

Fordelingseffekter af husleje- regulering: Målrettet som spredning

Af
Jakob Langager
De Økonomiske Råds sekretariat

Keywords: rent control, distribution, machine learning
JEL codes: C53, R32, L51, D39

Abstract*

In this paper, I analyze whether tenants' economic gain from rent reduction due to rent control is targeted low-income households. I investigate if households with low income have larger gains than high-income households, and if households with the same income have similar gains. To compute the gains, I estimate the counterfactual market rent for regulated units using a random forest model based on the rents in comparable unregulated units. I find that the average gain due to rent control is highest for high-income renting households, and that there are large differences in the gain for households with similar income. More than 95 percent of the total variation in the gains is caused by differences in the gains for households with similar income. Consequently, rent control is very imprecise and not targeted low-income households within the group of renters in the Danish population.

* Denne artikel bygger på en delanalyse i kapitel IV i DØR (2023) og dertilhørende baggrundsnotat. Jeg kunne ikke have lavet dette arbejde uden formændene for De Økonomiske Råd og mine kollegaer i DØRS, særligt Jesper Kühl, Anne Kristine Høj og Thomas Bjørner. Mange tak for det gode samarbejde, sparring og hjælp.

1. Indledning

Boligforhold har stor betydning på vores liv. Den geografiske placering er afgørende for, hvilke jobs vi kan varetage, den omgangskreds vores børn vokser op i og de rekreative muligheder i vores fritid. Boligens størrelse, indretning og kvalitet sætter ligeledes de grundlæggende rammer for vores liv og har betydning for vores livskvalitet. Endelig er huslejen den største udgiftspost for den gennemsnitlige husstand og udgør ca. 25 pct. af privatforbruget for husstande, som bor til leje, jf. DØR (2023).

Lejeboliger udgør en stor del af det samlede boligmarked i Danmark. Omkring 40 pct. af alle boliger er lejeboliger, og halvdelen af disse boliger er private udlejningsboliger, jf. DØR (2023). Huslejen ved indflytning er reguleret for de ca. 75 pct. af private lejeboliger, der er opført frem til 1991. Det betyder, at ca. 420.000 boliger er huslejereguleret, hvorfor lejen i de fleste af disse boliger er lavere end, hvis lejen var fastsat på markedsvilkår. Det indebærer, at ca. 800.000 personer bor i regulerede, private lejeboliger.

Et ofte fremført argument for huslejeregulering er, at det sikrer billigere boliger for husstande med lave indkomster – særligt i de største byer, hvor boligpriserne er særligt høje. Det er bl.a. blevet fremført som argument, at beboersammensætningen i de største byer som København og Aarhus bør afspejle et bredt udsnit af befolkningen, så beboerne ikke kun er de familier, som har råd til en bolig på markedsvilkår.

I denne analyse vil jeg undersøge, om den økonomiske gevinst ved lavere husleje er målrettet husstande med lav indkomst, og om husstande med samme (lave) indkomster har samme gevinst. Herved undersøges det, om alle lejere med relativt lave indkomster får glæde af de lavere huslejer.

Min hypotese er, at huslejeregulering ikke er målrettet husstande med lav indkomst. Reguleringen indebærer, at der er for mange, som ønsker at bo i de billige regulerede lejeboliger i forhold til udbuddet af boliger. Det betyder, at de regulerede lejeboliger fordeles ud fra eksempelvis ventelister, netværk eller udlejernes præferencer. Det tilsiger, at det til en vis grad kan være tilfældigheder, som afgør, om en husstand får tilbudt en billig lejebolig.

Jeg betegner den økonomiske gevinst ved lavere husleje som følge af huslejeregulering som *reguleringsgevinsten*. Reguleringsgevinsten giver en indikation på i hvor høj grad, huslejereguleringen nedsætter huslejen i en reguleret bolig.¹ Fokus i analysen er på de ældre, regulerede private lejeboliger, der er opført frem til 1991. Jeg medtager ikke almene boliger, der foruden huslejeregulering også er kendetegnet ved subsidiering i form af offentligt kapiteltilskud ved opførelse og til løbende udgifter.

1. Reguleringsgevinsten er dog alene udtryk for en økonomisk gevinst for de nuværende lejere og ikke et udtryk for en samfundsmæssig gevinst ved huslejeregulering.

2 FORDELINGSEFFEKTER AF HUSLEJEREGULERING

Reguleringsgevinsten er forskellen mellem den faktiske, regulerede husleje og den kontrafaktiske, markedsbestemte husleje i fravær af regulering. Der foreligger oplysninger om den faktiske husleje for 40 pct. af de private lejeboliger primært ud fra husstande, der modtager boligstøtte. Disse boliger er dog ikke repræsentative for alle lejeboliger, og eksempelvis er husstandsindkomsten forholdsvis lav. Jeg imputerer derfor huslejen for de resterende lejeboliger ud fra den gennemsnitlige husleje pr. kvm i bygningen og Heckmans to trins selektionsmodel, jf. DREAM (2012). Det giver et mål for den faktiske husleje i alle lejeboliger.

Den markedsbestemte husleje estimerer jeg ud fra huslejen i sammenlignelige uregulerede private lejeboliger.² Jeg bruger en *random forest*-model, da denne er velegnet til estimation, hvor formålet er prædiktion, jf. Hastie mfl. (2008) og afsnit 5.2. Modellen estimerer den markedsbestemte husleje pr. kvm. ud fra boligens fysiske og geografiske karakteristika, eksempelvis boligens størrelse, antal værelser, geografisk placering defineret ud fra postnumre og miljøfaktorer som støj, luftforurening samt nærhed til natur.

For at belyse, om reguleringsgevinsten tilfalder husstande med lav indkomst, undersøges fordelingen af reguleringsgevinsterne på tværs af og indenfor indkomstdeciler. Indkomstdecilerne er dannet ud husstandens placering i fordelingen af den ækvivalerede disponible indkomst for alle husstande i Danmark.

Analysen viser, at huslejereguleringen ikke synes at være målrettet husstande med lav indkomst. For det første er den årlige reguleringsgevinst i ældre, regulerede lejeboliger højere for husstande med høje indkomster. For det andet er der store forskelle i gevinsten for husstande med omtrent samme indkomst. Det betyder, at nogle lavindkomsthusstande har en meget stor gevinst, mens andre har en meget lille gevinst. Eksempelvis har knap 20 pct. af husstandene i anden indkomstdecil ingen reguleringsgevinst, mens knap ti pct. har en årlig gevinst på mindst 40.000 kr. Det betyder også, at langt størstedelen af de samlede forskelle i gevinsterne skyldes forskelle for husstande med omtrent samme indkomst. For det tredje er huslejeregulering ikke en effektiv måde at støtte boligforbruget for lejere med lav indkomst. Huslejereguleringen mindsker dog den samlede indkomstulighed i samfundet, fordi de fleste lejere har lave indkomster sammenlignet med ejerboligindehavere. Andre instrumenter, eksempelvis boligstøtten, ser ud til at være mere målrettet til at støtte boligforbruget blandt husstande med lave indkomster. Huslejeregulering kan derfor opfattes som at skyde med spredeshagl. Det gavner både husstande i og udenfor målgruppen, og blandt husstandene i målgruppen får nogle meget gavn af det, mens andre næsten ikke får gavn af det.

