

R&D Report
Research - Methods - Development
1998:1



Statistics Sweden

Statistiska centralbyrån

Nybyggnadsstatistik - en simuleringsstudie

Catarina Elffors

R&D Report
Research - Methods - Development 1998:1

Preliminär statistik: nybyggnadskostnader -en simuleringsstudie
(Statistics of new constructions - A simulation study based on preliminary data)

Statistics Sweden
1998

Från trycket	Februari 1998
Ansvarig utgivare	Lars Lyberg
Producent	Statistiska centralbyrån, utvecklingsavdelningen
Förfrågningar	Catarina Elffors, Statistiska centralbyrån BoR/BY-S, Box 24300, Box 24300, S-104 51 STOCKHOLM telefon 08-783 47 98 telefax 08-783 49 05

© 1998, Statistiska centralbyrån

ISSN 0283-8680
Printed in Sweden
SCB-Tryck, Örebro, 1998

INLEDNING

TILL

R & D report : research, methods, development / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1988-2004. – Nr. 1988:1-2004:2.

Häri ingår Abstracts : sammanfattningar av metodrapporter från SCB med egen numrering.

Föregångare:

Metodinformation : preliminär rapport från Statistiska centralbyrån. – Stockholm : Statistiska centralbyrån. – 1984-1986. – Nr 1984:1-1986:8.

U/ADB / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1986-1987. – Nr E24-E26

R & D report : research, methods, development, U/STM / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1987. – Nr 29-41.

Efterföljare:

Research and development : methodology reports from Statistics Sweden. – Stockholm : Statistiska centralbyrån. – 2006-. – Nr 2006:1-.

R&D Report
Research - Methods - Development
1998:1

Nybyggnadsstatistik - en simuleringsstudie

Catarina Elffors

Innehåll

1	Bakgrund	1
2	Population	1
3	Uppskattning av populationsstorleken	2
4	Simuleringsstudien	4
5	Resultat av simuleringsstudien på 1995 och 1996 års material	6
6	Slutsatser	8
6.1	Sammanfattning	8
6.2	Redovisning av preliminär statistik	9
6.3	Information om användning av statistiken	12
7	Referens	12
Bilaga 1	Formler	13
Bilaga 2	Kortaste och längsta konfidensintervall för given bortfallsstorlek och år	15

Preliminär statistik: nybyggnadskostnader

En simuleringsstudie

I denna rapport redovisas en simuleringsstudie av hur bortfall kan påverka några vanligen förekommande punktskattningar i nybyggnadskostnadsstatistiken. Simuleringsstudien belyser endast hur ett objektsbortfall (ännu ej inkomna enkäter) påverkar skattningarna. Det främsta syftet med simuleringsstudien är att ge en vägledning till om och när man kan redovisa preliminär statistik.

1 Bakgrund

Dataunderlaget till nybyggnadskostnadsstatistiken insamlas på två olika sätt, dels ett administrativt material (1992 och 1993 års regler), dels en enkät till byggherrarna (1993 års regler)¹. För objekt med beslut enligt 1992 års regler hämtas samtliga uppgifter till statistiken från det administrativa materialet. Det administrativa materialet för objekt med beslut enligt 1993 års regler är begränsat. Det är endast ett fåtal av uppgifterna som används. Merparten av uppgifterna för objekt med beslut enligt 1993 års regler hämtas från enkäten till byggherrarna.

Nybyggnadsprojekt med 1992 års regler blir allt mer sällsynta. På grund av en regeländring är det administrativa materialet troligen helt borta 1998. 1995 utgjorde det administrativa materialet 25% av nybyggnadsprojekten och 1996 15%.

Referenstidpunkt för nybyggnadskostnadsstatistiken är *påbörjandetidpunkten*². Enkäter skickas ut (månadsvis) till byggherrar till aktuella projekt. Tidpunkten för *beslut om statligt stöd* används härvid som *indikator* på att ett nybyggnadsprojekt är aktuellt. Beslut om statligt stöd kan komma både före och efter påbörjandet. Under pågående år kan vi därför räkna med både över- och undertäckning, då ramen är ofullständig med avseende på påbörjandetidpunkten.

Det finns i princip tre orsaker till förseningarna. En orsak kan hänföras till *indikatorvariabeln*, tidpunkt för beslut om statligt stöd. Beslut om statligt stöd kan komma långt efter påbörjandetidpunkten. Förseningarna kan också bero på långa svarstider (många påminnelser) och ofullständigt ifyllda enkäter (ofta omfattande kompletteringsarbete). Alla tre ovan nämnda faktorer påverkar tillgången till bearbetningsbart material.

Efterfrågan på aktuell statistik över nybyggnadskostnader är stor, både vad avser punktskattningar och förändringsskattningar (Byggnadsprisindex). *Användningen* av statistiken ställer olika *kvalitetskrav* på det statistiska underlaget. Kraven kan vara så långtgående att man bör invänta definitivt material. Med rätt användning och måttliga kvalitetskrav kan även preliminära resultat används. Frågan är bara *hur* preliminärt dataunderlag man kan acceptera med tanke på sekretessregler och användning.

Statistik över nybyggnadskostnader publiceras en gång per år i Statistiska meddelanden Bo 22. Denna statistik bygger på definitivt material med inget eller enstaka objektsbortfall. Publiceringen sker cirka ett halvt år efter redovisningsåret. Preliminär statistik publiceras under året i Byggindex.

I Statistiska meddelanden Bo 22 finns en utförlig beskrivning av undersökningen om nybyggnadskostnader. Undersökningen finns även beskriven på SCB:s hemsida på Internet, textdatabasen.

2 Population

Dataunderlaget till nybyggnadskostnadsstatistiken bygger på uppgifter om nybyggda bostäder med statligt stöd. Det är både uppgifter om kostnader ("priser") och teknisk information om

¹ 1992 och 1993 års regler syftar på reglerna för statligt stöd för nybyggnad av bostäder.

² Påbörjande av byggnadsarbeten såsom gjutning av källargolv, källarmurar, bottenplatta etc.

projektet ("byggnaden") som samlas in. Kostnaderna indelas i byggnads- och markkostnader. Produktionskostnaden uttrycker den totala kostnaden, dvs den sammanlagda mark- och byggnadskostnaden.

I den statistiska redovisningen är kostnaderna normerade, dvs kostnaderna redovisas per kvadratmeter total primär bruksarea och per lägenhet. Både den totala kostnaden för hela byggprojekt och den normerade kostnaden varierar. Orsakerna till variationen är många. Det är inte självklart, om man ser till den normerade kostnaden, att stora projekt är billigare än små. Kostnaderna varierar också mellan år, men skillnader i kostnader mellan år behöver inte bero på en kostnadsförändring utan kan lika gärna förklaras av att byggnadsprojekten skiljer sig åt mellan år.

I tabellerna 1 och 2 redovisas den empiriska fördelningen för normerade kostnader för flerbostadshus och gruppbyggda småhus 1995 och 1996. Bearbetningen är utförd på det definitiva materialet.

