Statistisk sentralbyrå
Statistics Norway

Revidert modell for beregning av boligformue: Resultater fra modellkjøring 2015 - 2024

Per Medby og Anita Dal

TALL SOM FORTELLER
NOTATER / DOCUMENTS
2026/37

I serien Notater publiseres dokumentasjon, metodebeskrivelser, modellbeskrivelser og standarder.

© Statistisk sentralbyrå

Publisert: 15. oktober 2026

Endret 6. november 2026: side 9 og 10

ISBN 978-82-587-2133-5 (elektronisk)
ISSN 2535-7271 (elektronisk)

Standardtegn i tabeller	Symbol
Ikke mulig å oppgi tall Tall finnes ikke på dette tidspunktet fordi kategorien ikke var i bruk da tallene ble samlet inn.	.
Tallgrunnlag mangler Tall er ikke kommet inn i våre databaser eller er for usikre til å publiseres.	..
Vises ikke av konfidensialitetshensyn Tall publiseres ikke for å unngå å identifisere personer eller virksomheter.	:
Desimaltegn	,

Forord

Statistisk sentralbyrå (SSB) har laget en metode for å beregne markedsverdier for boliger i Norge (Jf. Notater 2009/53). Beregningene har blant annet blitt brukt i SSBs mikrosimuleringsmodell LOTTE. Modellen anslår endringer i skatteinntekter og inntektsfordeling som følge av endringer i inntektsskatt og formuesskatt for personer. Beregningene brukes også av Skatteetaten for å beregne formuesverdier for boliger.

Modellen har nå blitt vesentlig revidert. Revidert modell har blitt utviklet av forskerne Erlend Eide Bø og Odd Erik Nygård, med bidrag fra seniorrådgiver Eirik Sjåvik.

Dataene som er brukt i dette notatet, er fra perioden januar 2015 til november 2024. Publikasjonen er utarbeidet av seniorrådgiver Per Medby og seniorrådgiver Anita Dal, og inngår i en serie som utgis årlig og oppdateres med aktuelle resultater og eventuelle metodejusteringer for hvert år. Ansvarlig seksjonssjef er Roger Jensen, Seksjon for eiendoms-, areal- og primærnæringsstatistikk.

Statistisk sentralbyrå, 2. oktober 2026

Per Morten Holt

Sammendrag

Formålet med publikasjonen er å beskrive resultatene fra den reviderte modellen for beregning av boligformue basert på data fra årgangene 2015–2024. Resultatene fra modellen brukes til å beregne markedsverdier for boliger i Norge. Resultatene med disse årgangene ble brukt til å beregne verdiene på skattekortet for 2026. Modellen bygger på ti år med boligomsetninger og beregner gjennomsnittlige kvadratmeterpriser for boliger, avhengig av størrelse, alder og geografisk plassering.

Modellen har nå blitt vesentlig revidert. Den viktigste endringen er at modellen nå bruker grunnkretser som geografisk enhet i stedet for kommuner. Dette gjør modellen mindre sårbar for endringer i kommunestrukturen, og mer treffsikker siden den tillater mer heterogenitet i lokale priser. Grunnkretsene grupperes til beregningsområder ved hjelp av en maskinlæringsalgoritme, og for hvert beregningsområde og hver boligtype gjøres en separat hedonisk regresjon.

Innhold

Forord **3**
Sammendrag **4**
1. Innledning **6**
2. Datagrunnlaget **7**
3. Om beregningene av boligverdier **8**
3.1. Grunnlaget for beregningene 8
3.2. Regresjonsmodellen 9
3.3. Eksempel på beregning av boligverdi 10
4. Modellytelse **11**
5. Resultater **15**
Referanser **16**
Vedlegg A: Eksempel på beregningsområde **17**

1. Innledning

SSB har hatt en modell som har beregnet gjennomsnittlige kvadratmeterpriser per boligtype og geografisk område, som kan brukes til å beregne markedsverdien på boliger (jf. Notater 2009/53). Modellen brukes også av Skatteetaten for å beregne formuesverdier på boliger. Fra og med 2014 har også flere kommuner tatt i bruk disse formuesverdiene ved utskrivning av eiendomsskatt.