Min analyse bidrager til de danske analyser, som har beregnet reguleringsgevinsterne og analyseret deres fordeling (DØR, 2001; DREAM, 2012; Bloze og

2. En lejebolig er ureguleret, hvis den er opført efter 1991, og beboerne har boet højst to år i den.

Skak, 2013). Analysens primære bidrag er at undersøge fordelingen af reguleringsgevinsterne på et horisontalt plan. Dvs. sammenligne gevinsterne for husstande med omtrent samme indkomst. DØR (2001) og DREAM (2012) fokuserer på forskelle i gevinsterne for husstande med forskellig indkomst og viser, at gevinsterne typisk er størst for de rigeste husstande. Jeg viser, at forskellene i reguleringsgevinsterne på tværs af alle husstande primært skyldes forskelle for husstande med samme indkomst – ikke forskelle på tværs af indkomstgrupper. Det skyldes formentlig, at tilfældigheder spiller en central rolle i, hvem som får de attraktive, billige lejeboliger. Denne tilfældighed er med til at underbygge, at huslejerregulering ikke er målrettet husstande med lav indkomst.

Sekundært bidrager denne analyse med en ny og mere præcis metode til at beregne reguleringsgevinsterne. Jeg beregner den kontrafaktiske, markedsbestemte husleje ud fra lejen i uregulerede lejeboliger vha. en machine learning-teknik. DØR (2001) og DREAM (2012) anvender salgspriser på ejerboliger til at estimere boligværdien af en lejebolig og derefter konvertere dette til en husleje. Jeg viser, at denne metode har større usikkerhed i sin estimation af reguleringsgevinsterne. Ved at bruge en machine learning-teknik følger jeg også metoden til at estimere boligværdier i flere nyere danske studier, jf. Nationalbanken (2023) og DREAM (2023). Jeg finder samme kvalitative resultater som DØR (2001) og DREAM (2012), og denne analyse validerer derfor deres resultater, omend de er forbundet med større usikkerhed.

Resten af papiret er opdelt på følgende måde. I afsnit 2 redegør jeg for reglerne for huslejerregulering, mens afsnit 3 beskriver teori for fastsættelsen af den markedsbestemte husleje. Afsnit 4 giver et overblik over det anvendte data, og i afsnit 5 beskriver jeg metoden til at beregne reguleringsgevinsterne. Jeg analyserer fordelingen af reguleringsgevinsterne i afsnit 6 og diskuterer disse resultater i afsnit 7. Papirets konklusion er i afsnit 8.

2. Regler om huslejerregulering i Danmark

Alle private lejeboliger i Danmark har en form for huslejerregulering, men overordnet afhænger graden af huslejerregulering, om boligen er opført frem til 1991 eller derefter.³ Lejeboliger opført frem til 1991 er strammest reguleret, idet både lejen ved nydulejning og den efterfølgende stigningstakst er reguleret. Der skelnes mellem *omkostningsbestemt leje* og *leje efter det lejedes værdi*.⁴ Ved førstnævnte regulering må huslejen ikke overstige omkostningerne,⁵ mens leje ved det lejedes

3. Se DØR (2023) for en detaljeret beskrivelse for reglerne om huslejefastsættelse.

4. Det er primært beliggenhedskommune, og om boligen er gennemgribende renoveret, der afgør, hvordan huslejen reguleres for boliger opført frem til 1991.

5. Der tillades et lille kapitalafkast.

4 FORDELINGSEFFEKTER AF HUSLEJEREGULERING

værdi fastsættes efter lejen for lejemål med tilsvarende lokal beliggenhed, art, størrelse, kvalitet, udstyr og vedligeholdelsestilstand. Lejen efter det lejedes værdi er typisk en del højere end den omkostningsbestemte husleje, jf. Transport- og boligministeriet (2019).

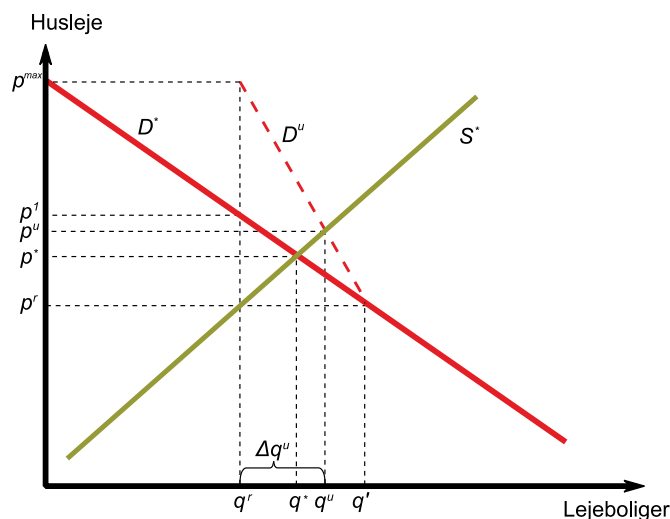
Lejeboliger opført efter 1991 har fri lejefastsættelse ved nyudlejning. Det betyder, lejer og udlejer frit aftale husleje på markedsvilkår, hvilket typisk fører til højere end den regulerede leje for boliger opført frem til 1991. Eksempelvis er den gennemsnitlige husleje pr. kvadratmeter for boliger opført efter 1991 ca. 50 pct. højere i København. Efter indflytningen må lejen stige med udviklingen i nettoprisindekset. Nettopriserne vokser typisk langsommere end den markedsbestemte husleje, jf. DØR (2023). Det betyder, at huslejen også er reguleret for lejeboliger opført efter 1991, idet huslejen i igangværende lejeforløb ikke må stige i samme grad som udviklingen i den markedsbestemte husleje. Efter fraflytning kan udlejer og en ny lejer frit aftale lejen for boligen, så lejen igen afspejler markedislejen.

3. Fastsættelse af markedsbestemt husleje

I en idealiseret situation, hvor lejeboligmarkedet er karakteriseret ved fuldkommen konkurrence og ingen markedsfejl, vil den markedsbestemte husleje ligge på et efficient niveau, så boligudbuddet svarer til efterspørgslen, svarende til p^* i figur 1. I det danske lejeboligmarked er der mindst tre årsager til, at den markedsbestemte husleje for de uregulerede lejeboliger kan være højere end under fuldkommen konkurrence.

For det første kan regulering af en del af lejeboligmarkedet medføre en højere leje på det uregulerede delmarked sammenlignet med huslejen på et helt ureguleret lejemarked (p^u vs. p^*). I en dansk kontekst betyder det, at huslejereguleringen for de ældre, regulerede private lejeboliger kan øge huslejen for boliger opført efter 1991, hvor huslejen ved nyudlejning kan aftales frit. Det skyldes, at en huslejeregulering med et bindende prisloft på huslejen, p^r , kan reducere udbuddet af regulerede lejeboliger til q^r , hvilket indebærer en overefterspørgsel efter lejeboliger, jf. figur 1. Efterspørgslen efter boliger på det uregulerede delmarked starter derfor ved q^r . De regulerede boliger fordeles ikke ud fra betalingsvillighed, og det medfører, at nogle af de lejere med høj betalingsvillighed fortrænges fra de regulerede lejeboliger og efterspørger uregulerede lejeboliger, jf. Bloze og Skak (2013). Denne gruppe af lejere har betalingsviljer for lejeboliger mellem p^r og p^{max} . Efterspørgselskurven på det uregulerede delmarked kan derfor vises ved den stiplede røde linje, D^u , der forbinder snitpunktet af p^{max} og q^r med snitpunktet af p^r og D^* , jf. Andersson og Söderberg (2012) og Mense mfl. (2023). I kombination med udbudskurven kan D^u indebære en ligevægtspris for uregulerede lejeboliger, p^u , som kan være højere end ligevægtsprisen i fravær af nogen form for regulering, p^* .

