

Hedonisk prisindeks med rolling time dummy-metoden

Metodedokument tilknyttet statistikdokumentationen [Andelsboligsalg](#).

Den hedoniske metode bruges til at korrigere for forskelle i kvalitet mellem de observerede handler. Rolling time dummy-metoden er en variant af den klassiske multiple tidsdummy-metode, men med et løbende (rullende) datavindue.

Trin-for-trin beregning:

1. Valg af datasæt

- Man samler transaktionsdata for de seneste fire kvartaler.
- Hver observation indeholder pris og en række karakteristika (størrelse, beliggenhed, byggeår osv.).

2. Estimering af hedonisk model

- En regressionsmodel estimeres for hver rullende periode, her en log-lineær model

$$\ln(P_i) = \beta_0 + \sum_k \beta_k X_{ik} + \sum_t \delta_t D_{it} + \epsilon_i \quad (1)$$

hvor:

- P_i = pris for andelsbolig i
- X_{ik} = andelsboligens karakteristika (fx størrelse, beliggenhed, byggeår)
- D_{it} = tidsdummyer, der måler den generelle prisudvikling
- δ_t = estimerer prisudviklingen over tid

3. Rolling window: Glidende datavindue

- I stedet for at bruge hele datasættet, opdateres modellen løbende med de fire seneste kvartaler.
- Når et nyt kvartal tilføjes, fjernes det ældste kvartal, så vinduet ruller fremad.

4. Prisudviklingen mellem de to seneste kvartaler

- Prisudviklingen mellem de to seneste kvartaler beregnes ved forskellen mellem de tilhørende tidsdummy-koefficienter. Ved at eksponentiere differencen opnås den relative prisændring:

$$\Delta I_{t-1:t} = e^{(\delta_t - \delta_{t-1})} \quad (2)$$

- Kun forskellen mellem de to nyeste dummy-koefficienter indgår i det officielle indeks, og tidligere perioder revideres ikke. Tidligere perioder opdateres kun, hvis der er modtaget yderligere indberetninger. For at opnå det aktuelle prisindeksniveau for periode t multipliceres denne ændring med prisindekset for forrige kvartal.

Fordele ved rolling time dummy-metoden:

- **Løbende opdatering** – modellen tager altid højde for de nyeste handler.
- **Undgår revisioner** – tidligere perioder ændres ikke, hvilket giver stabilitet.
- **Fanger markedsændringer hurtigere** – da modellen kun ser på de seneste data.

Metoden er altså en god balance mellem præcision og stabilitet i et dynamisk andelsboligmarked.

Prisindeks for andelsboliger består samlet set af følgende tre prisindeks:

- Hele landet
- Landsdel København By
- Resten af landet (hele landet uden landsdel København By) – offentliggøres ikke pga. volatilitet.

Det bemærkes, at prisindekset for "resten af landet" ikke offentliggøres særskilt på grund af stor volatilitet. Det indgår i stedet som en intern komponent i beregningen af det samlede landsdækkende prisindeks, hvor det anvendes til at sammenveje de to geografiske delindeks.

Brug af én samlet regressionsmodel:

For at få de mest præcise prisindeks er det valgt, at der laves én samlet regressionsmodel for de to regionale prisindeks vedr. alle værdiansættelsesmetoder ved at inkludere interaktionsled mellem tid og sted for både landsdel København By og resten af landet, se ligning (3):

$$\ln(P_i) = \beta_0 + \sum_k \beta_k X_{ik} + \sum_t \delta_t D_{it} + \gamma_H R_{iH} + \sum_t \theta_t^H (D_{it} \cdot R_{iH}) + \epsilon_i \quad (3)$$

hvor:

- P_i = pris for andelsbolig i
- X_{ik} = andelsboligens karakteristika (fx størrelse, beliggenhed, byggeår)
- D_{it} = tidsdummyer, der måler den generelle prisudvikling
- δ_t = estimerer prisudviklingen over tid
- R_{iH} = dummy for landsdel København By (1 hvis i København, 0 ellers)
- γ_H = niveauforskel mellem København og resten af landet
- θ_t^H = interaktionsled, der måler afvigelsen i prisudviklingen for København

Modelselektion: Hybridtilgang

For at finde den rette balance mellem modelkompleksitet og forklaringskraft, er følgende fremgangsmåde valgt:

1. Tidsdummyer og interaktionsled for de to seneste kvartaler tvinges med i modellen – så prisudviklingen altid estimeres.
2. Kontrolvariable vælges via modelselektion (SBC) – så kun de relevante for perioden bliver inkluderet (datadrevet model).
3. Tidsdummy-koefficienter og interaktionsled er udeladt for de to første kvartaler i estimeringsvinduet for at reducere antallet af forklarende variable.

Beregning af prisindeks med denne model:

Man kan nu beregne tre forskellige prisindeks vha. tidsdummy-variable:

1. Prisindeks for resten af landet (referencekategori)

$$I_t^{\text{resten af landet}} = e^{\delta_t} \cdot 100 \quad (4)$$

Her repræsenterer δ_t den gennemsnitlige prisudvikling for resten af landet.

2. Prisindeks for landsdel Byen København

$$I_t^{\text{Hovedstaden}} = e^{\delta_t + \theta_t^H} \cdot 100 \quad (5)$$

Her repræsenterer $\delta_t + \theta_t^H$ den gennemsnitlige prisudvikling for Landsdel Byen København.

3. Nationalt indeks (vægtet gennemsnit)

$$I_t^{\text{land}} = W_h \cdot I_t^{\text{Hovedstaden}} + W_r \cdot I_t^{\text{Resten af landet}} \quad (6)$$

, hvor W_h og W_r er andelen er vægtene baseret på værdien af periodens solgte boliger. I praksis betyder det, at vægtene prisopdateres i forhold til 4. kvartal året før og den aktuelle periode.

Bemærk: Det nuværende ejerboligindeks anvender vægte for landsdele baseret på ejendomsværdien af den samlede bestand af ejerboliger. Da vi ikke har oplysninger om værdien af den samlede bestand af andelsboliger, anvender vi i stedet ejendomsværdien af de omsatte andelsboliger som grundlag for vægtningen. Det svarer til ejerboligindekset inden for landsdele, hvor vægtene er baseret på ejendomsværdien af de solgte boliger.