

# STATISTIKENS FRAMSTÄLLNING

## Elevpaneler för longitudinella studier

### Ämnesområde

Utbildning och forskning

### Statistikområde

Befolkningens utbildning

### Produktkod

UF0501

### Referenstid

Vårterminen 2020

### Kontaktuppgifter

<b>Statistikansvarig myndighet</b>	Statistiska centralbyrån
<b>Kontaktinformation</b>	Tobias Toom
<b>E-post</b>	<a href="mailto:fornamn.efternamn@scb.se">fornamn.efternamn@scb.se</a>
<b>Telefon</b>	010-479 65 41

## Innehåll

1	Statistikens sammanhang.....	3
2	Undersökningsdesign .....	3
2.1	Målstorheter .....	3
2.2	Ramförfarande .....	3
2.3	Förfaranden för urval och uteslutning .....	3
2.3.1	Urvalsförfarande.....	3
2.3.2	Uteslutning från insamling (cut-off) .....	4
2.4	Insamlingsförfarande.....	4
2.4.1	Datainsamling .....	4
2.4.2	Mätning.....	4
2.4.3	Bortfallsuppföljning .....	5
2.5	Bearbetningar.....	5
2.6	Granskning.....	6
2.6.1	Granskning under direktinsamlingen .....	6
2.6.2	Granskning av mikrodata och insamlade statistikvärden.....	6
2.6.3	Granskning av makrodata .....	7
2.6.4	Granskning av redovisning .....	7
2.7	Skattningsförfarande .....	7
2.7.1	Principer och antaganden .....	7
2.7.2	Skattningsförfarande för målstorheter.....	7
2.7.3	Skattningsförfarande för tillförlitlighet.....	9
2.7.4	Röjandekontroll .....	9
3	Genomförande .....	9
3.1	Kvantitativ information .....	9
3.2	Avvikelser från undersökningsdesignen .....	9

## 1 Statistikens sammanhang

Statistiken beskriver hur förekomsten av olika utbildningsinsatser t.ex. stödundervisning och modersmålsundervisning har förändrats över tid och under elevernas tid i grundskolan.

Statistiken bygger på insamlade uppgifter från skolenheter för ett urval av elever (en panel) som följs från årskurs 3 och genom grundskolan.

I detta dokument beskrivs översiktligt design och genomförande av statistiken om Elevpaneler för longitudinella studier. Dokumentet kompletterar *koalitetsdeklarationen* för Elevpanelerna som finns tillgänglig på [www.scb.se/UF0501](http://www.scb.se/UF0501) under rubriken *Dokumentation*.

## 2 Undersökningsdesign

### 2.1 Målstorheter

De statistiska målstorheterna i statistiken är summor och andelar som visar hur vanligt förekommande utbildningsinsatserna är för elever i grundskolan. Med utbildningsinsatser menas åtgärdsprogram, stödåtgärder, modersmålsundervisning och svenska på andraspråk.

Målstorheterna avser förekomst av utbildningsinsatser för elever som fanns i svensk grundskola i årskurs 3 och som vid mättillfället fortfarande var inskrivna.

### 2.2 Ramförfarande

Ramen konstruerades utifrån Skolverkets register avseende elever i grundskolan 15 oktober 2013. Ramen består av skolenheter med minst en elev i årskurs 3 enligt grundskolans elevregister. Rampopulationen består av de elever som fanns registrerade på aktuella skolenheter. Från fem skolor där personuppgifter saknades för eleverna samlades klasslistor in, vilka tillsammans med registeruppgifterna låg till grund för det slutliga elevurvalet. Eleverna utgör observationsobjekten och skolenheterna är uppgiftskällor. På skolenheterna är det rektorn som tillfrågas om uppgifterna och kan själv lämna dem eller delegera uppgiftslämnandet till lämplig person. Alla uppgifter för observationsobjekten på den aktuella enheten lämnas vid samma tidpunkt.