Tabell 1 Empirisk fördelning av byggnadskostnad per kvadratmeter primär bruksarea 1995 och 1996. Flerbostadshus och gruppbyggda småhus. Avrundat till 100-tal kronor. Definitiva uppgifter.

År Hustyp	Antal objekt	Percentil								
		10	20	30	40	50	60	70	80	90
1995										
Flerbostadshus	79	7 300	8 100	8 400	9 000	9 400	9 900	10 500	11 300	12 200
Gruppbyggda småhus	270	6 100	7 000	7 600	7 900	8 200	8 400	9 000	9 500	10 300
1996										
Flerbostadshus	122	7 300	7 900	8 400	9 000	9 300	10 100	10 400	11 400	12 500
Gruppbyggda småhus	530	6 300	7 000	7 300	7 600	7 700	8 400	9 100	9 700	11 000

Tabell 2 Empirisk fördelning av produktionskostnad per kvadratmeter primär bruksarea 1995 och 1996. Flerbostadshus och gruppbyggda småhus. Avrundat till 100-tal kronor. Definitiva uppgifter.

År Hustyp	Antal objekt	Percentil								
		10	20	30	40	50	60	70	80	90
1995										
Flerbostadshus	79	8 100	9 000	9 500	10 100	10 500	11 200	11 900	12 500	13 500
Gruppbyggda småhus	270	7 700	8 800	9 700	10 000	10 400	10 600	11 100	11 800	12 700
1996										
Flerbostadshus	122	8 300	8 900	9 500	9 900	10 700	11 200	12 200	12 900	13 600
Gruppbyggda småhus	530	8 000	8 600	9 300	9 900	10 100	10 300	10 700	11 900	13 100

Variationen i materialet, som framgår av tabellerna 1 och 2 ovan, och det begränsade antalet nybyggnadsprojekt försvårar möjligheterna att redovisa preliminär statistik framför allt på detaljerad nivå.

3 Uppskattning av populationsstorleken

Det totala antalet objekt i undersökningspopulationen är okänt tills hela undersökningsperioden har förflutit.

Ramen för undersökningen utgörs av FENIX³ och ansökan om statligt stöd för bostadsbyggande. *Referenstidpunkten* är *påbörjandemånaden*. Inledningsvis används *beslutstidpunkten*

³ FENIX är egentligen ett beräkningssystem som Länsstyrelsen använder när de beräknar det statliga stödet för bostadsbyggande. De variabler som ingår i beräkningarna registreras härvid i en databas. Ägare till FENIX är Boverket.

som en *skattning av påbörjandetidpunkten*. Eftersom beslut om statligt stöd kan fattas både *före* och *efter* påbörjandet kan objekt preliminärt föras till felaktig redovisningsperiod under året. Uppgift om faktisk påbörjandetidpunkt erhålls först när enkäten besvarats. Om byggherren ej besvarat enkäten inom tre veckor skickas en *brevpåminnelse* ut. Efter ytterligare två⁴ veckor görs den första *telefonpåminnelsen*. I samband med denna kontakt erhålls vanligen en skattning av påbörjandetidpunkten.

Beslut efter påbörjandetidpunkten innebär temporär undertäckning under undersökningsperioden.

Under *ideala förhållanden* borde man kunna ha bearbetningsbara uppgifter i databasen *två* månader efter redovisningskvartalets utgång. Med ideala förhållanden avses att

- den uppskattade påbörjandetidpunkten i princip sammanfaller med beslutstidpunkten
- byggherren ofördröjligen fyller i enkäten
- få eller inga kompletteringar behövs

Redovisningsperioden bygger på påbörjandetidpunkt.

Det problem som *undertäckningen* kan förorsaka under pågående insamlingsperiod illustreras i tabellerna 3 och 4. Tabellerna visar tillgängligt dataunderlag under ideala förhållanden.

Tabell 3 Procentuell andel bearbetningsbara enkäter under ideala förhållanden 1995

Hustyp	Redovisningskvartal	Antal objekt	Procentuell andel bearbetningsbara enkäter under ideala förhållanden			
			1995 maj/juni	1995 aug/sept	1995 nov/dec	1996 feb/mars
Flerbostadshus	1995:I	25	60	80	80	100
	1995:II	13		31	69	100
	1995:III	17			82	100
	1995:IV	24				100
	1995:I-IV	79	24	39	61	100
Gruppbyggda småhus	1995:I	116	49	72	88	100
	1995:II	55		36	51	100
	1995:III	31			35	100
	1995:IV	68				100
	1995:I-IV	270	29	43	55	100

⁴ Den första telefonpåminnelsen görs numera redan en vecka efter det att brevpåminnelsen skickats ut. Förändringen infördes under senare delen av 1996.

Tabell 4 Procentuell andel bearbetningsbara enkäter under ideala förhållanden 1996

Hustyp	Redovisningskvartal	Antal objekt	Procentuell andel bearbetningsbara enkäter under ideala förhållanden			
			1996 maj/juni	1996 aug/sept	1996 nov/dec	1997 feb/mars
Flerbostadshus	1996:I	21	56	75	91	100
	1996:II	7		75	89	100
	1996:III	11			88	100
	1996:IV	6				100
	1996:I-IV	122	43	62	82	100
Gruppbyggda småhus	1996:I	56	33	74	92	100
	1996:II	82		71	88	100
	1996:III	12			64	100
	1996:IV	17				100
	1996:I-IV	530	30	67	86	100

Tabellerna 3 och 4 ger ett intryck av att undertäckningen är större under de två första kvartalen under ett år än under de senare. Detta är emellertid en synvilla och kan till viss del förklaras av att man "bryter" insamlingsåret vid månadsskiftet januari/februari året efter insamlingsåret. Tekniskt betyder detta att de objekt som finns i ramen vid månadsskiftet januari/februari året efter insamlingsåret och tillhör insamlingsåret utgör ramen (samtliga kända nybyggda bostadshus med statligt stöd). Datainsamlingen kan slutföras. Under april månad året efter insamlingsåret är det definitiva materialet klart.

4 Simuleringsstudien

Indikatorvariabeln (undertäckning), långa svarstider och ofullständigt ifyllda blanketter är de faktorer som påverkar tillgång till preliminärt dataunderlag.

De förseningar som uppkommer i enkätundersökningen genererar vid en given tidpunkt ett *objektsbortfall*. Det är dock viktigt att notera att objektsbortfallet endast är temporärt. Målsättningen är att få in alla enkäter så komplett ifyllda som möjligt, även om detta förorsakar många påminnelser och upprepade kontakter med uppgiftslämnarna. Det bör tilläggas att uppgiftsplikt gäller för denna undersökning.

Det kan, trots alla ansträngningar, förekomma partiellt bortfall. För att minimera det partiella bortfallet hämtas uppgifter från annat håll och som en sista utväg görs imputeringar med hjälp av uppgifter från liknande projekt. Denna undersökning behandlar dock inte det partiella bortfallet.

