,16	0,33	-0,04	0,16	-0,2	-0,07	-0,08	-0,14	-0,24	-0,23	-0,13	-0,18	-0,12	1	
Holmestrand (HS)	0,26	0,28	-0,13	0,01	-0,26	0,09	-0,03	-0,4	-0,05	-0,02	-0,02	-0,04	-0,07	-0,07	-0,04	-0,05	-0,03	-0,08	1

A.5: Estimeringsresultater for MKM

$\ln P(z_{tr}, z_{\varphi}, z_{kj})$	M1	M2	M3	M4
Observasjoner	36890	40362	36890	19461
R²	0,69	0,69	0,70	0,77
Konstant	0,104*** (0,569)	6,11*** (0,424)	10,92*** (1,06)	14,08*** (1,39)
Z_{tr(tog)}_i	-0,096*** (,0076)	-0,11*** (0,005)	-0,0278* (0,01744)	-0,036* (0,02)
Z_{tr(bil)}_i	-0,133*** (0,019)	-0,03*** (0,007)	0,0263 (0,0277)	0,14*** (0,031)
Z_{φ₁}_i	0,201*** (0,005)	0,21*** (0,005)	0,2017*** (0,0051)	-0,276* (0,137)
Z_{φ₂}_i	0,104*** (0,004)	0,10*** (0,004)	0,12022*** (0,0049)	0,029 (0,160)
Z_{φ₄}_i	0,04*** (0,005)	0,048*** (0,005)	0,04942*** (0,0053)	0,54*** (0,153)
w	1,15*** (0,043)	0,94*** (0,015)	0,2268*** (0,0856)	-0,101 (0,1177)
Z_{boa}	-	-	-	-0,007*** (0,0017)
Z_{bra}	-	-	-	0,0014 (0,001)
Z_{sov}	0,431*** (0,0039)	0,43*** (0,003)	0,4247*** (0,0039)	0,388*** (0,005)
Z_{alder}	-0,078*** (,0014)	-0,073*** (0,001)	-0,0780*** (0,001)	-0,08*** (0,0017)
Z_{arbled}_i	0,0007*** (0,00005)	-	-0,0002*** (0,000096)	-0,0007 (0,0001)
Z_M_i	-0,007*** (0,002)	-	0,0005 (0,0024)	-0,0005 (0,0027)
Z_{anm}_i	0,0002*** (0,00004)	-	-0,0001** (0,00008)	-0,0004*** (0,00009)
Z_{bhd}_i	-0,0002 (0,00007)	-	-0,00009 (0,00009)	0,0001 (0,0001)
Z_{BB}_i	-0,010*** (0,0024)	-	-0,0119*** (0,0029)	-0,004 (0,003)
Z_{ene}	0,257*** (0,004)	0,25*** (0,004)	0,2721*** (0,004)	0,28*** (0,005)
Z_{leil}	-0,114*** (0,004)	-0,106*** (0,004)	-0,1159*** (0,004)	-0,117*** (0,006)
Z_{to}	0,063*** (0,006)	0,062*** (0,005)	0,0734*** (0,005)	0,066*** (0,007)
1995	0,06*** (0,0123)	-0,40*** (0,011)	-7419*** (0,0429)	0,09*** (0,016)
1996	0,156*** (0,013)	-0,35*** (0,01)	-0,6436*** (0,0370)	0,22*** (0,023)
1997	0,258*** (0,015)	-0,27*** (0,009)	-0,5181*** (0,0316)	0,36*** (0,03)
1998	0,28*** (0,016)	-0,20*** (0,008)	-0,4013*** (0,0283)	0,45*** (0,038)
1999	0,378*** (0,018)	-0,158*** (0,008)	-0,3150*** (0,0231)	0,62*** (0,046)
2000	0,456*** (0,017)	-0,06*** (0,007)	-0,1602*** (0,0195)	0,71*** (0,046)
2001	0,404*** (0,02)	0,026*** (0,006)	-0,07990*** (0,0185)	0,83*** (0,052)
2002	0,33***	0,022***	-0,0206**	0,88***

	(0,022)	(0,005)	(0,0100)	(0,06)
<i>2004</i>	0,33***	-0,005	0,0518***	0,95***
	(0,025)	(0,005)	(0,0088)	(0,066)
<i>2005</i>	0,35***	-0,007***	0,0990***	1,01***
	(0,03)	(0,006)	(0,0161)	(0,076)
<i>2006</i>	0,58***	0,16***	0,1940***	1,06***
	(0,027)	(0,005)	(0,01358)	(0,067)
<i>Moss</i>	-	-	0,2498***	-
			(0,0163)	
<i>Eidsberg</i>	-	-	-0,1871***	-
			(0,0147)	
<i>Rygge</i>	-	-	0,1143***	-
			(0,0187)	
<i>Oppegård</i>	-	-	0,5017***	-
			(0,0457)	
<i>Bærum</i>	-	-	0,5814***	-
			(0,0593)	
<i>Skedsmo</i>	-	-	0,3741***	-
			(0,0442)	
<i>Nittedal</i>	-	-	0,2916***	-
			(0,0343)	
<i>Ullensaker</i>	-	-	0,2318***	-
			(0,0311)	
<i>Eidsvoll</i>	-	-	-0,0240	-
			(0,0196)	
<i>Drammen</i>	-	-	0,2246***	-
			(0,0230)	
<i>Vest</i>	0,024**	-	-	-
	(0,011)			
<i>Øst</i>	0,124***	-	-	-
	(0,007)			

