

STATISTIKENS FRAMSTÄLLNING

Inträdet på arbetsmarknaden efter gymnasieskolan

Ämnesområde

Utbildning och forskning

Statistikområde

Befolkningens utbildning

Produktkod

UF0512

Referenstid

2023

Kontaktuppgifter

Statistikansvarig myndighet	Statistiska centralbyrån
Kontaktinformation	Sofia Larsson, Annica Wallerå
E-post	Sofia.larsson@scb.se , annica.wallerå@scb.se
Telefon	010-479 63 02, 010-479 62 57

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.....	4
2.3.2	Uteslutning från insamling (cut-off)	4
2.4	Insamlingsförfarande.....	4
2.4.1	Datainsamling	4
2.4.2	Mätning.....	5
2.4.3	Bortfallsuppföljning	5
2.5	Bearbetningar.....	6
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	6
2.6.4	Granskning av redovisning	6
2.7	Skattningsförfarande	7
2.8	Principer och antaganden	7
2.8.1	Skattningsförfarande för målstorheter.....	7
2.8.2	Skattningsförfarande för tillförlitlighet.....	8
2.8.3	Röjandekontroll	8
3	Genomförande	8
3.1	Kvantitativ information.....	8
3.2	Avvikelser från undersökningsdesignen	8

1 Statistikens sammanhang

Inträdet på arbetsmarknaden efter gymnasieskolan beskriver inträdet på arbetsmarknaden för personer som tagit examen från gymnasieskolan. Statistiken ger en samlad bild av vad unga personer gör tre år efter avslutad gymnasieutbildning. Statistiken beskriver främst den huvudsakliga sysselsättning, arbetets överensstämmelse med utbildningen och om arbetet motsvarar utbildningsnivån. Dessutom beskrivs förberedelse för arbetslivet och utbildningens nytta i arbetet.

I detta dokument beskrivs upplägg och genomförande av den undersökning som resulterar i statistik om inträdet på arbetsmarknaden efter gymnasieskolan. Läs om statistikens kvalitet i kvalitetsdeklarationen som finns tillgänglig på www.scb.se/UF0512, under rubriken *Dokumentation*.

2 Undersökningsdesign

2.1 Målstorheter

Några av målstorheterna i undersökningen är sysselsättning tre år efter avslutad utbildning, överensstämmelse mellan arbetet och det område utbildningen var inriktad mot och hur utbildningen förberett för arbete och studier på högskola.

2.2 Ramförfarande

Undersökningens observationsobjekt är gymnasieavgångna läsåret 2019/20 som avgått med examen.

Register över gymnasieskolans betygsuppgifter användes för att ta fram den aktuella rampopulationen som omfattade ca 80 500 personer med examen.

Rampopulationen för undersökningen är personer som enligt registret ovan, slutfört gymnasiet våren 2020 med en examen. Rampopulationen avgränsades även till folkbokförda i Sverige 2022-12-31.

Såväl observationsobjekt som målobjekt utgörs av individer. Individerna som svarar i undersökningen utgör uppgiftskällor.

Utöver de uppgiftskällorna används SCB:s interna register RTB och Registret över befolkningens utbildning (UREG) för uppgifter om föräldrars utbildningsnivå, ålder, kön mm. Registren är individregister och uppgifter hämtas för varje individ.

2.3 Förfaranden för urval och uteslutning

2.3.1 Urvalsförfarande

Från Skolverkets register över gymnasieavgångna drogs ett urval av 14 998 personer varav 11 494 med examen från ett yrkesprogram och 3 504 med examen från ett högskoleförberedande program.

Urvalsdesignen är ett stratifierat obundet slumpmässigt urval utan återläggning.

I den aktuella undersökningsomgången genomfördes förändringar i urvalsdesignen för att möjliggöra en mer detaljerad statistikredovisning bland examinerade från yrkesprogram. Den totala urvalsstorleken allokterades (fördelades) med en större andel examinerade från ett yrkesprogram. För examinerade från yrkesprogram stratifierades urvalet utifrån program och kön, där urvalsstorleken fördelades icke-proportionellt för att uppnå en godtagbar precision bland så många grupper som möjligt. För examinerade från högskoleförberedande program stratifierades urvalet utifrån kön, nationell bakgrund och föräldrarnas utbildningsnivå, där urvalsstorleken allokterades proportionellt mot varje elevgrupps relativa storlek med undantag för en urvalsförstärkning bland grupper med utländsk bakgrund för att uppnå tillräcklig precision.