Figur 1
Markedsbestemt husleje i et todelt lejeboligmarked



Kilde: DØR (2023).

For det andet kan reguleringen af stigningstakten for lejeboliger med fri lejefastsættelse ved nyudlejning give udlejere et incitament til at sætte en højere husleje ved indflytning, jf. Basu & Emerson (2000). En lav stigningstakt indebærer på længere sigt et tab for udlejer og en tilsvarende besparelse for lejer. En højere husleje ved nyudlejning kan modgå dette tab for udlejer, og en lejer med en gennemsnitlig forventet lejeperiode vil være villig til at acceptere den høje husleje ved indflytning, da det modsvares af en lavere stigning i huslejen siden hen. Samlet set betyder det, at regulering af stigningstaksten i ellers uregulerede lejeboliger kan give en højere husleje ved nyudlejning, end hvis huslejen i igangværende lejeforløb kunne stige med udviklingen i den markedsbestemte husleje.

For det tredje kan udlejer have markedsmagt, der giver mulighed for at sætte en højere husleje end under fuldkommen konkurrence. Til forskel for de første to årsager handler denne pointe ikke om reguleringen af lejeboligmarkedet, men skyldes en potentiel markedsfejl. Antallet af boliger, der opfylder den enkelte lejers ønsker, kan være begrænsede, selvom lejeboligmarkedet som udgangspunkt består af mange potentielle boliger. Det kan resultere i, at der er et begrænset antal boliger at vælge imellem, hvis den boligsøgende har meget faste kriterier til f.eks. en lejeboligs indretning, beliggenhed og størrelse. Det kan give udlejer en vis grad af markedsmagt og dermed mulighed for at kræve en højere husleje, jf. Arnott (1989) og Raess og Ungern-Sternberg (2002). Ligeledes kan udlejer have markedsmagt, hvis enkelte udlejere ejer en stor del af lejeboligerne i et område. Det kan medføre højere husleje, hvis efterspørgslen efter boliger i et givet område

6 FORDELINGSEFFEKTER AF HUSLEJEREGULERING

er inelastisk. En analyse fra Nationalbanken (2022) viser en høj koncentration af ejerskab af lejeboliger ejet af virksomheder i en række kommuner og finder tegn på, at dette kan have ført til en lidt højere husleje i pågældende kommuner.

4. Data

Analysen bygger på registerdata fra Danmarks Statistik i perioden 2010-21.⁶ Jeg fokuserer på husstande i private lejeboliger og definerer dem som boliger beboet af en *lejer* og ejet af privatpersoner eller selskaber. Kollegier og offentligt ejede lejeboliger indgår ikke i analysen, da disse boliger er atypiske i forhold til det generelle lejemarked. Ligeledes indgår almene boliger ikke i analysen, der foruden huslejeregulering også er kendetegnet ved subsidiering i form af offentligt kapitaltilskud ved opførelse og til løbende udgifter.

De regulerede, private lejeboliger er opført frem til 1991 og består af boliger med omkostningsbestemt leje og leje efter det lejedes værdi, jf. afsnit 2. Ideelt set var disse lejeboliger opdelt i de to grupper, men ud fra mit datagrundlag er det ikke muligt troværdigt at inddele i de to reguleringstyper.⁷ Jeg betegner nyudlejede boliger opført efter 1991 som *uregulerede lejeboliger*, hvor lejen frit aftales uden nogen direkte regulering. En bolig er nyudlejet, hvis lejerne har boet i boligen i højst to år. Der ca. 560.000 private lejeboliger i Danmark. Heraf er ca. 420.000 opført frem til 1991, mens der er ca. 35.000 uregulerede lejeboliger.⁸

Den markedsbestemte husleje estimeres ud fra nyudlejede boliger opført efter 1991, jf. afsnit 5. Det kan dog medføre, at den estimerede, markedsbestemte husleje overvurderes ift. i en situation, hvor huslejen i alle lejeboliger blev fastsat på markedsvilkår, jf. afsnit 3. En følsomhedsanalyse indikerer dog, at dette primært vurderes at påvirke niveauet af gevinsterne og ikke fordelingen, jf. afsnit 6.2.

Husstandene inddeles i ti grupper baseret ud fra deres placering i fordelingen af den ækvivalerede disponible indkomst for alle husstande i Danmark. På den måde sammenlignes reguleringsgevinsten på tværs af husstande med forskellig indkomst og mellem husstande med omtrent samme indkomst. Indkomsten og alle andre monetære variable, fx huslejen og reguleringsgevinsten, er ækvivaleret. På den måde bliver den sammenlignelig på tværs af husstande med forskelli-

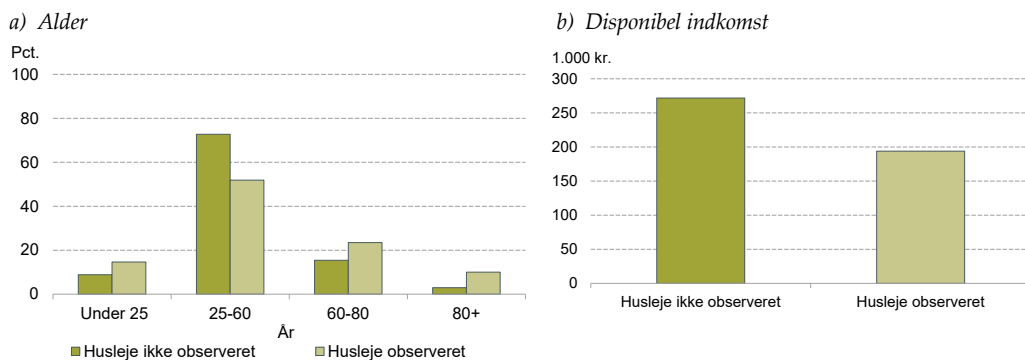
6. Det er ikke muligt ud fra data fra Danmarks Statistik at identificere, hvilke boliger som er udlejede i perioden 2019-21. Jeg antager derfor, at en bolig er en lejebolig i 2019-21, hvis den var en lejebolig i 2016-18. På den måde kan jeg identificere op til 90 pct. af alle lejeboligerne i 2019-21.
7. Data fra EjendomDanmark giver mulighed for at gøre det for ca. 5 pct. af de ældre private lejeboliger, jf. diskussionen i afsnit 7.2.
8. De resterende ca. 100.000 private lejeboliger er opført efter 1991, hvor beboerne har boet mindst to år i boligen.

ge størrelser, idet der tages samtidig højde for stordriftsfordele ved at bo flere personer i samme husstand.

