

Byggnadsprisindex, BPI

Byggnadsprisindex, BPI, är en outputprisindex som ska mäta den genomsnittliga prisutvecklingen för nyproducerade bostäder.

BPI beräknas som en kedjeindex med årliga länkar. Indexlänken är kvoten mellan en värdekomponent och en kvalitetskomponent. Värdekomponenten är ett enhetsvärdeindex av pris per kvadratmeter. Kvalitetskomponenten skattas med hjälp av en hedonisk regressionsfunktion där projektets läge utgör en viktig kvalitetsvariabel. Försök att använda mer detaljerad lägesbeskrivning har dock inte givit önskat resultat.

BPI avser flerbostadshus och gruppbyggda småhus. Dataunderlaget är litet vilket medför att kvalitetskomponenten skattas osäkert. En fråga är om det är bra att skatta den hedoniska funktionen med en mängd data från flera år.

I föreliggande PM redovisas analyser av data som hittills har använts vid beräkningen av BPI. Ett antal förslag som förbättrar precisionen i statistiken redovisas. Det finns fler variabler om projekten inom SCB som inte har analyserats här.

SCB ser gärna att Nämnden för Konsumentprisindex diskuterar frågor och förslag i denna PM.

Innehåll

Byggnadsprisindex, BPI	1
Syfte.....	2
Statistikens storheter.....	2
Utfall.....	10
Analys	11
Resultat.....	21
Pågående arbete.....	26
Slutsatser och förslag.....	27
Referenser.....	29
Appendix	30

Bakgrund

Byggnadsprisindex ska mäta den genomsnittliga prisutvecklingen för nyproducerade bostäder.

Kritik har framförts i ett antal utredningar. Borg et al. (2013) och Lind och Song (2011) kritiserar BPI för att inte tillräckligt relevanta variabler ingår i den hedoniska funktionen. Detta resulterar i att kvalitetsförbättringar i byggsektorn – avspeglade i färdigställda gruppbyggda småhus respektive flerbostadshus – inte fångas i modellen. Detta leder till att BPI blir missvisande eftersom prisutvecklingen överskattas – om kvalitetsutvecklingen är positiv. En konsekvens blir då att produktivitetsökningen i byggsektorn inte kan uppskattas korrekt, utan visas som lägre än vad den egentligen är.

Det är värt att notera att kritiken inte har förekommit enbart avseende den svenska (SCB) BPI, utan liknande kritik har framförts vad gäller motsvarande index i Kanada (Harrison, 2007), Storbritannien (Yu och Ive, 2008), USA (Chapman et al, 2009), och ett antal andra länder. De problem som det pekas på avseende BPI är alltså ett allmänt metodproblem.

I appendix 3 och 4 förtecknas ett antal citat ur dessa artiklar.

Ett antal utredningar har under åren genomförts av SCB för att försöka förbättra kvaliteten, bl a. Walestad (2007).

Syfte

Syftet med föreliggande PM är att beskriva hur BPI beräknas, nämna något om framförd kritik, redovisa skattningar av osäkerheter samt presentera förbättringsförslag.

Statistikens storheter

Denna text är en redigerad och förkortad version av kvalitetsdeklarationen som finns på scb.se.

Statistikens sammanhang

Uppgifter från kommuner och byggherrar samlas in och samutnyttjas till statistikprodukterna:

- Bygglövsstatistik för bostäder och lokaler www.scb.se/BO0701
- Nybyggnad av bostäder www.scb.se/BO0101
- Priser för nyproducerade bostäder www.scb.se/BO0201
- Byggnadsprisindex www.scb.se/PR0501.

Statistikens ändamål

Byggnadsprisindex ska mäta den genomsnittliga prisutvecklingen för nyproducerade bostäder, exklusive köp av tomtmark, rensad från kvalitetsutveckling.

På SCB produceras ett antal index relaterade till olika delar av bostadsbyggande, uppdelat på input och output. Byggnadsprisindex är ett så kallat outputindex, i den meningen att det ska spegla prisutvecklingen för den färdiga produkten, ett byggprojekt.

Statistikanvändares informationsbehov

Statistikens användare är främst:

- Departement för analys och uppföljning
- Konjunkturinstitutet för beräkningar av investeringar i bostäder samt för investeringsprognoser
- Boverket för uppföljning av kostnadsutvecklingen och för utvärdering av den förda bostadspolitiken samt för prognosarbete
- Statliga utredningar för studier av bostadsproduktionens förändring och utveckling på längre sikt
- Nationalräkenskaperna för beräkningar av investeringar i bostäder
- Byggmaterialindustrin för produktionsplanering och marknadsstudier
- Byggherrar och konsulter för att bedöma kostnader vid planering av nya bostadsobjekt.
- Forskning
- Allmän information
- Användarrådet för statistik över mark och bebyggelse som har inflytande över statistiken. Där ingår representanter från Naturvårdsverket, Socialdepartementet, Boverket, Sveriges Byggindustrier, Fastighetsägarna Sverige, Lantmäteriverket, Länsstyrelsen i Stockholms län, Sveriges Kommuner och Regioner, Kungliga Tekniska Högskolan, Hyresgästföreningen Riksförbundet, Sveriges Allmännyttiga Bostadsföretag och Sveriges lantbruksuniversitet.

Objekt och population

Objekt

Målobjekten är påbörjade nybyggnadsprojekt avseende ordinära flerbostadshus samt gruppbyggda småhus. Observationsobjekten är nybyggnadsprojekt som klassificeras som flerbostadshus eller gruppbyggt småhus via variabeln hustyp i frågeformuläret. Ett observationsobjekt kan bestå av ett eller flera flerbostadshus, gruppbyggda småhus eller en kombination av ett eller flera flerbostadshus och gruppbyggda småhus.

Population

Undersökningens intressepopulation är nybyggnadsprojekt påbörjade under den aktuella referensperioden.

Målpopulationen är nybyggnadsprojekt avseende ordinära flerbostadshus och gruppbyggda småhus som ska försäljas eller upplåtas med hyres- bostads- eller äganderätt. Avgränsningen till påbörjade projekt görs för statistiken ska vara aktuell. Med påbörjade projekt avses påbörjandet av de egentliga byggnadsarbetena (gjutning av källargolv, källarmurar, bottenplatta och dylikt).

Rampopulationen innehåller information om bygglov och kontaktuppgifter till byggherrar. Ett godkänt bygglov kan av olika skäl få ett fördröjt påbörjande, dessa hanteras så att de får ett framtida datum för påbörjande och återkontaktas vid framtida undersökningsomgångar.

I populationen ingår inte:

- Privatperson som bygger småhus
- Företag som säljer tomt med bygglov till en privatperson som anlitar någon annan att bygga själva huset (småhus)
- Projekt där den uthyrningsbara lokalarean (LOA-h) är större än en tredjedel av summan av bostadsarea (BOA) och uthyrningsbar lokalarea (LOA-h)
- Projekt där antalet specialbostäder (bostäder för särskilda boendegrupper som exempelvis specialbostad för äldre eller funktionshindrade samt studentbostad) utgör mer än 10 procent av det totala antalet bostäder
- Brandskadeärenden
- Påbyggnader
- Blandprojekt som t.ex. ombyggnad och nybyggnad där man inte kan särskilja kostnaderna för nybyggnad
- Projekt med flerbostadshus där innehållet är färre än tre bostäder
- Projekt som inte är ”ren” nybyggnad t.ex. förflyttning av hus.

Variabler

Intressevariabeln avseende totalt produktionspris ”borde vara” ett faktiskt slutpris. Eftersom byggnationerna precis har påbörjats när enkäten besvaras så finns inte den uppgiften tillgänglig, endast ett uppskattat slutpris finns tillgängligt. Detta gäller även för byggnadspris, markpris och moms. Därför definieras målvariablerna som dessa uppskattade priser.

Vidare insamlas uppgifter om antal lägenheter som totalt byggs i projektet, hustyp (flerbostadshus eller gruppbyggt småhus), lägenhetsarea/ bostadsarea samt moms för byggnad och bidrag. Momsen avser hela projektet men kan komma från både mark och byggnad. Moms för mark är ofta liten och är

förknippad med markarbeten. I beräkningar läggs hela momsen i priset för själva byggnaden.

Brutto- och nettopris skapas utifrån variabeln bidrag:

- Brutto avser projektets byggnadspris inklusive mervärdesskatt utan avdrag för bidrag.
- Netto avser projektets byggnadspris inklusive mervärdesskatt med avdrag för bidrag. Om inget bidrag erhållits för projektet är netto- och bruttopriset lika.

När ett nybyggnadsprojekt består av både gruppbyggda småhus och flerbostadshus används information om antal lägenheter för respektive hustyp för att klassificera objektet enligt mest-principen som antingen ett flerbostadshus eller gruppbyggt småhus.

För att beräkna kvalitetskomponenten i index för anbudspriser används flera variabler som reflekterar ”kvalitet”.

För flerbostadshus vill man förklara den logaritmerade kostnaden per kvadratmeter lägenhetsarea. För att kunna göra det används nedanstående variabler i modellen:

- Genomsnittlig köpeskilling per kvadratmeter på försålda småhus, per kommun
- Upplåtelseform (hyresrätt, bostadsrätt, äganderätt)
- Yteffektivitet = $(\text{Bostadsarea} + \text{uthyrningsbara lokalarean}) / \text{Bruttoarea}$
- Antal hygienrum per lägenhet
- Husform (lamellhus, loftgångshus, punkthus, övriga)
- Källare finns/finns ej
- Antal balkonger per lägenhet i genomsnitt
- Hiss finns/finns ej
- Antal garage eller parkeringsdäck per lägenhet i genomsnitt

För gruppbyggda småhus vill man förklara den logaritmerade kostnaden per kvadratmeter med följande variabler i modellen:

- Genomsnittlig köpeskilling per kvadratmeter på försålda småhus, per kommun
- Bostadsarea = $\text{Summa bostadsarea} / \text{summa antal lägenheter}$
- Antal hygienrum per lägenhet
- Husform (friliggande villa, kedjehus, radhus+övriga)
- Källare finns/finns ej
- Fasadbeklädnad (sten, puts+plåt+övrigt)
- Takmaterial (lertegel, plåt+papp+annat)
- Inredningsbar vind finns / finns ej
- Annat än en-planshus
- Upplåtelseform (hyresrätt, bostadsrätt, äganderätt)
- Antal garage eller parkeringsdäck i genomsnitt per lägenhet

Genomförande

BPI är en totalundersökning som samlar in information via elektroniska frågeformulär ställda till byggherrar. Uppgifter samlas in varje månad som en gemensam del av undersökningen Priser för nyproducerade bostäder (BO0201).

För att slutpriset ska reflektera värdet under referensåret görs en indexomräkning av de insamlade priserna där de skrivs ned med faktorprisindex för den tid byggnadsarbetet beräknas ta. Faktorprisindex är ett inputprisindex som beskriver prisutvecklingen för de enskilda insatsfaktorerna i byggprocessen. För en mer utförlig beskrivning kring detta se dokumentationen för Priser för nyproducerade bostäder på www.scb.se/BO0201.

Publiceringen av BPI innefattar fyra olika index: anbudspriser, pågåendepriser, funktionsprisindex samt kvalitetsindex beräknade för flerbostadshus och gruppsybyggda småhus.

Funktionsprisindex beskriver prisutvecklingen per bostadslägenhet. Det blir därför beroende av skillnader i dels byggtakten och dels hur priserna över tid för att få lägenheterna uppförda fluktuerar. Ingen kvalitetsjustering görs i detta index.

Index för **anbudspriser** ger en uppskattning av rena prisförändringar för nyproducerade bostäder. Hänsyn tas till förändrad kvalitet som skattas med hjälp av en hedonisk regressionsmodell och räknas bort från de observerade prisförändringarna.

Anbudsprisindex = Funktionsprisindex / Kvalitetsindex.

Kvalitetsindex är ett index som beskriver utveckling av kvalitet menat som kombination av dels den komponent som används för att justera anbudspriserna för kvalitetsförändringar i nyproduktion dels en komponent som beror på genomsnittlig ytförändring per lägenhet.

Index för **pågående priser** räknas fram per kvartal och år och ger en ungefärlig bild av kostnadsutvecklingen för nyproduktion som pågår under aktuellt kvartal (år). Beräkningen bygger på uppgifter om anbudspriser och antal pågående lägenheter från nuvarande och tidigare kvartal. Justeringar för kostnadsökningar för indexreglerade avtal görs med faktorprisindex för entreprenörens kostnader. Årsindex beräknas som ett medelvärde över kvartalsberäkningar.

De statistiska måtten är kedjeindex med årslänkar. En årslänk beskriver förändringen i genomsnittspriser mellan referensåret och föregående år. För anbudspriser och funktionsprisindex är måttet i form av kvoter av genomsnitt (pris per kvadratmeter, pris per lägenhet) för referensår jämfört med föregående år med respektive utan kvalitetsjustering.

Kedjeindex med årliga länkar

Detta avsnitt har i stor utsträckning "lånats" från Ribe (2014).

BPI beräknas som ett kedjeindex med årliga länkar, som kedjas till en serie indextal med indexbasår 1968=100.

Här beskrivs principerna för beräkningen av de årliga indexlänkarna P_t som avser prisutvecklingen från år $t - 1$ till år t , med tonvikt på kvalitetsjustering genom hedonisk regression enligt metoden adjustment index, kallad AD av Dalén (1992).