Stratumindelningen görs utifrån flera variabler. Urvalet stratifieras i första hand efter examen och ej examen. Gruppen med examen stratifieras utifrån kön (kvinna och man), programtyp (högskoleförberedande program och yrkesprogram), bakgrund (svensk bakgrund och utländsk bakgrund), föräldrars utbildningsnivå (minst en förälder har eftergymnasial utbildning och ingen förälder har eftergymnasial utbildning) och behörighet till högskolestudier (behörig och ej behörig). Gruppen som ej har examen stratifieras efter kön (kvinna och man) och programtyp (högskoleförberedande program och yrkesprogram).

Urvalsförfarandet gjordes med utgångspunkten att med en svarsandel om minst 35 procent kan redovisningen av ovanstående grupper göras med ett konfidensintervall på som mest plus/minus 5 procentenheter.

2.3.2 Uteslutning från insamling (cut-off)

Ingen del av målpopulationen utesluts från direktinsamlingen.

2.4 Insamlingsförfarande

2.4.1 Datainsamling

Data samlades in direkt från de individerna genom en enkät (webb/papper). 91 procent svarade på webben och 9 procent via pappersblankett. Vissa uppgifter om personer, så som ålder, kön och föräldrars utbildningsnivå hämtas från RTB och UREG.

Uppgiftslämnaren nådde webbformuläret via en uppgiftslämnarsida. På uppgiftslämnarsidan kunde respondenten läsa mer om undersökningen och dess syfte.

Under den planerade insamlingsperioden gjordes fyra utskick. Det första utskicket var ett informationsbrev om undersökningen som innehöll inloggningsuppgifter till webbformuläret och en hänvisning till uppgiftslämnarsidan. Två veckor senare skickades ett tack- och påminnelsekort. Ytterligare två veckor senare gjordes ett tredje utskick då pappersenkäten skickades. Två veckor senare skickades det sista påminnelsebrevet ut. Den ursprungliga datainsamlingsperioden pågick från 2023-04-20 till och med 2020-06-11. Då svarsinflödet var lägre än beräknat skickades ytterligare en påminnelse insamlingsperioden förlängdes till 2023-07-30. Då insamlingen avslutades hade svar inkommit från 4 730 individer vilket motsvarar en svarsfrekvens på 31,6 %.

2.4.2 Mätning

Frågeformuläret bestod av 25 frågor. Webbformuläret och pappersformuläret hade samma utformning. Webbformuläret innehöll inga kontroller i syfte att inte riskera att ifrågasätta uppgiftslämnaren. Enkäten har under 2023 genomgått en mätteknisk granskning med syftet att minimera risken för mätfel.

Formuläret inleds med en fråga om huvudsaklig sysselsättning som är gemensamma för alla. Därefter följer formuläret olika vägar beroende på vad respondenten svarat angående om huvudsaklig sysselsättning.

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 under *Inträdet på arbetsmarknaden*. Det går även i Kvalitetsdeklarationen för Inträdet på arbetsmarknaden hitta information om observationsvariablerna.

2.4.3 Bortfallsuppföljning

Bortfallet mäts som andel uteblivna svar av samtliga i urvalet exklusive övertäckningen. För att hålla nere bortfallet skickades fyra påminnelser: två tack- och påminnelsekort och två ytterligare påminnelser skickades till dem som inte svarat.

För att ytterligare minska bortfallet användes belöning i form av en värdecheck för de som svarat. Uppgiftslämnaren kunde läsa om det i missivet och på uppgiftslämnarsidan.

2.5 Bearbetningar

Bortfall respektive övertäckning kodas utifrån information från uppgiftslämnaren själv eller RTB. Övertäckning kan exempelvis härröra från att individen avlidit eller flyttat utomlands.

Det partiella bortfallet bearbetas inte genom imputering. För objektsbortfall görs uppräknings från de svarande samt kalibrering i samband med estimation, se avsnitt 2.7.

2.6 Granskning

2.6.1 Granskning under direktinsamlingen

Registrering av data från frågeformulär på papper skedde genom att inkomna formulär lästes in maskinellt (skanning).

Skanningsprogrammet kontrollerar att korrekta värden registrerades vid skanningen, till exempel om svar var otydligt markerade. De värden som föll ut som fel vid dessa kontroller rättades upp manuellt med stöd av skanningsprogrammet, genom så kallad verifiering.

2.6.2 Granskning av mikrodata och insamlade statistikvärden

För ett fåtal poster förekom dubbelmarkeringar, det vill säga om fler svarsalternativ fanns angivna för en fråga som bara fick ha ett svar. Rättning av dubbelmarkeringar skedde utifrån svaren på andra frågor i frågeformuläret eller slumpmässigt mellan de olika svarsalternativen. Ytterligare kontroller som gjordes var bland annat att endast valida värden förekom i materialet samt att så kallade hoppinstruktioner hade efterföljts och att rätt antal svar för varje fråga hade lämnats.

2.6.3 Granskning av makrodata

Eftersom många totaler och deltotaler återkommer i flera av de tabeller som tas fram och publiceras, jämfördes antalsuppgifterna mellan tabellerna.

