

# Folkomröstningar, valdeltagandeundersökningen

## 2003

ME0114

### Innehåll

<b>0</b>	<b>Allmänna uppgifter .....</b>	<b>2</b>
0.1	Ämnesområde .....	2
0.2	Statistikområde .....	2
0.3	SOS-klassificering .....	2
0.4	Statistikansvarig .....	2
0.5	Statistikproducent .....	2
0.6	Uppgiftsskyldighet .....	2
0.7	Sekretess och regler för behandling av personuppgifter .....	2
0.8	Gallringsföreskrifter .....	3
0.9	EU-reglering .....	3
0.10	Syfte och historik .....	3
0.11	Statistikanvändning .....	3
0.12	Uppläggning och genomförande .....	3
0.13	Internationell rapportering .....	4
0.14	Planerade förändringar i kommande undersökningar .....	4
<b>1</b>	<b>Översikt .....</b>	<b>5</b>
1.1	Observationsstorheter .....	5
1.2	Statistiska målstorheter .....	5
1.3	Utfloeden: statistik och mikrodata .....	6
1.4	Dokumentation och metadata .....	6
<b>2</b>	<b>Uppgiftsinsamling .....</b>	<b>7</b>
2.1	Ram och ramförfarande .....	7
2.2	Urvalsförfarande .....	7
2.3	Mätinstrument .....	8
2.4	Insamlingsförfarande .....	8
2.5	Databeredning .....	8
<b>3</b>	<b>Statistisk bearbetning och redovisning .....</b>	<b>10</b>
3.1	Skattningar: antaganden och beräkningsformler .....	10
	3.1.1 Sammanvägning av olika delurval .....	10
	3.1.2 Estimation utifrån RTB-variabler .....	12
	3.1.3 Estimation utifrån AKU-variabler .....	15
	3.1.4 Referenser .....	16
3.2	Redovisningsförfaranden .....	17
<b>4</b>	<b>Slutliga observationsregister .....</b>	<b>18</b>
4.1	Produktionsversioner .....	18
4.2	Arkiveringsversioner .....	18
4.3	Erfarenheter från senaste undersökningsomgången .....	18
	<b>Bilaga 1: Frågeformulär till Länsstyrelsen .....</b>	<b>19</b>

## 0 Allmänna uppgifter

### 0.1 Ämnesområde

*Ämnesområde: Demokrati*

### 0.2 Statistikområde

*Statistikområde: Folkomröstningar*

### 0.3 SOS-klassificering

*Tillhör (SOS)*



För undersökningar som ingår i Sveriges officiella statistik gäller särskilda regler när det gäller kvalitet och tillgänglighet, se Förordningen om den officiella statistiken (2001:100).

### 0.4 Statistikansvarig

*Myndighet/organisation:* Statistiska centralbyrån  
*Postadress:* Box 24300, 104 51 STOCKHOLM  
*Besöksadress:* Karlavägen 100, STOCKHOLM  
*Kontaktperson:* Richard Öhrvall  
*Telefon:* 08-506 941 58  
*Telefax:* --  
*E-post:* valstatistik@scb.se

### 0.5 Statistikproducent

*Myndighet/organisation:* Statistiska centralbyrån  
*Postadress:* Box 24300, 104 51 STOCKHOLM  
*Besöksadress:* Karlavägen 100, STOCKHOLM  
*Kontaktperson:* Richard Öhrvall  
*Telefon:* 08-506 941 58  
*Telefax:* --  
*E-post:* valstatistik@scb.se

### 0.6 Uppgiftsskyldighet

Uppgiftsskyldighet föreligger inte enligt lagen om den officiella statistiken (SFS 2001: 99). Enskilda individer lämnar inte uppgifter till statistiken.

### 0.7 Sekretess och regler för behandling av personuppgifter

*I myndigheternas särskilda verksamhet för framställning av statistik gäller sekretess enligt 24 kap. 8 § offentlighets- och sekretesslagen*

(2009:400). Vid automatiserad behandling av personuppgifter gäller reglerna i personuppgiftslagen (1998:204). På statistikområdet finns dessutom särskilda regler för personuppgiftsbehandling i lagen (2001:99) och förordningen (2001:100) om den officiella statistiken.

## 0.8 Gallringsföreskrifter

Gallring enligt lagen (2001:99) om vissa personregister för officiell statistik.

## 0.9 EU-reglering

Undersökningen berörs inte av någon EU-reglering.

## 0.10 Syfte och historik

Syftet med valdeltagandeundersökningarna är att på ett utförligt sätt producera statistik över deltagandet vid allmänna val. Undersökningarna genomförs med syftet att komplettera bilden av det totala och regionala valdeltagande genom att undersöka valdeltagandet i olika grupper i samhället.

Mer allmänt kan sägas att syftet är att allsidigt redovisa och ge en samlad bild av allmänna valen som en dokumentation av ett av de viktigaste inslagen i den svenska demokratin. Valstatistiken har i olika former producerats av SCB från den tid som serien *Bidrag till Sveriges officiella statistik* började utkomma på 1860-talet. Valdeltagandeundersökningar i samband med allmänna val har genomförts sedan år 1911.

Undersökningen har genomgått ett flertal större förändringar under årens lopp, exempelvis vad gäller upplägg och undersökningspopulation.