Metoden i den form som beskrivs här har tillämpats sedan redovisningsåret 2006. Dessförinnan tillämpades liknande men inte fullt jämförbara metoder.

Indexlänken P_t för prisutvecklingen från år $t - 1$ till år t beräknas som kvoten mellan två indextal V_t och K_t enligt:

$$(1) \quad P_t = \frac{V_t}{K_t}$$

där

t = Aktuellt år (redovisningsåret)

P_t = Prisindexlänken

V_t = Värdekomponenten

K_t = Kvalitetskomponenten

Indextalen V_t och K_t som används i beräkningen enligt formel (1) beräknas i sin tur enligt följande.

Värdekomponenten V_t är ett enhetsvärdeindex för priset per kvadratmeter beräknat enligt:

$$(2) \quad V_t = \frac{\bar{y}_t}{\bar{y}_{t-1}}$$

där

t = Aktuellt år

$t-1$ = Året före aktuellt år

\bar{y}_t = Medelpriset per kvadratmeter i projekten aktuellt år

\bar{y}_{t-1} = Medelpriset per kvadratmeter i projekten året före aktuellt år

Form av medelvärde. Överstreckssymbolen i uttrycken \bar{y}_t och \bar{y}_{t-1} anger beräkning av medelvärde per kvadratmeter av respektive pris. För småhus avses därvid kvadratmeter bostadsarea, och för flerbostadshus avses lägenhetsarea. Medelpriserna per kvadratmeter beräknas genom att motsvarande medelpriser per kvadratmeter inom byggnadsprojekten vägs ihop med projektens motsvarande areor.

Kvalitetskomponenten K_t har för sin beräkning utgångspunkten i en regressionsekvation (modell) av formen:

$$(3) \quad \ln Y = a + b_1x_1 + b_2x_2 + \dots + b_kx_k + \varepsilon$$

där

Y = Priset för ett projekt *per kvadratmeter*, den beroende variabeln

a = Intercept

b_i = Regressionskoefficienter

x_i = Kvalitetsvariabler

ε = Residual (restterm)

Semilogaritmisk hedonisk regressionsekvation är numera allmänt en vanlig praxis, av goda skäl (jfr Destatis, 2009; Triplett, 2006).

Regressionskoefficienterna b_i har sin tolkning i att de i exponentierad form e^{b_i} anger proportioner med vilka priset ändras per enhets ökning i kvalitetsvariabeln.

Regressionskoefficienterna skattas genom att regressionsekvationen (3) anpassas till observerade data för år $t - 1$. Utifrån de skattade regressionskoefficienterna beräknas kvalitetskomponentens indextal fram enligt:

$$(4) \quad K_t = \frac{\exp(b_{1,t-1}\bar{x}_{1,t} + b_{2,t-1}\bar{x}_{2,t} + \dots + b_{k,t-1}\bar{x}_{k,t})}{\exp(b_{1,t-1}\bar{x}_{1,t-1} + b_{2,t-1}\bar{x}_{2,t-1} + \dots + b_{k,t-1}\bar{x}_{k,t-1})}$$

Överstreckssymbolen i uttrycken av formen $\bar{x}_{i,s}$ anger medelvärde per kvadratmeter bostads- respektive lägenhetsarea, analogt med behandlingen av prisvariabeln i formel (2) enligt förklaringarna till denna. (Interceptet a_{t-1} behövs inte i formel (4), eftersom faktorn $e^{a_{t-1}}$ kan brytas ut ur täljare och nämnare och förkortas bort.)

Vid beräkningen av K_t används alltid regressionskoefficienterna för perioden före det aktuella året. Detta innebär att regressionskoefficienterna byts ut en gång per år. Ovanstående beskrivning motsvarar den av Walestad (2007).

Kommentar om principer bakom variabelval

Förklarande variabler ska fånga kvalitet

De förklarande variablerna ska beskriva byggnadernas kvalitet. De ska vara valda med den utgångspunkten, inte för att kunna förklara prisvariationerna optimalt. Egenskaper som beror av vad som för tillfället är populärt ska inte vara med.

Kvalitet i byggherre- respektive producentperspektiv

Som utvecklas mer av Byggnadsindexkommittén (SOU 1971:79) så kan byggherren (köparen) respektive byggproducenten (säljaren) ha olika perspektiv på vilka slags kvalitetsvariabler som är viktiga för nyproducerade bostäder.

För byggherren och ytterst de boende är kvalitet en fråga om nytta och funktionalitet i användningen av bostäderna, men för producenten är kvalitet sådana produkttegenskaper som går att ta betalt för. Begreppsmässigt är detta inte utan vidare samma sak och kunde potentiellt motivera olika val av kvalitetsvariabler.

BPI är ett producentprisindex och skulle därför idealt kvalitetsjustera ur ett producentperspektiv. I praktiken kan det dock vara vanskligt att helt hålla isär de två perspektiven. Som visas av Gordon (1990) och även berörs av Byggnadsindexkommittén (SOU 1971:79) så tenderar också de två perspektiven teoretiskt att resultera i likartade kvalitetsvärderingar under vissa förutsättningar om jämvikt och fungerande marknadsmekanismer. I någon mån kan det därför vara försvarligt att ta olika slags variabler som proxy för varandra.

Avvägning av antalet kvalitetsvariabler

Kvalitetsvariablerna får inte vara alltför många, för att inte medföra för stor osäkerhet; se Destatis (2009, Sect. 7.1, Annex 2). Detta är särskilt viktigt för år då det funnits få byggprojekt.

Prisvariabel som kvalitetsvariabel för lägesfaktorn

En av kvalitetsvariablerna avser genomsnittlig köpeskilling per kvadratmeter på försålda småhus i kommunen. Detta är ett sätt att beakta den viktiga lägesfaktorn. Läget är allmänt en av de viktigaste faktorerna bakom bostadspriser, och detta slår igenom även i priset exklusive tomtmarken på bostäder i nyproducerade hus. Byggnadsprojekten åren $t-1$ respektive t kan vara mer eller mindre markant olika belägna, och det gör det väsentligt att i kvalitetsjusteringen för indexlänken beakta dessa möjliga skillnader i läget.

Att på så sätt ha med en prisberoende variabel som kvalitetsvariabel kan verka något okonventionellt. Annars brukar kvalitetsvariablerna i hedonisk regression allmänt väljas så att de mäter fysiska karaktäristika. Den nämnda prisberoende kvalitetsvariabeln kan dock knappast leda till några anomalier här, och den verkar vara ett effektivt sätt att få med lägesfaktorn.

I Walestad (2007) diskuteras det geografiska lägets betydelse: ”Ett projekts geografiska läge har stor betydelse för slutpriset. Det finns relativt få byggprojekt i Sverige och dessa är koncentrerade till ett relativt lågt antal kommuner. Det finns 290 kommuner i Sverige och mellan 1995 och 2004 hade inga byggprojekt av den typen varit inkluderade i indexpopulationen i cirka 160 av dessa kommuner. Hjälpvariabeln vi har använt är det genomsnittliga kvadratmeterpriset för värdet av fastigheter per kommun för de senaste tre åren (2003-2005) och matchat dessa värden till varje projekt.”

Kommentar om principer bakom val av vägningsalternativ

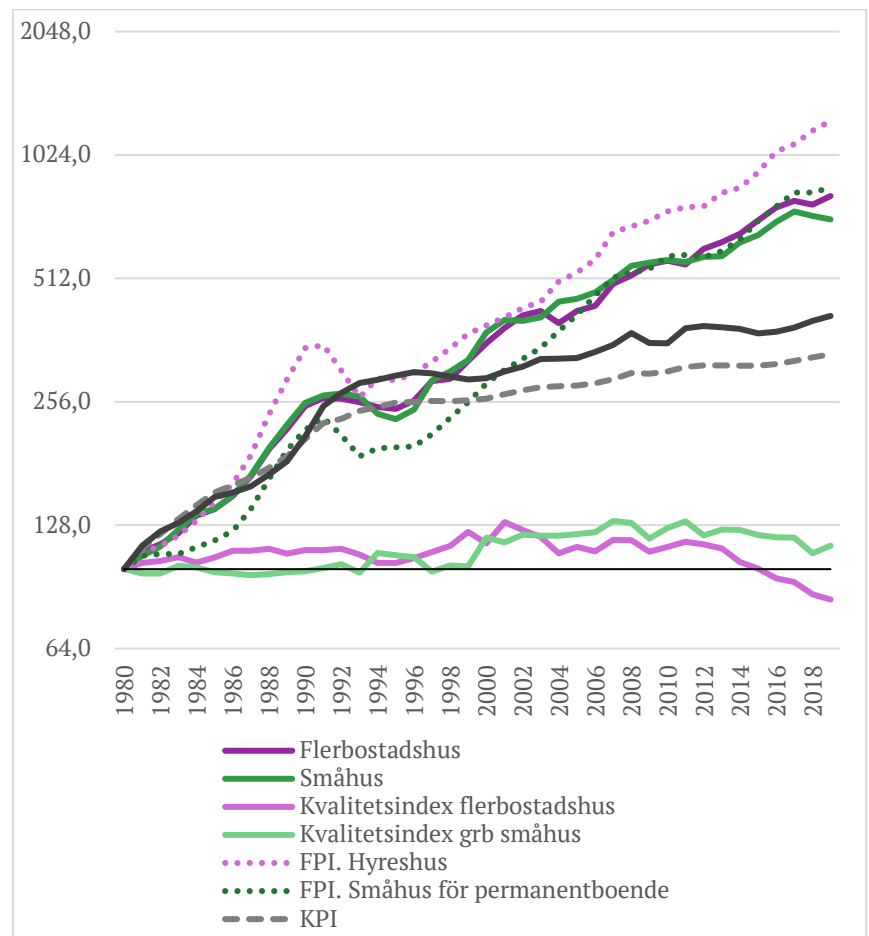
I skattningen av regressionskoefficienterna i formel (3) tillämpas inte någon vägning av observationerna, dvs bygggnadsprojekten. Byggnadsprojekten väger därmed sinsemellan lika tungt i skattningen av

regressionskoefficienterna. Detta motiveras med att koefficienterna då blir väntevärdesriktigt och effektivt skattade.

Vägning i proportion mot area eller pris kan ge dålig precision genom att låta stora projekt dominera för mycket, så att det effektiva antalet observationer minskar, och vore därför knappast lämpligt. Emellertid, om det skulle vara problem med att små projekt avviker markant från det allmänna mönstret, så kan det vara tänkbart att vid en metodöversyn eventuellt införa någon enkel regel som reducerar vikten för de minsta projekten, medan vikterna annars får förbli lika.

Utfall

Figur 1 visar byggnadsprisindex jämfört med konsumentprisindex (KPI) med basår 1980. Det kan noteras att kvalitetsindex för flerbostadshus har minskat stadigt sedan 2011.



Figur 1: Byggnadsprisindex (BPI), kvalitetsindex, faktorprisindex (FPI), konsumentprisindex (KPI) och KPI för boende. 1980 = 100.

Analys

Data

Det startas inte särskilt många projekt per år. Detta kan utgöra ett problem vad gäller statistikens kvalitet. I denna analys används följande två synsätt så att användaren själv ska kunna välja.

1. Det är en totalundersökning och då blir det ingen urvalsosäkerhet.
2. Det är inte de ändliga populationerna man vill göra inferens till, det är mer generellt hur priser sätts (superpopulationsansatsen).

Underlagen för BPI hade 3 respektive 6 procent bortfall år 2019.

Tabell 1: Antal byggprojekt i data per år 2013-2019.

Redovisningsår	Flerbostads hus	Gruppbyggda småhus
2013	305	388
2014	362	545
2015	440	566
2016	507	652
2017	464	585
2018	350	437
2019	331	400

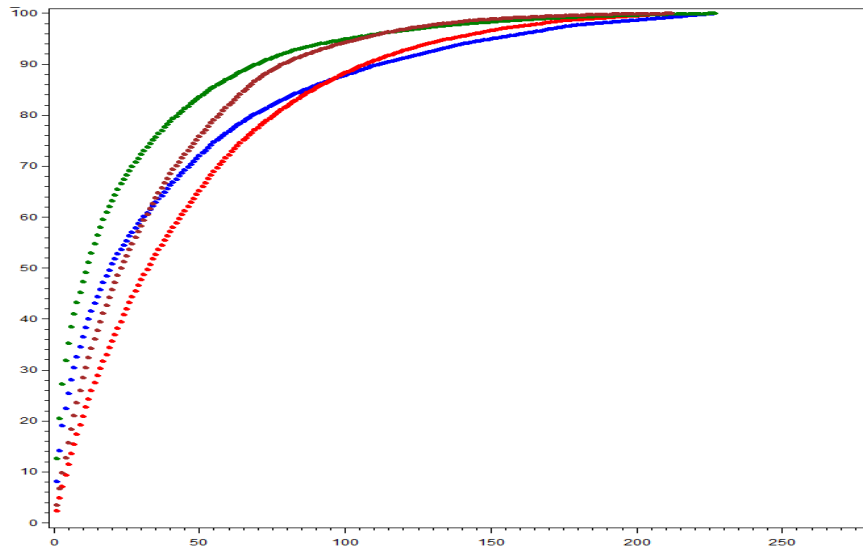
Lägesvariabeln

I detta avsnitt redovisas ett antal alternativa förklaringsvariabler till genomsnittlig köpeskillning för småhus per kommun, som har använts i de hedoniska regressionsmodellerna.