Modellen har nå blitt vesentlig revidert. Revisjonen skyldtes at det over tid ble avdekket svakheter i den opprinnelige modellen, særlig i forbindelse med kommunereformen. Kommuneinndelingen ble vesentlig endret per 1.1.2020. Det førte til at modellen traff dårligere enn før i noen regioner, fordi kommuner i forskjellige soner ble slått sammen. På denne bakgrunn fikk SSB i oppdrag av Finansdepartementet å vurdere en verdsettelsesmodell som i mindre grad er basert på administrative grenser slik at den ikke påvirkes av eventuelle endringer i kommunestruktur. Resultatene fra oppdraget er publisert i Notater 2022/11[1], Bø m.fl. (2022) og Notater 2023/3.[2] Beregninger viser at den reviderte modellen har høyere treffsikkerhet enn den opprinnelige modellen.

Grunnlagsdata i modellen er, som i den opprinnelige modellen, boligtransaksjoner fra Finn.no over en tiårsperiode. Boligverdi beregnes videre, på samme måte som i den opprinnelige modellen, separat for hver kombinasjon av boligtype og region (såkalte beregningsområder). Modellen er basert på faktiske variasjoner i prisnivå. Den største forskjellen i forhold til tidligere er at den reviderte modellen bruker grunnkretser som geografisk enhet i stedet for kommuner. Formålet med å dele kommuner inn i grunnkretser er å lage små, stabile geografiske enheter. Norge er per 01.01.2024 delt inn i 14 101 grunnkretser, mens det er 357 kommuner. For å gruppere grunnkretser til enheter med tilstrekkelig antall observasjoner brukes en maskinlæringsalgoritme.

[1] Forprosjekt om boligformuesmodellen
[2] Nye beregningsområder i boligformuesmodellen

2. Datagrunnlaget

Datagrunnlaget for beregningene består av omsatte boliger på markedsplassen Finn.no i perioden januar 2015 – november 2024.

I behandlingen av datamaterialet er det brukt visse kriterier for å skille ut observasjoner som med stor sannsynlighet er feilaktige.

Boligomsetninger med omsetningsverdi på under 350 000 kroner filtreres ut. Det samme gjelder boliger i byer med kvadratmeterpriser under 10 000 kroner. Videre blir det stilt krav til kvadratmeterpriser og areal ved at modellen bare benytter boliger som tilfredsstiller følgende kriterier:

- *Eneboliger*: Boligareal mellom 50 og 550 kvadratmeter og kvadratmeterpris mellom 5 000 og 150 000 kroner.
- *Småhus*: Boligareal mellom 40 og 350 kvadratmeter og kvadratmeterpris mellom 5 000 og 150 000 kroner.
- *Blokkleiligheter*: Boligareal mellom 12 og 350 kvadratmeter og kvadratmeterpris mellom 8 000 og 200 000 kroner.

Vi filtrerer også vekk boliger vi ikke får punktfestet med koordinat.

Etter filtrering på ovennevnte kriterier består datagrunnlaget inn i modellen av 881 472 salg, fordelt på 255 984 eneboliger, 120 738 småhus og 504 750 leiligheter.

3. Om beregningene av boligverdier

I dette kapittelet vises hovedtrekkene ved modellen som beregner boligverdier. Modellen er beskrevet mer i detalj i Notater 2023/3. [3] Boligverdien beregnes med utgangspunkt i informasjon om boligtype, boligens areal, beliggenhet og alder. Datagrunnlaget for beregningene er som nevnt alle boliger omsatt på Finn.no i en bestemt tidsperiode.

3.1. Grunnlaget for beregningene

- Boligene er delt i tre boligtyper: eneboliger, småhus og leiligheter.
- Sammenhengen mellom boligverdi og boligens karakteristikker estimeres med utgangspunkt i en logaritmisk funksjonsform.

Beregningsområdene er basert på grunnkretser som geografiske grunnenheter i stedet for kommuner. Formålet med å dele kommuner inn i grunnkretser er å ha små stabile geografiske enheter. Grunnkretser består av et geografisk sammenhengende område som er mest mulig ensartet når det gjelder natur og næringsgrunnlag, kommunikasjonsforhold og bygningsmessig struktur. Fordelen ved å bruke grunnkretser i forhold til en inndeling som f.eks. kunne ha bygd på rutenett, er at grunnkretser er basert på homogene områder. De følger i stor grad naturlige geografiske grenser som elver, motorveier, endringer i boligstruktur o.l.