### 2.3 Förfaranden för urval och uteslutning

#### 2.3.1 Urvalsförfarande

Urvalet till årets undersökningsomgång i panelen drogs inför starten av panelen år 2014. Urvalet drogs som stratifierat urval av skolenheter som hade elever i årskurs 3 den 15 oktober 2013 med

urvalssannolikhet proportionell mot redovisat antal elever enligt grundskolestatistiken för läsåret 2013/2014. Stratifiering gjordes efter typ av skolhuvudman:

1. Skolor med kommunen som huvudman
2. Skolor med annan huvudman/friskolor

Urvalssannolikheterna fastställdes så att 15 procent av urvalet skulle avse elever i friskolor. Eftersom endast ca 10 procent av samtliga elever i årskurs 3 går i en friskola innebär det en väsentligt förhöjd urvalssannolikhet för friskolor

Eftersom friskolorna har i genomsnitt färre elever per skola än de kommunala skolorna blev resultatet av urvalsdragningen att 6 procent av de kommunala skolorna och 9 procent av de berörda friskolorna ingår i urvalet.

### **2.3.2 Uteslutning från insamling (cut-off)**

Uteslutna från målpopulationen är alla sameskolor. Dessa hade vid urvalsdragningen totalt 19 elever.

## **2.4 Insamlingsförfarande**

### **2.4.1 Datainsamling**

Statistiken baseras på uppgifter insamlade direkt från skolorna. Svaren kompletteras med registeruppgifter från SCB:s utbildningsregister och Registret över totalbefolkningen (RTB) samt Skolverkets skolenhetsregister.

Direktinsamling sker via frågeformulär på webben som fylls i av respektive skola. I formuläret förtrycks elevernas namn och personnummer, samt klassbeteckning föregående terminer, vilket gör att det framgår för vilka elever som uppgifterna ska fyllas i.

Alternativet hade varit insamling via telefon men givet de efterfrågade uppgifternas karaktär bedömdes frågeformulär på webben vara en bättre och mer kostnadseffektiv lösning.

Inför den första kontakten efter att en ny panel upprättats sänds informationsbrev till berörda skolenheter och genom dessa även till målsmän för samtliga elever. Varje vår fram till att majoriteten av eleverna förväntas ha gått ut årskurs 9 skickas sedan informationsbrev med inloggningsuppgifter ut till skolans rektor.

### **2.4.2 Mätning**

Vid uppgiftsinsamlingen från skolorna, som genomförs under våren för referensåret, ställs frågor om förekomst av olika utbildningsinsatser t.ex. stödundervisning och modersmålsundervisning.

Förtryckta uppgifter i frågeformuläret är elevens namn och personnummer samt klassbeteckning föregående vårtermin och höstterminen aktuellt läsår. Alla uppgifter i webbformuläret, inklusive de som är förtryckta kan uppdateras om de är felaktiga eller saknas.

Enkäten är uppbyggt så att endast alternativ som är aktuella ska kryssas i, dvs. man ska inte kryssa i att ett alternativ inte är aktuellt. Det finns också möjlighet att lämna en kommentar för varje elev om något skulle vara fel.

Om eleven inte längre finns på den aktuella skolan efterfrågas även följande uppgifter.

- Lämnat skolan (*Orsak anges*)
- Lämnat skolan tidpunkt
- Ny grundskola (skolenhetskod)
- Ny grundskolekommun (kommunkod)

Information om de slutliga observationsvariablerna och statistikens detaljerade innehåll finns på SCB:s webbplats. Där beskrivs alla variabler och värdemängder m.m. Dokumentationen finns på [www.metadata.scb.se](http://www.metadata.scb.se) under *Elevpaneler*.

Det går även i kvalitetsdeklarationen för Elevpaneler hitta information om observationsvariablerna.

#### **2.4.3 Bortfallsuppföljning**

Objektsbortfallet utgörs av elever där skolenheter inte besvarat enkäten eller där eleven inte ingår i skolans rapportering, t.ex. på grund av att eleven flyttat och den nya skolan ej kunnat nås eller att elevens vårdnadshavare anmält att eleven inte skall ingå i undersökningen. Påminnelsebrev skickades ut vid tre tillfällen. Dessa åtgärder är fasta. Telefonpåminnelser görs i slutet av insamlingen vid behov. Denna åtgärd är adaptiv.

#### **2.5 Bearbetningar**

Bortfall respektive övertäckning kodas utifrån information från uppgiftslämnaren själv eller RTB. Det kan t.ex. vara när en elev rapporterats flyttat utomlands eller enligt RTB har skyddad identitet. I det första fallet kodas eleven bort som övertäckning och i det andra som bortfall och de ingår inte i redovisningen.