Kommuner

Bostadsbyggandet är naturligtvis mest förekommande i stora kommuner. 19 kommuner hade 50 procent av projekten avseende flerbostadshus 2015 – 2019 (blå kurva i figur 2 nedan). 12 kommuner räckte för att täcka 50 procent av lägenhetsarean i projekten (grön). För gruppbyggda småhus är koncentrationen något lägre (röd och brun).

87 kommuner hade minst ett byggprojekt varje år 2013-2019.



Figur 2: Kumulativ fördelning av projekt per kommun. 2015 - 2019

Blå = antal flerbostadshus, röd = antal gruppsybyggda småhus,
grön = lägenhetsyta i flerbostadshus, brun = lägenhetsyta i gruppsybyggda småhus

På en grov nivå, utan hänsyn till utsikt, bullriga miljöer, kommunikationer mm, är kommunen en lägesvariabel. En idé är att beräkna en BPI per kommun och sedan väga ihop dessa. Enklare är då att använda en hedonisk modell med en så kallad dummyvariabel per kommun. Det "kostar" dock många frihetsgrader. Detta har prövats med data för 2014 - 2019 sammanaget, se avsnitt Resultat nedan.

Genomsnittlig köpeskillning per kommun

En av kvalitetsvariablerna som alltid har använts i BPI är genomsnittlig köpeskillning per kvadratmeter för småhus i kommunen. För flerbostadshus sker så få försäljningar per år att det blir en volatil statistikserie. Därför används statistik för småhus.

Det är viktigt att köpeskillingen avser ett enda år alternativt ett medelvärde för flera år, annars fångas skillnader mellan år för y-variabeln fastighetspris upp av x-variabeln köpeskillning – om man har en regressionsmodell med flera år och tidsdummyvariabler.

En nackdel med genomsnittlig köpeskillning per kvadratmeter och kommun är att detta är en statistik med viss statistisk osäkerhet, vilken fortplantar sig i den hedoniska modellens skattningar.

I den hedoniska modellen, med y-variabeln logaritmerad fastighetspris per kvadratmeter, skulle det intuitivt vara lämpligt att x-variabeln genomsnittlig köpeskillning också är logaritmerad. Rent numeriskt visar det sig att detta inte ger högre förklaringsgrad, R^2 , så även fortsättningsvis används variabeln köpeskillning utan logaritmering som det alltid varit.

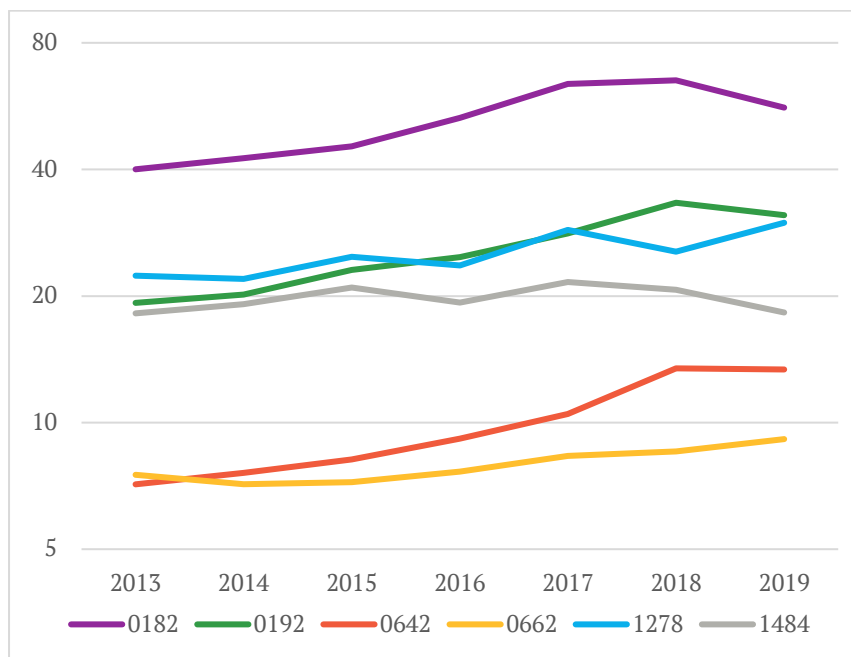
I analysen har det också testats om det finns kommuner som avviker i relationen mellan genomsnittlig köpeskilling och byggnadspris. Utifrån befolkning, yta, medelinkomst, kommunalskatt, politisk majoritet m.m. har fyra kluster skapats. Åtta höginkomstkommuner i Stockholms och Skåne län bildar ett kluster, där byggnadspriserna för flerbostadshus ligger 19 – 22 procent högre än i 146 glesbygdskommuner, trots att variabeln genomsnittlig köpeskilling per kommun är med som förklarande variabel.

Överlag är förklaringsgraden för regressionsanalyserna högre för flerbostadshus än för gruppbyggda småhus, som högst 43 procent mot 22 procent. Tittar man bara på lägesvariabeln är förklaringsgraden 28 procent respektive 17 procent.

Den genomsnittliga köpeskillingen per kommun baseras på administrativa uppgifter från Lantmäteriets fastighetsprisregister. Fastighetsprisregistret täcker hela landet och omfattar alla typer av fastigheter. Kopplat till de fastigheter som ingår i överlåtelsen finns även egenskapsuppgifter från bland annat fastighetstaxeringen.

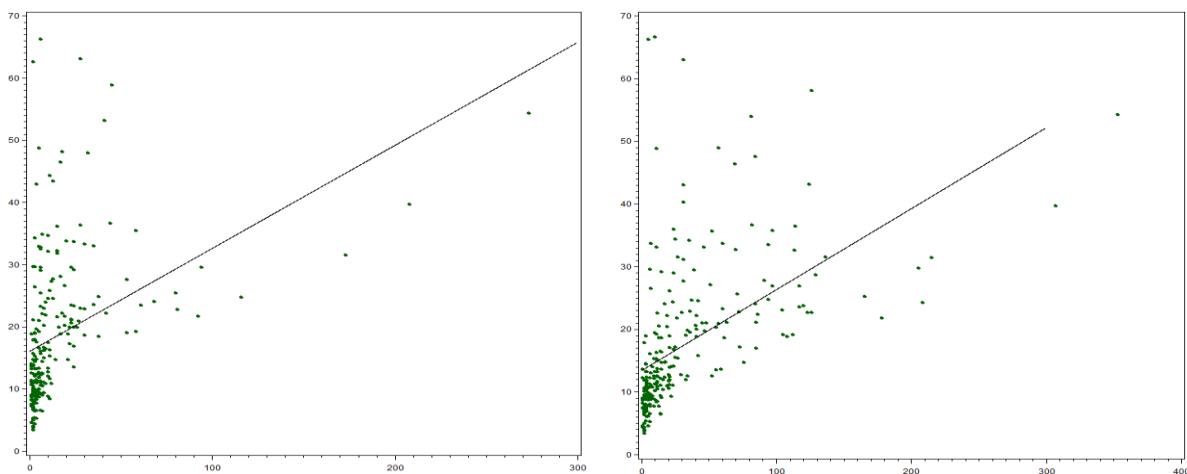
Målpopulationen består av de beviljade överlåtelser som registrerats i Lantmäteriets fastighetsprisregister under den specifika referensperioden. Köpeskilling är det belopp i SEK som erlagts för det som har överlåtit. I genomsnitt har tre fjärdedelar av kommunerna cirka 80 försäljningar eller fler per år. Det görs ingen kvalitetsvärdering så en del av variationen mellan kommuner förklaras av husens kvalitet i olika avseenden men såklart beror priser på i vilka kommuner fastigheterna ligger.

Figur 3 visar data från SSD för sex medvetet valda kommuner för sju år. Genomsnittlig köpeskilling per kvadratmeter och kommun varierar mellan knappt 8 tusen kronor och drygt 60 tusen kronor. Genomsnittlig köpeskilling varierar också kraftigt mellan åren. Även om statistiken är en totalundersökning av alla sålda småhus så är skillnader mellan åren knappast en indikator på byggnadsprisernas utveckling.



Figur 3: Genomsnittlig köpeskillning per kvadratmeter på försålda småhus. Medvetet val av sex kommuner. Per kommun 2013-2019

Figur 4 visar vilken samvariation som finns mellan genomsnittlig köpeskillning och kommunstorlek mätt med antal byggprojekt.



Figur 4: Genomsnittlig köpeskillning per kommun mot antal byggprojekt. 2016-2019. Den vänstra avser flerbostadshus och den högra avser gruppsybyggda småhus

Andra kommunala variabler med potential

Finns det någon annan kommunal egenskap med relation till byggnadspriser per kvadratmeter som har stadigare nivåer? Genomsnittlig bruttoinkomst per kommun 2018 och genomsnittlig hyra per kvadratmeter enligt undersökningen Hyror i bostäder, HiB, har provats. Det visar sig att dessa alternativ inte är bättre än den genomsnittliga köpeskillingen, men inte mycket sämre heller.

Skatteverkets värdeområden

Inom fastighetstaxeringen delar Skatteverket in landet i drygt 2 000 geografiskt avgränsade områden, så kallade värdeområden. Dessa består av fastigheter med liknande förutsättningar. Områdena följer inga administrativa gränser utan avgränsas av vägar, vatten och liknande. För fastigheter inom dessa områden samlar Skatteverket in uppgifter om uthyrda ytor och hyresintäkter för hyreshusmen även totala ytor och motsvarande hyror som skulle ha varit. Statistiken är inte av högsta kvalitet, bland annat avser uppgifterna fyra år, utan att vi vet vilket. För statistikprodukten Hyror i bostadslägenheter, HiB, har analyser visat att Skatteverkets genomsnittshyror per kvadratmeter för dessa värdeområden har högt förklaringsvärde för variationen i hyror i lägenheter i Sverige.

Resultat av analyser med olika lägesvariabler

Skatteverkets hyror per kvadratmeter i värdeområden har prövats som förklaring av byggnadspriser per kvadratmeter. Ett problem har varit att matcha byggprojekt till värdeområde. Fastighetsbeteckningen har en namndel och en sifferdel, av typen ”Gläntan 5”. Byggprojektens fastighetsnamn har matchats med fastigheter som är kopplade till ett värdeområde och därvid har nästan hälften av byggprojekten matchas till ett värdeområde.

Tabell 3a: Jämförelse av olika lägesvariabler. Regressionsanalys med data 2014-2019.

Lägesvariabel	Adjusted R ²	
	Flerbostads- hus	Gruppbyggda småhus
290 dummy-variabler för kommuner	46,59	36,71
4 dummyvariabler för 4 kluster av kommuner	40,87	23,31
Genomsnittlig köpeskillning per kommun	45,34	27,96
Genomsnittlig köpeskillning per kommun samt 4 kommunkluster	45,45	27,86
Log(Bruttoinkomst per inv 2018)	42,39	25,45
Genomsnittlig hyra per kvm enligt HiB	41,84	23,79
Ingen lägesvariabel	36,14	18,93

Tabell 3b: Jämförelse av två lägesvariabler. Regressionsanalys med data där matchningen med värdeområden 2016-2019 lyckats till 49 procent. T-värden.

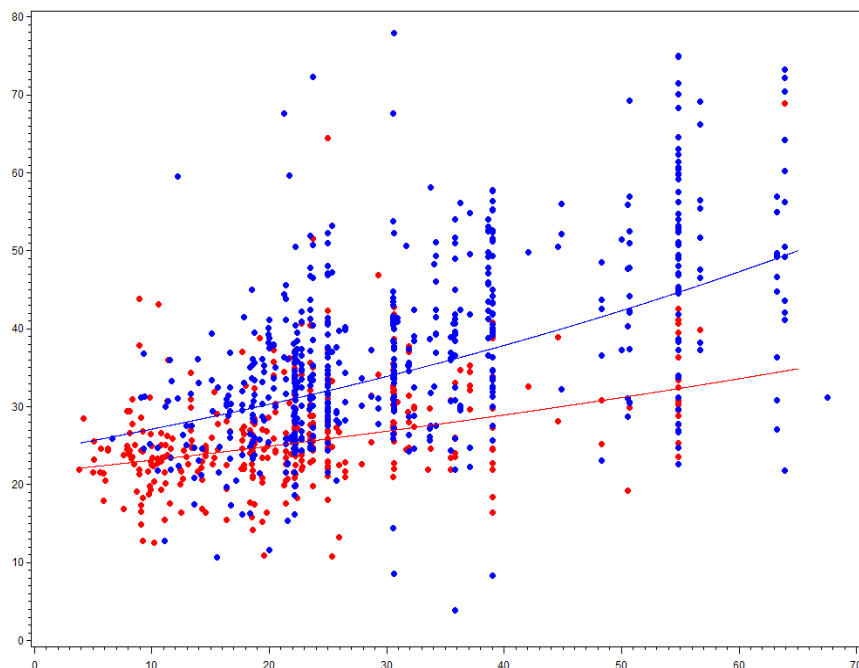
Lägesvariabel	Flerbostadshus		Småhus	
	Brutto	Netto	Brutto	Netto
log(Genomsnittlig köpeskillning per kommun)	10,94	11,52	6,82	6,89
log(Genomsnittlig hyra för värdeområde)	1,88	2,10	4,84	4,82

Den första slutsatsen är att i byggnadsprisindex är hyran per kvadratmeter i Skatteverkets värdeområden inte en bra variabel för att förklara variationer.