Grunnkretser er små enheter (det er 14 101 grunnkretser i Norge per 01.01.2024). En enkelt grunnkrets mangler derfor ofte et tilstrekkelig antall boligtransaksjoner for å kunne estimere en regresjonsmodell, selv en med få forklaringsvariabler. Dette gjelder særlig mange landlige grunnkretser som har få observasjoner av leiligheter og småhus. Slike grunnkretser må derfor slås sammen til større enheter. Slike grupper av sammenslåtte grunnkretser er beregningsområdene i modellen. Vi har definert at et beregningsområde minimum må innbefatte 500 omsetninger.

Vi bruker en enkel maskinlæringsalgoritme kalt beslutningstre (Decision tree) for å sette sammen grunnkretser som ligner på hverandre. Algoritmen setter sammen grunnkretser i beregningsområder slik at det kvadrerte avviket mellom kvadratmeterpriser for enkeltobservasjoner og gjennomsnittsprisen innad i beregningsområdene blir minimert. Det vil si at likhet mellom grunnkretser måles ved gjennomsnittlig kvadratmeterpris for gitt boligtype.

Forklaringsvariablene er geografisk plassering (lengdegrad, breddegrad, gjennomsnitt av lengdegrad og breddegrad for omsetninger innad i grunnkretsen). En algoritme deler opp observasjonene i datasettet i flere trinn, basert på forklaringsvariablene for å skape beregningsområder som best mulig predikerer kvadratmeterprisene. Oppdelingen fortsetter til videre oppdeling vil føre til beregningsområder med færre observasjoner enn det fastsatte minimum (500).

Grupperingen gjøres separat for forskjellige boligtyper (eneboliger, småhus og leiligheter). Til sammen ble det i kjøringen opprettet 1 272 ulike beregningsområder, svært varierende i størrelse. Ikke overraskende er det færrest beregningsområder for småhus, siden antall omsetninger i denne gruppen er minst.

[3] Nye beregningsområder i boligformuesmodellen

Tabell 3.1. Antall beregningsområder fordelt på boligtyper.

Boligtype	Antall beregningsområder
Enebolig	395
Småhus	178
Leilighet/blokk	699
Totalt	1272

Kilde: Statistisk sentralbyrå

Det største beregningsområdet totalt gjelder boligtypen småhus og inkluderer totalt 644 grunnkretser. Til sammenligning er det 62 beregningsområder for leiligheter som kun inkluderer 1 grunnkrets. Disse beregningsområdene finnes bare i de store byene, Trondheim (2), Tromsø (1), Stavanger (1), Bergen (1) og Oslo (57).

For hvert enkelt beregningsområde gjøres så en hedonisk regresjon. Det vil si at man har en separat regresjon for hvert enkelt beregningsområde for hver boligtype. Å ha mange separate regresjoner gir større fleksibilitet fordi alle parameterverdiene i modellen kan være forskjellige for hver kombinasjon av boligtype og beregningsområde.

Den grunnleggende avveiningen ved aggregering av grunnkretser er at man ønsker å skape områder som er så homogene som mulig (noe som tilsier små enheter) samtidig som man er avhengig av et visst antall omsetninger i den hedoniske regresjonen (noe som tilsier større enheter). Vi krever at det er minimum 500 omsetninger i en gruppe av grunnkretser, for å få et tilfredsstillende antall observasjoner for den hedoniske regresjonen. Det resulterer i mange, geografisk små beregningsområder i sentrale strøk, og geografisk større beregningsområder i mindre sentrale strøk. Slik utnytter man den store mengden informasjon (antall transaksjoner) i byer til å bedre predikere lokale prisnivåer.

3.2. Regresjonsmodellen

Totalt sett blir det 1272 regresjoner for hele landet. For hvert beregningsområde estimeres følgende funksjon:

(1) $\ln(kvadratmeterpris) = b_0 * konst + b_1 * \ln(p_rom) + b_2 * alder_g + b_3 * grunnkrets_i + b_4 * år_t + \varepsilon$

(1) viser at logaritmen til kvadratmeterprisen er en funksjon av logaritmen til arealet målt ved p-rom (ln (p-rom)), aldersgruppe ($alder_g$) og hvilken grunnkrets boligen ligger i ($grunnkrets_i$). Årsdummyene inngår for å ta høyde for virkninger av prisutvikling over tid. Prediksjonene beregnes for år 2024 som er referansekategorien for årsdummyene. Konstantleddet viser skjæringspunktet med x-aksen. b_i-ene er parameterverdier (koeffisienter). ε er et residualledd.