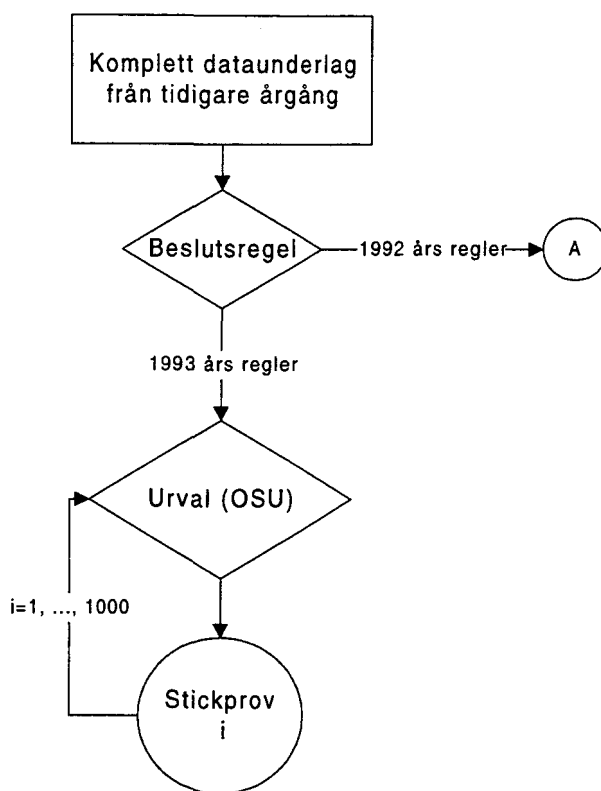
Simuleringsstudien utgår från en komplett ram⁵. De ofullständigheter i ramen som genererar undertäckning studeras inte här. Det är endast det temporära objektsbortfallet som studeras och vilka effekter det får om man gör bearbetningar på hela redovisningsperioden (år).

Bortfallet drabbar *hela projekt*, genom att enkäten inte besvaras i tid. Egentligen är det byggherren, dvs *uppgiftslämnaren*, som *genererar bortfallet*. Vi antar att bortfallet är procentuellt lika stort bland projekt med flerbostadshus som bland projekt med gruppbyggda småhus. Antagandet bygger på att vi för närvarande inte har någon anledning att tro att bortfallet skulle vara beroende av hustyp. Det finns heller inga kända indikationer på att bortfallet är selektivt, t ex drabbar viss typ av projekt. Därför utgår vi ifrån att bortfallet är *slumpmässigt*.

I undersökningen utnyttjas 1995 och 1996 års definitiva material. Dataunderlaget kommer från nybyggnadskostnadsregistret. Från *enkätundersökningen* simuleras⁶ för vardera året tre bort-

⁵ Särndal et al [1] Remark 2.11.3 och kapitel 7.9.1

fallsstorlekar, 20%, 30% respektive 40% bortfall. För given bortfallsstorlek simuleras 1 000 stickprov. Stickproven dras från samma population, dvs med återläggning *mellan* stickprov. Urvalen dras som obundet slumpmässiga urval utan återläggning för hustyperna flerbostadshus och gruppbyggda småhus.



Figur 1 Flöde för simulering

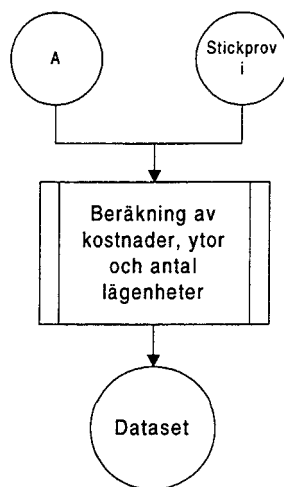
Det av urvalsdragningen resulterade *stickprovet* tillsammans med uppgifter från det *administrativa materialet* utgör ett *simulerat preliminärt material*.

Det genererade preliminära materialet skall illustrera en tänkbar bortfallssituation. Precis som under pågående undersökningsperiod är projekten av olika art, dvs både "stora" och "små". Sannolikheten att ett eller flera "extrema" projekt skall komma med i ett eller flera stickprov skall i princip svara mot sannolikheten att uppgiftslämnaren för ett dylikt projekt besvarar enkäten i tid.

Med hjälp av det preliminära materialet skattas några parametrar. Bearbetningarna utförs i simuleringsprogrammet enligt samma beräkningsprincip (se bilaga 1, formler) som gäller för den definitiva statistiken som redovisas i bl a SM Nybyggnadskostnader. Det är i nybyggnadsstatistiken centrala och efterfrågade parametrar såsom kostnader per ytenhet och lägenhet samt yta per lägenhet som skattas. Bearbetningar utförs för de mest efterfrågade redovisningsgrupperna.

⁶ Simuleringen utförs med hjälp av ett SAS-program som utarbetats av Catarina Elffors i samarbete med Anders Kraftling, SCB.

Variabler:	Redovisningsgrupper:
kostnader:	– hustyp
– byggnadskostnad	– region (6 st) ⁷
– produktionskostnad	– region (2 st) ⁸
– markkostnad	– byggherrekategori (flerbostadshus)
yta:	– upplåtelseform (gruppbyggda småhus)
primär bruksarea	– region och byggherrekategori (flerbostadshus)
antal lägenheter	– region och upplåtelseform (gruppbyggda småhus)



Figur 2 Flöde bearbetningar

5 Resultat av simuleringstudien på 1995 och 1996 års material

Simuleringsstudien ger en uppskattning av osäkerheten i skattningarna när man använder ett preliminärt material. Varje stickprov simulerar en realiserad bortfallssituation. Nivån på den skattade parametern kan vara rimlig, dvs överensstämna någorlunda med beräkningen på det definitiva materialet. Men man kan också finna stora avvikelser mellan skattningen på det preliminära materialet och en beräkning med hjälp av det definitiva materialet. Variabiliteten uttrycks både såsom variationskoefficienter och 95-procentiga konfidensintervall.

Simuleringen som redovisas i denna rapport är tidsödande. Själva simuleringen av 1 000 stickprov med åtföljande bearbetningar i varje stickprov tar cirka två timmar. Den efterföljande bearbetningen av framför allt variansskattningar (SAS-program) går däremot mycket snabbt. Sammanställningen av resultatet av bearbetningarna bör heller inte ta så lång tid i anspråk. Det kan vara praktiskt att samla resultatet av bearbetningarna i en EXCEL-bok framför allt för att enkelt kunna ta fram underlag till bedömning av om man kan redovisa preliminär statistik. Trots att efterarbetet av simuleringen egentligen inte är tidskrävande är det knappast försvarbart att gång efter annan genomföra simuleringarna för olika bortfallsstorlekar.