Valdeltagandeundersökningar har genomförts i samband med nationella folkomröstningar 1957, 1980, 1994 och 2003. Folkomröstningar regleras av en folkomröstningslag (SFS 1979:369), vilken är en ramlag som kompletteras av en särskild lag för varje folkomröstningstillfälle.

## 0.11 Statistikanvändning

Huvudanvändare är de politiska partierna, massmedia, den samhällsvetenskapliga forskningen och den politiskt intresserade allmänheten. Detta inkluderar även internationella organisationer och forskare i andra länder. Resultaten används för att belysa den demokratiska processen, det politiska livet och medborgarnas politiska engagemang.

Valstatistiken utgör en grund för redovisningar av viktiga skeenden och förhållanden i svensk politik och demokrati.

## 0.12 Uppläggning och genomförande

För att kunna redovisa valdeltagandet i olika grupper är undersökningen uppbyggd av ett antal urval. För att allsidigt beskriva valdeltagandet bland svenska medborgare folkbokförda i Sverige används dels

arbetskraftsundersökningens urval i åldersgruppen 18-74, dels ett kompletterande urval bestående av äldre personer (utan övre åldersgräns). För att beskriva valdeltagande bland svenska medborgare som bor utomlands (utlandssvenskar) och i Sverige röstberättigade utländska medborgare dras separata urval. För mer information om hur urvalen dragits samt urvalsstorlekar, se avsnitt 2.2.1 Urval.

SCB framställer blanketter som skickas ut till alla länsstyrelser. På blanketterna markerar länsstyrelserna om urvalspersonerna har röstat i respektive val samt om de har förtidsröstat. Uppgifter om valdeltagande hämtar länsstyrelserna från röstlängderna. Därefter skickar länsstyrelserna in blanketterna till SCB som skannar in materialet, bearbetar det och slutligen framställer statistik.

### **0.13 Internationell rapportering**

Ingen internationell rapportering sker.

### **0.14 Planerade förändringar i kommande undersökningar**

Inga konkreta förändringar är planerade.

# 1 Översikt

## 1.1 Observationsstorheter

### Målpopulationer

Röstberättigade i folkomröstningen om euron 2003

### Variabler

Röstberättigade

Röstande

Valdeltagande

Dessutom finns även variabeln ”röstat i förtid”, som avser om personen har röstat vid något annat ställe än i vallokal (vanligen i förtid även om man kan rösta på annat ställe också på valdagen). I ordinarie omgångar av valdeltagandeundersökningen går information om förtidsröstning inte att koppla till ett specifikt val, utan endast till själva valtillfället. När det gäller folkomröstningen 2003 är det dock möjligt att göra den kopplingen. Detta då valtillfället endast avser ett val (med undantag för de kommuner som anordnade en lokal folkomröstning samtidigt som den nationella folkomröstningen om euron).

När det gäller övriga variabler ovan görs skattningar för olika redovisningsgrupper inom dem. För att kunna definiera redovisningsgrupper används olika bakgrundsvariabler. Undersökningsobjekten i valdeltagandeundersökningen är individer.

## 1.2 Statistiska målstorheter

Nedan anges vilka redovisningar som publicerades vid offentliggörandet av undersökningens resultat. Nya redovisningsgrupper som följer den tabellplan som använts i bland annat valdeltagandeundersökningen vid allmänna valen 2010 har senare tagits fram för publicering på SCB:s webbplats. Skillnaderna mellan dessa tabellplaner är dock små.

Objektgrupp		Variabel	Mått
Population	Indelning i redovisningsgrupper		
Röstberättigade i folkomröstningen om euron, samtliga	Kön, ålder, region, medborgarskapsland, födelseland, svensk/utländsk bakgrund, blivit svenska medborgare/alltid varit det, medborgarskapsdatum, civilstånd, sammanboendeform, inkomst, utbildningsnivå, arbetskraftsstatus, yrke, socioekonomisk grupp, anställningssektor och facklig	Röstberättigade, Röstande, Valdeltagande	Antal i 1000-tal, andel i procent, felmarginal

	centralorganisation		
Röstberättigade i folkomröstningen om euron, utlandssvenskar	Kön, ålder och region.	Röstberättigade, Röstande, Valdeltagande	Antal i 1000-tal, andel i procent, felmarginal

### 1.3 Utflöden: statistik och mikrodata

Valdeltagandeundersökningens resultat presenteras i form av resultatredovisningar som en del av Sveriges officiella statistik.

Redovisningen sker framför allt i form av tabeller och diagram. För en mer ingående beskrivning av redovisningen, se avsnitt 3.2

*Redovisningsförfaranden*. Förutom den ordinarie redovisningen görs olika specialbearbetningar och analyser som presenteras i olika former.

Mikrodata finns tillgängligt för ytterligare bearbetningar och analyser. För mer information om hur statistiken tillgängliggörs, se avsnitt 3.2

*Redovisningsförfaranden*.

### 1.4 Dokumentation och metadata

Valdeltagandeundersökning finns dokumenterad i enlighet med SCB:s dokumentationssystem. Dokumentationen finns tillgängliggjord på SCB:s webbplats. Där finns dokumentation av undersökningens kvalitet, vilket bland annat innefattar redovisningar av bortfall, med mera.

Dokumentationen innefattar även dokumentation av metadata för de register som tas fram inom ramen för undersökningen.