Ved hjelp av de estimerte parameterverdiene fra regresjonsanalysene for kvadratmeterpriser og omformuleringer av (1) er det mulig å beregne verdien på den enkelte bolig. Boligverdien (B) kan regnes ut ved hjelp av (2):

$$(2) B = \exp(\text{konstantledd} + (\text{log_areal} * \ln(p_rom)) + \text{alder} + \text{grunnkrets_koef}) * p_rom * \text{korreksjonsfaktor}$$

I formel (2) representerer symbolene som følger:

- B er boligverdien som beregnes
- konstantledd, et fast tall i hvert beregningsområde
- log_areal er resultatet av b_1*ln(areal). Denne viser betydningen av arealet på boligen.

- P_rom er boligens areal
- Alder er resultatet av b_2*$alder_g$. Denne viser betydningen av hvilken aldersgruppe boligen tilhører
- grunnkrets_koef er resultatet av b_3*$grunnkrets_i$. Denne viser betydningen av grunnkretsen boligen ligger i
- korreksjonsfaktor er et fast tall i hvert beregningsområde

Korreksjonsfaktoren er framkommet ved hjelp av residualene fra regresjonsanalysen, dvs. formelen MSE/2. (MSE= Mean Squared error). Korreksjonsfaktoren korrigerer for skjevhet når man transformerer logaritmiske data tilbake til den opprinnelige skalaen (Miller, 1984).

3.3. Eksempel på beregning av boligverdi

Følgende eksempel viser hvordan boligverdien B beregnes for en enebolig i Østfold, der arealet (P-rom) er lik 175 kvm, boligen ligger i grunnkrets 31200101 (Skantebygda) i Rakkestad kommune og ble bygd i 1999. Boligen tilhører aldersgruppe 3 (20-34 år).

For eneboliger i denne grunnkretsen i Østfold med disse egenskapene finner vi følgende parameterverdier (koeffisienter) i Statistikkbanken (tabell 14737):

konstantledd = 12,711569, Log(areal) =-0,456693, alderskoeffisient = -0,222063, grunnkretskoeffisient = -0,001856 og korreksjonsfaktor = 1,021809

Boligverdien vil da kunne regnes ut slik:

$$\begin{aligned} B = \exp((12{,}711569) + (-0{,}456693 \cdot \ln(175)) \\ + (-0{,}222063) + (-0{,}001856)) \\ \cdot 175 \\ \cdot 1{,}021809 \end{aligned}$$

$$\begin{aligned} B = \exp(10{,}12893) \\ \cdot 175 \\ \cdot 1{,}021809 \end{aligned}$$

$$\exp(10{,}12893) = 25\,057{,}5385$$

Dette tallet multipliseres så med arealet og en korreksjonsfaktor slik at boligverdien for en enebolig på 175 kvadratmeter i den aktuelle grunnkretsen i Rakkestad blir:

$$25\,057{,}5385 \cdot 175 \cdot 1{,}021809 = 4\,480\,703$$

Boligverdiene som disse beregningene gir, kan tolkes som den gjennomsnittlige markedsverdien til alle boliger innenfor en bestemt boligtype, med ett bestemt boligareal, en bestemt beliggenhet og som har en bestemt alder.

4. Modellytelse

Et vanlig mål på ytelsen i prediksjoner er det såkalte MAPE. MAPE står for Mean Absolute Percentage Error (på norsk *gjennomsnittlig absolutt prosentvis feil*). Den angir den gjennomsnittlige prosentvise forskjellen mellom den verdien modellen predikerer og faktisk oppnådd pris. Det er et mål på treffsikkerheten til en modell; jo lavere verdi på MAPE, desto mer presise er prediksjonene. Alle tall er basert på kryssvaliderte data. Kryssvalidering vil si at datasettet deles opp i flere mindre deler (eller grupper), her ti. Modellen trenes på ni av delene og testes på den gjenværende tiendedelen. Prosessen gjentas slik at alle delene får fungere som testsett én gang. Resultatene slås sammen for å gi et mer nøyaktig bilde av modellens kvalitet. Da unngår man å måle eventuell overtilpasning i modellen.