Resultatet av den här redovisade simuleringstudien kan användas såsom en vägledning för när eller om man kan redovisa preliminär statistik. Vilken statistik som kan redovisas beror på val av

- parameter
- redovisningsgrupp(er)
- kvalitetskrav

⁷ Storstadsregioner (3 st) och länsregioner (3 st)

⁸ Storstads- och länsregioner

Redovisning av preliminär statistik medför vissa risker. Simuleringsstudien visar på en osäkerhet som ibland kan vara mycket stor. Konfidensintervall som är bredare än 1 000 kronor/m² och 500 000 kronor/lägenhet förekommer både på byggnads- och produktionskostnaderna. I bilaga 2 redovisas kortaste och längsta konfidensintervall (95%) för kostnader per kvadratmeter och lägenhet samt lägenhetsyta (total primär bruksarea). Konfidensintervall beräknas för hela landet och de redovisningsgrupper som redovisas på sidan 5. Nedan i tabellerna 5-8 redovisas på *riksnivå* definitiva uppgifter och intervallskattningar för kostnader per kvadratmeter och lägenhet samt lägenhetsyta (total primär bruksarea).

Tabell 5 Kostnader per kvadratmeter primär bruksarea. Definitiva uppgifter och konfidensintervall (95%). Flerbostadshus 1995 och 1996. Avrundat till 100-tal kronor.

År	Byggnadskostnad per m ²		Produktionskostnad per m ²	
	Undre gräns	Övre gräns	Undre gräns	Övre gräns
1995				
Definitiva uppgifter	10 100		11 300	
Bortfall (%)				
20	10 000	10 400	11 200	11 600
30	10 100	10 500	11 200	11 800
40	10 100	10 600	11 200	11 900
1996				
Definitiva uppgifter	10 700		11 700	
Bortfall (%)				
20	10 500	11 000	11 500	12 000
30	10 500	11 100	11 500	12 100
40	10 500	11 200	11 500	12 200

Tabell 6 Kostnader per kvadratmeter primär bruksarea. Definitiva uppgifter och konfidensintervall (95%). Gruppbyggda småhus 1995 och 1996. Avrundat till 100-tal kronor.

År	Byggnadskostnad per m ²		Produktionskostnad per m ²	
	Undre gräns	Övre gräns	Undre gräns	Övre gräns
1995				
Definitiva uppgifter	8 400		10 100	
Bortfall (%)				
20	8 400	8 600	10 100	10 300
30	8 400	8 600	10 100	10 400
40	8 400	8 700	10 100	10 500
1996				
Definitiva uppgifter	9 000		10 600	
Bortfall (%)				
20	9 000	9 100	10 600	10 800
30	9 000	9 200	10 600	10 900
40	9 000	9 300	10 600	11 000

Tabell 7 Lägenhetsyta (total primär bruksarea) och kostnader per lägenhet. Definitiva uppgifter och konfidensintervall (95%). Flerbostadshus 1995 och 1996. Avrundat till 100-tal kronor.

År	Lägenhetsyta		Byggnadskostnad per lägenhet		Produktionskostnad per lägenhet	
	Undre gräns	Övre gräns	Undre gräns	Övre gräns	Undre gräns	Övre gräns
1995						
Definitiva uppgifter	78		784 000		874 000	
Bortfall (%)						
20	77	80	770 000	821 000	858 000	917 000
30	77	81	771 000	837 000	859 000	937 000
40	77	82	774 000	853 000	862 000	955 000
1996						
Definitiva uppgifter	80		835 000		915 000	
Bortfall (%)						
20	79	82	812 000	875 000	891 000	955 000
30	79	83	809 000	890 000	886 000	969 000
40	79	84	809 000	906 000	886 000	985 000

Tabell 8 Lägenhetsyta (total primär bruksarea) och kostnader per lägenhet. Definitiva uppgifter och konfidensintervall (95%). Gruppbyggda småhus 1995 och 1996. Avrundat till 100-tal kronor.

År	Lägenhetsyta		Byggnadskostnad per lägenhet		Produktionskostnad per lägenhet	
	Undre gräns	Övre gräns	Undre gräns	Övre gräns	Undre gräns	Övre gräns
1995						
Definitiva uppgifter	106		895 000		1 078 000	
Bortfall (%)						
20	104	109	882 000	925 000	1 057 000	1 118 000
30	104	109	883 000	935 000	1 057 000	1 131 000
40	104	110	885 000	946 000	1 059 000	1 146 000
1996						
Definitiva uppgifter	104		936 000		1 107 000	
Bortfall (%)						
20	103	106	931 000	959 000	1 095 000	1 132 000
30	103	106	933 000	969 000	1 095 000	1 142 000
40	103	106	937 000	978 000	1 096 000	1 152 000

Anm. 1996, 40% bortfall: Byggnadskostnad per lägenhet beräknat på det definitiva materialet ligger inte i intervallet.

6 Slutsatser

6.1 Sammanfattning

Undertäckning och objektsbortfall är de främsta hindren för en tidig preliminär redovisning av nybyggnadskostnadsstatistik. Både undertäckning och objektsbortfall är temporära feltyper.

Att under pågående insamlingsår för varje aktuell period känna till den verkliga populationsstorleken kan vara ett problem med nuvarande indikatorvariabel. Den temporära undertäckningen är det största problemet. Med hjälp av erfarenhet från tidigare år och bättre information i ramen bör man kunna uppskatta populationsstorleken.

Från den i avsnitt 3 redovisade populationsberäkningen på 1995 och 1996 års material kan man dra följande slutsatser

- för de två första kvartalen kan man knappast rekommendera en tidig preliminär redovisning ens om bortfallet är aldrig så litet
- om kostnaderna ökar under året (inflation) kommer en sammanslagning av data för flera redovisningsperioder sannolikt innebära en underskattning

Små redovisningsgrupper⁹ kommer alltid att utgöra ett problem. Om kunden insisterar på en viss redovisningsnivå kan man på grund av sekretessregler tvingas att dölja information i de redovisningsgrupper där det preliminära dataunderlaget är otillräckligt. Man kan som alternativ till detta

- rekommendera redovisning på en högre aggregerad nivå
- ha riktade insatser på de små redovisningsgrupperna för att påskynda uppgiftslämnarna

En bättre indikator för påbörjandetidpunkt än beslutstidpunkten skulle underlätta¹⁰ för uppgiftslämnarna och förhoppningsvis därigenom även påskynda uppgiftslämnandet.

Skattningar baserade på preliminärt material på variabeln markkostnad bör man om möjligt avstå ifrån.

6.2 Redovisning av preliminär statistik

Efterfrågan på aktuell statistik över nybyggnadskostnader är stor. Ju tidigare man kan få tillgång till aktuella uppgifter desto bättre. Därför är även preliminär statistik intressant.

Vårt dilemma är dock att vi inte alltid kan tillmötesgå användarnas önskemål om aktuella, dvs tidiga uppgifter. Problemet ligger bland annat i *när* vi får kännedom om påbörjade nybyggnadsprojekt (ramen) och den *tid* som datainsamlingen tar (se 3 Uppskattning av populationsstorleken).

All statistik som redovisas skall kvalitetredovisas. Den preliminära statistiken utgör inget undantag. Kvalitetsinformation är en vägledning till tolkning och användning av statistiken. Uppgift om osäkerhet är till viss hjälp, men även information om förhållandet mellan beräkningar baserade på preliminärt och definitivt material kan fungera som vägledning.

