

# Bilaga 5.

# Utnyttjade urvalsundersökningar

För *Barn och deras familjer 2003* används tre urvalsundersökningar (surveyer): SCB, Undersökningarna om levnadsförhållanden (ULF) 2002 och 2003; SCB, Arbetskraftsundersökningarna (AKU) 2003; SCB och Hushållens ekonomi (HEK) 2002.

I dessa surveyer är urvalsperson och vår undersökningperson, barn, inte densamma. Man gör ett urval av vuxna som ger uppgifter om de barn som finns i hushållet. Detta påverkar olika sannolikhetsfördelningar i främst två avseenden:

Det **ena** är att barnet får en förhöjd sannolikhet att komma med i urvalet. I ULF kan barnet komma med antingen genom att mamman ingår i urvalet eller att pappan gör det. Därmed dubbleras sannolikheten för att komma med i studien för barn som bor med bägge föräldrarna. Olika estimatorer måste korrigeras för detta.

I AKU däremot väljer man att redovisa antingen "alla mammors barn" plus barn till ensamstående pappor eller vice versa. I vardera fallet har varje barn samma urvalssannolikhet som mamma alternativt pappa. Det blir inte fråga om dubbla urvalssannolikheter.

Det **andra** är att skattningarna av proportioner av någon egenskap hos barn genomgående är kvotskattningar. Basen, totala antalet barn i urvalet, är nämligen slumpmässigt på grund av varierande barnantal i familjerna. Totala antalet barn i urvalet kan sägas vara summan av olikstora "kluster". Vid skattning av varianser för populationstotaler påverkas emellertid inte beräkningarna av denna anledning. Uppräkningstal inom ULF kan direkt användas bara med den korrigeringen att de divideras med två när barnen har sammanboende föräldrar. Uppräkningstalen i HEK är redan anpassade till

redovisning av hushåll och samma uppräkningsstal kan användas för barnstatistik. I AKU används samma uppräkningsstal som för mamman respektive pappan.

## Arbetskraftsundersökningen (AKU)

Arbetskraftsundersökningarna (AKU) utförs varje månad. Från undersökningen beräknas månads-, kvartals-, årsgenomsnitt. Syftet med undersökningarna är att beskriva de aktuella sysselsättningsförhållandena och utvecklingen på arbetsmarknaden.

### *Undersökningspopulation*

Undersökningspopulationen är alla i Sverige folkbokförda personer som fyllt 16 men ej 65 år.

### *Urvalsram*

SCB:s register över totalbefolkningen (RTB).

### *Urval*

Urvalet består av tre separata urval, ett för varje månad i kvartalet. Vart och ett av urvalen, som för varje månad består av ca 19 500 personer, roteras på så sätt att en åttondel byts ut mellan två på varandra följande undersökningstillfällen. För varje urval inträffar detta alltså med tre månaders mellanrum. Personer i urvalet intervjuas med andra ord en gång per kvartal och sammanlagt åtta gånger under en tvåårsperiod, varefter de byts ut mot nya.

Urval dras i slutet av första kvartalet varje år för att täcka det kommande årets behov av urvalspersoner. Vid urvalsdragningen sker en indelning i strata efter län, kön, medborgarskap (svenskt, utländskt) och sysselsättningen (sysselsatt, ej sysselsatt) enligt sysselsättningsregistret (SCB). Härigenom erhålls 192 strata. Inom

varje stratum dras ett systematiskt sannolikhetsurval. Inom strata är därvid registret sorterat med avseende på ålder. För barnstatistiken används ett helt års urval.

#### *Insamlingsmetod*

Telefonintervju används. I vissa fall, t.ex. för urvalspersoner som saknar telefon, görs besöksintervju.

#### *Bortfall*

Bortfallet var 16,3 procent 2003. Hjälpinformation från SCB:s sysselsättningsregister, SCB:s register över totalbefolkningen (RTB) och AMS arbetssökanderegister används för att minska bortfallsfelet.

#### *Barnstatistik*

Vid framställningen av statistik om barn låter man urvalspersonernas barn utgöra studieobjekt. Urvalspersonen kan antingen vara mamman eller pappan. Intervjupersonen uppger hur många barn som finns i familjen samt när barnen är födda (år, mån). Statistiken baseras på uppgifter från mamman eller pappan.

För närmare information om AKU se "Beskrivningen av statistiken" på AKU:s webbplats: [www.scb.se/aku](http://www.scb.se/aku).