Variabeln *stödundervisning* avser om eleven under vårterminen 2020 haft stödundervisning i form av anpassad studiegång, särskild undervisningsgrupp och/eller specialundervisning på annat sätt och härleds från dessa variabler.

Uppgift om kön, samt född i Sverige/ utomlands hämtades från RTB i samband med att panelen startades år 2014. Den sistnämnda uppgiften härleddes då från födelseland. Uppgift om föräldrars utbildning hämtades från Registret över befolkningens utbildning (Utbildningsregistret) i samband med publiceringen år 2014. Aktualitet var 2013-12-31. I de fall uppgift saknades i den registerversionen användes år 2015 versionen med aktualitet 2014-12-31. I de fall uppgift saknades i båda dessa versioner användes versionen med aktualitet 2015-12-31 följt av 2016-12-31, 2017-12-31 eller 2018-12-31. Om uppgift saknades i alla dessa tidigare versioner användes version 2019-12-31.

Variabeln skolhuvudman hämtades från Skolenhetsregistret med aktualitet maj 2020.

Variabeln årskurs avser elevernas skolår eller den årskurs eleverna antas befinna sig i utifrån att de fanns i årskurs 3 den 15 oktober 2013.

För varje svarande elev har en vikt beräknats. Vikterna har beräknats utifrån urvalsdesignen samt antaganden om objektsbortfall och täckningsfel. Läs mer i avsnitt 2.7.2 *Skattningsförfarande för målstorheter*.

## **2.6 Granskning**

De insamlade uppgifterna granskas under insamlingen, efter avslutad insamling samt i aggregerad form.

### **2.6.1 Granskning under direktinsamlingen**

Insamlingen skedde med hjälp av en webblankett i vilken det finns inbyggda kontroller. Uppgiftslämnaren får vid ifyllandet uppfelmeddelanden i de fall frågeformuläret inte fyllts i på ett korrekt eller logiskt sätt. Ibland måste felen rättas upp för att uppgiftslämnaren ska kunna skicka in blanketten och ibland räcker det med att uppgiftslämnaren lämnar en kommentar.

### **2.6.2 Granskning av mikrodata och insamlade statistikvärden**

När uppgiftslämnaren skickat in lämnade uppgifter har dessa sedan bearbetats, granskats och rättats. SCB kontrollerar inkomna uppgifter och i de fall ofullständiga uppgifter har lämnats, återkontakts uppgiftslämnaren.

Sedan kontrolleras exempelvis att alla värden som förekommer på de olika variablerna är godkända. Granskningen är mer djupgående för de variabler som ingår i den publicerade statistiken.

I de fall det förekommer oklarheter kring vilket värde en individ ska ha används ibland grundskolans elevregister samt RTB för samgranskning.

### 2.6.3 Granskning av makrodata

En granskning på makronivå görs av de variabler som ingår i den publicerade statistiken. Jämförelser görs med föregående års statistik och eventuella förändringar över tid undersöks mer ingående. Om en förändring skett används när det är möjligt grundskolans elevregister för samgranskning.

### 2.6.4 Granskning av redovisning

Alla tabeller och diagram läggs ut på intern server, och granskas igen innan de publiceras. Det kontrolleras att antals- och andelsuppgifter i SCB:s statistikdatabas samt i diagrammen på webbplatsen och i statistiknyheten stämmer överens med grunddata. Rubriker och förklaringar till tabeller och diagram granskas så att de är korrekta. Det sker även granskning efter publiceringen på webbplatsen.

## 2.7 Skattningsförfarande

### 2.7.1 Principer och antaganden

I syfte att skatta olika målstorheter i undersökningen används en modellassisterad estimator (regressionsestimator).

Bortfall och täckningsfel förekommer i undersökningen, vilket kan leda till ökad osäkerhet i skattningarna. Detta hanteras genom att använda hjälpinformation vid estimationen. Syftet med att använda hjälpinformation är att förbättra kvaliteten i skattningarna.

Estimatorn justerar för det täckningsfel som observerats under insamlingen. Kompensation för övertäckning görs genom antagandet att övertäckningen i hela urvalsramen är lika stor som den observerade övertäckningen bland svarande elever. Ingen kompensation för eventuell undertäckning görs.

