

# DET STATISTISKA REGISTRETS FRAMSTÄLLNING OCH KVALITET

## Registret över befolkningens utbildning

Här beskrivs registret över befolkningens utbildning som innehåller den folkbokförda befolkningen vid årets slut och den högsta av klarade utbildningsnivå t.o.m. vårterminen det aktuella året.

### **Produktkod**

UF0537

### **Referenstid**

Den folkbokförda befolkningen 2022-12-31 och deras högsta avklarade utbildningsnivå t.o.m. vårterminen 2022.

## Innehåll

<b>Det statistiska registrets framställning .....</b>	<b>4</b>
F1 Det statistiska registrets sammanhang .....	4
F2 Utformning av framställningen .....	4
F2.1 Det statistiska registrets innehåll i stora drag .....	4
F2.2 Datakällor .....	4
F2.3 Tidsaspekter .....	4
F2.4 Insamlingsförfarande .....	5
F2.4.1 Datainsamlingsmetoder och leverantörer .....	5
F2.4.2 Mätning .....	6
F2.4.3 Brister i leveranser .....	6
F2.5 Bearbetningar med granskning .....	6
F2.5.1 Kodning .....	6
F2.5.2 Dubblettkontroller .....	7
F2.5.3 Rimlighetskontroller .....	8
F2.5.4 Imputeringar .....	8
F2.5.5 Modellbaserade beräkningar .....	8
F2.5.6 Jämförelser med andra register och datakällor .....	9
F3 Genomförande .....	10
F3.1 Kvantitativ information .....	10
F3.2 Avvikelser från utformningen .....	10
<b>Det statistiska registrets kvalitet .....</b>	<b>11</b>
K1 Relevans .....	11
K1.1 Ändamål och informationsbehov .....	11
K1.1.1 Registrets ändamål .....	11
K1.1.2 Registeranvändares informationsbehov .....	11
K1.2 Registrets innehåll .....	11
K1.2.1 Objekt och population .....	11
K1.2.2 Variabler .....	11
K1.2.3 Referenstider .....	12
K2 Tillförlitlighet .....	12
K2.1 Tillförlitlighet totalt .....	12
K2.2 Osäkerhetskällor .....	13
K2.2.1 Täckning .....	13
K2.2.2 Mätning .....	14
K2.2.3 Bortfall .....	15
K2.2.4 Bearbetningar .....	15
K2.2.5 Modellantaganden .....	16
K2.3 Preliminärt register jämfört med slutligt .....	16
K3 Aktualitet och punktlighet .....	16
K3.1 Framställningstid .....	16
K3.2 Frekvens .....	16
K3.3 Punktlighet .....	16

K4	Tillgänglighet och tydlighet.....	16
K4.1	Tillgång till registret .....	16
K4.2	Informationsspridning.....	16
K4.3	Dokumentation .....	17
K5	Jämförbarhet och sammanvändbarhet.....	17
K5.1	Jämförbarhet över tid.....	17
K5.2	Sammanvändbarhet med andra register.....	19
<b>Allmänna uppgifter</b>	<b>.....</b>	<b>20</b>
U1	Sekretess och personuppgiftsbehandling.....	20
U2	Bevarande och gallring .....	20
U3	Uppgiftsskyldighet.....	20
U4	EU-reglering och internationell rapportering.....	20
U5	Historik.....	20
U6	Kontaktuppgifter .....	20
<b>Bilaga 1</b>	<b>Frågeformulär, Utbildning i annat land än Sverige .....</b>	<b>22</b>

## Det statistiska registrets framställning

### F1 Det statistiska registrets sammanhang

Registret över befolkningens utbildning (UREG) tillhör den sfär i registersystemet som har objekttypen individ. Populationen utgörs av den folkbokförda befolkningen enligt basregistret Registret över totalbefolkningen (RTB). UREG innehåller uppgifter om personernas högsta avklarade utbildningsnivå, utbildningens inriktning, när personen avslutade sin utbildning, från vilken källa utbildningsuppgiften hämtats, födelseland, födelselän, folkbokföringskommun och personens ålder.

Syftet med att ta fram registret är att med det som underlag kunna framställa registerbaserad statistik om befolkningens utbildningsnivå och utbildningsinriktning.