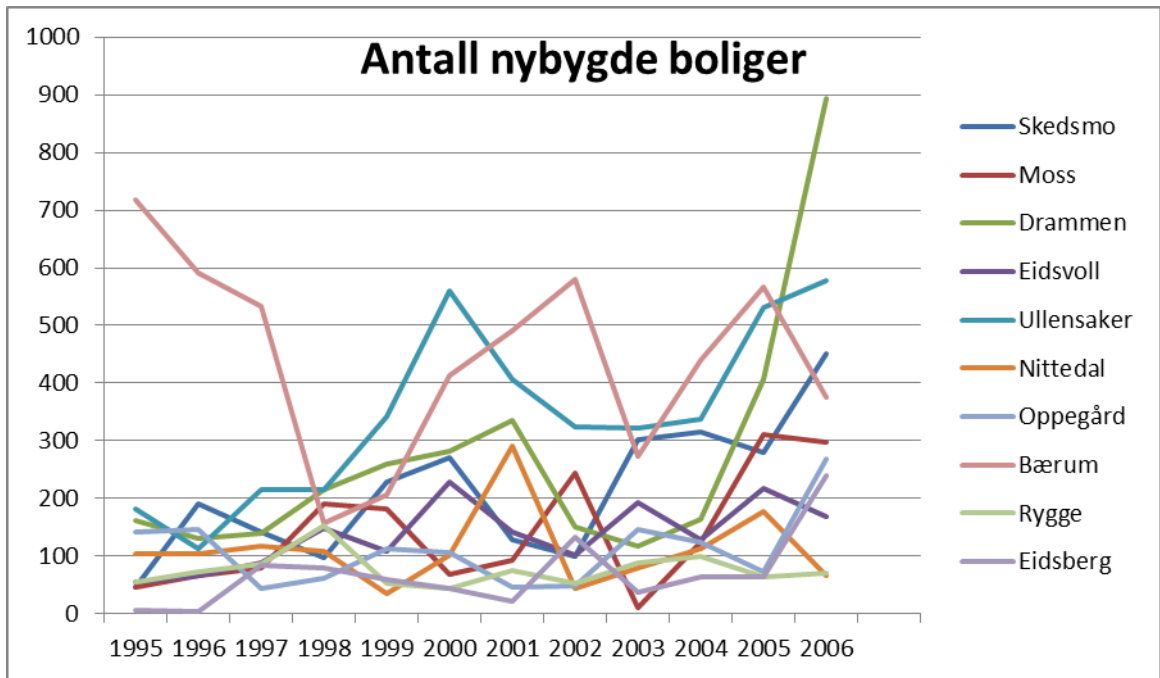
Standardfeilene er representert i parentesene, *** p < 0,01, ** p < 0,05, * p < 0,1

A6: Estimeringsresultater for difference in difference

$\ln P(z_{tr}, z_{\varphi}, z_{kj})$	M5	M6	M7
<i>Observasjoner</i>	36890	14378	36890
R^2	0,67	0,67	0,65
<i>Konstant</i>	2,85*** (0,26)	-3,18*** (0,5)	-3,39*** (0,20)
$Z_{tr(tog)_i}$	-0,08*** (0,008)	-0,117*** (0,02)	0,017** (0,007)
$Z_{tr(bil)_i}$	0,023*** (0,008)	0,123*** (0,017)	0,105*** (0,007)
$Z_{\varphi_1 i}$	0,87*** (0,02)	1,37*** (0,038)	1,34*** (0,016)
$Z_{\varphi_2 i}$	0,21*** (0,005)	0,142*** (0,014)	0,21*** (0,005)
$Z_{\varphi_4 i}$	0,11*** (0,005)	0,030** (0,014)	0,109*** (0,005)
w	0,03*** (0,005)	0,015 (0,013)	0,035*** (0,005)
Z_{boa}	-	-	-
Z_{bra}	-	-	-
Z_{sov}	0,435 (0,004)	0,39*** (0,0066)	0,436*** (0,004)
Z_{alder}	-0,077*** (0,0015)	-0,06*** (0,0022)	-0,076*** (0,001)
Z_{arbled_i}	-0,0003*** (0,00002)	-0,0009*** (0,00004)	-0,0002*** (0,00003)
Z_{M_i}	0,0135*** (0,0022)	-0,0122*** (0,0028)	-0,003 (0,002)
Z_{anm_i}	-0,00005 (0,00004)	-0,0001 (0,00006)	0,00006 (0,00004)
Z_{bhd_i}	0,0013*** (0,00005)	0,00071*** (0,00007)	0,0009*** (0,00005)
Z_{BB_i}	-0,003 (0,0023)	0,01** (0,004)	0,005** (0,002)
Z_{ene}	0,25*** (0,005)	0,28*** (0,007)	0,25*** (0,005)
Z_{leil}	-0,1*** (0,005)	-0,12*** (0,007)	-0,108*** (0,005)
Z_{to}	0,06*** (0,006)	0,12*** (0,009)	0,062*** (0,006)
≤ 1999	0,26*** (0,007)	-	0,15*** (0,012)
≤ 1996	-	0,07*** (0,013)	-
<i>Skedsmo</i>	0,05*** (0,014)	0,14*** (0,0192)	0,24*** (0,032)
Δ reisetid	-0,03** (0,014)	0,03* (0,015)	-0,08** (0,0318)

Standardfeilene er representert i parentesene, *** p < 0,01, ** p < 0,05, * p < 0,1

A.7 Antall nybygde boliger



A.8 Jernbanelinjer NSB