### **Hushållens ekonomi (HEK)**

Undersökningen om Hushållens Ekonomi har som syfte att belysa inkomstförhållanden för hela Sveriges befolkning. Redovisning sker med avseende på såväl hushåll som individer. Med hushåll menas kosthushåll. Ett kosthushåll består av personer som bor i samma bostad och har gemensam ekonomi och kosthushållning. I redovisningen från undersökningen används begreppet helårshushåll som är de hushåll där alla vuxna individer (18 år eller äldre) varit registrerade i folkbokföringen både 1 januari och 31 december samma år.

#### *Undersökningspopulation*

Alla hushåll med personer 18 år eller äldre.

#### *Urvalsram*

SCB:s register över totalbefolkningen (RTB).

#### *Urval*

Urvalsramen år 2002 innehöll följande strata:

- personer 18-74 år
- personer 75 år eller äldre
- de 5% som har de högsta inkomsterna
- de 5% som har de lägsta inkomsterna
- vårdtagare=de med vårdersättning
- handikappade=de med handikappersättning

Urvalet drogs som ett stratifierat urval med OSU (obundet slumpmässigt urval). Det totala urvalet år 2002 var ca 18 800. Urvalsstorleken har under åren 1993-2002 varierat mellan 13 000-19 000.

#### *Insamlingsmetod*

Uppgiftsinsamlingen sker dels direkt genom telefonintervju, dels genom användande av olika myndigheters register.

#### *Bortfall*

Bortfallet var cirka 32 procent år 2002. För personer som ingår i bortfallet finns dock samtliga registeruppgifter. Saknade uppgifter om familjesammansättning ersätts med uppgifter från befolkningsregister.

#### *Barnstatistik*

Det ursprungliga urvalet omdefinieras till att gälla en barnpopulation. De barn som ingår i urvalet för barnstatistiken är de som ingår i de utvalda familjeenheter enligt vad urvalspersonen själv rapporterar. I 2002 års HEK finns omkring 7 000 barn i åldrarna 0-17 år bland svarande helårshushåll.

### **Undersökningar om levnadsförhållanden (ULF)**

SCB genomför löpande årliga undersökningar om levnadsförhållanden. Undersökningarna är avsedda att belysa välfärden i nuläge och över tid, samband mellan problem på olika områden samt skillnader mellan olika befolkningsgrupper. Undersökningen innehåller en basuppsättning frågor som återkommer varje år. Därutöver finns ett antal fördjupningsområden som inkluderas enligt ett rullande schema och där varje område återkommer vart åttonde år och då för en tvåårsperiod.

### Undersökningspopulation

Urvalet görs huvudsakligen bland personer i åldrarna 16-84 år. Barn 10-18 år i urvalspersonernas hushåll medtas i undersökningen men till dem ställs andra frågor än dem urvalspersonerna får.

### Urvalsram

SCB:s register över totalbefolkningen (RTB).

### Urval

Varje år dras ett systematiskt urval efter ålder, 7 000 - 8 000 personer. Från urvalet avlägsnas personer som varit med i ULF under de senaste sju åren. En del av urvalet utgörs av en panel, som ingått i ULF vid föregående fördjupningsomgång. Panelen kompletteras med immigranter och personer som "vuxit in" i populationen (16-23 år).

Vanligtvis redovisas resultaten för dubbelårgångar, vilka har ett urval på omkring 15 000 personer. Till en dubbelårgångs urvalspersoner hör omkring 6 500 barn i åldrarna 0-17 år.

### Insamlingsmetod

Huvudsakligen besöksintervjuer men i viss utsträckning telefonintervjuer. Därtill kommer en omfattande registerinsamling.

### Bortfall

Bortfallet var 25 procent år 2002. När skattningar tas fram används poststratifiering i syfte att minska bortfallsfelet. Strata bildas efter kombinationer av kön, ålder, H-region och civilstånd.

### Barnstatistik

Det ursprungliga urvalet omdefinieras till att gälla en barnpopulation. De barn som ingår i urvalet för barnstatistiken är de som ingår i urvalspersonernas familjer enligt vad urvalspersonen själv rapporterar. Observationerna kan sålunda komma från intervjuer antingen med pappan eller med mamman.

### Urvals- och bortfallsfel

Samtliga urvalsundersökningar är behäftade med slumpfel som beror på stickprovsdrag-

ningen. Felens storlek är beroende på vilken urvalsdessign och urvalstorlek som har använts.

Alla utvalda personer i ett stickprov vill eller kan inte delta i undersökningen. Bortfallet minskar stickprovets storlek och därmed blir slumpfelet större än vad det annars skulle bli. Om bortfallet är slumpmässigt behöver det i övrigt inte påverka resultatet. Men bortfallet är ofta mer eller mindre selektivt. Låg svarsfrekvens inom vissa vitala grupper kan snedvrída undersökningsresultatet.