Tabell 4.1 viser at MAPE er 12,53 prosent for alle boliger. Vi ser at prediksjonsevnen er dårligst for eneboliger med en MAPE på 18,04, mens den er på 10,18 for blokkleiligheter og 10,69 for småhus. En MAPE på 12,53 betyr at estimatene i gjennomsnitt har et avvik på 12,53 prosent fra de faktiske verdiene. Enkeltobservasjoner med høyt avvik kan trekke opp gjennomsnittet. Derfor er det nyttig også å se på medianfeilen. Mdape er et mål som viser medianverdien av avviket. Vi ser at median avvik er på 8,7 prosent. Det er på 12,61 for eneboliger, 7,81 forsmåhus, og 7,57 for leiligheter. PM 50 er et mål som viser hvor mange boliger sim modellen treffer innenfor et intervall på +- 50 prosent. Om lag 98 prosent av boligene faller innenfor dette intervallet. PM 20 er et mål som viser hvor mange boliger modellen treffer innenfor et intervall på +- 20 prosent. Om lag 83 prosent av boligene faller innenfor dette intervallet. I den opprinnelige modellen var P20 langt lavere, om lag 73 prosent (Medby 2025, 2026). P20 er 88,71 for blokkleiligheter, 87,51 for småhus og 69,86 for eneboliger. Modellen treffer dårligst for eneboliger og best for blokkleiligheter. At prediksjonene er best for leiligheter og dårligst for eneboliger, er et trekk vi kjenner igjen også fra den opprinnelige modellen. Dette trolig fordi det er større heterogenitet blant eneboliger.

PM10 viser hvor mange boliger modellen treffer innenfor et intervall på +- 10 prosent. Om lag 56 prosent av boligene faller innenfor dette intervallet.

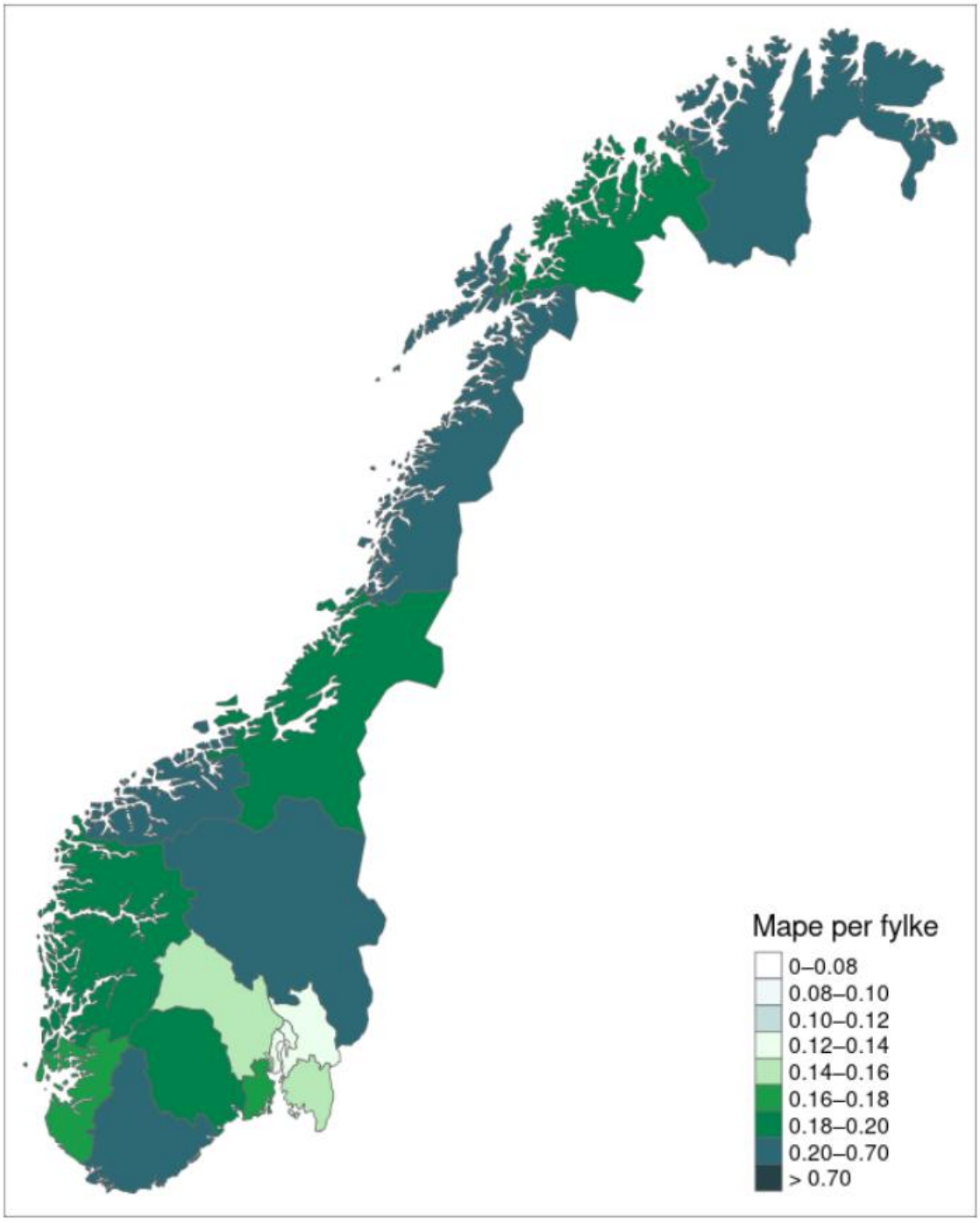
Tabell 4.1. Mål på forholdet mellom den estimerte markedsprisen over den faktisk observerte omsetningsprisen. 2015-2024. Alle boligtyper samlet. Prosent.