## 2 Uppgiftsinsamling

### 2.1 Ram och ramförfarande

Valdeltagandeundersökningen 2003 är en urvalsundersökning som består av fem urvalsdelar: ett urval från SCB:s arbetskraftsundersökning (AKU), ett kompletterande urval av äldre personer, ett urval med utländska medborgare och ett urval med utlandssvenskar. För mer information om urvalsdelarna, se avsnitt 2.2 Urvalsförfarande.

Den övergripande ramen som används i undersökningen är den preliminära röstlängden som tas fram 30 dagar innan valdagen. I röstlängden finns de personer som är röstberättigade i respektive val. Alla de urval som används i undersökningen antingen dras ur eller matchas mot röstlängden.

Urvalsramen för AKU utgörs av SCB:s register över totalbefolkningen (RTB) kompletterat med uppgifter om sysselsättningsstatus från Sysselsättningsregistret (SREG). Urvalsramen för äldreurvalet och urvalet av förstagångsväljare i riksdagsval är också RTB. Dessa urval matchas den preliminära röstlängden. Urvalsramen för utlandssvenskar är den preliminära röstlängden. Urvalet av utländska medborgare dras från den preliminära röstlängden efter att den har kompletterats med information om medborgarskapsland. Den informationen hämtas från RTB.

### 2.2 Urvalsförfarande

För att täcka in samtliga röstberättigade baseras valdeltagandeundersökningen på ett antal delurval.

Valdeltagandeundersökningen år 2003 består av 4 delurval. Dessa beskrivs nedan:

1. Urvalen från Arbetskraftsundersökningens undersökningsomgångar i augusti, september och oktober 2003. Urvalen från Arbetskraftsundersökningarna innefattar till skillnad från många andra undersökningsomgångar även röstberättigade utländska medborgare. Totalt sett omfattar den här urvalsdelen 63 100 röstberättigade i Sverige i åldrarna 18-74 år. För mer information om urvalsförfarandet i Arbetskraftsundersökningarna, se information om den undersökningen på SCB:s webbplats, [www.scb.se](http://www.scb.se).
2. Ett urval om 3 000 svenska medborgare 75 år och äldre som har dragits från Registret över totalbefolkningen (RTB). Urvalet har matchats mot den preliminära röstlängden och de personer som inte fanns med i den preliminära röstlängden har betraktats som övertäckning. Urvalet är ett systematiskt urval från personnummersorterad RTB.
3. Ett urval om 9 300 i Sverige röstberättigade utländska medborgare som har dragits från den preliminära röstlängden. Innan urvalet har dragits har, från RTB, information om medborgarskapsland matchats på den preliminära röstlängden. Medborgarskapsländerna har grupperats i de två grupperna: 1) EU-länder, Norge och Island

- (4 700); 2) Övriga länder (4 600). Urvalet är draget som ett obundet slumpmässigt urval inom strata.
4. Ett urval om 1 200 röstberättigade svenska medborgare boende utomlands (utlandssvenskar) som dragits från den preliminära röstlängden. Urvalet är draget som ett obundet slumpmässigt urval.

Urvalsdelarna är i huvudsak ömsesidigt uteslutande, men det finns urvalsdelar som överlappar varandra. Det gäller urvalsdel 1 och 3 som båda innehåller utländska medborgare 18-74 år. Detta har dock hanterats genom negativ samordning för att undvika att urvalsdel 3 innefattar personer som ingår i urvalsdel 1.

## 2.3 Mätinstrument

Mätinstrumentet består, förutom arbetskraftsundersökningens intervjublankett, av blankett som skickas till länsstyrelserna för avprickning mot röstlängden (se bilaga).

## 2.4 Insamlingsförfarande

Underlaget till statistiken inhämtas huvudsakligen genom att SCB sänder länsstyrelserna underlag i form av listor på urvalspersoner. Länsstyrelserna går igenom de avprickade röstlängderna och markerar om personen röstat eller inte i respektive val. Dessutom noterar länsstyrelserna om urvalspersonen enligt röstlängden har förtidsröstat eller inte.

De avprickade blanketterna skannas in och logiska kontroller görs av datamaterialet. Vid saknade uppgifter eller vid ologiska värden görs återkontakter med berörd länsstyrelse och materialet rättas därefter upp. Samtliga länsstyrelser skickar in blanketter som är helt ifyllda. Undersökningen saknar bortfall i det avseendet. För bakgrundsvariabler som hämtas från Arbetskraftsundersökningen finns ett bortfall som motsvarar det i den undersökningen. Bortfallet i det avseendet uppgår till drygt 16 procent.

## 2.5 Databeredning

Det är endast ett fåtal variabler som samlas in i valdeltagandeundersökningen. I samband med skanningen görs vissa logiska kontroller, som exempelvis att en person som har förtidsröstat också har röstat i folkomröstningen, se avsnitt 2.4 *Insamlingsförfarande*. Variabler som hämtas från Arbetskraftsundersökningen (AKU) har kontrollerats i samband med datainsamlingen för den undersökningen, se dokumentation för den undersökningen på SCB:s webbplats.

Variabler från SCB:s Registret över totalbefolkningen (RTB) matchas på redan i samband med urvalsdragningen. Detta då vissa variabler, som civilstånd och medborgarskapsland, kan förändras över tid. Efter avslutad datainsamling och efter att datamaterialet har skannats och kontrollerats



matchas ytterligare variabler från olika register på. Det gäller variabler avseende inkomst från SCB:s Inkomst- och taxeringsregistret (IoT) och variabler avseende utbildning från SCB:s utbildningsregister. Dessutom hämtas variabler som avser intervjuvar i AKU för den urvalsdelen som hämtats därifrån och matchas på.