Huslejeoplysningerne stammer fra to kilder. For det første bygger huslejen i perioden 2010-21 primært på boligstøtteregisteret. Boligstøtteregistret bygger på data fra Udbetaling Danmark. I 2018 er huslejedata mangelfulde, da udbetalingerne overgik fra KMD til ATP, og 2018 indgår af denne årsag ikke i analysen. For det andet suppleres huslejeoplysningerne i årene 2014-21 med data fra EjendomDanmark, som har indsamlet huslejen for et større udsnit af deres medlemsorganisationer.

Jeg observerer huslejen for ca. 40 pct. af de private lejeboliger ud fra boligstøtteregistret og EjendomDanmarks data. Husstandene i disse boliger er ikke repræsentative for alle husstande i private lejeboliger, jf. figur 2. Husstandene i disse boliger har i gennemsnit en lavere indkomst. Dertil er de enten ældre eller yngre end husstandene, der bor i boliger uden observeret husleje. Det skyldes, at huslejen i private lejeboliger primært observeres, hvis husstanden modtager boligstøtte, der i højere grad gives til husstande med lavere indkomst og folkepensionister, jf. DØR (2023).

Figur 2
Husstandskarakteristika opdelt på om husleje observeres i data



Anm.: Figuren i panel a) viser fordelingen af husstandene ud fra alderen på den ældste person i husstanden, mens figuren i panel b) viser den gennemsnitlige husstandens samlede disponible indkomst.

Kilde: Egne beregninger på baggrund af registerdata og data fra EjendomDanmark.

Huslejerne imputeres derfor i de resterende 60 pct. af lejeboligerne med samme metode som i DREAM (2012). For det første imputeres huslejen som den gennemsnitlige kvadratmeterleje i bygningen. Det gøres for ca. halvdelen af boligerne uden observeret husleje, da metoden kræver, at mindst én anden bolig i bygningen har observeret husleje. For det andet og for den resterende halvdel, som bor i bygninger uden andre boligstøttemodtagere, beregnes huslejen i to trin ved

8 FORDELINGSEFFEKTER AF HUSLEJEREGULERING

hjælp af Heckmans selektionsmodel (Heckman, 1979). I første trin estimerer jeg sandsynligheden for, at en husstand modtager boligstøtte. I andet trin indgår denne sandsynlighed som en forklaringsfaktor i en OLS-regression, der estimerer huslejen ud fra en række karakteristika om boligen og dens beliggenhed. På den måde kan der (til en vis grad) tages højde for selektionen, idet jeg direkte kontrollerer for sandsynligheden for at modtage boligstøtte. Analysens konklusioner ændres dog ikke ved kun at inkludere husstande, hvor huslejen observeres, jf. afsnit 6.2.

5. Metode til beregning af reguleringsgevinsterne

Dette afsnit beskriver, hvordan de såkaldte reguleringsgevinster beregnes. Jeg definerer reguleringsgevinsten (RG) som forskellen mellem den estimerede markedsbestemte husleje (hl^{mb}) og den faktiske, regulerede husleje (hl):

$$RG = hl^{mb} - hl$$

Den faktiske husleje observeres eller imputeres, jf. afsnit 4. Den markedsbestemte husleje estimeres på baggrund af huslejen i uregulerede lejeboliger, hvor lejen kan opfattes som markedsbestemt.

5.1. Estimation af markedsbestemt husleje

For at beregne den markedsbestemte husleje for regulerede lejeboliger estimerer jeg en model ud fra sammenlignelige uregulerede lejeboliger. Jeg bruger en *random forest*-metode, da formålet er at prædiktere den markedsbestemte husleje og ikke at finde årsagssammenhængen mellem huslejen og forklaringsfaktorerne. Her er random forest-metoden (og lignende machine learning-metoder) velegnet ift. lineære unbiased estimators som OLS, jf. Hastie mfl. (2008) og afsnit 5.2.

Random forest-metoden opstiller ikke en *a priori* funktionsform for sammenhængen mellem den markedsbestemte husleje og forklaringsfaktorerne. I stedet opstilles et meget stort antal modeller, som hver især prædikterer den markedsbestemte husleje, med forskellige udpluk af forklaringsfaktorer i hver model. I hver model udvælges de faktorer, som er vigtigst for at prædiktere den markedsbestemte husleje. Den endelige estimation opgøres herefter som et simpelt gennemsnit over alle de opstillede modellers prædiktioner. Se Athey og Imbens (2019) og Hastie mfl. (2008) for en dybdegående beskrivelse af random forest metoden.

Konkret estimerer jeg den markedsbestemte husleje pr. kvadratmeter ud fra boligernes fysiske og geografiske karakteristika, jf. tabel 1. Disse karakteristika er blandt andet boligens størrelse, antal værelser, geografisk placering defineret ud

fra postnumre og miljøfaktorer som støj, luftforurening samt nærhed til natur.⁹ Disse karakteristika er valgt ud fra de faktorer, som anses for vigtigst i at bestemme boligens værdi, jf. Freeman (1993) og Sopranzetti (2015).

Tabel 1
Variable i random forest-modellen

Outcome variabel	Husleje pr. kvm
Forklarende variable:	
<i>Kontinuerte variable</i>	Boligens størrelse, antal etager i bygningen, antal badeværelser, antal værelser, antal toiletter, afstand til kyst, større søer og skove, luftforurening, støj, ejendomsskatten og kalenderår.
<i>Kategoriske variable (dummy for hver kategori)</i>	Elevator i opgang, varmeinstallation, opvarmingsmiddel, køkkenforhold og postnummer ^a .

a) Der er ikke en dummy for hvert postnummer, da der i stedet bruges *sufficient representation*-tilgangen, jf. Johannemann mfl. (2019) og Nationalbanken (2023).

Analysen kan ikke tage højde for betydningen af boligens opførelsesår på den markedsbestemte husleje, fordi de ældre, regulerede lejeboliger er opført frem til 1991, mens de uregulerede lejeboliger er opført efter 1991. Jeg tager delvist højde for dette ved at skalere den markedsbestemte husleje ud fra effekten af opførelsesår på salgsprisen for ejerboliger.¹⁰ Se DØRS (2023) for en detaljeret gennemgang af metoden.

5.2. Prædiktionskvalitet

De estimerede sammenhænge mellem huslejen og de forklarende faktorer ved random forest-metoden er svære at gennemskue. Eksempelvis viser random forest-metoden ikke marginale effekt af de forskellige forklarende variable ligesom en OLS-model. Det kan gøre det sværere at vurdere, om modellen er velspecificeret, og hvor god modellen er til at prædiktere den markedsbestemte husleje.¹¹

9. Det kan være problematisk at inkludere områdevariable (her postnumre) som fixed effects i en random forest-estimator, jf. Johannemann mfl. (2019) og Hastie mfl. (2008). Der tages derfor højde for boligens beliggenhed vha. *sufficient representation*-tilgangen, jf. Johannemann mfl. (2019) og Nationalbanken (2023). Konkret bruger jeg *means encoding*, hvor jeg beregner gruppenspecifikke gennemsnit af alle de øvrige forklarende variable for hvert postnummer. Disse gennemsnitsvariable tilføjes i modellen i stedet for fixed effects.
10. Et robusthedstjek viser, at *fordelingen*, men ikke *niveauet*, af reguleringsgevinsten er robuste for, hvordan der kontrolleres for boligens opførelsesår, jf. DØRS (2023).
11. DØRS (2023) argumenterer dertil for, at modellen synes velspecificeret på baggrund af, hvilket variable som er vigtigst i at prædiktere den markedsbestemte husleje, og sammenlængningen mellem den estimerede markedsbestemte husleje og udvalgte inputvariable.