	Alle boliger	Eneboliger	Småhus	Leiligheter
Mape	12,53	18,04	10,69	10,18
Mdape	8,70	12,61	7,81	7,57
PM10	55,89	41,36	60,59	62,15
PM20	83,07	69,86	87,51	88,71
PM50	98,05	95,22	98,92	99,30

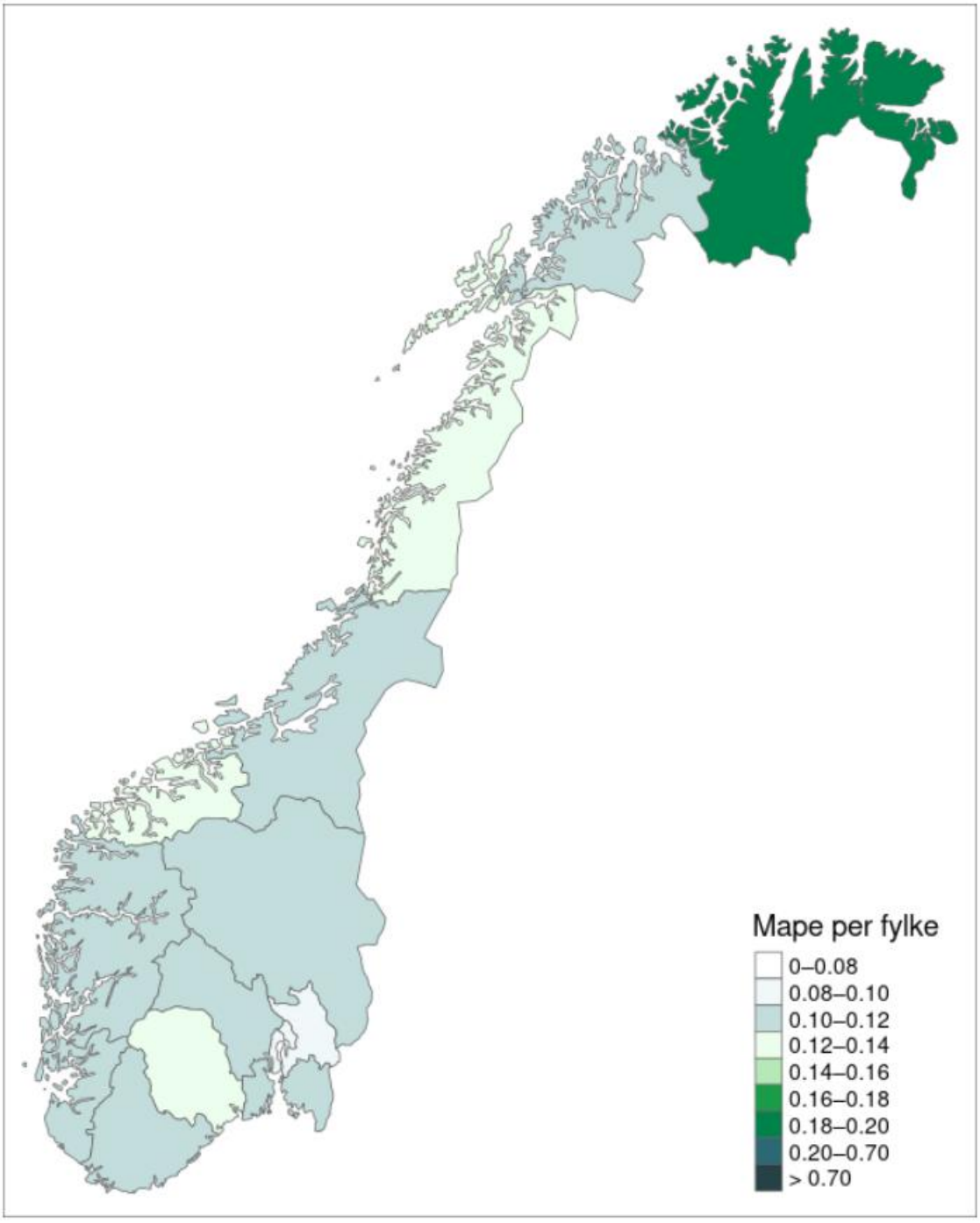
Kilde: Statistisk sentralbyrå

Når det gjelder den geografiske treffsikkerheten, viser den seg ikke overraskende å være høyest i sentrale strøk. Figurene 4.1-4.3 viser MAPE per boligtype på fylkesnivå. Figuren illustrerer at MAPE er lavest for eneboliger. Vi ser at Oslo-området har lavest MAPE (best