Bortfallet är vanligtvis lägre bland personer som har barn än bland personer som inte har barn. När man räknar upp urvalet till totalnivå erhålls därför systematiskt en överskattning av antalet barn. Samtidigt ger den högre svarsfrekvensen i barnfamiljer jämförelsevis god säkerhet vad gäller olika fördelningar inom populationen barn och barnfamiljer.

### Skattningar

De tre undersökningarna AKU, ULF och HEK har alla en mer eller mindre komplicerad design.

AKU har stratifierat urval (OSU i strata) med efterstratifiering (samt paneler)

ULF har systematiskt urval (approximativt OSU) med efterstratifiering

HEK har stratifierat urval av typen nätverksurval

### Skattning av proportioner

Följande estimator för en proportion är applicerbar vid ett enkelt OSU av vuxna urvalspersoner och där man utnyttjar detta urval för estimering av proportionen barn. Förutsättningen för en sådan skattning är att man inhämtat upplysning om urvalspersonernas antal barn.

$$p = \frac{\sum_{i=1}^n \frac{1}{\pi_i} a_i}{\sum_{i=1}^n \frac{1}{\pi_i} m_i} \quad (1)$$

$a_i$  = antalet barn med viss egenskap i familj i

$m_i$  = antalet barn i familj i

$n$  = stickprovsstorlek (antal familjer med barn)

$\pi_i$  = inklusionssannolikhet för familj  $i$   
 $N$  = antal familjer med barn i rampopulationen

”Familj  $i$ ” väljs genom ”urvalsperson  $i$ ”. I ULF och HEK är inklusionssannolikheterna reducerade eftersom en familj kan väljas på flera sätt (t.ex. genom båda sammanboende föräldrarna). Att en familj i realiteten skulle väljas ut två gånger är osannolikt och bortses från.

Estimatoren för totala antalet barn med en viss egenskap utgörs endast av täljaren i (1). Eftersom den stokastiska nämnaren ej ingår förenklas variansberäkningen. Vidare kan sägas att inversen av inklusionssannolikheterna, de s.k. ”uppräkningsstalen” (bortsett från reduktionen enligt ovan), är generella för alla i undersökningen dvs. även barnlösa.

### Skattning av varianser för proportioner och totaler

Vid en skattning av variansen för en proportion måste hänsyn tas till det varierande barnantalet per familj. Variansen kan uttryckas för ett enkelt OSU,

$$v(p) = \frac{1-f}{n} \frac{\sum_{i=1}^n \left(\frac{m_i}{\pi_i}\right)^2 \left(\frac{a_i}{m_i} - p\right)^2}{\left(\frac{m_i}{\pi_i}\right)^2 (n-1)} \quad (2)$$

där  $f = \frac{n}{N}$

Observera att medelvärdet i nämnaren gäller kvoten inom parentesen.

Variansen i formel (2) kan alltså användas vid ett enkelt OSU. Men de olika undersökningarna har en mer komplicerad design och skattningen av varianser blir generellt mycket mer komplicerad (åtminstone för AKU och HEK).

Varianser användes för beräkning av konfidensintervall.

Konfidensintervall (95 -procentigt):

$$p \pm 1,96\sqrt{v(p)}$$

### AKU

För AKU har standardavvikelser inte beräknats. Man kan räkna med att konfidensintervallen är

ungefär hälften av dem som gäller för liknande skattningar från ULF.

### HEK

HEK har ett nätverksurval där skattningar av varianser för proportioner kan göras för alla personer i hushållet och därmed även för barn. Här har det ordinarie Heksystemet för variansskattningar utnyttjats<sup>1</sup>. Det ger en lösning på kvotskattningsproblemen och ger också en väldefinierad estimator som tar hänsyn till övriga ”design”-egenskaper. 95-procentiga konfidensintervall ges för ett antal centrala skattningar. Konfidensintervallen ges i direkt anslutning till tabellerna i kapitlet om familjernas ekonomi.

### ULF

Inom ULF-systemet har man också infört en möjlighet till korrigerig för ”nätverkseffekten” och varianserna beräknas med hjälp av en kvotskattning<sup>2</sup>. 95-procentiga konfidensintervall ges för ett antal centrala skattningar i anslutning till tabeller i olika kapitel.

<sup>1</sup> SCB, Extrakt systemet

<sup>2</sup> SCB, Tanjasystemet