Jeg kan dog teste, hvor præcis modellen er til at prædiktere den markedsbestemte husleje for uregulerede lejeboliger. Det gøres ved at sammenligne den observerede (markedsbestemte) husleje pr. kvm for uregulerede lejeboliger med den estimerede markedsbestemte husleje pr. kvm ud fra random forest-metoden.

Prædiktionskvaliteten måles ved at sammenligne den estimerede husleje med den faktiske husleje i et sample af 30 pct. af uregulerede lejeboliger, som ikke indgik i samplet til at estimere modellerne,¹² dvs. en out-of-sample prædiktionskvalitet. Konkret vurderer jeg prædiktionskvaliteten ud fra R^2 samt andelen af estimerede markedsbestemte huslejer indenfor 10 og 20 pct. af den faktiske (markedsbestemte) husleje. Alle mål ville være 100 pct., hvis modellen helt nøjagtig kunne prædiktere den markedsbestemte husleje.

Prædiktionskvaliteten forekommer overordnet set høj, jf. tabel 2. I den foretrukne specifikation (baseline-modellen) er R^2 ca. 89 pct., og i 83 (94) pct. af tilfældene er den prædikterede markedsbestemte husleje indenfor 10 (20) pct. af den faktiske husleje. Anvendes i stedet en OLS-metode, er prædiktionskvaliteten markant lavere. R^2 er 72 pct., mens andelen inden for 10 pct. er 59 pct. Det viser, at random forest metoden er bedre til at prædiktere den markedsbestemte husleje end en OLS-model.

Tabel 2
Prædiktionskvalitet af random forest-modellen

	R^2	Andel indenfor 10 pct.	Andel indenfor 20 pct.
	Baseline model		
Baseline, random forest	0,89	0,83	0,94
OLS	0,72	0,59	0,85
Ejerboligmetoden (RF)	0,13	0,37	0,66

Anm.: Første række viser prædiktionskvaliteten for baseline modellen, jf. afsnit 3.1 og 3.2. Anden række estimerer modellen vha. en OLS-model i stedet for en random forest-model. Tredje række estimerer huslejen ud fra salgsprisen på ejerboliger. Der bruges en random forest model til estimationen. Tabellen bygger på 30 pct. af nyudlejede boliger opført efter 1991, som ikke indgik i estimationen, altså en out-of-sample prædiktionskvalitet. $R^2 = 1 - \frac{MSE}{var(husleje)}$, hvor MSE er summen af kvadrerede fejlede. Andel indenfor 10 (20) pct. viser, hvor stor en andel af observationerne som er indenfor 10 (20) pct. af den faktiske husleje. OLS-modellen medtager de samme inputvariable som random forest modellen.

Kilde: Egne beregninger på baggrund af registerdata og data fra EjendomDanmark.

12. Prædiktionskvaliteten kan kun testes på nyudlejede boliger opført efter 1991 og ikke de uregulerede lejeboliger, da den markedsbestemte husleje ikke kendes for disse boliger.

Estimation ud fra husleje i stedet for ejerboligpriser

Som tidligere beskrevet estimerer jeg den markedsbestemte husleje ud fra huslejen i uregulerede lejeboliger. På den måde estimeres den markedsbestemte husleje mere præcist, end hvis huslejen var estimeret den ud fra værdien på ejerboliger, som er den anvendte metode i bl.a. DØR (2001) og DREAM (2012). Ved denne »ejerboligmetode« estimeres først værdien af en lejebolig på baggrund af værdien af sammenlignelige ejerboliger. Værdien af lejeboligen konverteres herefter til en »markedsbestemt« husleje ved at antage, at værdien af lejeboligen skal forrentes med et (kalibreret) afkast på den bedste alternative investering. Denne metode skaber større usikkerhed, og det ses ved, at R^2 kun er 13 pct., mens andelen inden for 10 pct. er 37 pct., jf. tabel 2. Det skyldes både, at det er mere usikkert at estimere ejerboligværdien snarere end huslejen, og der er en vis usikkerhed ved at beregne det korrekte afkast på den bedste alternative investering.

6. Resultater

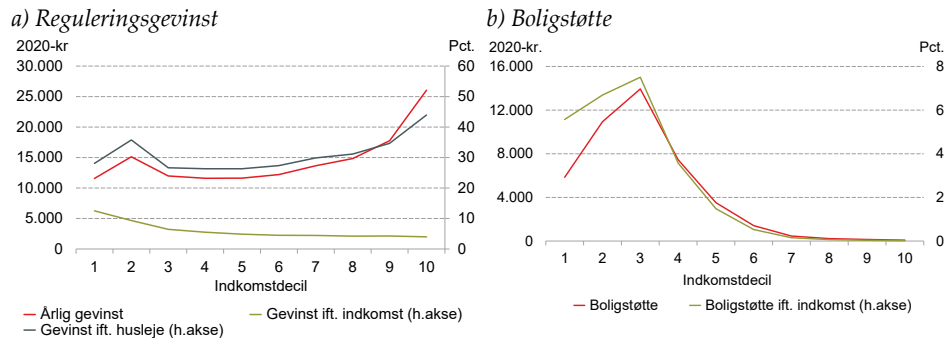
I dette afsnit viser jeg fordelingen af reguleringsgevinsterne med fokus på fordelingen på tværs af og indenfor de indkomstdeciler.¹³ På den måde undersøges det, om huslejeregulering er målrettet husstande med lav indkomst.

6.1. Forskelle på tværs af og indenfor indkomstdeciler

Den gennemsnitlige årlige ækvivaleret reguleringsgevinst ligger mellem 12 og 25.000 kr., hvilket svarer til 25-45 pct. af huslejen, jf. figur 3a. Den absolutte gevinst er størst i den højeste indkomstdecil, hvor den er dobbelt så stor som gevinsten i den laveste indkomstdecil. Det afspejler blandt andet, at de attraktive ældre lejeboliger med lavere husleje ofte tilfalder husstande med høj indkomst, blandt andet fordi disse husstande i højere grad bor i Københavnsområdet. Billedet er omtrent det samme, når gevinsten opgøres relativt til huslejen, hvor gevinsten udgør knap 30 pct. af huslejen i den laveste indkomstdecil og ca. 45 pct. i den højeste gruppe. Reguleringsgevinsten relativt til husstandenes disponible indkomst er dog størst for husstande med lave indkomster. I gennemsnit er gevinsten omkring 12 pct. af indkomsten i den laveste indkomstdecil og ca. 4 pct. af indkomsten i den højeste gruppe. Det betyder, at reguleringsgevinsten udgør en mindre del af husstandenes indkomst, jo højere indkomsten er.