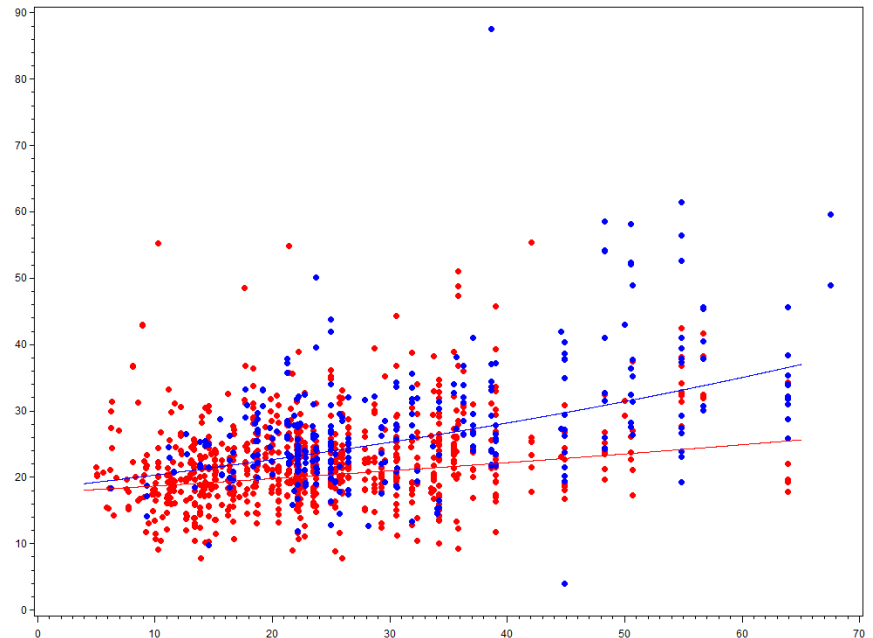
Slutsats två är att den genomsnittliga köpeskillingen för småhus per kommun är ett bra sätt att i en enda variabel beskriva projektens läge i regressionerna.

Illustration av variationen

I nedanstående punktdiagram visas den spridning i byggnadspriser som finns mellan kommuner i olika nivåer av genomsnittlig köpeskillning per kvadratmeter. För flerbostadshus är det också stor skillnad mellan upplåtelseformerna hyresrätt och bostadsrätt. Tillsammans har dessa två variabler förklaringsgraden R^2 runt 38 procent för flerbostadshus och 23 procent för gruppbyggda småhus för åren 2015 – 2018 men betydligt sämre för 2019.



Figur 5a: Byggnadspris per kvadratmeter mot genomsnittlig köpeskillning på försålda småhus per kommun. Flerbostadshus 2017. Rött = hyresrätter, blått = bostadsrätter.



Figur 5b: Byggnadspris per kvadratmeter mot genomsnittlig köpeskillning på försålda småhus per kommun. Gruppbyggda småhus 2017. Rött = hyresrätter, blått = bostadsrätter.

Vägt eller inte vägt?

Den hittillsvarande beräkningsmetoden innebär att *Värdekomponenten* V_t är kvoten av vägda medelvärden av pris per kvadratmeter och är i den meningen att medelvärdena beräknats som summa priser dividerat med summa areor.

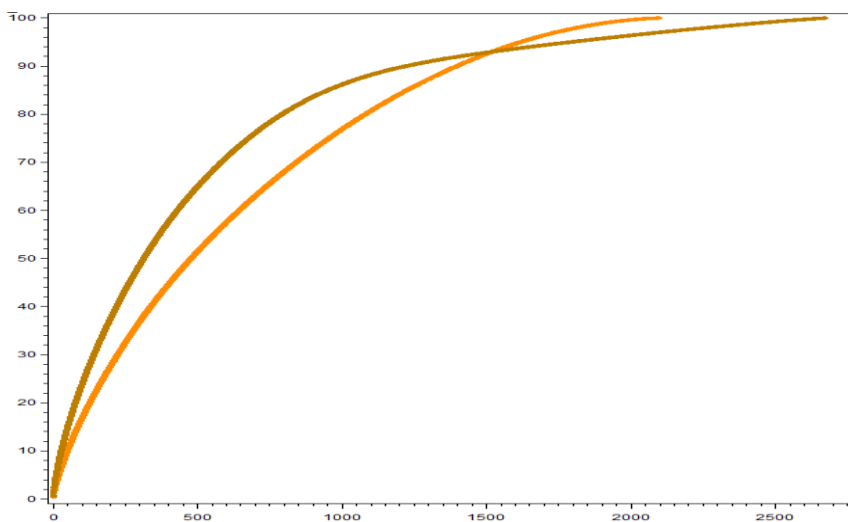
Likaledes beräknas medelvärdena av de kvalitetsvariablerna, vägt med areor per år, $\bar{x}_{i,S}$.

Den hedoniska funktionen, å andra sidan, skattas med enskilda projekts genomsnittliga pris per kvadratmeter som beroende variabel. Här görs ingen vägning med projektens storlek. Om det är stor variation i projektens storlek (areor) så gör det skillnad med ovägt och vägt.

Att inte väga i skattningen av regressionskoefficienterna motiveras av att precisionen då blir avsevärt bättre, samtidigt som skattningen där blir approximativt väntevärdesriktig under konventionsantagandet att regressionskvationen inklusive residualtermen är sann. Att väga i medelvärdesberäkningen av V och K är nödvändigt för att index ska bli plutokratiskt. Med oviktade medelvärden över projektmängden får man ett "projektokratiskt" index.

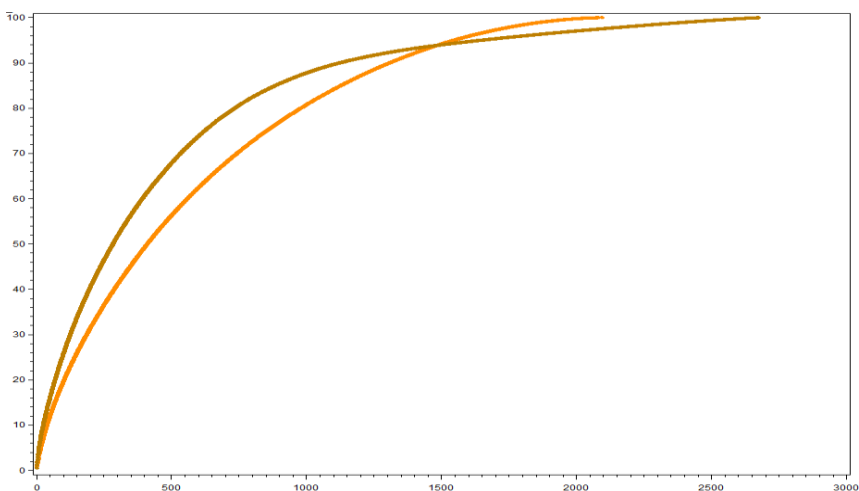
23 procent av projekten täckte 50 procent av totala lägenhetsarean för flerbostadshus. 11 procent av projekten för gruppbyggda småhus täckte 50

procent av totala bostadsarean, 28 procent av projekten täckte 80 procent av arean, se figur 6a.



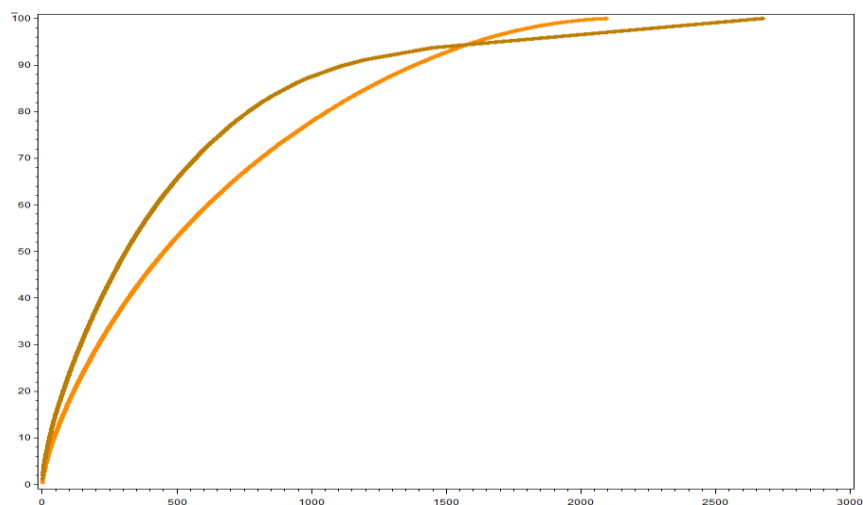
Figur 6a: Kumulativ fördelning av projektens areor 2015 - 2019. Orange = flerbostadshus, ljusbrun = gruppbyggda småhus.

Samma mönster kan ses för projektens priser i figur 6b som för projektens areor i figur 6a.



Figur 6b: Kumulativ fördelning av projektens priser 2015 - 2019. Orange = flerbostadshus, ljusbrun = gruppbyggda småhus.

Även för antal lägenheter i figur 6c kan samma mönster ses som i figur 6a och 6b.



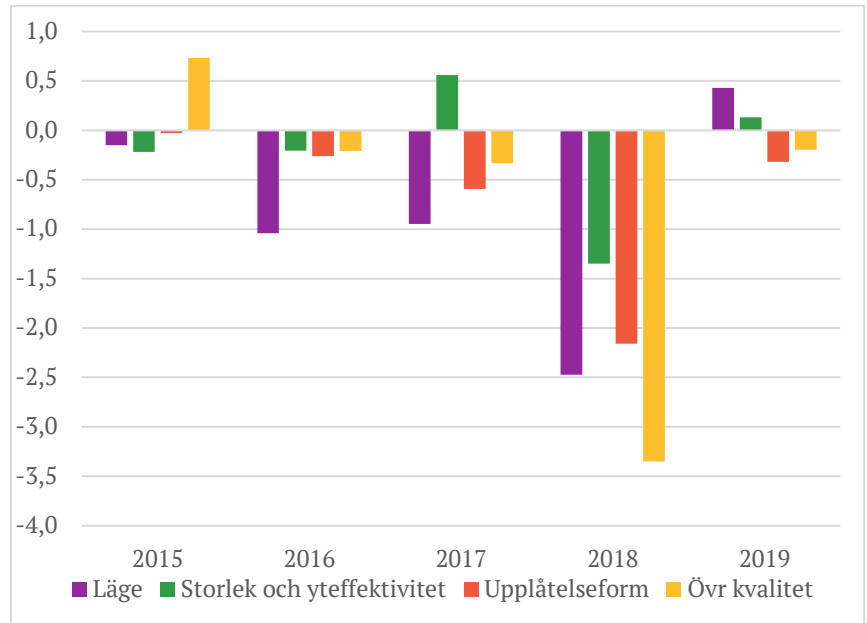
Figur 6c: Kumulativ fördelning av projektens antal lägenheter 2015 – 2019. Orange = flerbostadshus, ljusbrun = gruppbyggda småhus.

Kvalitetsindex

Kvalitetsindex beräknat utifrån kvalitetskomponenten i BPI publiceras i Sveriges statistiska databaser, SSD. Detta motsvarar det som för KPI kallas ”Implicit Quality Index”, IQI.

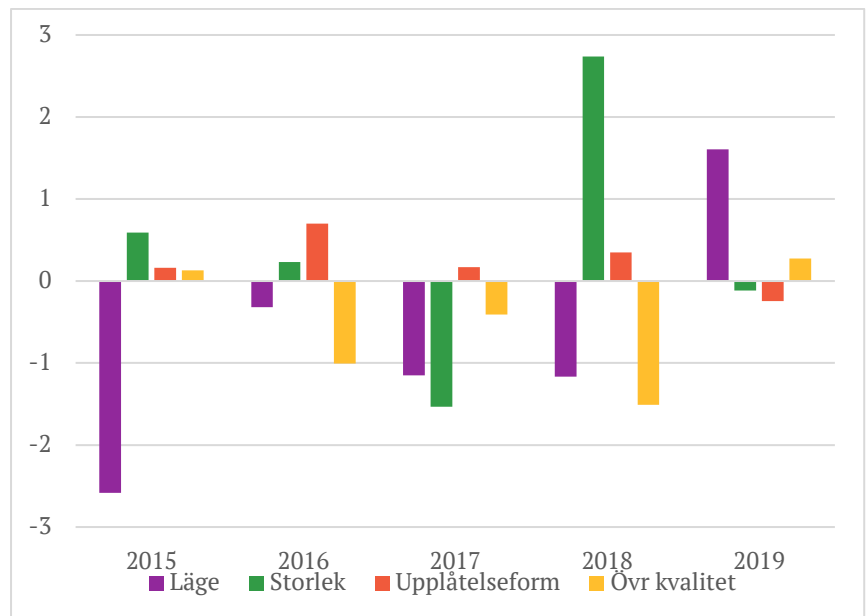
Sedan den stora omläggningen 2007-2008 har den hedoniska modellen inte ändrats i uppsättningen av kvalitetsvariabler. Åren innan detta laborerade man med olika variabler varje år i den modell som användes då, vilket gjorde tolkningen av kvalitetsindex svårare.

I vår analys har förklaringsvariablerna, och därmed kvalitetsindex, delats upp i fyra grupper. Den analys som redovisas här nedan är inte helt konsistent med publicerade resultat då några fler förklarande variabler har inkluderats i modellerna. Variabelgruppen Läge består av genomsnittlig köpeskilling per kommun. Variabelgruppen Storlek består av antal kvadratmeter, antal våningar och yteffektivitet. Yteffektivitet definieras som bostadsarean plus annan uthyrningsbar lokalarea som andel av bruttoarean. Upplåtelseform för bostaden/bostäderna (äganderätt, bostadsrätt och hyresrätt) får vara en egen grupp av förklaringsvariabel. Variabelgruppen Övrigt består av resten.