Till statistiken över befolkningens utbildning hör också vård och utveckling av Svensk utbildningsnomenklatur (SUN) knutet, som är den officiella svenska klassifikationen av utbildningsnivå- och utbildningsinriktning.

Statistik tas fram som underlag för utbildningsplanering på både nationell och regional nivå. Statistik över befolkningens utbildning i olika regioner behövs bl.a. för en effektiv fördelning av samhällets utbildningsresurser.

Registret används inom SCB:s uppdragsverksamhet både för tabelluppdrag och på mikronivå till forskare. Utbildningsnivå från registret används även flitigt i bortfallsanalyser och skattningsförfaranden i många av SCB:s undersökningar.

### F2 Utformning av framställningen

#### F2.1 Det statistiska registrets innehåll i stora drag

UREG är ett individregister som innehåller uppgifter om personens högsta utbildning och ligger till grund för statistik över befolkningens utbildning.

#### F2.2 Datakällor

UREG baseras på flera olika källregister plus en större enkätundersökning riktad till nyinvandrade utrikes födda som saknar uppgift om högsta utbildning i registret. En insamling görs också av examinerade från vissa eftergymnasiala utbildningar som inte rapporteras in via den reguljära högskole- och yrkeshögskolestatistiken. Källregistren hämtas från både andra myndigheter, externa källor, och interna källor inom SCB. Från de olika källregistren hämtas uppgifter om personernas avklarade utbildningar. Objekttypen är individ i samtliga källor som ingår i UREG. De insamlade utbildningsuppgifterna från samtliga källor sparas i en bruttofil. Från bruttofilen väljs en utbildning för varje person enligt de prioriteringsregler som beskrivs i avsnitt K2.2.4.

#### F2.3 Tidsaspekter

Utbildningsregistret framställs en gång per år. Uppgifter om genomgångna utbildningar rapporteras kontinuerligt in till SCB från landets skolor och utbildningsanordnare. Dessa indata bearbetas separat under produktionsåret för att sedan ingå i den slutliga registerproduktionen i början av redovisningsåret. Enkäten till nyinvandrade utrikes födda skickas ut i november varje år och insamlingen avslutas i början av mars, redovisningsåret.

## F2.4 Insamlingsförfarande

### F2.4.1 Datainsamlingsmetoder och leverantörer

Efter förberedelser under hösten så sker det grundläggande arbetet inför den årliga uppdateringen av UREG, under november till april. Det innebär insamling och bearbetning av uppgiftskällor som sammanställs till en ny version av registret. Sammanlagt är det cirka ”aktiva” 40 källor som ligger till grund för den årliga uppdateringen av registret. Utöver det finns det ungefär lika många källor som tidigare ingått i uppdateringen men som inte längre ger något årligt tillskott (t.ex. folk- och bostadsräkningarna, löneregister och utbildningsformer som utgått som kompletterande utbildning, KU, och gymnasieskolans specialkurser). Ett flertal av källorna kommer från andra myndigheter. Uppgiftsskyldighet föreligger enligt förordningen om den officiella statistiken ([2001:100](#)), vilket innebär att myndigheter ska lämna de uppgifter som behövs för framställning av officiell statistik.

Till varje myndighet eller annan typ av uppgiftslämnare skickas en databegäran som innehåller uppgifter om vad leveransen ska innehålla, population och referens-tidpunkt och information om hur uppgifterna ska levereras.

Den primära insamlingsmyndigheten, i avseendet störst datamängd, är Arbetsförmedlingen. Populationerna i indata till SCB består dels av personer som varit inskrivna vid arbetsförmedlingarna t.o.m. ett visst datum, dels av personer som genomgått en arbetsmarknadsutbildning som omfattat minst 140 dagar under en viss period.

Andra viktiga leverantörer av data till UREG är de externa källor som ligger till grund för skolstatistiken. Uppgiftslämnarna är skolor inom olika skolformer och data samlas in och tas om hand av SCB:s datainsamlingsavdelning.

Insamlingen till SCB:s Högskoleregister är extern och sker via LADOK från högskolor och universitet till SCB.