I samband med den statistiska bearbetningen skapas en del härledda variabler. Det gäller dels variabler som definierar aktuella redovisningsgrupper (se avsnitt *1.2 Statistiska målstorheter*), dels variabler som identifierar tillhörighet med avseende på den hjälpvektor som används vid kalibrering (se avsnitt *3.1 Skattningar: antaganden och beräkningsformler*). Vid de statistiska beräkningarna tas ett antal olika vikter fram och läggs till datamaterialet. Det gäller dels vikter som behövs för de skattningar som ligger till grund för de resultat som publiceras (se avsnitt *3.1 Skattningar: antaganden och beräkningsformler*), dels kalibrerade vikter som kan användas vid olika former av analyser (för information om kalibreringen se avsnitt *3.1 Skattningar: antaganden och beräkningsformler*). Det behandlade datamaterialet lagras sedan i form av ett slutgiltigt dataset som läggs i en databas i SQL-miljö.

### 3 Statistisk bearbetning och redovisning

Publiceringen av resultat från valdeltagandeundersökningen i samband med 2003 års folkomröstning om euron skedde den 31 mars 2004 i ett Statistiskt meddelande. Resultat från valdeltagandeundersökningen ingick även i rapporten *Folkomröstningen om euron*, vilken publicerades den 21 oktober 2004. Efter dessa publiceringar har SCB reviderat estimation och tabellplaner för valdeltagandeundersökningar i samband med allmänna val. Med anledning av det har även resultatframställningen för valdeltagandeundersökningen i samband med 2003 års folkomröstning setts över och reviderats vad gäller estimation och tabellplan.

Dessa revideringar har inte inneburit några större förändringar. Estimationen baseras i båda fallen på kalibrering mot kända totaler. Den beskrivning som ges här avser dock den reviderade estimationen då denna ligger till grund för de variabler som ingår i det dataset som SCB har dokumenterat. För beskrivning av metoden som ligger till grund för den ursprungliga publiceringen hänvisas till rapporten *Folkomröstningen om euron* (vilken finns tillgänglig via SCB:s webbplats).

Det reviderade estimationsförfarandet har använts för att ta fram resultat enligt en reviderad tabellplan som underlättar jämförelser mellan år (se avsnitt 3.2).

#### 3.1 Skattningar: antaganden och beräkningsformler

Urvalet till valdeltagarundersökningen består av ett AKU-urval och av tre tilläggsurval. AKU-urvalet innefattar röstberättigade medborgare i åldrarna 18-74 år. Tilläggsurvalen innefattar urval av utländska medborgare, utlandssvenskar samt äldre (75 år och äldre). Skattningar beräknas för redovisningsgrupper som dels definieras av variabler från RTB (Registret över totalbefolkningen) och dels av variabler från AKU.

För skattning i redovisningsgrupper definierade av RTB-variabler kan hela urvalet användas medan för skattningar i redovisningsgrupper definierade av AKU-variabler kan endast de svarande i AKU-delen av urvalet användas. Skattningar för de två typerna av redovisningarna (RTB- respektive AKU-variabler) görs på olika sätt eftersom den tillgängliga informationen, både variabler och observationer, skiljer sig åt.

##### 3.1.1 Sammanvägning av olika delurval

Två av delurvalen avser delvis samma population, se avsnitt 2.1 *Ram och ramförfarande*. Utländska medborgare är dragna både i AKU och i tilläggsurvalet av utländska medborgare. För att skattningarna ska bli korrekta måste vikterna för de utländska medborgarna i de båda delurvalen korrigeras med en faktor som är  $\leq 1$ .

För redovisning av RTB-variabler görs först en hopvägning av AKU-urvalen och därefter en hopvägning av alla urval. För redovisning baserad på AKU-variabler används endast de hopvägda AKU-urvalen.

### 3.1.1.1 Sammanvägning av års- och delurval i AKU

Urvalet från AKU består av urvalen till de tre mätmånaderna augusti, september och oktober 2003. AKU-urvalet består i sin tur av två urval, dels ordinarie AKU som omfattar åldrarna 18-64 dels ett tilläggsurval som omfattar åldrarna 65-75 år. Bägge urvalen är sammansatta av tre årsurval dragna åren 2001, 2002 och 2003.

Hela AKU-urvalet måste vägas ihop för att de ska kunna räknas upp till en population som motsvarar Sveriges befolkning i åldrarna 15-74 år.

Låt  $n_i$  vara antalet personer i urvalet som är dragna år  $i$  där  $i=1,2,3$  motsvarar urvalsåren 2001, 2002 och 2003. Varje individ tilldelas en årsvikt,  $p_i^{(ar)}$ , som anger hur stor andel urvalsåret  $i$  bidrar med till det totala urvalet. Årsvikten bildas genom:

$$p_i^{(ar)} = n_i / \sum_{i=1}^3 n_i$$

I ordinarie AKU ingick ett separat tilläggsurval i Malmö kommun i åldrarna 16-64 år. Dessa måste vägas ihop med ordinarie AKU. Låt  $n_q^A$  vara antalet i grundurvalet som tillhör kategori  $q$ , där  $q$  betecknar de som tillhör Malmö kommun och låt  $n_q^B$  vara motsvarande antal i tilläggsurvalet i Malmö kommun. För varje urvalsår  $i$ , bildas vikten:

$$v_{iqk} = \begin{cases} n_{iq}^A / (n_{iq}^A + n_{iq}^B) & \text{om individ } k \in \text{kategori } q \text{ i ordinarie AKU} \\ 1 & \text{om individ } k \notin \text{kategori } q \text{ i ordinarie AKU} \\ n_{iq}^B / (n_{iq}^A + n_{iq}^B) & \text{om individ } k \in \text{tilläggsurvalet} \\ 1 & \text{om individ } k \in \text{AKU - urvalet 65 - 74 år} \end{cases}$$