Figur 7a: Läge, storlek, upplåtelseform och övrig kvalitet. Flerbostadshus 2015 – 2019.

De största effekterna 2015 – 2019 för flerbostadshus var att andelen bostadsrätter minskade från 2017 till 2018 och samtidigt var andelen projekt i kommuner med lägre köpeskilling högre. Det finns ett signifikant samband mellan dessa två egenskaper, att bostadsrätter byggs i kommuner med höga priser, men det innebär knappast några problem för indexberäkningen, möjligen i uttolkningen av vad som förklarar ”kvaliteten”.



Figur 7b: Läge, storlek, upplåtelseform och övrig kvalitet. Gruppbyggda småhus 2015 – 2019.

Variabelgruppen Storlek avser projektens storlek med avseende på antal kvadratmeter, antal våningar (endast flerbostadshus) och yteffektivitet. Storlek har varierat relativt kraftigt mellan åren med effekt på priset. Det ser ut som en tanke att riktningen på effekten varje år är den motsatta för flerbostadshus och gruppbyggda småhus. För såväl flerbostadshus som gruppbyggda småhus är läget (kommunen) en viktig faktor för priset.

Resultat

Översyn av kvalitetsvariabler

Flera år av data har här använts i en multilateral modell utan vägning för att begränsa antalet skattade modeller. Data avseer 2014 – 2018, 2015 – 2019 och 2016 – 2020.

Tabell 4a: Kvalitetsvariabler i hedoniska modeller. Flerbostadshus 2018 – 2020. T-värden i variansanalysen.

	2020	2019	2018
Intercept	50,19	49,95	49,79
Genomsnittlig köpeskilling per värdeyta, ovägt medelvärde	19,06	19,53	21,60
Bostadsrätter	17,61	16,44	16,02
Äganderätter ¹	0,92	-0,19	-0,24
Yteffektivitet, två decimaler	-7,63	-6,40	-6,73
Antal hygienrum per lägenhet, två decimaler	0,26	1,34	1,43
Loftgångshus	0,49	0,70	1,90
Punkthus	-0,49	-0,38	0,35
Övriga	0,28	0,18	2,00
Källare finns/finns ej i projektet	1,93	1,53	1,46
Antal balkonger per lägenhet, två decimaler	1,89	1,57	1,77
Hiss finns/finns ej i projektet	5,47	7,09	8,38
Antal garage eller parkeringsdäck per lägenhet, två decimaler	1,91	3,02	2,99
Ar(t)	2,24	-0,39	0,79
Ar(t-2)	0,24	-0,90	-1,52
Ar(t-3)	-0,54	-2,40	-5,05
Ar(t-4)	-1,97	-5,74	-6,34

I den nu använda modellen för flerbostadshus kan man i variansanalysen se att flertalet av de ingående variablerna inte är signifikanta. Endast

¹ Dummyvariabeln Ägarrätter är inte signifikant på grund av att det endast funnits fem projekt som avser ägarlägenheter.

variablerna genomsnittlig köpeskillning, bostadsrätt, yteffektivitet samt tillgången till hiss är signifikanta.

Variabeln Punkthus har haft olika tecken de tre åren.

Projektstorlek, mätt med total lägenhetsarea, har testats att inkluderas i modellen. Projektstorlek visar sig vara en signifikant variabel. Ju större total area desto lägre pris per kvadratmeter.

Genomsnittlig lägenhetsarea har också testats att inkludera i modellen. Denna visar sig vara starkt signifikant.

Även variabeln Antal våningsplan har testats i modellen. Denna variabel är naturligt korrelerad med Hiss. Hus med minst 4 våningar har hiss. Ungefär hälften av husen med två och tre våningar har hiss. Antal våningar är signifikant, ju högre hus desto högre pris per kvadratmeter, men detta kan inte beskrivas linjärt eller loglinjärt. Flerfamiljshus med en, två och tre våningar bör separeras från flerfamiljshus med fyra våningar eller fler med hjälp av dummyvariabler, där flerfamiljshus med fyra våningar eller fler används som referens. Bland hus med två och tre våningar har hiss ungefär lika stor prishöjande effekt.

I den nu använda modellen för gruppbyggda småhus kan man i variansanalysen se att ett par av de ingående variablerna inte är signifikanta. Variablerna Kedjehus samt Fasadmaterial puts, plåt, övrigt är inte signifikanta. Se tabell 4b.

Riktigt "dåliga" variabler är de som haft olika tecken de tre åren. En sådan är Inredningsbar vind.

Sedan finns det variabler som är signifikanta vissa år men inte andra. Exempel på sådana variabler är Källare och "Annat än en-planshus".

Projektsstorlek, mätt med total lägenhetsarea har även testats i modellen för gruppbyggda småhus. Här är den dock inte signifikant.

**Tabell 4b: Kvalitetsvariabler i hedoniska modeller.
Gruppbyggda småhus 2018 – 2020. T-värden i variansanalysen.**

	2020	2019	2018
Intercept	110,14	118,93	117,50
Bostadsarea per lägenhet, två decimaler	-16,54	-17,32	-17,45
Genomsnittlig köpeskillning per värdeyta, ovägt medelvärde	18,06	20,79	21,91
Antal hygienrum per lägenhet, två decimaler	3,51	4,68	4,96
Kedjehus	0,16	0,65	0,22
Radhus o övriga	-2,23	-2,69	-3,98
Källare finns/finns ej i projektet	1,70	3,34	3,91
Fasadmaterial Sten	3,49	3,22	4,80
Fasadmaterial puts, plåt, övrigt	-1,29	-0,49	-0,06
Takmaterial Lertegel	2,43	2,20	2,09
Takmaterial Plåt, Papp, annat	7,23	6,60	7,27
Inredningsbar vind finns/finns ej i projektet	-0,52	-0,31	0,69
Annat än en-planshus	3,45	2,60	1,59
Bostadsrätter	4,06	5,86	5,85
Hyresrätter inkl. kooperativa hyresrätter	-4,62	-3,28	-3,06
Antal garage eller parkeringsdäck per lägenhet, två decimaler	5,12	5,14	4,15
Ar(t)	0,54	-2,56	5,67
Ar(t-2)	2,43	-5,66	2,14
Ar(t-3)	-2,55	-3,56	-2,15
Ar(t-4)	-1,14	-7,42	-2,82

Skattningar av varianser

Random groups metoden har tillämpats för att skatta varianser för BPI. Metoden beskrivs bl a av Särndal, Swensson & Wretman(1997). Varianserna för BPI har skattats utan ändlighetskorrektioner, dvs som om BPI vore en indikator på prissättning snarare än hur de faktiska projektens priser har varierat mellan åren. Konfidensintervallens bredd blir ungefär ± 4 procentenheter.

Icke-parametriskt skattade varianser med Dependent Random groups

Beteckningar enligt Särndal, Swensson & Wretman (1997)

Om man betraktar vår population av projekt som ett urval från en superpopulation har man alltså ett urval s per år. Detta urval delas slumpmässigt upp i A disjunkta mängder $s_I - s_A$. Om det hade funnits stratifiering eller liknande hade man sett till att varje delurval skulle få ungefär likadana sammansättningar.

Låt $\hat{\theta}_1 - \hat{\theta}_A$ vara skattningar av en sökt parameter θ , var och en baserad på $1/A$ av den totala urvalsstorleken. I vårt fall är den årliga BPI-länken den parametern.

Bilda medelvärdet av alla skattningar, som också är en skattning

$$\hat{\theta}_{\text{DRG}} = \frac{1}{A} \sum_{\alpha=1}^A \hat{\theta}_{\alpha}$$

En varians estimator ges nu av

$$\hat{V}_{\text{DRG}} = \frac{1}{A(A-1)} \sum_{\alpha=1}^A (\hat{\theta}_{\alpha} - \hat{\theta}_{\text{DRG}})^2$$

Eftersom populationerna (urvalen) är ganska små när 15-20 parametrar i regressionsmodellerna ska skattas väljs $A = 2$. Samtidigt skapas 2 000 variansskattningar som som ett medelvärde beräknas av. Detta är en variant som har använts i andra sammanhang och brukar kallas Repeated Random Groups.

Med denna metod skattas varianser för såväl BPI som dess beståndsdelar Värdekomponenten och Kvalitetsjusteringen.

De hedoniska regressionskoefficienterna skattas dels som i BPI med data för basåret och dels med fyra år av data och i detta fall används även tre dummyvariabler för år (vilka sedan inte används).

Modells kattning av varians

Från anovan vid regressionsanalyserna med den alternativa metoden Tids-dummy variable index, kallad DV av Dalén (1992), hämtas parametern stder. Metoden har tillämpats 1) bilateralt med endast de två år som jämförs och 2) multilateralt med fem år av data, varav de två sista mäter den sökta prisutvecklingen. Dessa index kan sägas vara i konflikt med vad som sagts ovan i avsnitt "Vägt eller ovägt" om vilket alternativ som är teoretiskt mest korrekt.

Skattningar

Variansen för BPI förväntas vara mindre än för värdekomponenten beroende på att kvalitetsjusteringen ska göra projekten för de två åren mer "lika". För gruppbyggda småhus i tabell 5c ser man inte denna effekt de två senaste åren. Den hedoniska modellen har för låg förklaringsgrad och skapar därmed mer variation än den reducerar.

Det framgår av tabell 5a och 5c att skattningar av de hedoniska koefficienterna med flera år av data inte nämnvärt förbättrar precisionen i BPI med i övrigt traditionell beräkning med värdekomponent och kvalitetsjustering.

Tabell 5a: Ickeparametriskt skattade varianser för BPI flerbostadshus med random groups metoden.

Basår – aktuellt år	1 års regression			4 års regression		
	BPI	Värde- kompo- nenten	Kvalitets- juste- ringen	BPI	Värde- kompo- nenten	Kvalitets- juste- ringen
2014 – 2015	3,4	5,5	1,6			
2015 – 2016	2,7	5,3	1,7			
2016 – 2017	2,7	4,4	1,5	2,6	4,4	1,4
2017 – 2018	4,1	4,1	2,1	3,9	4,1	1,5
2018 – 2019	3,8	4,9	1,2	3,8	4,9	1,7

Tabell 5b: Modellbaserat skattade varianser för BPI flerbostadshus. Modellskattningar för tids-dummy-index med två respektive fem år av data

	2 år	5 år
2014 – 2015	2,3	
2015 – 2016	2,2	2,0
2016 – 2017	2,6	2,2
2017 – 2018	3,7	2,9
2018 – 2019	3,7	3,5

Tabell 5c: Ickeparametriskt skattade varianser för BPI gruppbyggda småhus. Random groups metoden.

Basår – aktuellt år	1 års regression			4 års regression		
	BPI	Värde- kompo- nenten	Kvalitets- juste- ringen	BPI	Värde- kompo- nenten	Kvalitets- juste- ringen
2014 – 2015	3,7	4,3	1,7			
2015 – 2016	2,2	3,2	0,8			
2016 – 2017	3,2	4,6	1,0	3,2	4,6	1,0
2017 – 2018	5,7	5,0	3,3	5,7	5,0	3,3
2018 – 2019	6,2	4,9	2,7	6,3	4,9	2,8

Tabell 5d: Modellbaserat skattade varianser för BPI gruppbyggda småhus. Modellskattningar för DV-index med två respektive fem år av data.

	2 år	5 år
2014 – 2015	1,6	
2015 – 2016	1,6	1,5
2016 – 2017	1,8	1,6
2017 – 2018	2,6	2,2
2018 – 2019	3,8	3,0

Pågående arbete

Ett parallellt arbete som pågår vid sidan av denna utredning är ett projekt kallat Slutprisprojektet. I det här projektet har SCB under hösten 2020 skickat ut ca 1 000 enkäter till de byggherrar som lämnade fullständiga svar och som använts till produktionen av statistiken för år 2016 eller 2017 samt erhållit slutbesked (klarbesked) från kommunen senast januari 2020. I enkäten har dessa byggherrar fått svara på frågor kring hur kostnaderna / priserna för projektet har sett ut när nybyggnadsprojektet blivit klart.

Vid föreliggande rapports skrivande har inga resultat tagits fram ännu.

Slutsatser och förslag

Indexmetod

1. **Vägt eller ovägt.** Den hittillsvarande beräkningsmetoden innebär att *Värdekomponenten* är kvoten av två års vägda medelvärden av pris per kvadratmeter. Likaledes beräknas medelvärdena av de förklarande variablerna, vägt med projektens areor per år. De hedoniska funktionerna skattas med projektens genomsnittliga pris per kvadratmeter som beroende variabel. Här görs ingen vägning med projektens storlek. Dock kan area ingå som förklarande variabel. Om det är stor variation i projektens storlek (areor) så kan det göra skillnad på parameter-skattningarna med ovägt och vägt. Är det bra som SCB gör?