treffsikkerhet) for alle boligtyper. Også de andre storbyområdene merker seg ut ved lavere MAPE enn resten av landet. For eneboliger kommer Finnmark, Nordland, Møre og Romsdal, Innlandet og Agder dårligst ut. For småhus har Finnmark lavest MAPE, mens det for blokkleilighetenes del er Møre og Romsdal som kommer dårligst ut.

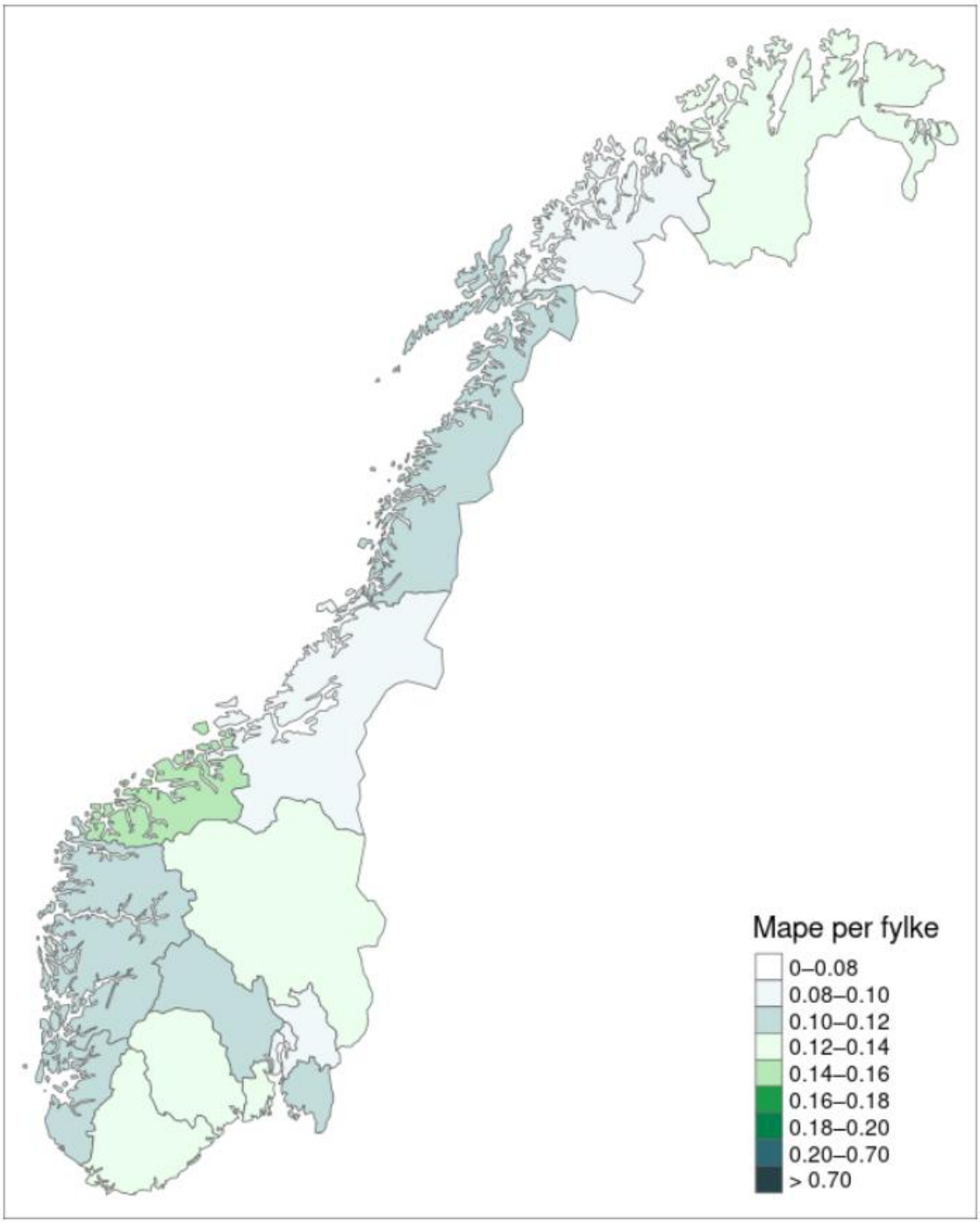
Figur 4.1. Prosentvis fordeling av forholdet mellom den estimerte markedsprisen over den faktisk observerte omsetningsprisen. 2015-2024.Eneboliger per fylke.