Slutligen bildas vikter,  $u$ , som väger ihop AKU-urvalen. Årsvikten  $p_i^{(ar)}$  multipliceras med motsvarande v-vikt och kallas för  $u$ -vikt, dvs.

$$u_k = p_i^{(ar)} \times v_{iqk}.$$

### 3.1.1.2 Sammanslagning av överlappande urvalsdelar

Två av de urvalsdelar som finns i 2003 års valdeltagandeundersökning avser delvis samma population och behöver vägas ihop för att skattningen ska avse populationen röstberättigade på ett korrekt sätt.

Röstberättigade utländska medborgare i åldrarna 18-74 år förekommer både i AKU och i tilläggsurvalet av utländska medborgare. Dessa grupper vägs ihop enligt följande. Låt:

$n_{q_1}^A$  vara antalet i AKU-urvalet som tillhör kategori  $q_1$ , där  $q_1$  betecknar utländska medborgare i åldrarna 18-64 år med rösträtt i valet.

$n_{q_1}^B$  vara antalet i tilläggsurvalet av utländska medborgare som tillhör kategori  $q_1$ , där  $q_1$  betecknar utländska medborgare i åldrarna 18-64 med rösträtt i valet.

$n_{q_2}^C$  vara antalet i AKU-urvalet som tillhör kategori  $q_2$ , där  $q_2$  betecknar utländska medborgare i åldrarna 65-74 år med rösträtt i valet.

$n_{q_2}^D$  vara antalet i tilläggsurvalet av utländska medborgare som tillhör kategori  $q_2$ , där  $q_2$  betecknar utländska medborgare i åldrarna 65-74 år med rösträtt i valet.

$$v_{kq} = \begin{cases} n_{q_1}^A / (n_{q_1}^A + n_{q_1}^B) & \text{om individ } k \in \text{kategori } q_1 \text{ i AKU - urvalet} \\ n_{q_1}^B / (n_{q_1}^A + n_{q_1}^B) & \text{om individ } k \in \text{kategori } q_1 \text{ i tilläggs - urvalet} \\ n_{q_2}^C / (n_{q_2}^C + n_{q_2}^D) & \text{om individ } k \in \text{kategori } q_2 \text{ i AKU - urvalet} \\ n_{q_2}^D / (n_{q_2}^C + n_{q_2}^D) & \text{om individ } k \in \text{kategori } q_2 \text{ i tilläggs - urvalet} \\ 1 & \text{om individ } k \notin \text{kategori } q_1 \text{ eller } q_2 \text{ i vare sig} \\ & \text{AKU - urvalet eller tillägg - survalet} \end{cases}$$

$v_{kq}$  multipliceras med  $u$ -vikten, dvs.  $q_k = u_k \times v_{kq}$  där  $u_k$  är  $u$ -vikten för individ  $k$ . Vikterna  $q_k$  används vid beräkning av de slutgiltiga vikterna, se nästkommande avsnitt.

### 3.1.2 Estimation utifrån RTB-variabler

I detta avsnitt ges en kortfattad beskrivning av en GREG-estimator som är en slags regressionsestimator, vilken används vid estimation där redovisningsgrupper är definierade utifrån RTB-variabler. För en mer utförlig beskrivning av regressionskattningar hänvisas till Särndal C-E., Swensson B. och Wretman J. (1992).

#### 3.1.2.1 Regressionsestimatorn för RTB-variabler

Den generaliserade regressionsestimatorn, GREG, är en skattningsprocedur som utnyttjar hjälpinformation i estimationsfasen. Idén med att utnyttja hjälpinformation bygger på att hjälpvariabler samvarierar med undersökningsvariabeln. Utnyttjandet av hjälpinformation syftar till att reducera urvals- och bortfallsfelet. Regressionsestimation innebär att för urvalsobjekten  $k$  observeras  $(y_k, \mathbf{x}_k)$  där  $y_k$  är ett mätvärde från undersökningen medan  $\mathbf{x}_k$  är en vektor med hjälpinformation vars populationstotal  $\mathbf{t}_x = \sum_U \mathbf{x}_k$  är känd.

Från en population  $U$  bestående av  $N$  individer dras ett slumpmässigt urval  $s$  av storleken  $n_s$  enligt designen  $p(\cdot)$  som är sådan att alla individer har en sannolikhet att komma med i urvalet som är känd och större än noll.

I valdeltagandeundersökningen innebär designen  $p(\cdot)$  bl.a. att populationen delas in i  $H$  strata, där stratum  $h$  innehåller  $N_h$  individer. Inom varje stratum  $h$  dras ett slumpmässigt urval av storleken  $n_h$  så att alla individer inom strata har samma sannolikhet att komma med i urvalet. Om bortfall förekommer kan data om  $y$ -variabeln endast samlas in för en delmängd av storleken  $m_h$ . I valdeltagandeundersökningens RTB-del förekommer inget bortfall, varför  $m_h = n_h$ .