2. **Skattning av hedoniska koefficienter med många år.** Som demonstreras i tabeller 5a och 5c skulle det vara möjligt att skatta de hedoniska regressionerna med mer data genom att inte endast använda data för år t-1 utan alla år t-x till och med år t. Utvecklingen mellan år t-1 och år t är

$$P_t = \frac{V_t}{K_t} \quad \text{där} \quad V_t = \frac{\bar{y}_t}{\bar{y}_{t-1}} \quad \text{och} \quad K_t \text{ är kvalitetsskillnaden värderad.}$$

Utvecklingen mellan år t-x och t blir

$$\frac{\bar{y}_{t-x+1}}{\bar{y}_{t-x}} \cdot \frac{\bar{y}_{t-x+2}}{\bar{y}_{t-x+1}} \cdot \dots \cdot \frac{\bar{y}_t}{\bar{y}_{t-1}} \cdot K_{t-x+1} \cdot K_{t-x+2} \cdot \dots \cdot K_t = \frac{\bar{y}_t}{\bar{y}_{t-x}} \cdot K_{t-x+1} \cdot K_{t-x+2} \cdot \dots \cdot K_t$$

dvs kvoten av medelpriser okorrigerat och korrigerat med produkten av alla hedoniska korrigeringar där var och en av dessa baseras på ett år av data. Om tillfälligheter (slumpen) spelar ett spratt för ett år så korrigeras det inte nästa år, de är "smått oberoende".

Vårt förslag är att det varje år ska hedoniska regressioner skattas med data för ett glidande fönster av flera år, ju längre, desto bättre. Detta är särskilt viktigt i tider av låg konjunktur då antalet projekt per år kan bli väldigt litet. Notabelt är dock resultatet i tabellerna 5a och 5c där det inte framgår att varianserna blir mindre.

3. **Tids-dummy-variabel-metoden.** De skattade varianserna med tids-dummy-metoden är något lägre än de med justeringsmetoden. Variansskattningarna är ännu något lägre när man tar med fem år av data i en multilateral index, bestående av året som ska jämföras med året innan (bas) samt tre år innan dess. Man får över tid stabilare skattningar men ett mindre färskt index, mindre följsamt med prisutvecklingen.

4. **Vägning vid tids-dummy-variabel-metoder.** Det hittillsvarande förfarandet innebär att de hedoniska koefficienterna skattas ovägt medan medelvärden av priser per kvadratmeter och hedoniska

egenskaper har beräknats vägt. Om en tids-dummy-metod väljs, så ska den beräknas vägt med projektvärde för att ge korrekt plutokratiskt index.

Kvalitetsvariabler

5. **Regionala effekter.** Det finns starkt signifikanta skillnader i byggnadspris per kvadratmeter mellan kommuner. Det kan noteras att Skatteverkets cirka två tusen värdeområden inte bidrar till förklaring av variationen i priser. Denna information är ju underlag för fastighets-taxeringen och har visat sig förklara hyror i hyreshus. Styr byggnadspriser av vilken kommun husen byggs men inte av läge inom kommun (centralt, förort, glesbygd, nya områden, förtätningar)?

Markpriset, inklusive beredning av marken, för projektet ingår inte i byggnadspriset men variabeln finns registrerad hos SCB. Föreliggande utredning har inte haft tid att studera om det vore den ultimata "lägesvariabeln", är det tänkbart?

6. **Dåligt förklarande variabler.** Förklarande variabler som under flera år inte är signifikanta och i synnerhet de variabler som under en följd av år har olika tecken tas bort.

7. **Vilka kvalitetsvariabler fattas?** Vilka ytterligare variabler bör ingå för att förklara variationer i byggnadspriser? För flerbostadshus har konstaterats att projektstorlek och genomsnittlig lägenhetsstorlek är starkt signifikanta och hittills inte använts.

BPI är en outputindex och därför är det relevant och väsentligt att kvalitetsjustera för variabler som indikerar produktionsprestationer. Det kan vara fördyrande omständigheter för byggande som nya av lagar och förordningar, ändrade EU-regler, miljöhänsyn, arbetarskydd, arbete med grund som pålning mm. Missar i detta kan ha gjort BPI överskattande. Denna risk kan behöva vara en prioriterad fråga i en eventuell översyn av BPI.

8. **Är projektstorlek en kvalitetsvariabel?** Om man rensar priserna för olika projektstorlek i den hedoniska modellen så kommer ändrade projektstorlekar inte att visa sig när användarna i nästa steg analyserar produktivitetsutveckling.

Statistikens kvalitet

9. **Inferens till ändlig population eller superpopulation?** Bör SCB i kvalitetsdeklarationen presentera skattningar av medelfel till superpopulation? (Till ändlig population inte relevant för BPI:s användningar.)

10. **Kvalitetsdeklaration.** Eftersom kvaliteten inte kan förbättras så mycket är det desto viktigare att statistiken får en i möjlig mån lättläst kvalitetsdeklaration, särskilt avseende tillförlitlighet. Reservation behöver dock göras för osäkerheten i kännedomen om osäkerheten, genom att populationen av byggprojekt är exceptionellt flyktig från år till år.

Referenser

Borg, L., Warsame, A., Song, H-S. och Lind, H. (2013). *Kvalitetsförändringar och produktivitetmätning: Hur stort är egentligen mätfelet?*, Avd f Bygg- och fastighetsekonomi, Inst f Fastigheter och Byggnad, KTH Sept 2013.

Dalén, J. (1992). *Operationalising a hedonic index in an official price index program: personal computers in the Swedish Import Price Index*, SCB R&D Report 1992:15 R & D Report 1992:15. [Operationalising a hedonic index in an official price index program – personal computers in the Swedish import price index \(scb.se\)](#)

Destatis (2009). *Handbook on the application of quality adjustment methods in the Harmonised Index of Consumer Prices*, Statistics and Science Vol. 13, Wiesbaden: Federal Statistical Office of Germany.

Gordon, R. J. (1990). *The Measurement of Durable Goods Prices*, a selection from an out-of-print volume from the National Bureau of Economic Research. (PDF) [The Measurement of Durable Good Prices \(researchgate.net\)](#)

Lind, H. och Song, H-S. (2011). *Dålig produktivitet utveckling i byggindustrin: Ett faktum eller mätfel??*, Institutionen för fastigheter och Byggnad, Avd f Bygg- och fastighetsekonomi, Kungliga Tekniska Högskolan. Slutrapport 2011-12-11.

Ribe, M. (2014). *Regressionsmetoden i Byggnadsprisindex (BPI)*, Utkast 2 till dokumentation, SCB, finns i appendix A5.

Särndal, Swensson & Wretman(1997). *Model Assisted Survey Sampling*, Springer Series in Statistics

SOU 1971:79. *Byggnadsindex för bostäder*, Byggnadsindexkommittén

Triplett, J. (2006). *Handbook on Quality Adjustment of Price Indexes for Information and Communication Technology Products*, OECD Directorate for Science, Technology and Industry, Paris: OECD.

Walestad, M. (2007). *Final report for "Draft Grant Agreement Nr 44401.2005.004-2005.312*, SCB:s slutrapport i ett projekt som genomförs i samarbete med Eurostat

Appendix

A1: Fyra metoder enligt Byggnadsindexkommittén (SOU 1971:79)

1.3.5 Metoder för beräkning av byggnadsindex inom bostadsområdet (kapitel 5)

1.3.5.1 Beräkning av byggnadsprisindexar

I fyra av de diskuterade metoderna -typhusmetoden, kvalitetsklassmetoden, kvalitets-elementmetoden samt regressionsmetoden - baseras beräkningarna på uppgifter som direkt avser slutprodukten. Dessa metoder förutsätter tillgång till pris- eller kostnadsobservationer för hela hus eller projekt. Två metoder - faktorprismetoden och kostnads-komponentmetoden - avser beräkningar som utförs på grundval av uppgifter från tidigare produktionsled, avseende antingen de inputfaktorer (material, arbetskraft och kapital) som används i byggnadsproduktionen eller vissa mellanprodukter. I båda dessa fall måste man göra kompletterande beräkningar av produktivitetsförändringens effekt på slutproduktens pris eller kostnad.

Typhusmetoden innebär, att beräkningar baseras på uppgifter avseende ett eller flera utvalda typhus. Den motsvarar det förfarande som i allmänhet tillämpas vid konstruktion av prisindexar på andra områden, där t. ex. en varugrupp får representeras av ett fåtal representantvaror.

Kvalitetsklassmetoden innebär, att man korsklassificerar det föreliggande observationsmaterialet med utgångspunkt i för kvalitetsvärderingen väsentliga variabler, t. ex. lägenhetsytans storlek, det genomsnittliga antalet rum per lägenhet, ytan för andra lokaler i huset m. m. För varje cell som erhålls vid korsklassificeringen beräknas indextal såsom kvoten mellan de två situationernas genomsnittspriser för de bostadshus som tillhör cellen. Dessa indextal sammanvägs sedan till delindexar och totalindexar.

Kvalitetsselementmetoden innebär, att man specificerar ett erforderligt antal kvalitetskaraktéristika, vilka man på något sätt värderar, dvs. för vilka man bestämmer ett värde i kronor per enhet (per m², per styck etc). De enhetsvärden vilka utnyttjas för beräkning av låneunderlag och pantvärde i samband med bostadslångivningen är exempel på värderingar av detta slag.

Regressionsmetoden innebär, att man genom en statistisk skattningsmetod utviner de för beräkningarna erforderliga värderingarna ur själva det material för vilket prisindexberäkningar skall utföras. Byggnadspriset förutsattes kunna uttryckas som en enkel funktion av olika för indexberäkningarna relevanta kvalitetsegenskaper hos huset. Dessa antas därvid kunna uttryckas numeriskt. Funktionen skattas med konventionella statistiska metoder, varefter indextal beräknas med utgångspunkt i denna.

Faktorprismetoden. Faktorprisindex har i stor utsträckning kommit att direkt användas såsom ersättare för en byggnadsprisindex. Som framhållits leder detta i en tidsindex till ett systematiskt fel, eftersom man försummar effekten på prisutvecklingen av den produktivitetsförändring som sker. Om man fristående kunde beräkna produktivitetsutvecklingen, skulle det i princip vara möjligt att skatta en byggnadskostnadsindex genom att utgå från en faktorprisindex och korrigera denna med hänsyn till produktivitetsutvecklingen. För att erhålla en byggnadsprisindex skulle man därutöver behöva ha information om och korrigera för vinstutvecklingen. Skall faktorprismetoden vara framkomlig förutsätter det att man kan göra acceptabla förenklande antaganden beträffande produktivitets- och vinstutvecklingen.

Kostnads-komponentmetoden innebär, att man genomför prisindexberäkningar icke direkt för de ursprungliga produktionsfaktorerna utan för på lämpligt sätt avgränsande kostnads-komponenter i det färdiga huset. En sådan metod tillämpas i Västtyskland.

På grundval av denna inledande genomgång av olika principiellt möjliga metoder kommer kommittén till slutsatsen, att kvalitetselementmetoden och regressionsmetoden eller kombinationer av dessa inbördes och med kvalitetsklassmetoden bäst svarar mot de föreliggande kraven på en byggnadsprisindex och är både möjliga och värda att i första hand prövas.

A2: Förklarande variabler enligt Byggnadsindexkommittén (SOU 1971:79)

1.3 Sammanfattning av utredningen

1.3.6 Förslag till byggnadsindexar för nyproduktion av bostäder (kapitel 6)

1.3.6.3 Variabelurval för byggnadsprisindex

I avsnitt 6.4.3 med komplettering i bilaga 8 redovisas detaljerade listor - en för flerfamiljshus och en för gruppbyggda småhus - över de förklarande variabler som föreslås bli använda i regressionsberäkningarna. Förslaget omfattar för flerfamiljshus följande grupper av variabler.

- i) Utrymmes- och planlösningsstandard
- ii) Utrustningsstandard
- iii) Omslutningsytor, drifts- och underhållsekonomi
- iv) Anpassbarhet
- v) Tomtstandard
- vi) Bilplatser
- vii) Övriga variabler

Den sistnämnda gruppen består huvudsakligen av klassificeringsvariabler med underrubrikerna.

- a) Geografisk belägenhet
- b) Projektbelägenhet
- c) Projektstorlek
- d) Hustyp
- e) Förekomst av paining samt olika grundläggningssätt
- f) Byggherrekategori
- g) Platsbygge-monteringsbygge
- h) Entreprenadform, upphandlingsform, prisbestämningssätt

Variabelurvalet för småhus överensstämmer i huvuddragen med variabelurvalet för flerfamiljshus.

A3: Några citat ur Lind, H. och Song, H-S. (2011)

[Sidan 3] Den viktigaste slutsatsen i rapporten är att det finns systematiska fel i de index som används för omräkning av löpande priser till fasta priser i byggindustrin. Dessa fel leder till att prisutvecklingen överskattas, vilket i sin tur innebär att den reala produktivitetens utvecklingen underskattas. Hur mycket produktivitetens utvecklingen underskattats går inte att veta eftersom det inte samlas in data som skulle göra en sådan omräkning möjlig.

[Sidan 3] När det gäller *Byggnadsprisindex* för bostäder är problemet att indexet bara tar hänsyn till ett fåtal kvalitetsdimensioner och att indexet alltså inte fångar upp alla kvalitetsförändringar.

[Sidan 21] Det finns i princip två alternativ till den hedoniska metoden när det gäller byggindustrin: En "direkt bedömningsmetod" och en "typhusmetod".

[Sidan 21] Typhusmetoden innebär att en grupp av experter varje år får bedöma hur mycket ett specifikt typhus skulle kosta att producera. Därmed kan kostnadsutvecklingen vid given kvalitet fås fram. Om det verkliga priset stigit mer än kostnaden för att bygga typhuset tolkas det som att kvaliteten på de byggda husen stigit i relation till typhuset.