Kontroller gjordes även mellan tidigare undersökningar för att kontrollera rimligheten av värdena. Statistiken kan jämföras med tabeller för tidigare år från Inträdet på arbetsmarknaden, men det ska göras med försiktighet eftersom det finns olikheter mellan undersökningarna som kan ge skillnader i statistiken.

2.6.4 Granskning av redovisning

Inför publicering granskades samtliga framställda tabeller.

SCB kontrollerar att alla tabeller finns med och att ingen av dem är tom eller innehåller obegripliga värden, till exempel interna koder. Rubriker och förklaringar till tabeller granskas så att de är korrekta. Slutligen utfördes kontroller av överensstämmelse där samma siffervärde eller text används på flera ställen.

2.7 Skattningsförfarande

Målstorheterna i undersökningen skattas med hjälp av en medellassisterad estimator i form av en kalibreringsestimator (GREG). I skattningsförfarandet används hjälpinformation i form av registervariabler för att i möjligaste mån reducera de systematiska fel som uppkommit på grund av bortfall samt för att reducera skattningarnas varians.

2.8 Principer och antaganden

Följande hjälpvariabler har använts:

1. Kön * Föräldrarnas utbildningsnivå
2. Kön * Programtyp
3. Kön * Bakgrund
4. Meritvärde
5. Program
6. Programtyp * Sysselsättningsstatus

Dessa variabler har använts med primärt syfte att minska de skevheter som bortfallet kan föra med sig. Under antagandet att variablerna samvarierar med svarsbenägenheten kan eventuella skevheter reduceras.

Vikterna kompenserar för objektsbortfallet men inte för det partiella bortfallet. Kalibreringen (punktskattningar och konfidensintervall) och de tabeller som presenteras har tagits fram med hjälp av SAS och av SCB:s skattningsprogram ETOS.

2.8.1 Skattningsförfarande för målstorheter

Kalibreringsestimatorn är en skattningsprocedur som utnyttjar hjälpinformation, d.v.s. för populationen eller urvalet sedan tidigare kända variabler. Den bakomliggande idén är att hjälpinformation som samvarierar med undersökningsvariabler och/eller svarsbenägenhet har en god förmåga att reducera skevheter som uppkommer på grund av bortfall. Dessutom kan sådan hjälpinformation bidra till att reducera skattningarnas varians. Hjälpinformationen finns beskriven i avsnitt 2.7.1.

En kalibreringsestimator för en populationstotal $Y = \sum_U y_k$ kan skrivas

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

där r är svarsmängden, y_k är värdet på undersökningsvariabeln y för individ k och w_k är uppräkningsvikten för individ k . För varje individ i urvalet har vi tillgång till en vektor $\sum_U \mathbf{x}_k$ med hjälpinformation på populationsnivå. Denna används för att beräkna uppräkningsvikterna w_k enligt

$$w_k = 1 + \left(\sum_U \mathbf{x}_k - \sum_r d_k \mathbf{x}_k \right)' \left(\sum_r \mathbf{x}_k \mathbf{x}_k' \right)^{-1} \mathbf{x}_k$$

där d_k är designvikten för individ k (genererad av urvalsdesignen)

Dessa vikter kommer att uppfylla kalibreringsvillkoret

$$\sum_r w_k \mathbf{x}_k = \sum_U \mathbf{x}_k$$

och är således kalibrerade till hjälpinformationen på populationsnivå.

2.8.2 Skattningsförfarande för tillförlitlighet

Konfidensintervall är beräknade enligt en metod som i 95 procent täcker in det sanna värdet (under antagande att systematiska fel är försumbara) och beräknas enligt formeln:

$$\text{Punktskattning} \pm 1,96 \times \text{estimatorns medelfel}$$

där siffran 1,96 är konfidenskoefficienten som anger antalet standardavvikelser från medelvärdet.

2.8.3 Röjandekontroll

För att skydda enskilda individers svar redovisas endast statistiken på aggregerad nivå. Inga grupper redovisas som bygger på färre än 30 svar och inga celler redovisas som bygger på mindre än 3 svar.

3 Genomförande

3.1 Kvantitativ information

Undersökningens rampopulation bestod av ca 80 500 personer varav ca 55 400 med examen från ett högskoleförberedande program och ca 25 100 personer med examen från ett yrkesprogram. Urvalet bestod av 3 504 med examen från ett högskoleförberedande program och 11 494 med examen från ett yrkesprogram.

Objektsbortfallet för undersökningen var 59,7 procent för personer med examen från ett högskoleförberedande program och 71,1 procent för de med examen från ett yrkesprogram.

3.2 Avvikelser från undersökningsdesignen

Inga avvikelser har gjorts.