Under hösten förbereds också den årliga enkäten Utbildning i annat land än Sverige (UAL), som riktar till nyinvandrade utrikes födda personer som saknar utbildningsuppgift i UREG. Utskicket av enkäten sker i mitten av november för att det skall vara en så aktuell population av nyinvandrade som möjligt och få en insamlingsperiod som är så lång som möjlig för att generera många enkätsvar som kan ingå i uppdateringen av UREG. Utsändningen består av ett frågeformulär med fem frågor om den utländska utbildningen och ett missiv på 20 sidor för de vanligast förekommande språken bland de nyinvandrade. Efter det första utskicket skickas två påminnelser med nytt frågeformulär och missiv vid båda tillfällena. Det finns två sätt att svara på enkäten, via webb eller pappersenkät. Utsändning, insamling och skanning av den enkät som används utförs av datainsamlingsavdelningen, medan alla bearbetningar, som exempelvis kodning, utförs av UREG:s personal.

Till uppdateringen av UREG 2022 uppgick antalet utskickade enkäter till 58 900. Samtliga, drygt 18 000 eller 31 procent av enkätpopulationen, med en digital brevlåda, Kivra eller motsvarande, fick enkäten digitalt i alla utsändningar utom en där enkäten kunde besvaras via pappersenkät. För den resterande gruppen som inte hade en digital brevlåda i november 2022 så genomfördes ett experiment med en webbboffensiv som innehöll ett 4-sidigt missiv med information om undersökningen på svenska och engelska och en hänvisning till webben för översättning av informationen på 16 språk. Det innebar att 8 000 i två av tre utsändningar fick utsändningen med inloggningsuppgifter och resterande, 32 000, fick UAL som

traditionellt pappersutskick i två av tre utsändningar och en utsändning som webboffensiv. Utvärderingen av experimentet kan man läsa i en separat rapport. Gruppen med digital brevlåda fick en tredje påminnelse i slutet av undersökningsperioden, i början av februari 2023, vilket av kostnadsskäl inte gick ut till dem som saknade digital brevlåda. Det var första gången som vi gick ut med en tredje påminnelse i UAL.

Förändringarna som gjordes i UAL 2022 innebar också att en del ändringar och kompletteringar gjorts i missivet för webboffensiven och översättningen av de 16 olika språken i UAL har uppdaterats på några ställen och anpassats till dagens mallar. Några förändringar i själva enkäten med fem frågor om personens utländska utbildning har inte gjorts. Fr.o.m. UAL 2022 är det möjligt för alla som får enkäten postalt att utöver inloggningsuppgifter även möjligt att logga in med BankID. Mer information om de mättekniska förändringarna i UAL finns att läsa i en separat rapport.

Svarsandelen var ungefär 33 procent, vilket är drygt 1 procentenhet lägre än föregående år. Jämfört med åren 2007 och 2008 så har svarsfrekvensen gått ner betydligt med 18 respektive 15 procentenheter.

#### **F2.4.2 Mätning**

Registrets kvalitet bestäms huvudsakligen av de administrativa system som registret bygger på. Databildningen i de administrativa systemen skiljer sig som regel från SCB:s datainsamlingar. Även om man inom de administrativa systemen har uppgiftslämnare som fyller i blanketter, så är villkoren för uppgiftslämnandet helt annorlunda och styrs av administrativa regler och lagar.

Från SCB går det årligen ut en enkät, se bilaga 1 och 2, till nyinvandrade utrikes födda i åldern 20-59 år som saknar uppgift om högsta utbildning i UREG. Enkäten är årlig och för närvarande översatt till de 16 vanligaste språken bland de som invandrat till Sverige.

#### **F2.4.3 Brister i leveranser**

SCB har regelbundna kontakter med olika uppgiftskällor, och det finns framtagna rutiner för hur leveranser ska genomföras.

Vid beställningen av externa data från myndigheter, t.ex. Universitets- och högskolerådet (UHR), Socialstyrelsen, Arbetsförmedlingen, och Myndigheten för yrkeshögskolan (MYH), specificerar UREG vilka variabler som ska levereras, vilken populationen ska vara samt referenstidpunkt. På detta sätt försöker SCB säkerställa att rätt data levereras. Vilka bearbetningar och granskning som utförs av den externa dataleverantören har SCB liten kunskap om.