13. Analysens konklusioner ændres ikke ved opdeling i indkomstpercentiler i stedet for deciler, jf. DØRS (2023).

Figur 3
Reguleringsgevinster og boligstøtte fordelt på tværs af indkomstdeciler



Anm.: Figurene viser den gennemsnitlige reguleringsgevinst opgjort på lejeboligtype og indkomstdecil. Reguleringsgevinsten og den disponible indkomst er ækvivaleret. Indkomstdecilerne er dannet ud fra indkomstdecilerne beregnet ud fra alle husstande i Danmark uanset boligform, men kun husstande, som bor i lejeboliger, er inkluderet i figuren.

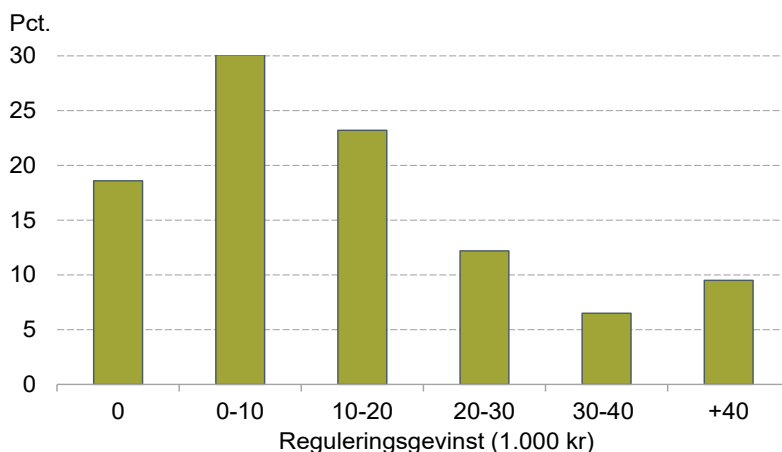
Kilde: Egne beregninger på baggrund af registerdata og data fra EjendomDanmark.

Ovenstående billede er i overensstemmelse med fordelingen af reguleringsgevinsterne i DØR (2001) og DREAM (2012). DØR (2001) viser, at gevinsterne i private lejeboliger i 1999 er dobbelt så store for husstande i den højeste indkomstdecil ift. de fleste husstande med lave indkomster.¹⁴ Det samme finder DREAM (2012) for gevinsterne i 2010, ligesom DREAM også viser, at gevinsterne relativt til indkomsten er højest for husstande med lave indkomster. Der er altså bred empirisk belæg for, at den absolutte gennemsnitlige reguleringsgevinst blandt lejerne er størst for husstandene med de højeste indkomster, mens gevinsterne relativt til indkomsten aftager med indkomsten.

Det er ikke kun på tværs af indkomstdecilerne, der er store forskelle i reguleringsgevinsterne. Indenfor hver indkomstdecil er der også store forskelle, jf. figur 4. De 10 pct. af husstandene med de største gevinster i anden indkomstdecil har ca. 40 pct. af de samlede gevinster for alle husstande i anden indkomstdecil svarende til årlige reguleringsgevinster på mindst 40.000 kr. Til sammenligning er der knap 20 pct. af husstandene i anden indkomstdecil, som ikke har en reguleringsgevinst.

14. DØR (2001) finder dog, at gevinsterne i den laveste indkomstdecil er forholdsvis høj, så der snarere er en u-formet sammenhæng mellem reguleringsgevinsterne og husstandsindkomsten.

Figur 4
Fordeling af reguleringsgevinster indenfor 2. indkomstdecil



Anm.: Figuren viser fordelingen af ækvivaleret reguleringsgevinster i huslejen for husstande i 2. indkomstdecil. Husstande med negative gevinster indgår i nulsøjlen Indkomstdecilerne er dannet ud fra indkomstdecilerne beregnet ud fra alle husstande i Danmark uanset boligform.

Kilde: Egne beregninger på baggrund af registerdata og data fra EjendomDanmark.

Det såkaldte Theil-indeks kan dekomponere forskellene i reguleringsgevinsterne i forskellene på tværs af og indenfor indkomstdecilerne.¹⁵ Herved kan jeg undersøge, om forskellene i gevinsterne primært skyldes forskelle indenfor eller på tværs af indkomstdecilerne. I ældre, regulerede private lejeboliger afspejler ca. 97 pct. af forskellene i reguleringsgevinsterne forskelle indenfor indkomstdecilerne, jf. DØRS (2023).¹⁶ Forskellen i reguleringsgevinsterne kan altså næsten udelukkende tilskrives forskelle blandt husstande med samme indkomst. Ca. 1/5 af forskellen indenfor indkomstdeciler skyldes geografiske forskelle i gevinsterne, der er særligt høje i København. Der er dog stadig store forskelle i gevinsterne indenfor indkomstdecilerne blandt husstande boende i samme områder, jf. afsnit 6.2.

15. Theil-indekset er detaljeret beskrevet i DØRS (2023).

16. Ca. 95 pct. af forskellene i reguleringsgevinsterne skyldes stadig forskelle indenfor indkomstdecilerne, selvom husstandene opdeles i 100 grupper i stedet for ti.

Sammenligning med boligstøtte

Boligstøtte er en anden måde at støtte husstandenes boligforbrug, hvor størrelsen af støtten bl.a. afhænger af husstandens indkomst. Det er derfor ikke overraskende, at den gennemsnitlige boligstøtte generelt er højest for husstande med lav indkomst, jf. figur 3b. Det gælder både den absolutte boligstøtte og boligstøtten relativt til indkomsten. Dertil fremgår det, at husstande med høje indkomster ikke modtager boligstøtte. Det viser, at boligstøtten i højere grad tilfalder husstande med lav indkomst sammenlignet med gevinsterne ved huslejeregulering. Derudover får husstande med samme indkomst i højere grad samme boligstøtte.

6.2. Følsomhedsanalyse

Resultaterne i forrige afsnit er robuste overfor diverse metodeændringer og ændringer i opgørelsen. I dette afsnit viser jeg dette for to af disse ændringer, mens øvrige følsomhedsanalyser kan ses i DØRS (2023).

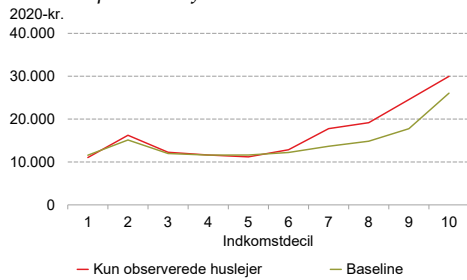
I baseline-modellen imputeres den faktiske husleje for ca. 60 pct. af lejeboligerne, jf. afsnit 4. Det giver en større usikkerhed om den faktiske husleje for disse boliger. Fordelingen af reguleringsgevinster ændres dog ikke det store ved kun at medtage boliger, hvor huslejen direkte observeres gennem boligstøtteregistret eller data fra EjendomDanmark, jf. figur 5a. Gevinsterne bliver endnu større for højindkomsthusstandene, og der er stadig store forskelle på gevinsterne for husstande med samme lave indkomster.