Förhållandet mellan beräkningar baserade på preliminärt och definitivt material kan uttryckas med hjälp av en enkel indikator

$$\frac{\text{preliminärt resultat}}{\text{definitivt resultat}} \cdot 100 \quad (1)$$

Av erfarenhet vet vi att uppgifter för första kvartalet är av särskilt stort intresse. Det är också det första kvartalet som bereder oss det största problemet genom undertäckning i databasen (se 3 Uppskattning av populationsstorleken tabell 3 och 4) och sent besvarade enkäter (temporärt bortfall).

I diagram 1-4 illustreras vid olika tidpunkter förhållandet mellan preliminärt och definitivt resultat med hjälp av indikatorvariabeln (1) för parametrarna **antal lägenheter** och **produktionskostnad per kvadratmeter total primär bruksarea** (PK/m²). Uppgifterna avser första kvartalet 1995 och 1996. Under senare delen av 1996 genomfördes förändringar i påminnelsearbetet (se 3 Uppskattning av populationsstorleken). Denna förändring tillsammans med ökad kännedom om undersökningen bland byggherrarna kan ha positiv inverkan på undersökningen, dvs de preliminära resultaten konvergerar snabbare mot de definitiva.

Simuleringsstudien genomfördes på bortfallsstorlekarna 20, 30 och 40 procent. Storleken på bortfallet valdes på känn. Avsikten var att "bortfallet" skulle motsvara några tidpunkter i datainsamlingsarbetet. Med erfarenhet av datainsamlingen för 1995 och 1996 kan vi "avslöja" tidpunkterna för första kvartalet. För första kvartalet 1995 skulle 40% bortfall svara mot tillgängligt material den sista december 1995. För motsvarande period 1996 har man 40% bortfall i slutet av september 1996. För gruppbyggda småhus är situationen obetydligt bättre.

⁹ Redovisningsgrupper med ett förväntat litet antal observationer.

¹⁰ Vid påbörjandetidpunkten, referenstidpunkten, bör byggherren ha underlaget till enkäten mest aktuellt.

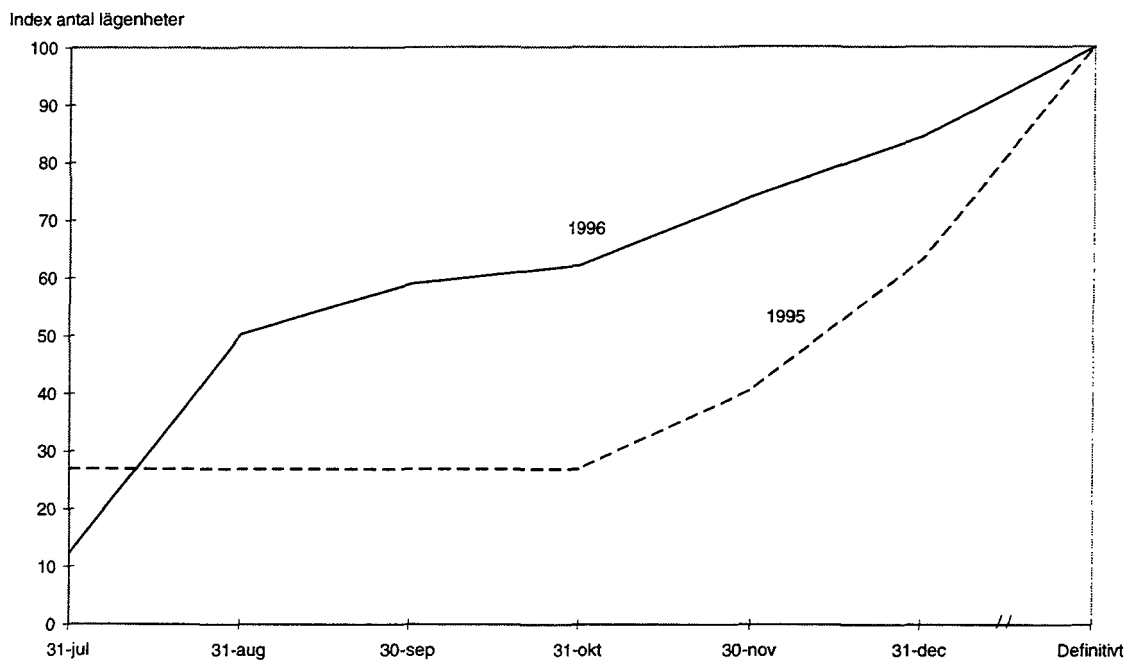


Diagram 1 Flerbostadshus första kvartalet 1995 och 1996. Antal lägenheter. Preliminära uppgifter i förhållande till definitiva.

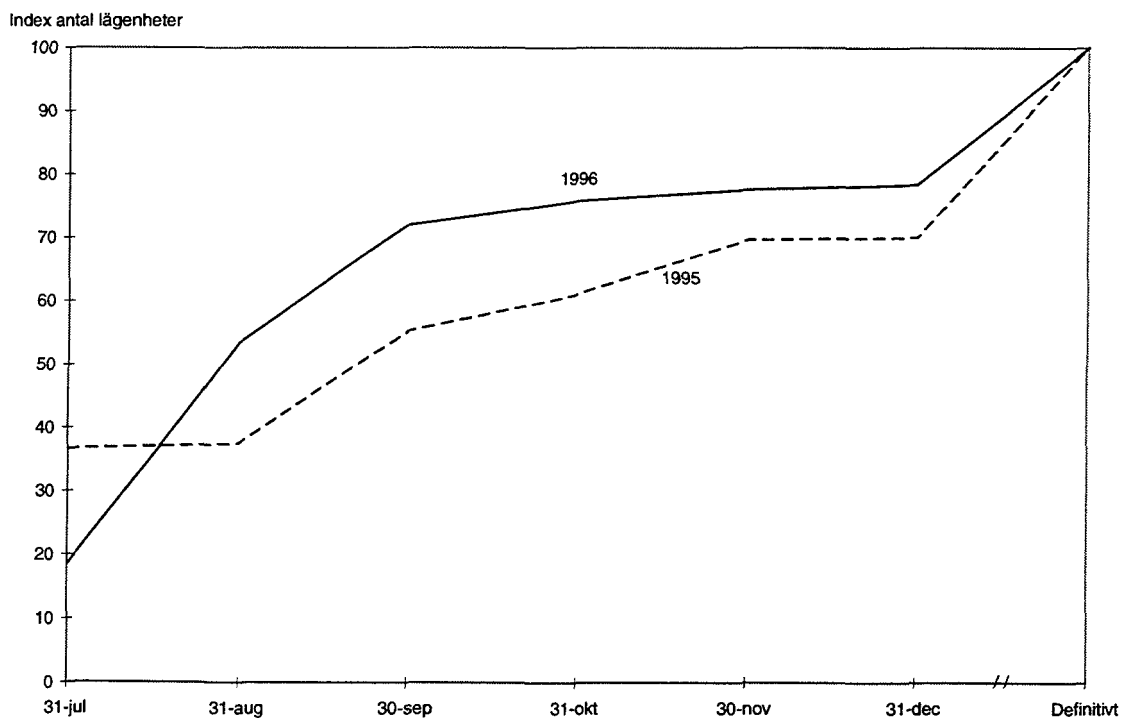


Diagram 2 Gruppbyggda småhus första kvartalet 1995 och 1996. Antal lägenheter. Preliminära uppgifter i förhållande till definitiva.