Inklusionssannolikheten,  $\pi_k$ , för objekt  $k$  är sannolikheten att objekt  $k$  ingår i urvalet. Under OSU ges inklusionssannolikheten för objekt  $k$  i stratum  $h$  av:

$$\pi_k = \frac{n_h}{N_h}$$

där  $N_h$  är antalet i stratum  $h$  i populationen och  $n_h$  är urvalsstorleken i stratum  $h$ . Designvikten definieras som inversen till inklusionssannolikheten. Under OSU är designvikten för objekt  $k$ :

$$d_k = \frac{N_h}{n_h}$$

Genom att utnyttja GREG-estimatoren transformeras designvikten  $d_k$  till en kalibrerad vikt. När den kalibrerade vikten används för att skatta totalerna för hjälpvariablerna erhålls de exakta populationstotalerna som är givna i hjälptotalen.

En regressionsestimator för totalen  $t_y = \sum_U y_k$  kan skrivas i formen:

$$\hat{t}_y = \sum_s w_k y_k \quad (3.1.1)$$

$\hat{t}_y$  = Skattning av en total, t.ex. antalet röstberättigade.

$s$  = Urvalsmängden, vid bortfall ersätts  $s$  med svarsmängden  $r$ .

$y_k$  = Värdet av variabeln  $y$  för objekt  $k$ .

$w_k$  = Uppräkningsvikt som beror av både urvalsdessignen och hjälpvektorn

$\mathbf{x}_k$

$$w_k = g_k \times d_k$$

$$g_k = 1 + (\mathbf{t}_x - \hat{\mathbf{t}}_x)' \left( \sum_r \frac{\mathbf{x}_k \mathbf{x}_k' q_k}{\pi_k} \right)^{-1} \mathbf{x}_k q_k \quad (3.1.2)$$

$g_k$  kan ses som en korrektionsfaktor till  $d_k$  som kan minska urvalsfelet om hjälpinformationen i  $\mathbf{x}$  samvarierar med  $y$ .

$\mathbf{t}_x = (t_{x1}, \dots, t_{xj}, \dots, t_{xJ})'$  är en vektor av längden  $J$  som innehåller kända totaler från t.ex. ett register.

$\hat{\mathbf{t}}_x = (\hat{t}_{x1}, \dots, \hat{t}_{xj}, \dots, \hat{t}_{xJ})'$  är en vektor som innehåller skattningar av elementen i vektorn  $\mathbf{t}_x$  där skattningen av varje element  $t_{xj}$  ges av  $\hat{t}_x = \sum_s d_k x_k$

$\mathbf{x}_k = (x_{1k}, \dots, x_{jk}, \dots, x_{Jk})'$  är en vektor av längden  $J$  där  $J$  är antalet hjälpvariabler.

$q_k$  är en känd konstant, se avsnitt

### 3.1.1.1 Sammanvägning av års- och delurval i AKU

Variansen för  $\hat{t}_y$  skattas med:

$$\hat{V}(\hat{t}_y) = \sum_h \frac{N_h^2}{n_h} \left(1 - \frac{n_h}{N_h}\right) s_{eh}^2; \quad s_{eh}^2 = \frac{1}{n_h - 1} \sum_{s_h} \left( g_k e_k - \frac{\sum_{s_h} g_k e_k}{n_h} \right)^2 \quad (3.1.3)$$

där  $s_h$  är urvalsmängden i stratum  $h$ , när bortfall förekommer ersätts  $s_h$  med svarsmängden  $r_h$  och  $n_h$  ersätts med  $m_h$ .

$$e_k = y_k - \hat{\mathbf{B}}' \mathbf{x}_k, \quad \hat{\mathbf{B}} = \left( \sum_s d_k \mathbf{x}_k \mathbf{x}_k' q_k \right)^{-1} \sum_s d_k \mathbf{x}_k y_k q_k \quad (3.1.4)$$

I statistikpubliceringen redovisas också mått som ger information om osäkerheten i skattningarna, som beror på att de baserar sig på ett urval (och inte hela populationen). Måtten består av 95-procentiga konfidensintervall som definieras på följande sätt:

$$\hat{t}_y \pm 1,96 \sqrt{\hat{V}(\hat{t}_y)}$$

Såväl  $\hat{t}_y$  som  $\hat{V}(\hat{t}_y)$  beräknas i alla redovisningsgrupper definierade av RTB-variabler med SAS-programmet CLAN97, se Andersson och Nordberg (1998).

### 3.1.2.2 Konstruktion av hjälpvektor för RTB-variabler

Varje individ tilldelas värden på hjälpvektorn,  $\mathbf{x}_k$ . Både hjälpvektorn och dess totaler hämtas från den definitiva röstlängden. Tabell 1 nedan visar vilka hjälpvariabler som har använts.

I folkomröstningen är hjälpvektorn konstruerad utifrån variablerna *kön*, *åldersklass*, *medborgarskap*, *bosättningsort*, *röstat/ej röstat* och *län*. Hjälpvektorn har formen *kön*×*åldersklass*×*svensk medborgare i Sverige/utländsk medborgare/utlandssvensk* + *röstat/ej röstat*×*län* och innehåller 72 element (2×5×3+2×21). De första 30 positionerna består av 29 nollor och en etta där ettan pekar ut vilken kombination av kön, åldersklass, medborgarskap/bosättnings-ort som individ  $k$  tillhör. De 42

sista positionerna pekar ut i vilket län individ  $k$  tillhör samt om personen har röstat eller inte.