Som beskrevet i afsnit 3 overvurderes estimationen af reguleringsgevinsterne muligvis. Der er dog ikke tegn på, at denne overvurdering påvirker fordelingen af gevinsterne. Man vil forvente overvurderingen er størst i Københavnsområdet, hvor huslejereguleringen reducerer huslejen mest og dermed har den største effekt. Fordelingen af gevinster i Københavnsområdet er dog i overensstemmelse med fordelingen på landsplan, jf. figur 5b. Gevinsten er størst for husstande med høj indkomst, og der er store forskelle i gevinsterne for husstande med samme lave indkomster. Ikke overraskende er niveauet for gevinsterne højere end på landsplan. Det kan ikke udelukkes, at en del af de høje gevinster i Københavnsområdet skyldes større grad af overvurdering, men det må primært tilskrives den højere markedsbestemte husleje. Der er dog stadig over 20 pct. af husstandene, som har en gevinst under 10.000 kr. om året. Det er i overensstemmelse med, at en mulig overvurdering kan påvirke *niveauet* af gevinsterne, men ikke *fordelingen* af dem.

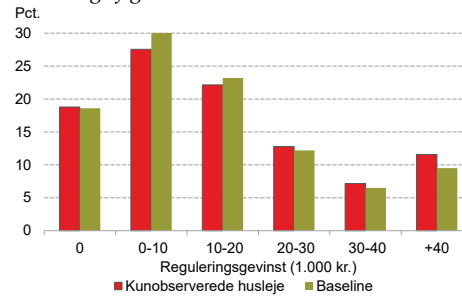
Figur 5
Følsomhedsberegninger

a) Kun boliger med observerede huslejer

Forskelle på tværs af indkomstdeciler

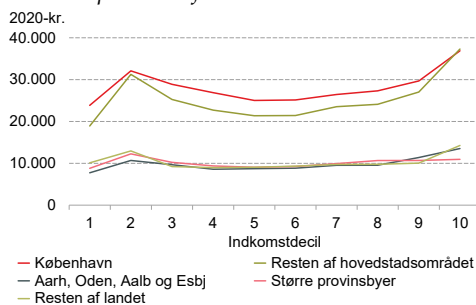


Fordeling af gevinster i 2. decil

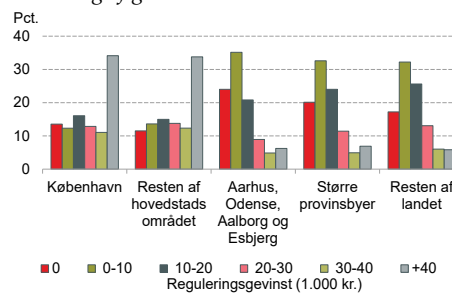


b) Gevinster indenfor for bygrupper

Forskelle på tværs af indkomstdeciler



Fordeling af gevinster i 2. decil



Anm.: De venstre figurer viser den gennemsnitlige reguleringsgevinst i hver indkomstdecil. De højre figurer viser fordelingen af reguleringsgevinster i huslejen for husstande i 2. indkomstdecil. Husstande med negative gevinster indgår i nul søjlen. I panel a) defineres beliggenheden ud fra kommunen eller skoledistriktet i stedet for postnumre (baseline). I panel b) medtages kun lejeboliger, hvor huslejen observeres. Reguleringsgevinsten og den disponible indkomst er ækvivaleret. Indkomstdecilerne er dannet ud fra indkomstdecilerne beregnet ud fra alle husstande i Danmark uanset boligform, men kun husstande, som bor i lejeboliger, er inkluderet i figuren. København inkluderer også Frederiksberg. Større provinsbyer omfatter kommuner, hvor 6.-20. største byer ligger.

Kilde: Egne beregninger på baggrund af registerdata og data fra EjendomDanmark.

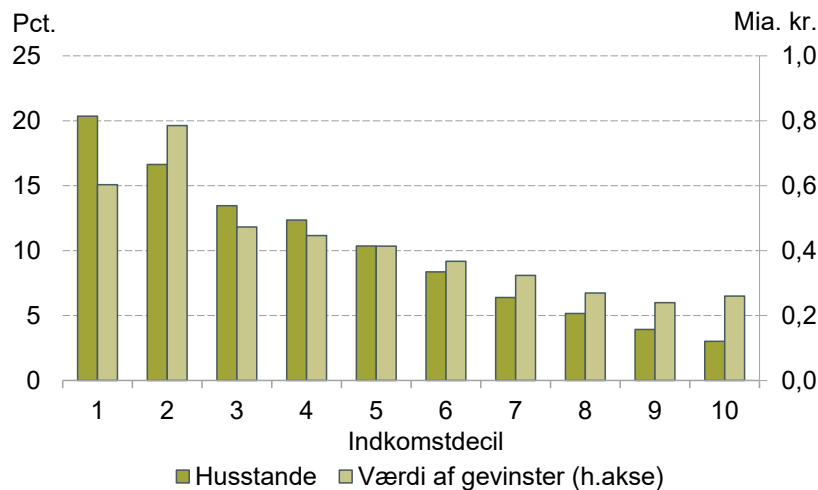
7. Diskussion

Resultaterne i forrige afsnit viser, at de høje reguleringsgevinster blandt lejerne især tilfalder højindkomsthusstandene, samt at der er store forskelle i gevinsterne for husstande med samme indkomst. Huslejeregulering synes derfor ikke at være målrettet husstande med lave indkomster. I dette afsnit diskuterer jeg denne konklusion og en række andre aspekter af analysen.

7.1. Omfordeling for lejere og hele befolkningen

Resultaterne i forrige afsnit fokuserer på fordelingen af reguleringsgevinsterne for husstande, som bor til leje. Jeg tager dermed ikke højde for, at langt størstedelen af lejerne har lave indkomster, hvor eksempelvis ca. 50 pct. af lejerne i ældre, regulerede lejeboliger tilhører de tre laveste indkomstdeciler, jf. figur 6. Husleje-reguleringen kan derfor godt samlet set være til gavn for husstande med lave indkomster, da langt flere af disse husstande bor til leje og derved får glæde af den lavere husleje.

Figur 6
Andel lejere og reguleringsgevinster fordelt på indkomstdeciler



Anm.: Figuren viser fordelingen af husstande og den samlede værdi af reguleringsgevinsterne i hver indkomstgruppe i regulerede, private lejeboliger. Der vil være 10 pct. i hver indkomstdecil, hvis indkomstfordelingen for disse husstande svarerede til indkomstfordelingen for alle husstande. Den samlede værdi af gevinsterne svarer til summen af gevinster for alle husstande i hver decil. Indkomstdecilerne er dannet ud fra alle husstande i Danmark uanset boligform, men kun husstande, som bor i lejeboliger, er inkluderet i figuren. Indkomst og gevinster er ækvivaleret.

Kilde: Egne beregninger på baggrund af registerdata.

Den samlede værdi af reguleringsgevinsterne er også størst for husstande i de laveste indkomstdeciler, selvom den gennemsnitlige gevinst er lavere for disse husstande. Eksempelvis er den samlede gevinst knap 2 mia. kr. i de tre laveste indkomstdeciler, mens den kun er knap 0,8 mia. kr. i de tre øverste deciler, jf. figur 6. Det betyder, at indkomstuligheden falder, når jeg medregner reguleringsgevinsten i den disponible indkomsten. Det er dog et beskedent fald, hvor eksempelvis Gini-koefficienten falder fra 0,301 til 0,299. Ved at lave samme øvelse med boligstøtten falder Gini-koefficienten lidt mere end ved reguleringsgevinsterne, men den samlede udbetalte boligstøtte er halv størrelse af den samlede værdi af regu-

leringsgevinsterne. Det illustrerer, at huslejereguleringen mindsker indkomstuligheden, men at eksempelvis boligstøtte er mere en mere effektiv måde til at tilgodese husstande med lav indkomst.