### **F2.5 Bearbetningar med granskning**

#### **F2.5.1 Kodning**

Uppgifterna om utbildning från källregistren kodas enligt Svensk utbildningsnomenklatur, (SUN 2020). SUN 2020 är en standard för klassificering av enskilda utbildningar samtidigt som den ger ett system för sammanförande till grupper av utbildningar. SUN 2020 ersätter den tidigare versionen SUN 2000. Mer om Svensk utbildningsnomenklatur (SUN), se [Svensk utbildningsnomenklatur \(SUN\)](https://www.scb.se/utbildning/utbildningsnomenklatur) ([scb.se](https://www.scb.se)).

När utbildningsregistret får in data från sina olika leverantörer ingår i vissa fall utbildningsnivå och inriktning redan i indata, dvs. det är redan kodat hos leverantören. Ett exempel på detta är MYH:s data. I andra fall är materialet inte kodat men innehåller information som gör att SCB kan SUN-koda utbildningar, som exempelvis UHR:s data över personer med bedömda utländska utbildningar. I indata ingår en beskrivning av utbildningen och uppgift om utbildningens längd. SCB utgår från dessa uppgifter och kodar manuellt utbildningsnivå och inriktning. Men för de flesta källor görs kodningen maskinellt genom en matchning mot en hjälpnycckeltabell som innehåller giltiga värden för utbildningsuppgifterna. Det förekommer också att typen av källa avgör vilken nivå och inriktning som samtliga personer från källan ska ha och då sätts samma koder på alla personerna. Ett exempel på det är Jordbruksverkets uppgifter om legitimerade veterinärer och källan avgångna från grundskolans årskurs 9.

Under år 2013 så har utbildningsnivå enligt den senaste internationella standarden ISCED 2011 implementerats i UREG. I ett projektarbete bestämdes en nivåkod enligt ISCED 2011 för varje nivåkod enligt dåvarande SUN 2000 och en nyckel skapades. Varje person har en utbildningsnivå enligt SUN 2020, för åren 2018 eller tidigare enligt SUN 2000, och en nivå enligt ISCED 2011 i registret. Uppgift om nivåkod enligt ISCED 2011 finns i UREG 2010 och framåt.

Under år 2015 har ISCED-F 2013, fields of education (utbildningsinriktning), implementerats i SCB:s Utbildningsregister. ISCED-F 2013 har tagits fram genom en nyckel mellan SUN2000Inr (4-siffernivån) och ISCED-F 2013 på detailed fields (4-siffernivån), vilket är den mest detaljerade redovisningen av ISCED-F 2013. Det finns även en framtagen nyckel mellan ISCED-F 2013 och SUN2020Inr (4-siffernivån) och för åren 2019 och senare i UREG redovisas ISCED-F 2013 baserat på SUN 2020. ISCED-F 2013 har införts från och med version 2010 av UREG.

### **F2.5.2 Dubblettkontroller**

Dubblettkontroller görs i varje indatakälla, både på personnummer och på den kombination av variabler som ligger till grund för val av högsta utbildning.

De variabler som används vid valet är utbildningsnivå, examensår, examenstermin, källkod, källår och utbildningsinriktning.

I första hand väljs personens utbildning som har högst utbildningsnivå. Om personen har flera utbildningar med samma utbildningsnivå används i tur och ordning högsta (senaste) examensår/examenstermin, högsta källkod (bäst källkvalitet), högsta (senaste) år för inläggning i registret. I sista hand väljs utbildningen med det ”högsta” värdet i utbildningsinriktning, ex. väljs 342z före 340a.

För att ta fram UREG:s population matchas de valda posterna mot RTB, folkbokförda den 31/12 aktuellt år.

I registerproduktionen körs också en procedur som tar fram personernas senaste personnummer. Uppdateringar av personnummer görs utifrån resultatet av proceduren. Dubblettshantering kan behövas efter uppdateringen. Personer med samordningsnummer kontrolleras mot RTB och i de fall det finns en koppling till ett riktigt personnummer används detta i UREG.