[Sidan 21] Grunddata kommer från enkäter till byggherrar. SCB får via olika kanaler information om nya projekt och skickar en enkät till byggherren. Det kan noteras att:

- Kostnader inrapporteras i förväg: enkättidpunkten är när bottenplattan är klar.
- Bland kostnader finns egna kostnader som byggherrekostnader och vinst som förefaller något subjektiva.

Även om detta kan ge fel i "nivån" på kostnaderna i enskilda fall och kanske en viss bias nedåt i tider med stigande priser, och omvänt vid sjunkande priser, så är det svårt att tro att det faktum att man mäter i förväg och inkluderar svärmätta egna kostnader, skulle påverka trenden. Eftersom det inte finns något som tyder på en sådan trend, dvs att dessa fel skulle bli större och större över tiden, tar vi inte vidare upp dessa omständigheter.

[Sidan 23] Det som t ex inte kontrolleras för är:

- andra kvalitetsaspekter hos lägenheten (typ utrustning i kök och badrum, golvmaterial etc)
- andra kvalitetsrelaterade aspekter på huset (t ex låg energiförbrukning, gröna byggnadsmaterial, utrymmen för sopsortering.)
- samhällsrelaterade kvalitetsaspekter (t ex om lägenheten är mer handikappanpassad, om huset byggts med bättre säkerhetsregler).

[Sidan 24] Slutsatsen av detta är rimligen att dagens byggnadsprisindex inte på något sätt kontrollerar för kvalitetsförändringar och därmed ger den ingen information om den "rena" prisutvecklingen. *Att jämföra detta index med t ex KPI och säga att priserna stigit snabbare i byggindustrin är därmed felaktigt eftersom man inte vet i hur hög grad de stigande priserna beror på att produkten har högre kvalitet.* Vi har inga data om den kvalitetsjusterade prisutvecklingen och om vi tror att kvaliteten i dagens hus är bättre i åtminstone några av de dimensioner som nämnts ovan jämfört med hus byggda för t ex 10 eller 15 år sedan så innebär det att byggnadsprisindex överskattar den "sanna" prisutvecklingen.

A4: Några citat ur Borg, L. et al. (2013)

[Sidan 2] Det finns en rad metoder för att skilja ut prisförändringar som beror av allmän inflation och vad som beror av kvalitetsförändringar, t ex

1. Expertmetod: En grupp experter får helt enkelt bedöma hur mycket som kvaliteten i genomsnitt förändras under året och hur stor den rena prisförändringen är.
2. Pris/kostnad för typobjekt (typhus): Här ska pris/kostnad för exakt samma produkt bedömas för de båda åren. Även här spelar expertbedömningar en stor roll.
3. "Hedonisk" metod: Detta innebär att ett statistiskt samband mellan kvalitet och pris tas fram utifrån observerade priser och egenskaper. Med hjälp av detta samband kan man uppskatta vad "förra årets hus" skulle kosta idag och därmed rensa bort den rena inflationen från mätningen.

[Sidan 4] Fyra olika kvalitetsdimensioner (Lind & Song 2012) kan identifieras, även om det finns vissa överlappningar mellan dem.

1. Kvalitet i själva produktionsprocessen, t ex metoder som minskar risk för olyckor. Ett bra exempel på detta är mer omfattande skyddsanordningar i samband med underhållsinsatser på vägar.
2. Kvalitet ur ett samhällsperspektiv, dvs egenskapsförändringar som beror av krav från samhället (stat och kommun). Ett exempel är t ex krav på bättre tillgänglighet för personer med funktionshinder men krav på energihushållning som går längre än vad som är företagsekonomiskt lönsamt.
3. Kvalitet ur ett livscykelperspektiv. En fastighetsägare kan välja en lite dyrare konstruktion för att på det sättet minska framtida drift- och underhållskostnader. Stigande energipriser kan t ex motivera installation av mer avancerade värmeåtervinningssystem och att man bygger med bättre isolering.
4. Kvalitet ut ett direkt kundperspektiv, vilket kan handla om att från parkettgolv i hela lägenheten, helkaklade badrum men också att man bygger viltstaket på längre sträckor av vägnätet.

[Sidan 10] Jämför vi utvecklingen av byggkostnadsindex med utvecklingen av konsumentprisindex framkommer att byggnadskostnadsindex stigit med ungefär 2 procent mer per år än konsumentprisindex. Av dessa 2 procent skillnad skulle alltså enligt beräkningarna ovan knappt hälften bero på felmätningar av den "rena" prisutvecklingen eftersom man inte tagit hänsyn till kvalitetsförändringarna.

A5: Ribe, M. (2014). Regressionsmetoden i Byggnadsprisindex (BPI), Utkast 2 till dokumentation, SCB

Detta dokument beskriver och kommenterar principerna för regressionsmetoden som tillämpas för Byggnadsprisindex (BPI), enligt läget i produktionen för redovisningsåret 2014. Dokumentet avses komplettera övrig dokumentation och beskrivning för BPI från SCB (2014).

Kedjeindex med årliga länkar

BPI beräknas som ett kedjeindex med årliga länkar, som kedjas till en serie indextal med indexbasår 1968=100.

Detta dokument beskriver principerna för beräkningen av de årliga indexlänkarna P_t som avser prisutvecklingen från år $t - 1$ till år t , med tonvikt på momentet med kvalitetsjustering genom hedonisk regression.

Metoden i den form som beskrivs här har tillämpats sedan redovisningsåret 2006. Dessförinnan tillämpades liknande men inte fullt jämförbara metoder.

Beräkningar som tillämpas för årslänk till BPI

Indexlänken P_t för prisutvecklingen från år $t - 1$ till år t beräknas som kvoten mellan två indextal V_t och K_t enligt:

$$(1) \quad P_t = \frac{V_t}{K_t}$$

där

t = Aktuellt år (redovisningsåret)

P_t = Prisindexlänken

V_t = "Värdekomponenten", index för prisutveckling enligt nedan

K_t = "Kvalitetskomponenten", index för kvalitetsutveckling enligt nedan.

Indextalen V_t och K_t som används i beräkningen enligt formel (1) beräknas i sin tur enligt följande. Dessa beräkningar görs utifrån byggnadsprojekt som observationsobjekt.

Värdekomponenten V_t är ett enhetsvärdeindex för priset per kvadratmeter beräknat enligt:

$$(2) \quad V_t = \frac{\bar{y}_t}{\bar{y}_{t-1}}$$

där

t = Aktuellt år

$t-1$ = Året före aktuellt år

\bar{y}_t = Medelpriset per kvadratmeter i projekten aktuellt år

\bar{y}_{t-1} = Medelpriset per kvadratmeter i projekten året före aktuellt år

Form av medelvärde. Överstrecksymbolen i uttrycken \bar{y}_t och \bar{y}_{t-1} anger beräkning av medelvärde per kvadratmeter av respektive pris. För småhus avses därvid kvadratmeter bostadsarea, och för flerbostadshus avses lägenhetsarea. Medelpriserna per kvadratmeter beräknas genom att motsvarande medelpriser per kvadratmeter inom byggnadsprojekten vägs ihop med projektens motsvarande areor.

Kvalitetskomponenten K_t har för sin beräkning utgångspunkten i en regressionssekvation (modell) av formen:

$$(3) \quad \ln Y = a + b_1 x_1 + b_2 x_2 + \dots + b_k x_k + \varepsilon$$

där

Y = Priset för ett projekt *per kvadratmeter*, den beroende variabeln

a = Intercept

b_i = Regressionskoefficienter

x_i = Kvalitetsvariabler

ε = Residual (restterm)

Regressionskoefficienterna skattas genom att regressionssekvationen (3) anpassas till observerade data för år $t - 1$. Utifrån de skattade regressionskoefficienterna räknas kvalitetskomponentens indextal fram enligt:

$$(4) \quad K_t = \frac{\exp(b_{1,t-1} \bar{x}_{1,t} + b_{2,t-1} \bar{x}_{2,t} + \dots + b_{k,t-1} \bar{x}_{k,t})}{\exp(b_{1,t-1} \bar{x}_{1,t-1} + b_{2,t-1} \bar{x}_{2,t-1} + \dots + b_{k,t-1} \bar{x}_{k,t-1})}$$

Överstrecksymbolen i uttrycken av formen $\bar{x}_{i,s}$ anger medelvärde per kvadratmeter bostads- resp. lägenhetsarea, analogt med behandlingen av prisvariabeln i formel (2) enligt förklaringarna till denna. (Interceptet a_{t-1} behövs inte i formel (4), eftersom faktorn $e^{a_{t-1}}$ kan brytas ut ur täljare och nämnare och förkortas bort.)

Vid beräkningen av K_t används alltid regressionskoefficienterna för perioden före det aktuella året. Detta innebär att regressionskoefficienterna byts ut en gång per år. Ovanstående beskrivning motsvarar den av Walestad (2007).

Terminologi. Ett par potentiellt tvetydiga uttryck: Uttrycket "reference period" kan i statistikredovisning betyda aktuell period och i indexsammanhang betyda basperiod. Ordet "modell" kan i samband med prisstatistik betyda produktvariant (såsom husmodell) och i statistisk metodik betyda regressionssekvation.

Nuvarande metod i relation till ursprungligt förslag från Byggnadsindexkommittén (SOU 1971:79) – med kommentarer

Den ovan beskrivna metoden följer till grundidé och syfte det ursprungliga förslaget från Byggnadsindexkommittén (SOU1971:79), den statliga offentliga utredning om BPI som presenterades 1971. Båda metoderna bygger på en *hedonisk* ansats med regressionsanalys

för att justera för skillnader i kvalitet. I en del detaljer skiljer sig dock den nuvarande metoden från den nämnda utredningens ursprungliga förslag.

Indexformel och regressionsekvation enligt Byggnadsindexkommittén (SOU 1971:79)

Regressionsekvationen (3) motsvaras i det ursprungliga förslaget av följande regressionsekvation:

$$(5) \quad Y' = a + b'_1 x_1 + b'_2 x_2 + \dots + b'_k x_k + \varepsilon'$$

där

Y' = Priset för ett projekt *per lägenhet*, den beroende variabeln

och övriga beteckningar analogs med dem i formel (3). Den nuvarande indexberäkningen genom formel (1), (2) och (4) motsvaras i det ursprungliga förslaget av den enda formeln:

$$(6) \quad P_t = \frac{a'_t + b'_{1,t} \bar{x}_{1,t-1} + b'_{2,t} \bar{x}_{2,t-1} + \dots + b'_{k,t} \bar{x}_{k,t-1}}{a'_{t-1} + b'_{1,t-1} \bar{x}_{1,t-1} + b'_{2,t-1} \bar{x}_{2,t-1} + \dots + b'_{k,t-1} \bar{x}_{k,t-1}}$$

som direkt ger indexlänken ur regressionskoefficienterna från regressionsekvationen (5).

Av beskrivningarna ovan följer att skillnader mellan den nuvarande metoden och det ursprungliga förslaget föreligger i väsentligen dessa detaljer:

- i. Annan beroende variabel – nu pris per kvadratmeter, inte per lägenhet
- ii. Annan form av regressionsekvation – nu med logaritmerat pris
- iii. Annan form av hedonisk metod – inom indexlänken behålls nu regressionskoefficienternas, och inte kvalitetsvariablernas, värden.

Dessa punkter utvecklas i de följande underavsnitten.

Annan prisvariabel – kommentar

Nuvarande metoden utgår från pris per kvadratmeter.

Ursprungliga förslaget utgår från pris per lägenhet.

Förklaring. Att utgå från pris per kvadratmeter som nu kan ses om ett sätt att effektivt utnyttja hjälpinformationen i prisets naturliga begreppsmässiga koppling till arean. Genom att denna koppling hålls utanför regressionsanalysen på det nämnda sättet så undviks att den stör skattningen av de viktiga regressionskoefficienterna för kvalitetsvariablerna. Sådana störningar skulle lätt kunna uppstå genom små avvikelser i data från regressionsekvationens linjaritet i högra ledet. Förklaringsgraden (R-kvadrat) är inte jämförbar mellan de två alternativen och ger därför inte relevant ledning i valet mellan dessa.

Annan form av regressionsekvation – kommentar

Nuvarande metoden utgår från en ”loglinjär” eller ”semilogaritmisk” regressionsekvation (3) där prisvariabeln i vänstra ledet är logaritmerad.

Ursprungliga förslaget utgår från en rent linjär regressionsekvation (5), utan någon transformering av prisvariabeln i vänstra ledet.

Förklaring. Semilogaritmisk hedonisk regressionsekvation är numera allmänt en vanlig praxis, av goda skäl (jfr Destatis, 2009; Triplett, 2006). Regressionskoefficienterna (b_i) i den semilogaritmiska ekvationen och de (b'_i) i den rent linjära har olika slags innebörd.

Talen b_i har sin tolkning i att de i exponentierad form e^{b_i} anger proportioner med vilka priset ändras per enhets ökning i kvalitetsvariabeln. Talen b'_i däremot anger belopp i kronor för prisets ökning per enhets ökning i kvalitetsvariabeln.

Metoden med semilogaritmisk ekvation är i en mening mera robust mot inflation än den med rent linjär ekvation. Om alla priser skulle multipliceras med samma faktor (i andra ord förändras med samma procentsats) efter att koefficienterna skattades, så överensstämmer ändå koefficienternas (utom interceptets) skattade värden med dem som skulle fås om de skattades om på aktuella priser. Samma egenskap medför också mindre fara för systematiskt fel genom att skatta koefficienterna på data från en tidigare period.