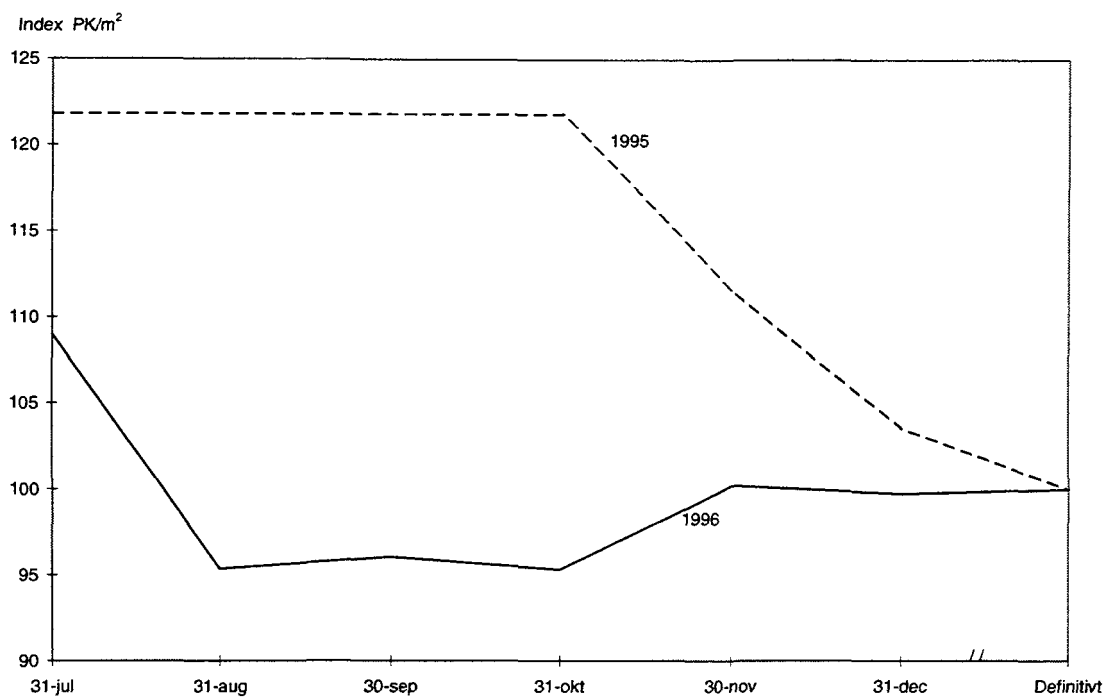


Diagram 3 Flerbostadshus första kvartalet 1995 och 1996. Produktionskostnad per kvadratmeter total primär bruksarea. Preliminära uppgifter i förhållande till definitiva.

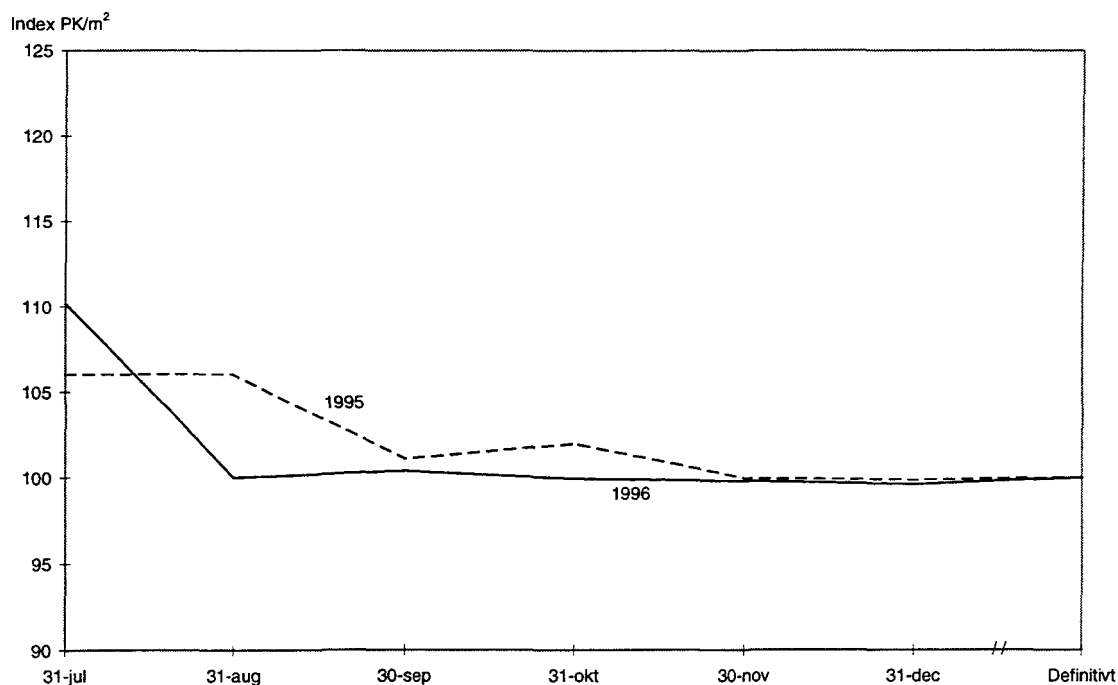


Diagram 4 Gruppbyggda småhus första kvartalet 1995 och 1996. Produktionskostnad per kvadratmeter total primär bruksarea. Preliminära uppgifter i förhållande till definitiva.

Preliminär statistik kan redovisas som intervallskattningar. Härvid kan man utnyttja resultat (uppskattning om osäkerhet) från de genomförda simuleringsstudierna. Ett alternativ till intervallskattningarna är att redovisa avrundade tal, t ex

- kostnad per kvadratmeter på 100 eller 1 000 kronor när
- kostnad per lägenhet på 1 000 eller 10 000 kronor när

6.3 Information om användning av statistiken

Uppgifter om kostnader per kvadratmeter och kostnader per lägenhet redovisas även som **tids-serier**. Denna redovisning inbjuder till felanvändning, dvs att belysa kostnadsförändringar med hjälp av kostnad per kvadratmeter eller kostnad per lägenhet. En slutsats av diagrammen 3 och 4 ovan är att det kan vara mycket vanskligt att använda kostnad per kvadratmeter eller per lägenhet för att uppskatta kostnadsförändringar speciellt när den aktuella uppgiften bygger på preliminär statistik. *Kostnadsförändringar* belyses bäst med hjälp av *byggnadsprisindex*.