**Tabell 1. Hjälpvariabler i folkomröstningen**

Variabel	Kategorier
Kön	Man, Kvinna
Ålder	18-29, 30-49, 50-64, 65-74, 75- år
Medborgarskap	Svensk medborgare i Sverige, Utländsk medborgare, Utlandssvensk
Län	21 län
Valdeltagande	Röstat/ej röstat i folkomröstningen

### 3.1.3 Estimation utifrån AKU-variabler

Vissa redovisningsgrupper definieras med hjälp av variabler i AKU, t.ex. arbetskraftsstatus. I de fallen baserar sig viktberäkningen och skattningarna på den del i svarmängden i AKU som tillhör målpopulationen i vår undersökning, dvs. 18-74 år och röstberättigad i folkomröstningen. Här används en tvåfas-ansats där första fasen är det urval som sker från befolkningen till AKU och den andra fasen utgörs av de som har svarat i AKU. Då uppgifterna om valdeltagandet omfattar samtliga individer i urvalet i första fasen kan det konstateras att det är betydligt lägre valdeltagande bland bortfallet i AKU än bland de som svarat i AKU. Denna information används i estimationen för att minska bortfallsfelet.

I tvåfas-ansatsen görs två kalibreringar. Den första fasens kalibrering använder hjälpinformation vars totaler är kända för hela populationen, dvs. antalet röstberättigade. Eftersom skattningen här avser valdeltagande i åldersgruppen 18-74 år kan inte informationen om antalet röstande användas i hjälpvektorn. Den andra fasens kalibrering använder hjälpinformation vars totaler finns för hela urvalet. Andra fasens kalibrering tar även hänsyn till de vikter som första fasens kalibrering genererat.

AKU är i sig sammansatt av tre delurval som är dragna under tre år. Dessa måste vägas ihop för att de ska räknas upp till en population som motsvarar Sveriges befolkning i åldrarna 15-74 år. I första fasen är detta redan gjort, läs mer om det i avsnittet

*3.1.1.1 Sammanvägning av års- och delurval i AKU.* För fas två görs detta med exakt samma tillvägagångssätt med enda skillnaden att här vägs endast samman de svarande,  $m$ , i AKU. Antalen  $n_{ij}$ ,  $i=1, 2, 3$ ,  $j=1, 2$ , ersätts alltså med  $m_{ij}$ .

### 3.1.3.1 Kalibreringsestimatoren för AKU-variabler

Vid estimationen utnyttjas två uppsättningar hjälpinformation, dels den del som är känd för hela populationen, dels den del som bara är känd för urvalet. Varje uppsättning hjälpinformation kan utnyttjas för att beräkna  $g$ -vikter av samma typ som i formel (3.1.2).

Låt  $g_{1k}$  vara vikten som erhålls genom att utnyttja hjälpinformationen i fas ett, och  $g_{2k}$  var vikten som erhålls genom att utnyttja hjälpinformationen i fas två. Den slutliga vikten erhålls då som,

$$w_k = d_k \times g_{1k} \times g_{2k}.$$

Variansestimatoren består av summan av två komponenter: den första från fas ett och den andra från fas två. Komponenterna är funktioner av residualerna som erhålls från användningen av hjälpinformationen i respektive fas.

En mer ingående beskrivning av estimationsproblemet i denna situation finns i Estevao och Särndal (2002). För estimation har programvaran SAS används (sas-macrot som inom SCB benämns TOPECA har använts).

### 3.1.3.2 Konstruktion av hjälpvektor för AKU-variabler

Ålder är här klassindelad i följande fyra grupper då AKU endast omfattar befolkningen upp till 74 år; 18-29, 30-49, 50-64, 65-74.

Hjälppvariablerna för skattningarna av deltagande i valet för AKU-variablerna består i fas ett av hjälpvektorn *röstberättigade per län + kön × åldersklass × svensk medborgare/utländsk medborgare* för röstberättigade medborgare i åldrarna 18-74 år. Antal element i denna hjälpvektor är 29 ( $21+2 \times 4 \times 2$ ). I fas två används samma hjälpvariabler som i fas ett, tillsammans med informationen om personen har röstat eller ej, för alla som kom med i urvalet till AKU. Hjälpvektorn i fas två består av: *röstberättigad per län + kön × åldersklass × röstat* och innehåller 37 element ( $21+2 \times 4 \times 2$ ).

Antalet röstande används som hjälpinformation vid estimation baserad på RTB-variabler, men kan inte användas här då den informationen inte är känd för den subpopulation som är aktuell här, dvs. röstberättigade medborgare folkbokförda i Sverige i åldrarna 18-74 år.

### 3.1.4 Referenser

Särndal, C-E., Swensson, B. och Wretman, J., (1992). *Model Assisted Survey Sampling*. New York: Springer Verlag.

Andersson, C. och Nordberg, L. (1998). *A User's Guide to CLAN97 – a SAS program for computation of point- and standard error estimates in sample surveys*. Stockholm: Statistiska centralbyrån.



Estevao, V.M. och Särndal, C.E. (2002). "The Ten Cases of Auxiliary Information for Calibration in Two-Phase Sampling", *Journal of Official Statistics*, nummer 18, sidorna 233-255.