### 2.7.2 Skattningsförfarande för målstorheter

Utifrån estimatorns egenskaper samt vald hjälpinformation, beräknas en vikt för varje svarande elev. Vikten kan beskrivas med formeln:

$$w_k = d_k^* v_k$$

där  $w_k$  = den totala vikten för elev  $k$

$d_k^*$  = designvikt justerad för bortfall av skolor

$v_k$  = justeringsvikt, baserad på hjälpinformation

Designvikten genereras från urvalet. Designvikten för elev  $k$  är densamma som skolans (den skola eleven tillhörde vid urvalstillfället) designvikt och kan beskrivas med formeln:

$$d_k = \frac{\sum_{U_h} x_i}{n_{lh} * x_i}$$

där

$\sum_{U_h} x_i$  = totalt antal elever i årskurs 3, stratum  $h$

$x_i$  = antal elever på skolan i årskurs 3 vid urvalstillfället

$n_{lh}$  = antal skolor i urvalet, stratum  $h$

Designvikten justeras för bortfall inom strata enligt:

$$d_k^* = \left( \frac{n_{lh}}{m_{lh}} \right) d_k$$

där  $m_{lh}$  är antal deltagande skolor (skolor som skickat in alla eller en delmängd av sina elever), stratum  $h$

I syfte att minska urvalsfel och bortfallsskevheter utnyttjas därefter hjälpinformation. Vi bildar en hjälpvektor  $\mathbf{x}_k$ , som anger till vilka kategorier av

*stratum + kön + åtgärdsprogram + födelseland + deltagande i svenska som andraspråk*

som elev  $k$  tillhör. Från register framställer vi hjälptotalerna  $\sum_U \mathbf{x}_k$ ,  
Vi utnyttjar denna hjälpinformation i en regressionsestimator.

Justeringsvikten beräknas enligt

$$v_k = 1 + (\sum_U \mathbf{x}_k - \sum_r d_k^* \mathbf{x}_k)' (\sum_r d_k^* \mathbf{x}_k \mathbf{x}_k')^{-1} \mathbf{x}_k$$

För att skatta en total, exempelvis antal elever med åtgärdsprogram, summeras vikterna för de elever i svarsmängden som har åtgärdsprogram:

$$\hat{Y} = \sum_r w_k y_k$$

där



$w_k$  = den totala vikten för elev  $k$

$y_k$  = 1 om elev  $k$  har åtgärdsprogram, 0 annars

Skattning av en kvot, exempelvis andel elever med åtgärdsprogram, görs genom att skatta varje total för sig och sedan dividera totalerna.

### 2.7.3 Skattningsförfarande för tillförlitlighet

Konfidensintervallen är beräknade enligt en metod som i 95 procent täcker in det sanna värdet och beräknas enligt formeln

punktskattning  $\pm 1,96 * (\text{skattning av estimatorns medelfel})$

där såväl punkt- som medelfelsskattning beräknas med hjälp av ett av SCB egenutvecklat program, ETOS2.

### 2.7.4 Röjandekontroll

Ingen röjandekontroll har gjorts utifrån bedömningen att inga skyddsmetoder behövs. Undersökningen är en urvalsundersökning och statistiken som publiceras är inte på en så pass detaljerad nivå att det bedömts föreligga risk för röjande.

## 3 Genomförande

### 3.1 Kvantitativ information

Ramen bestod av 3 670 skolor som hade 106 723 elever i årskurs 3.

Urvalet omfattade 243 av de totalt 3 670 skolor som hade elever i årskurs 3. Urvalet omfattade 190 kommunala skolor och 53 friskolor. Skolorna i urvalet hade enligt urvalsramen totalt 9 798 elever i årskurs 3, varav 1 508 elever (drygt 15 procent) går i friskolor. Senare konstaterades det korrekta antalet vara 9 794 elever.

Våren 2020 identifierades 225 stycken av urvalets 9 794 elever som övertäckning.

Det totala objektsbortfallet för undersökningen avseende våren 2020 var ca 26 procent (designvägd svarsandel).

### 3.2 Avvikelser från undersökningsdesignen

Inga avvikelser har gjorts.