### **F2.5.3 Rimlighetskontroller**

Rimlighetskontroller görs under hela bearbetningsprocessen, både i indata och i det slutliga observationsregistret. Rimlighetskontrollerna görs ex. på orimlig utbildningsnivå i förhållande till personens ålder, frekvenser på utbildningsnivåer jämfört med föregående år och i varje källa kontrolleras populationens storlek jämfört med tidigare år.

Rimlighetskontroller görs i varje enskild källa under arbetets gång. Kontroller görs av valida värden i utbildningsnivå och utbildningsinriktning. Rimligheten kontrolleras också med avseende på om de utbildningsnivåer som förekommer i källan är valida för just den källan. Liknande kontroller görs också för utbildningsinriktning. Rimlighetskontroller görs också för att upptäcka fel i och orimliga kombinationer av utbildningsnivå- och inriktning. I källorna görs också dubblettkoll på nyckeln i UREG.

Examensår med tidigare årtal eller samma årtal som när personen är född kontrolleras och sätts till okänt värde. Personernas examenstermin sätts i dessa fall också till okänd. Ogiltiga personnummer med avseende på bokstäver eller blanksteg kontrolleras.

### **Enkäten till nyinvandrade utrikes födda**

De inkomna enkätsvaren granskas i flera olika skeden av bearbetningen. I ett tidigt skede av insamlingen granskas den för att se om någon fråga verkar ha missuppfattats av många eller om det finns frågor som genererar många s.k. dubbelmarkeringar. Det senare innebär att uppgiftslämnaren markerat flera svar trots att frågan bara ska besvaras med ett alternativ.

När personen markerat flera svar på fråga 2 och 3, se bilaga 1 och 2, väljs alternativet med högsta antal år/längst utbildning. Skulle uppgiftslämnaren angett ett årtal på fråga 4 som är otydligt, orimligt eller använt sig av bokstäver så kodas det som okänt, \*\*\*\*. På fråga 5 avgör personal inom UREG enligt fastställda kriterier vilket alternativ som ska väljas om flera svarsalternativ markerats. Om personen lämnat svar på både webb och papper eller vid flera tillfällen väljs svaret som inkommit först. Någon rimlighetsbedömning mellan svaren på de olika frågorna görs inte.

På webblanetten går det endast att fylla i ett svarsalternativ per fråga.

Personal inom UREG följer inflödet i invandrenkäten i ett avregistreringssystem. Informationen används för att göra jämförelser med tidigare års svarsfrekvenser, fördelning av svar mellan papper och webb, andelen svar som inkommit innan första respektive andra påminnelsen och ge underlag till ansvariga chefer hur inflödet i årets undersökning ser ut.

### **F2.5.4 Imputeringar**

Inga imputeringar görs.

### **F2.5.5 Modellbaserade beräkningar**

Det förekommer att målvariabeln utbildningsnivå approximeras med hjälp av andra variabler. För källa SFI, svenska för invandrare, översätts variabeln antalet år i utbildning från hemlandet till en utbildningsnivå enligt SUN 2020. För några källor översätts avklarade poäng, t.ex. källa Komvux, eller lästa poäng, t.ex. källorna samlat betygskodokument/studiebevis och gymnasieintyg, till en utbildningsnivå enligt SUN 2020.



Valet av högsta utbildningsnivå för en individ utgår från ett antal prioriteringsregler. Se mer detaljer under K.2.2.4.

#### **F2.5.6 Jämförelser med andra register och datakällor**

UREG är grundkällan till uppgift om högsta utbildning samt inriktning vilket innebär att det inte finns någon annan producent att göra jämförelser med. I arbetet med indatakällorna görs jämförelser med externa leverantörers publicerade statistik på deras hemsidor. Populationen kontrolleras mot Registret över totalbefolkningen (RTB).