### 3.2 Redovisningsförfaranden

I den ordinarie publiceringen redovisas samtliga punktskattningar för grupperna män, kvinnor och alla.

De punktskattningar som redovisas i form av tabeller är;

- a) Röstberättigade i 1000-tal
- b) Röstande i 1000-tal
- c) Röstande i procent av röstberättigade (valdeltagandet)

Skattningen i c) redovisas med tillhörande 95%-igt konfidensintervall.

Den ordinarie publiceringen av resultaten skedde den 31 mars 2004 i ett Statistiskt meddelande. Resultat från valdeltagandeundersökningen ingick även i rapporten *Folkomröstningen om euron*, vilken publicerades den 21 oktober 2004. Dessutom publicerades pressmeddelanden i samband med att dessa publikationer offentliggjordes. De publicerade resultaten ingår i Sverige officiella statistik.

Tabellplanen för SCB:s valdeltagandeundersökningar har efter publiceringen av resultaten från undersökningen i samband med 2003 års folkomröstning setts över och reviderats. Med anledning av detta har en ny uppsättning resultattabeller tagits fram för publicering på SCB:s webbplats. Dessa tabeller liknar till stor del de tabeller som publicerades i samband med offentliggörande av undersökningens resultat, men de ger förbättrade möjligheter till jämförelser med resultat från SCB:s valdeltagandeundersökningar i avseende deltagande i val till Europaparlament, riksdag, landstings- och kommunfullmäktige.

Förutom denna resultatredovisning kommer materialet att användas för olika specialbearbetningar och analyser. Resultaten från sådana bearbetningar kan komma att redovisas på olika sätt beroende på vad som för det givna fallet är mest lämpligt.

## 4 Slutliga observationsregister

### 4.1 Produktionsversioner

I det här dokumentet (SCBDOK) har framtagningen av nedanstående slutliga observationsregister beskrivits.

<b>Register</b>	Folkomröstningar, valdeltagande
<b>Registervariant</b>	Valdeltagande
<b>Registerversion</b>	2003

Fortsatt dokumentation, av registrens detaljerade innehåll, finns på SCB:s webbplats. Där beskrivs alla variabler och värdemängder m.m. Dokumentationen hittar du här: <https://www.h2.scb.se/metadata> . Klicka dig fram med hjälp av namnen på Register, Registervariant och Registerversion som är angivna i ovanstående tabell.


### 4.2 Arkiveringsversioner

Ej aktuellt.

### 4.3 Erfarenheter från senaste undersökningsomgången

Arbetet gick enligt plan.

## Bilaga 1: Frågeformulär till Länsstyrelsen

SCB		Statistiska centralbyrån		Folkomröstning EMU 2003		Markera	
		Statistics Sweden				så <input checked="" type="checkbox"/> 	
				1			
Län	Kom	Dist	Nr i längd	Personnummer	Ej i längd <input type="checkbox"/>	Röstat med fönsterkuvert eller med brevöst (P, V el. B)	<input type="checkbox"/>
08	21	0101	248	000000-1234			
MÅNSSON MONA						Deltagit: Ja <input type="checkbox"/>	Nej <input type="checkbox"/>
BRUKSG 19				321177			
570	75	FÅGELFORS				Not:	
Län	Kom	Dist	Nr i längd	Personnummer	Ej i längd <input type="checkbox"/>	Röstat med fönsterkuvert eller med brevöst (P, V el. B)	<input type="checkbox"/>
08	21	0101	251	000000-1234			
ANDRESSON ANDERS						Deltagit: Ja <input type="checkbox"/>	Nej <input type="checkbox"/>
TRÅNSHULT 102				108773			
579	92	HÖGSBY				Not:	
Län	Kom	Dist	Nr i längd	Personnummer	Ej i längd <input type="checkbox"/>	Röstat med fönsterkuvert eller med brevöst (P, V el. B)	<input type="checkbox"/>
08	21	0101	284	000000-1234			
GUNNARSSON GUN						Deltagit: Ja <input type="checkbox"/>	Nej <input type="checkbox"/>
EKEDALSVÄGEN 1				133634			
570	75	FÅGELFORS				Not:	
Län	Kom	Dist	Nr i längd	Personnummer	Ej i längd <input type="checkbox"/>	Röstat med fönsterkuvert eller med brevöst (P, V el. B)	<input type="checkbox"/>
08	21	0101	458	000000-1234			
JANSSON JAN						Deltagit: Ja <input type="checkbox"/>	Nej <input type="checkbox"/>
SKÄRVETVÄGEN 1				137305			
570	75	FÅGELFORS				Not:	
Län	Kom	Dist	Nr i längd	Personnummer	Ej i längd <input type="checkbox"/>	Röstat med fönsterkuvert eller med brevöst (P, V el. B)	<input type="checkbox"/>
08	21	0101	464	000000-1234			
PERSSON PER						Deltagit: Ja <input type="checkbox"/>	Nej <input type="checkbox"/>
TJÄRDALSVÄGEN 2				128593			
570	75	FÅGELFORS				Not:	
Län	Kom	Dist	Nr i längd	Personnummer	Ej i längd <input type="checkbox"/>	Röstat med fönsterkuvert eller med brevöst (P, V el. B)	<input type="checkbox"/>
08	21	0101	486	000000-1234			
ULLANDER ULLA						Deltagit: Ja <input type="checkbox"/>	Nej <input type="checkbox"/>
KORNVÄGEN 6				318699			
570	75	FÅGELFORS				Not:	