Tabellerna där tidsserier redovisas bör därför förse med en not där man varnar för felanvändning, eller bättre, anger hur uppgifterna kan användas. Man bör dessutom ange lämpliga alternativ för den som vill studera kostnadsförändringar.

7 Referens

[1] Särndal, Swensson och Wretman (1992): Model Assisted Survey Sampling.

I programmeringsarbetet har även SAS Language and Procedures Usage 2 Version 6 (1991) (kapitel 12 Taking Stratified Random Samples) utnyttjats.

Formler

Variabler

Ej redovisningsgrupper i tabeller

Variabelnamn i BYGGSAM beräkningarna		Variabelnamn i SAS-
ANTLGH	Antal lägenheter	LGH
BRAP	Total primär bruksarea (m ²)	BRAP
BRAPBLGH	Bostadslägenhetsarea (m ²)	BRAPB
BRAPLGH	Lägenhetsarea (m ²)	BRAPL
MARKK	Markkostnad (tkr)	MARK
BK	Byggnadskostnad (tkr)	BK
PK	Total produktionskostnad (tkr)	PK

Beräkningar: tabeller

De beräkningar som förekommer i tabellerna är *totaler* och *vägda medelvärden* för olika redovisningsgrupper.

Totaler

$$\sum_i LGH_i \quad \text{Antal lägenheter} \quad (1)$$

$$\sum_i BRAP_i \quad \text{Total primär bruksarea} \quad (2)$$

Vägda medelvärden

Ytor

$$\frac{\sum_i BRAP_i}{\sum_i LGH_i} \quad \text{Total primär bruksarea per lägenhet} \quad (3)$$

Markkostnad

$$\frac{\sum_i MARK_i}{\sum_i BRAP_i} \quad \text{Markkostnad/m}^2 \text{ total primär bruksarea} \quad (4)$$

$$\frac{\sum_i \text{MARK}_i}{\sum_i \text{LGH}_i} \cdot \frac{\sum_i \text{BRAPB}_i}{\sum_i \text{BRAPL}_i} \quad \text{Markkostnad/lägenhet (flerbostadshus)}^1 \quad (5)$$

$$\frac{\sum_i \text{MARK}_i}{\sum_i \text{LGH}_i} \quad \text{Markkostnad/lägenhet (gruppbyggda småhus)} \quad (6)$$

Byggnadskostnad

$$\frac{\sum_i \text{BK}_i}{\sum_i \text{BRAP}_i} \quad \text{Byggnadskostnad/m}^2 \text{ total primär bruksarea} \quad (7)$$

$$\frac{\sum_i \text{BK}_i}{\sum_i \text{LGH}_i} \cdot \frac{\sum_i \text{BRAPB}_i}{\sum_i \text{BRAPL}_i} \quad \text{Byggnadskostnad/lägenhet (flerbostadshus)}^2 \quad (8)$$

$$\frac{\sum_i \text{BK}_i}{\sum_i \text{LGH}_i} \quad \text{Byggnadskostnad/lägenhet (gruppbyggda småhus)} \quad (9)$$

Produktionskostnad

$$\frac{\sum_i \text{PK}_i}{\sum_i \text{BRAP}_i} \quad \text{Produktionskostnad/m}^2 \text{ total primär bruksarea} \quad (10)$$

$$\frac{\sum_i \text{PK}_i}{\sum_i \text{LGH}_i} \cdot \frac{\sum_i \text{BRAPB}_i}{\sum_i \text{BRAPL}_i} \quad \text{Produktionskostnad/lägenhet (flerbostadshus)}^3 \quad (11)$$

$$\frac{\sum_i \text{PK}_i}{\sum_i \text{LGH}_i} \quad \text{Produktionskostnad/lägenhet (gruppbyggda småhus)} \quad (12)$$

¹ Beräkningen bygger på antagandet att lokalerna kostar lika mycket att bygga som bostäderna.

$\frac{\sum_i \text{BRAPB}_i}{\sum_i \text{BRAPL}_i}$ är en korrektionsfaktor. Skattning av kostnad per lägenhet korrigeras *inte* vid redovisning i

Byggindex.

² Se fotnot 1

³ Se fotnot 1

Kortaste och längsta konfidensintervall för given bortfallsstorlek och år

**Tabell 1 Kortast och längst konfidensintervall (95%). Flerbo-
stadshus 1995 och 1996. Kostnader per kvadratmeter
primär bruksarea. Avrundat till 100-tal kronor.**

År	Byggnadskostnad per m ²		Produktionskostnad per m ²		Markkostnad per m ²	
	Undre gräns	Övre gräns	Undre gräns	Övre gräns	Undre gräns	Övre gräns
Simulerat min/max bortfall %						
1995						
20 min	10 000	10 400	8 700	9 000	*	*
20 max	12 700	13 800	11 500	12 600	1 000	1 400
30 min	10 700	11 100	10 800	11 300	*	*
30 max	12 700	14 000	9 100	10 500	1 000	1 400
40 min	<u>10 700</u>	<u>11 200</u>	10 800	11 300	*	*
40 max	8 800	10 400	8 900	10 700	900	1 500
1996						
20 min	8 700	9 100	9 700	10 000	900	1 100
20 max	9 900	11 300	9 800	11 200	1 100	1 800
30 min	8 700	9 100	9 700	10 100	900	1 100
30 max	9 600	11 600	9 600	11 500	900	1 800
40 min	8 700	9 200	10 000	10 500	700	800
40 max	9 400	11 700	9 500	11 700	800	1 800

* Liten variation som ger ett extremt kort konfidensintervall.

Understruket intervall är ett av flera lika långa intervall.

**Tabell 2 Kortast och längst konfidensintervall (95%). Grupp-
byggda småhus 1995 och 1996. Kostnader per kvadrat-
meter primär bruksarea. Avrundat till 100-tal kronor.**

År	Byggnadskostnad per m ²		Produktionskostnad per m ²		Markkostnad per m ²	
	Undre gräns	Övre gräns	Undre gräns	Övre gräns	Undre gräns	Övre gräns
Simulerat min/max bortfall %						
1995						
20 min	8 500	8 700	10 900	11 100	<i>2 100</i>	<i>2 200</i>
20 max	7 900	9 300	8 800	10 600	800	1 500
30 min	8 500	8 800	10 900	11 100	<i>2 000</i>	<i>2 200</i>
30 max	7 600	9 700	8 400	11 000	600	1 600
40 min	<i>8 600</i>	<i>8 800</i>	10 900	11 200	<i>2 000</i>	<i>2 100</i>
40 max	7 100	10 200	7 900	11 800	500	1 800
1996						
20 min	<i>9 000</i>	<i>9 100</i>	9 900	10 200	1 100	1 200
20 max	6 400	8 100	7 900	8 800	700	1 600
30 min	<i>9 000</i>	<i>9 200</i>	<i>10 000</i>	<i>10 300</i>	1 500	1 700
30 max	6 200	8 200	9 700	10 800	600	1 700
40 min	<i>9 000</i>	<i>9 300</i>	<i>10 000</i>	<i>10 400</i>	1 500	1 700
40 max	5 900	8 200	7 600	9 000	500	1 900

Intervallen som är markerade med kursiv stil övertäcker inte det "sanna" värdet. Med sant värde avses det resultat som man får om man utgår ifrån hela det definitiva materialet.