### F3 Genomförande

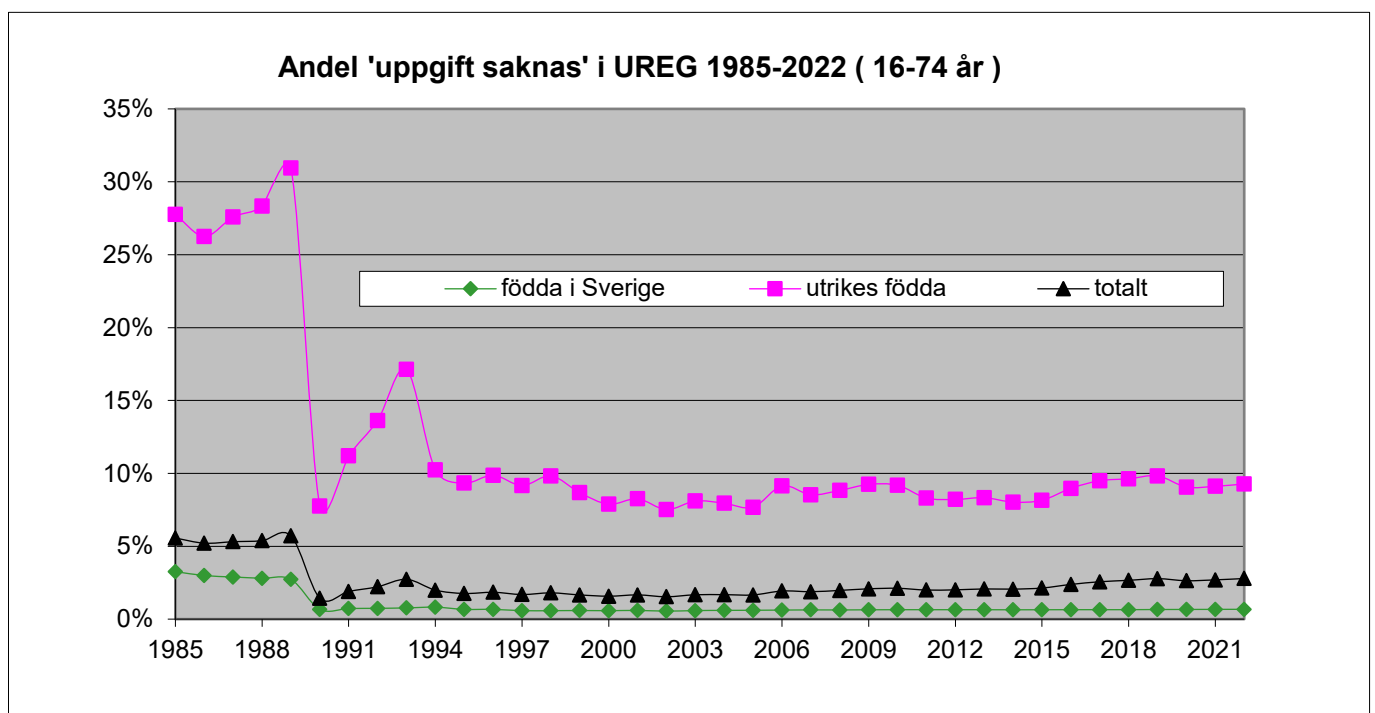
#### F3.1 Kvantitativ information

Utbildningsregistret 2022 innehåller den folkbokförda befolkningen den 31/12 2022, vilket var cirka 10,5 miljoner personer.

För personer i åldern 16-74 år saknade 2,8 procent helt och hållet uppgift om utbildningsnivå i registret.

Motsvarande andel för utrikes födda är 9,3 procent, och 0,7 procent för inrikes födda.

Diagram 1. Andel ”uppgift saknas” i Utbildningsregistret 1985-2022



#### F3.2 Avvikelser från utformningen

Inga påtagliga avvikelser har gjorts från utformningen eller upplägget jämfört med tidigare år.

## Det statistiska registrets kvalitet

### K1 Relevans

#### K1.1 Ändamål och informationsbehov

##### K1.1.1 Registrets ändamål

Registrets huvudsyfte är att utgöra underlag för att kunna framställa registerbaserad statistik om befolkningens utbildningsnivå och utbildningsinriktning.

Registret används inom uppdragsverksamheten på SCB, både för tabelluppdrag och på mikronivå till forskare. Utbildningsnivå från registret används även frekvent i bortfallsanalyser och skattningsförfaranden i många av SCB:s undersökningar.