Figur 4.2. Prosentvis fordeling av forholdet mellom den estimerte markedsprisen over den faktisk observerte omsetningsprisen. 2015-2024. Småhus per fylke.



Figur 4.3. Prosentvis fordeling av forholdet mellom den estimerte markedsprisen over den faktisk observerte omsetningsprisen. 2015-2024. Blokkleiligheter per fylke.



5. Resultater

Resultatene fra samtlige regresjonsanalyser ligger tilgjengelig i Statistikkbanken, tabell 14737: https://www.ssb.no/statbank/table/14737. På grunn av desimalavrunding kan koeffisientene fra regresjonsmodellene som ligger i Statistikkbanken, ha et marginalt avvik i forhold til de eksakt beregnede koeffisientene som blant annet Skatteetaten har fått tilsendt.

Se side 10 for et illustrerende eksempel på hvordan koeffisientene fra regresjonene kan benyttes til å beregne markedsverdien for en konkret bolig.

Referanser

Bø, Erlend Eide, Per Medby, Odd Erik Nygård og Mona Takle (2022): «Forprosjekt om boligformuesmodellen», SSB Notater 2022/11, Statistisk sentralbyrå.

Bø, Erlend E., Per Medby, Odd Erik Nygård og Mona Takle (2022): «Hedonisk verdsettelse av bolig med fleksibel geografisk inndeling», Tidsskrift for boligforskning, 5 (2): 70-87

Bø, Erlend Eide og Odd Erik Nygård (2023): «Nye begningsområder i boligformuesmodellen», SSB Notater 2023/3, Statistisk sentralbyrå.

Medby, Per (2025): «Modell for beregning av boligformue. Oppdatert med tall for 2024». SSB Notater 2025/10, Statistisk sentralbyrå.

Medby, Per (2026): «Modell for beregning av boligformue. Oppdatert med tall for 2025». SSB Notater 2026/10, Statistisk sentralbyrå.

Miller, Don**:** (1984). "Reducing Transformation Bias in Curve Fitting". *The American Statistician*: Vol. 38, No. 2, pp. 124-126.

Thomassen, Arild og Melby, Ingrid (2009): «Beregning av boligformue», SSB Notater 2009/53, Statistisk sentralbyrå.

Vedlegg A: Eksempel på beregningsområde

Figur A1 viser et eksempel på inndeling av beregningsområder for eneboliger i Hamar kommune. Det er til sammen 7 ulike beregningsområder i kommunen for denne boligtypen. Merk at beregningsområdene også går på tvers av kommunegrensene.

Figur A.1. Beregningsområder eneboliger i Hamar kommune.