Tabell 3 Kortast och längst konfidensintervall (95%). Flerbostadshus 1995 och 1996. Lägenhetsyta (total primär bruksarea) och kostnad per lägenhet. Kostnader avrundade till 1000-tal kronor.

År	Lägenhetsyta		Byggnadskostnad per lägenhet		Produktionskostnad per lägenhet		Markkostnad per lägenhet	
	Undre gräns	Övre gräns	Undre gräns	Övre gräns	Undre gräns	Övre gräns	Undre gräns	Övre gräns
Simulerat min/max bortfall %								
1995								
20 min	<u>73</u>	<u>75</u>	683 000	727 000	1 062 000	1 077 000	*	*
20 max	82	109	1 028 000	1 479 000	1 143 000	1 624 000	110 000	151 000
30 min	<u>73</u>	<u>76</u>	1 009 000	1 027 000	1 060 000	1 078 000		
30 max	81	113	1 014 000	1 550 000	1 127 000	1 697 000	106 000	154 000
40 min	73	76	1 007 000	1 026 000	1 058 000	1 077 000	*	*
40 max	82	116	1 020 000	1 612 000	1 133 000	1 759 000	104 000	156 000
1996								
20 min	<u>79</u>	<u>82</u>	598 000	627 000	661 000	696 000	74 000	84 000
20 max	58	71	708 000	934 000	787 000	1 026 000	88 000	142 000
30 min	<u>79</u>	<u>83</u>	598 000	636 000	660 000	706 000	72 000	85 000
30 max	<u>79</u>	<u>94</u>	676 000	978 000	752 000	1 072 000	80 000	145 000
40 min	<u>79</u>	<u>84</u>	598 000	646 000	658 000	716 000	69 000	86 000
40 max	<u>78</u>	<u>96</u>	657 000	1 015 000	733 000	1 113 000	80 000	139 000

* Liten variation som ger ett extremt kort konfidensintervall.

Understruket intervall är ett av flera lika långa intervall.

Tabell 4 Kortast och längst konfidensintervall (95%). Gruppbyggda småhus 1995 och 1996. Lägenhetsyta (total primär bruksarea) och kostnad per lägenhet. Kostnader avrundade till 1000-tal kronor.

År	Lägenhetsyta		Byggnadskostnad per lägenhet		Produktionskostnad per lägenhet		Markkostnad per lägenhet	
	Undre gräns	Övre gräns	Undre gräns	Övre gräns	Undre gräns	Övre gräns	Undre gräns	Övre gräns
Simulerat min/max bortfall %								
1995								
20 min	88	90	1 060 000	1 085 000	1 328 000	1 362 000	<i>265 000</i>	<i>280 000</i>
20 max	87	117	771 000	975 000	888 000	1 082 000	85 000	139 000
30 min	88	90	1 061 000	1 093 000	1 323 000	1 365 000	<i>257 000</i>	<i>277 000</i>
30 max	80	124	727 000	1 021 000	849 000	1 122 000	74 000	149 000
40 min	<u>125</u>	<u>127</u>	1 063 000	1 101 000	1 248 000	1 301 000	<i>249 000</i>	<i>272 000</i>
40 max	74	130	696 000	1 054 000	818 000	1 154 000	65 000	157 000
1996								
20 min	<u>103</u>	<u>106</u>	931 000	955 000	1 097 000	1 120 000	121 000	133 000
20 max	92	116	595 000	909 000	739 000	1 000 000	80 000	156 000
30 min	<u>103</u>	<u>106</u>	<i>935 000</i>	<i>965 000</i>	1 097 000	1 126 000	121 000	136 000
30 max	86	123	556 000	947 000	693 000	1 048 000	70 000	168 000
40 min	98	101	<i>940 000</i>	<i>976 000</i>	1 098 000	1 132 000	120 000	137 000
40 max	80	127	505 000	961 000	648 000	1 065 000	63 000	184 000

Intervallen som är markerade med kursiv stil övertäcker inte det "sanna" värdet. Med sant värde avses den skattning som man får om man utgår ifrån det totala dataunderlaget.

Understruket intervall är ett av flera lika långa intervall.

Förteckning över utkomna R&D Reports

R&D Reports är en för IT-enheten och U/STM gemensam publikationsserie, som 1988-01-01 ersatte de tidigare "gula" och "gröna" serierna. I serien ingick fram till årsskiftet 1992-93 även **Abstracts** (sammanfattning av metodrapporter från SCB).

Reports published 1996 and onwards:

- 1996:1 On Sampling with Probability Proportional to Size (*Bengt Rosén*)
(grön)
- 1996:2 Bortfallsbarometern nr 11 (*Antti Ahtiainen, Stefan Berg, Margareta Eriksson, Åsa Greijer, Dan Hedlin, Monica Rennermalm, Anita Ullberg*)
(grön)
- 1996:3 Regression Estimators in Theory and in Practice (*Tomas Garås*)
(grön)
- 1996:4 Quality Aspects of a Modern Database Service (*Pat Dean, Bo Sundgren*)
(gul)
- 1996:5 Metadata: A Quality Element in Official Statistics - the Swedish Approach
(*Bo Sundgren, Pat Dean*)
(gul)
- 1996:6 Our Legacy to Future Generation - Using Databases for Better Availability and Documentation (*Gösta Guteland, Erik Malmborg*)
(gul)
- 1997:1 Bortfallsbarometern nr 12 (*Anti Ahtiainen, Stefan Berg, Mats Bergdahl, Fredrik Granström, Dan Hedlin, Lena Otterskog, Monica Rennermalm*)
(grön)
- 1997:2 Quality Concept for Official Statistics - Entry in the forthcoming update of the Encyclopedia of Statistical Sciences, Wiley & Sons (*Eva Elvers, Bengt Rosén*)
(grön)
- 1998:1 Nybyggnadsstatistik - en simuleringsstudie (*Catarina Elffors*)
(grön)

090693-002101- 1

CATARINA ELFFORS

BR/BY-S

ISSN 0283-8680

Tidigare utgivna *R & D Reports* kan beställas genom Ingvar Andersson, SCB, U/SIB, Box 24300, 115 81 STOCKHOLM (telefon 08-783 41 47, fax 08-783 45 99, e-post ingvar.andersson@scb.se).

R & D Reports listed above as well as issues from 1988-1994 can - in case they are still in stock - be ordered from Statistics Sweden, att. Ingvar Andersson U/SIB, Box 24300, S-115 81 STOCKHOLM (telephone +46 8 783 41 47, fax +46 8 783 45 99, e-mail ingvar.andersson@scb.se).