##### K1.1.2 Registeranvändares informationsbehov

Statistik tas fram som underlag för utbildningsplanering på både nationell och regional nivå. Statistik över befolkningens utbildning i olika regioner behövs bl.a. för en effektiv fördelning av samhällets utbildningsresurser.

#### K1.2 Registrets innehåll

Målpopulationen utgörs av de personer som skulle ha varit folkbokförda i Sverige vid registrets referenstidpunkt. De viktigaste målvariablerna är högsta utbildningsnivå för avklarade utbildningar fram till och med vårterminen för referensåret samt inriktning för högsta avklarade utbildningsnivå.

##### K1.2.1 Objekt och population

Registrets målobjekt och observationsobjekt är individer. Det finns ingen skillnad mellan intresseobjekt och registrets mål- och observationsobjekt, även intresseobjektet utgörs av individer.

Målpopulationen utgörs av de personer som skulle ha varit folkbokförda i Sverige vid registrets referenstidpunkt, om de lagar, förordningar och andra regler som gäller folkbokföringen hade efterlevts utan fel.

Intressepopulationen omfattar i de flesta fall även personer som befinner sig i landet utan uppehållsrätt, uppehållstillstånd eller motsvarande. Det skulle vara mycket resurskrävande att inkludera dem i målpopulationen och bedömningen är att kvaliteten på registret inte påverkas nämnvärt av att utelämna dem.

##### K1.2.2 Variabler

De viktigaste målvariablerna i Utbildningsregistret är:

- Högsta utbildningsnivå (kodar enligt SUN 2020) för avklarade utbildningar fram till och med vårterminen för registrets referensår
- Utbildningsinriktning för högsta utbildning (kodar enligt SUN 2020)
- Utbildningsgrupp (ett aggregat av högsta utbildningsnivå och utbildningsinriktning)

Målvariabeln högsta utbildningsnivå skapas utifrån en prioritering av en persons samtliga kända avklarade utbildningar. Observationsvariabeln är således avklarade utbildningar.

Utbildningsnivå samvarierar starkt med svarsbenägenhet i statistiska undersökningar, vilket innebär att variabeln används i stor utsträckning som hjälpvariabel i statistiska undersökningar på SCB. Exempel på undersökningar som använder utbildningsnivå i skattningsförfarandet är Undersökningen om levnadsförhållanden (ULF) och Partisynpatiuundersökningen (PSU).

Utbildningsnivå används även för granskning av andra register och direktinsamlade data. Exempelvis använder yrkesregistret och lönestatistiken utbildningsnivå vid granskningsarbetet.

### **K1.2.3 Referenstider**

Utbildningsregistret avser den folkbokförda befolkningen vid respektive årsskifte, och avklarad utbildning till och med vårterminen före aktuellt årsskifte. Exempel: årgång 2022 avser befolkningen 2022-12-31 och avklarad utbildning till och med vårterminen 2022.