7.2. Gevinster ved omkostningsbestemt leje og det lejedes værdi

Som beskrevet i afsnit 4 har jeg undersøgt gevinsterne for private lejeboliger, der både er reguleret med omkostningsbestemt leje og leje efter det lejedes værdi. Omkostningsbestemt husleje er dog en væsentlig mere restriktiv regulering, hvorfor huslejen også er lavere i boliger med omkostningsbestemt leje ift. det lejedes værdi, jf. Transport og Boligministeriet (2019). Det betyder, at der må være stor forskel i reguleringsgevinsterne alt efter om, man bor i en bolig med omkostningsbestemt leje eller det lejedes værdi. Det kan potentielt forklare en del af de store forskelle i gevinsterne for husstande med samme indkomst.

I hovedanalysen kan jeg ikke skelne mellem, om en ældre, reguleret lejebolig har omkostningsbestemt leje eller leje efter det lejedes værdi. Ud fra data fra EjendomDanmark kan jeg dog for godt 5 pct. af de private lejeboliger opført frem til 1991 observere typen af huslejeregulering.¹⁷ Ved at tage højde for forskelle i reguleringsgevinsterne i hhv. boliger med omkostningsbestemt leje og leje efter det lejedes værdi falder forskellen i gevinsterne indenfor samme indkomstdecil med ca. 25 pct., så de udgør knap 75 pct. af de samlede forskelle. Blandt husstande i anden indkomstdecil, som bor i en omkostningsbestemt lejebolig, har knap 25 pct. en reguleringsgevinst under 10.000 kr., mens 35 pct. har en gevinst over 40.000 kr. Det indikerer, at en del af forskellene indenfor indkomstdecilerne skyldes forskelle i reguleringstypen, men der stadig er store forskelle i gevinsterne indenfor indkomstdecilerne for husstande med samme reguleringstype.

8. Konklusion

Store dele af det private lejeboligmarked er reguleret, hvilket giver lejerne betydelige besparelser i huslejen. Adgang til billigere boliger til lavindkomsthusstande er et centralt argument for den nuværende huslejeregulering, og i dette papir har jeg derfor undersøgt, i hvor høj grad den økonomiske gevinst ved lavere husleje tilfalder husstande med lave indkomster.

Analysen viser, at huslejereguleringen ikke synes at være målrettet husstande med lav indkomst af tre primære årsager. For det første er besparelsen i huslejen, den såkaldte *reguleringsgevinst*, højere for husstande med høje indkomster. Eksempelvis har de rigeste lejere en reguleringsgevinst, der dobbelt så stor som de

17. Fordelingen af reguleringsgevinster – både på tværs og indenfor indkomstdecilerne – er ca. i overensstemmelse med fordelingen af gevinster for alle private lejeboliger. Derfor synes resultaterne for disse boliger til en vis grad til at kunne generaliseres til alle private lejeboliger.

fattigste lejere. For det andet er der store forskelle i gevinsten for husstande med omtrent samme indkomst. Det betyder eksempelvis, at knap 20 pct. af lavindkomsthusstande ingen gevinst har, mens knap 10 pct. har en årlig gevinst på mindst 40.000 kr. For det tredje er huslejeregulering ikke den mest effektive måde at mindske den samlede indkomstulighed. Selvom huslejeregulering mindsker uligheden marginalt, er andre instrumenter, som eksempelvis boligstøtten, mere effektive, da de er mere målrettet husstande med lav indkomst.

Litteratur

- Andersson, R. og B. Söderberg (2012): Elimination of Rent Control in the Swedish Rental Housing Market: Why and How? *Journal of Housing Research*, 21 (2), s. 159-181.
- Arnott, R. (1989): Housing Vacancies, Thin Markets and Idiosyncratic Tastes. *Journal of Real Estate Finance and Economics*, 2, s. 5-30.
- Athey, S., og G. Imbens (2019): *Machine Learning Methods Economists Should Know About*. March 2019.
- Atkinson, A.B. (1975): *The Economics of Inequality*. Oxford University Press.
- Basu, K. og P.E. Emerson (2000): The Economics of Rent Control. *The Economic Journal*, 110, s. 939-962.
- Bloze, G, og M. Skak (2013): Rent Control and Misallocation. *Urban Studies* 50(10), s. 1988-2005.
- DREAM (2012): Kristensen, J.B. Konsekvenser af huslejeregulering på det private udlejningsmarkedet – En mikroøkonomisk analyse af 2000'erne, DREAM arbejdsrapport.
- DREAM (2023): Hansen, J. Z. og A. Ø. Iversen. Prisen på ejerboliger 1992-2021 – En husstandsbaseeret undersøgelse af den geografiske variation i boligprisudviklingen, DREAM analyse.
- DØR (2001): Det Økonomiske Råds formandskab. *Dansk Økonomi, forår 2001*
- DØR (2023): De Økonomiske Råds formandskab. *Dansk Økonomi, efterår 2023, kapitel IV*.
- DØRS (2023): De Økonomiske Råds sekretariat. Baggrundsnotat til analyse om huslejeregulering og fordeling – afsnit IV.5 i *Dansk Økonomi, efterår 2023*. Notatet kan tilgås via dors.dk.
- Freeman, A.M. (1993) Property value models, i *The measurement of environmental and resource values*, Washington, Resources for the future, 367-420.
- Hastie, T., Tibshirani, R., og J. Friedman (2008): *The Elements of Statistical Learning – Data Mining, Inference and Prediction*. 2nd edition.
- Heckman, J. J. (1979): Sample Selection as a Specification Error. *Econometrica*, 47, s. 153-161.

- Johannemann, J., Hadad, V., Athey, S., og St. Wager (2019): Sufficient Representation for Categorical Variables.
- Mense, A., C. Michelsen og K.A. Kholodilin (2023): Rent Control, Market Segmentation, and Misallocation: Causal Evidence from a LargeScale Policy Intervention. *Journal of Urban Economics*, 134.
- Nationalbanken (2022): Andersen, P., T. Krause og D. Gorea. Market concentration and the Danish rental market. Nationalbanken.
- Nationalbanken (2023): Andersen, P., A. Kuchler, P. V. Borgensgaard og B. B. Meyer. Housing wealth and consumption during Covid-19. Economic memo.
- Raess, P. og T. von Ungern-Sternberg (2002): A model of regulation in the rental housing market. *Regional Science and Urban Economics*, 32, s. 475–500.
- Sopranzetti, B.J. (2015): Hedonic Regression Models I Lee, E. C.-F og J.C. Lee (red.): *Handbook of Financial Econometrics and Statistics*. Springer.
- Theil, H. (1967): *Economics and Information Theory*. North-Holland Publishing Company.
- Transport og Boligministeriet (2019): Ekspertgruppens rapport om boligreguleringslovens § 5, stk. 2.