Annan form av hedonisk metod – kommentar

Nuvarande metoden bygger på en regressionsanalys på länkasperiodens priser, för att justera bort kvalitetsutvecklingen ur ett enhetsvärdeindex. Metoden är en form av *hedonic re-pricing* (hedonisk prisjustering) enligt Destatis (2009, Sect. 7.1, Annex 1).

Ursprungliga förslaget bygger på två separata regressionsanalyser som avser dels länkasperioden, dels aktuell period, för att värdera en fast korg av kvalitetsegenskaper. Metoden är en form av *hedonic characteristics price index method* enligt Destatis (2009, Sect. 7.1, Annex 1).

En motsvarande skillnad finns i vilka antaganden om trögt föränderliga faktorer som metoderna utgår från:

Nuvarande metoden utgår från antagandet att värderingen av kvalitetsegenskaper, i relativa tal, förändras trögt.

Ursprungliga förslaget utgår från antagandet att (proportionerna i) korgen av kvalitetsegenskaper förändras trögt.

Förklaring. Formulerat på detta sätt framstår antagandet bakom nuvarande metoden som mer plausibelt för BPI jämfört med det motsvarande antagandet bakom det ursprungliga förslaget. Byggtekniken och synen på byggnaders kvalitet kan förändras mycket på lång sikt, men knappast mycket i korta perspektiv. Däremot kan mixen av kvalitetsegenskaper förändras snabbt, beroende på vilka slags byggprojekt som är aktuella för tillfället och var de är lokaliserade.

Utvikning. För *fastighetsprisindex*, avseende befintliga byggnader, kan däremot *hedonic characteristics price index method* vara lämplig, eftersom stocken av befintliga byggnader är trögt föränderlig.

Indexmetoderna enligt Laspeyres och Paasche

I den form som beskrevs ovan är i en mening båda metoderna av "Laspeyres-typ", genom att dels regressionskoefficienterna i formel (4) för nuvarande metoden, dels kvalitetsvariabelvärdena i formel (6) för ursprungliga förslaget, avser länkbaserperioden $t - 1$. Motsvarande "Paasche-typ" är teoretiskt möjlig, med t istället för $t - 1$ på de nämnda ställena. Byggnadsindexkommittén berör detta, och även ytterligare varianter är teoretiskt sett möjliga för indexberäkningen. I den nuvarande metoden med semilogaritmisk regressionsekvation vore knappast så mycket vunnet på att frånga Laspeyrestypen.

Kommentar om principer bakom variabelval

Förklarande variabler ska fånga kvalitet

De förklarande variablerna ska beskriva byggnadernas kvalitet. De ska vara valda med den utgångspunkten, inte för att kunna förklara prisvariationerna optimalt. Egenskaper som beror av mode ska inte vara med, eftersom modet kan växla fram och tillbaka. Det skulle då leda fel att behandla förändringar åt alla håll som bestående kvalitetsförbättringar.

Kvalitet i byggherre- respektive producentperspektiv

Som utvecklas av Byggnadsindexkommittén (SOU 1971:79) så kan byggherren, köparen, respektive byggproducenten, säljaren, ha olika perspektiv på vilka slags kvalitetsvariabler som är viktiga för nyproducerade bostäder.

För byggherren och ytterst de boende är kvalitet en fråga om nytta och funktionalitet i användningen av bostäderna, men för producenten är kvalitet sådana produkttegenskaper som går att ta betalt för. Begreppsmässigt är detta inte utan vidare samma sak och kunde potentiellt motivera olika val av kvalitetsvariabler.

BPI är ett producentprisindex och skulle därför idealt kvalitetsjustera ur ett producentperspektiv. I praktiken kan det dock vara vanskligt att helt hålla isär de två perspektiven. Som visas av Gordon (1990) och även berörs av Byggnadsindexkommittén (SOU 1971:79) så tenderar också de två perspektiven teoretiskt att resultera i likartade kvalitetsvärderingar under vissa förutsättningar om jämvikt och fungerande marknadsmekanismer. I någon mån kan det därför vara försvarligt att ta olika slags variabler som proxy för varandra.

Avvägning av antalet kvalitetsvariabler

Kvalitetsvariablerna får inte vara alltför många, för att inte medföra för stor osäkerhet; se Destatis (2009, Sect. 7.1, Annex 2).

Prisvariabel som kvalitetsvariabel för lägesfaktorn

En av kvalitetsvariablerna avser genomsnittligt fastighetspris per kvadratmeter i kommunen (jfr Bilaga 1). Detta är ett sätt att beakta den viktiga lägesfaktorn. Läget är allmänt en kanske viktigaste faktor bakom bostadspriser, och detta slår igenom även i priset exklusive tomtmarken på bostäder i nyproducerade hus. Byggnadsprojekten åren $t - 1$ respektive t kan vara mer eller mindre markant olika belägna, och det gör det väsentligt att i kvalitetsjusteringen för indexlänken beakta dessa möjliga skillnader i läget.

Att på så sätt ha med en prisberoende variabel som kvalitetsvariabel kan verka något okonventionellt. Annars brukar kvalitetsvariablerna i hedonisk regression allmänt väljas så att de mäter fysiska karaktäristika. Den nämnda prisberoende kvalitetsvariabeln kan dock knappast leda till några anomalier här, och den verkar vara ett effektivt sätt att få med lägesfaktorn.

Kommentar om principer bakom val av vägningsalternativ

I skattningen av regressionskoefficienterna i formel (3) tillämpas inte någon vägning av observationerna, byggnadsprojekten. Byggnadsprojekten väger därmed sinsemellan lika tungt i skattningen av regressionskoefficienterna. Detta motiveras med att regressionskoefficienterna då blir väntevärdesriktigt och effektivt skattade, enligt Gauss-Markovs sats (se t.ex. Wilks, 1962, Sect. 10.3, eller Wikipedia, 2014), under antagandet att regressionskvationen är sann. Resultat av Nordberg (1989) visar att slutsatsen gäller även vid vissa slags avvikelser från antagandena bakom regressionskvationen.

Teoretiskt skulle en eventuell modellavvikelse i form av heteroskedasticitet kunna kompenseras med en för detta särskilt utformad vägning. Men i praktiken tillkommer då osäkerhet i beräkningen av vikterna, vilket mer eller mindre kan förta en eventuell potentiell vinst i noggrannhet, samtidigt som komplexiteten ökar. Vägning i proportion mot area eller pris kunde ge dålig precision genom att låta stora projekt dominera för mycket, så att det effektiva antalet observationer minskar, och vore knappast lämplig. Emellertid, om det skulle vara problem med att små projekt kan avvika markant från det allmänna mönstret, så kunde det vara tänkbart att vid en metodöversyn eventuellt införa någon enkel regel som reducerar vikten för de minsta projekten, medan vikterna annars får förbli lika.

Referenser

- Byggnadsindexkommittén (SOU 1971:79), Byggnadsindex för bostäder, Delbetänkande, Statens offentliga utredningar.
- Destatis (2009), Handbook on the application of quality adjustment methods in the Harmonised Index of Consumer Prices, Statistics and Science Vol. 13, Wiesbaden: Federal Statistical Office of Germany.
- Gordon, R.J. (1990), The Measurement of Durable Goods Prices, Chicago: Chicago University Press.
- Nordberg, L. (1989), Generalized linear modeling of sample survey data, Journal of Official Statistics, 5, 223-239.
- SCB (2014), Byggnadsprisindex 2012, PR0501, SCBDOK 3.2.
- Triplett, J. (2006), Handbook on Quality Adjustment of Price Indexes for Information and Communication Technology Products, OECD Directorate for Science, Technology and Industry, Paris: OECD.
- Walestad, M. (2007), Översyn av Byggnadsprisindex, Slutrapport för Draft Grant Agreement, SCB.
- Wikipedia (2014) English, Gauss-Markov theorem. [Läst 2014-12-11.]
- Wilks, S.S. (1962), Mathematical Statistics, New York: Wiley.

Bilaga 1: Variabler

Nedan listas de variabler som används i regressionsberäkningarna. Efter Walestad (2007).

Prisvariabel

Y-variabeln definieras som det pris som byggherren får betala för byggnaden i ett bostadsprojekt. Priset för tomten m.m. ingår inte. I priset ska vara inräknat allt sådant som:

- Material och arbetskraft som används för att bygga huset (entreprenadkontraktet)
- Byggherrekostnaden
- Projektering
- Anslutningsavgifter
- Moms
- Vinst eller förlust

Det är det logaritmerade priset per kvadratmeter som används i regressionssekvationen. Kvadratmeter ska då för flerbostadshus avse lägenhetsarea och för gruppbyggda småhus avse bostadsarea.

Kvalitetsvariabler för Flerbostadshus

- Fastighetspriser = Genomsnittliga fastighetspriser per kvadratmeter värdearea per kommun. Värdearea är $BOA + 20$ procent av BIA. BOA=Bostadsarea och BIA= Biarea.
- Upplåtelseform br = Upplåtelseform bostadsrätt (hyresrätt default)
- Yteffektivitet = $((BOA+LOA)/Bruttoarea (BTA))$
- Hygienrum = Antal hygienrum per lägenhet. Ett hygienrum med en toalett och handfat har värdet 0,5, ett hygienrum med en toalett och handfat + ett hygienrum med dusch/bad har värdet 1,5 och två hygienrum med dusch/bad har värdet 2.
- Husform 2 = Husform loftgångshus kod 3
- Husform 3 = Husform punkthus kod 4
- Husform 4 = Husform övrigt (sammanbyggda tvåbostadshus kod 5, terrasshus kod 6, annan typ kod 7, stjärnhus kod 8, icke-bostadshus med lägenheter kod 9, punkthus med 2 våningar kod 10, tvåbostadshus sammanbyggda med enbostadshus kod 11) och Husform 1 default
- Källare = Källare finns/finns ej
- Balkong = Antal balkonger per lägenhet
- Hiss = Hiss finns/finns ej
- Garage = Antal garage eller parkeringsdäck per lägenhet

Referensvärden:

- Upplåtelseform hr = Upplåtelseform hyresrätt
- Husform 1 = Husform lamellhus kod 2

Kvalitetsvariabler för Guppbyggda småhus

- Bostadsarea = Summa bostadsarea/summa antal lägenheter
- Fastighetspriser = Genomsnittliga fastighetspriser per kvadratmeter värdearea per kommun. Värdearea är BOA + 20 procent av BIA.
- Hygienrum = Antal hygienrum per lägenhet. Ett hygienrum med en toalett och handfat har värdet 0,5, ett hygienrum med en toalett och handfat + ett hygienrum med dusch/bad har värdet 1,5 och två hygienrum med dusch/bad har värdet 2.
- Husform 2 = Kedjehus kod 3
- Husform 3 = Radhus kod 2, parhus kod 4, friliggande tvåbostadshus kod 5, annan typ kod 6, fyrklöverhus kod 7, enbostadshus sammanbyggda med tvåbostadshus kod 8 (Husform 1 default)
- Källare = Källare finns/finns ej
- Fasadmaterial 2 = Betong (enbart) kod 2, gasbetong (enbart) kod 3, tegel (enbart) kod 4, kalksandsten (enbart) kod 5, flera material (övervägande betong) kod 20, flera material (övervägande gasbetong) kod 30, flera material (ej trä, puts eller plåt övervägande tegel) kod 40, tegel+trä (övervägande tegel) kod 41, tegel+puts (övervägande tegel) kod 46, tegel+plåt (övervägande tegel) kod 48, flera material (ej trä, puts eller plåt övervägande kalksandsten) kod 50, kalksandsten+trä (övervägande kalksandsten) kod 51, kalksandsten+puts (övervägande kalksandsten) kod 56, kalksandsten+plåt (övervägande kalksandsten) kod 58
- Fasadmaterial 3 = Puts (enbart) kod 6, flera material (ej trä övervägande puts) kod 60, puts+trä (övervägande puts) kod 61, plåt (enbart) kod 8, flera material (övervägande plåt) kod 80, flera material (övervägande annat) kod 90, uppgift saknas kod 99, fasadmaterial (Fasadmaterial 1 är default)
- Takmaterial 2 = Lertegel kod 3
- Takmaterial 3 = Plåt kod 2, papp kod 4, annat kod 5, uppgift saknas kod 9 (Takmaterial 1 är default)
- Vind = Inredningsbar vind finns/finns ej
- Antal plan 2 = Sluttningshus kod 4, sluttnings- + ½-planshus kod 8, sluttnings- + enplanshus kod 9 1½-plans hus kod 2, tvåplanshus kod 3 och 2½-plans hus kod 7, annat hus kod 5, etagehus kod 6, (Antal plan 1 default)
- Upplåtelseform br = Upplåtelseform bostadsrätt
- Upplåtelseform hr = Upplåtelseform hyresrätt (äganderätt default)
- Garage = Antal garage eller parkeringsdäck per lägenhet

Referensvärden:

- Husform 1 = Friliggande villa kod 1
- Fasadmaterial 1 = Trä (enbart) kod 1, trä+tegel (övervägande trä) kod 14, trä+kalksandsten (övervägande trä) kod 15, trä+puts (övervägande trä) kod 16
- Takmaterial 1 = Betongtakpannor kod 1
- Antal plan 1 = Enplanshus kod 1
- Upplåtelseform ar = Upplåtelseform äganderätt