## **K2 Tillförlitlighet**

### **K2.1 Tillförlitlighet totalt**

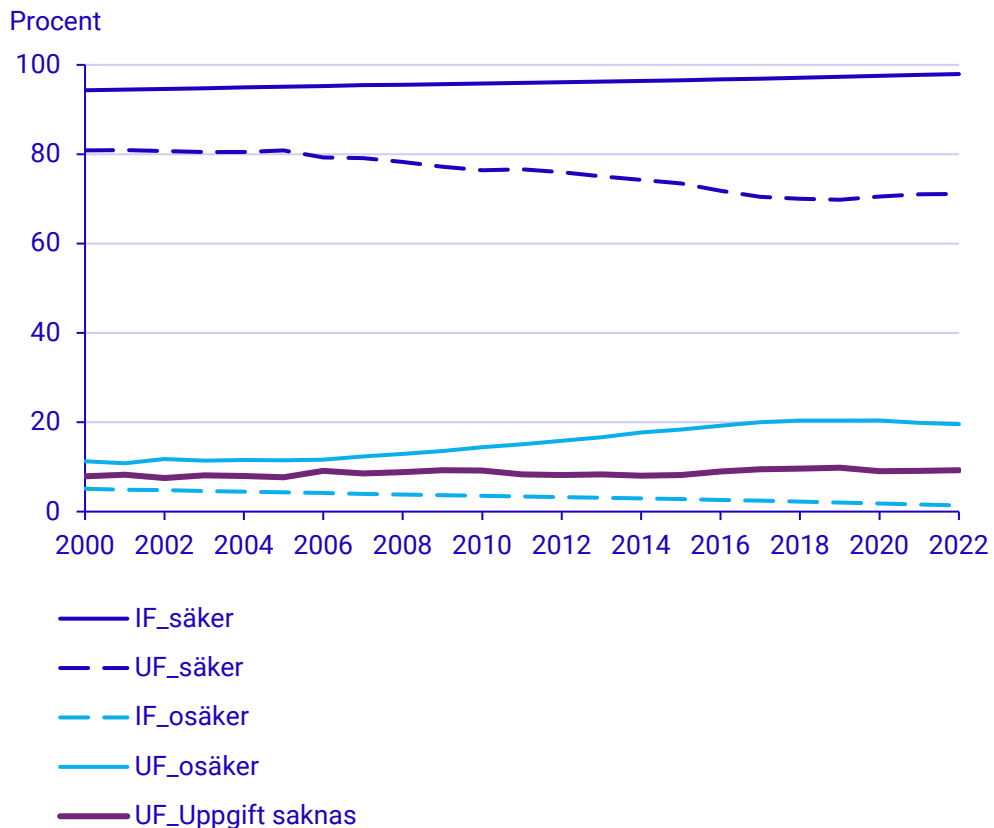
Sammantaget bedöms registret vara mycket tillförlitligt. Statistikuppgifter om utbildningsnivå och utbildningsinriktning hämtade från registret bör i de flesta fall vara av hög kvalitet. Registret används i stor utsträckning av andra statistikproducenter på SCB och andra aktörer (exempelvis myndigheter och forskare), vilket innebär att större brister i registret sannolikt skulle upptäckas.

Kodningen av utbildningsnivå och inriktning bör i de flesta fall av vara av hög kvalitet. Det gäller framför allt för inrikes födda, där nästan 98 procent har uppgift från en säker källa. Även för utrikes födda bör kvaliteten i många fall vara god, men här är det en lägre andel som har uppgifter från en säker källa (drygt 70 procent). För de osäkra källorna är kvaliteten ibland sämre, med uppgifter som i vissa fall inte medger en detaljerad kodning av nivå och inriktning.

Under perioden 2000-2022 har andelen inrikes födda med en säker källa i UREG ökat från 94 till 98 procent medan andelen med en säker källa i gruppen utrikes födda minskat från 81 procent år 2000 till 71 procent i UREG 2022. Den ökade andelen med en osäker källa bland de utrikes födda beror på en ökad invandring under de senaste åren och de nyinvandrade har i högre utsträckning en osäker källa för högsta utbildning framför allt under de första åren i Sverige. I UREG 2022 har knappt 20 procent av de utrikes födda en osäker källa för högsta utbildning och bland de inrikes födda är motsvarande andel 1,4 procent.

Av de utrikes födda med en säker källa i UREG 2022 har 28 procent källa "Utbildning i annat land än Sverige". Precisionen i nivå och inriktningskoden är inte lika god där som en uppgift från ett reguljärt statistikregister eller legitimationsregister.

### Andel med säker respektive osäker källa efter födelseland, inrikes och utrikes födda, i UREG 2000–2022 (16-74 år)



Ett undantag från den höga tillförlitligheten är när statistik ska tas fram för grupper där registret i hög utsträckning saknar uppgifter om högsta utbildningsnivå. Det kan då vara svårt för användaren att få en bild av hur utbildningsnivån ser ut för gruppen som saknar uppgift om utbildning. Grupper som i hög utsträckning saknar uppgift om utbildning är utrikes födda och då framför allt personer som nyligen invandrat till Sverige.

Även alltför detaljerad statistikredovisning av utbildningsnivå och utbildningsinriktning kan vara förknippad med betydande fel. Ett exempel på detta är uppgifter om utbildningsnivå för små geografiska områden (små kommuner eller delar av kommuner). Ett annat exempel är statistik på detaljerad nivå avseende utbildningsinriktning.

#### K2.2 Osäkerhetskällor

Den källa som bidrar mest till osäkerheten i statistiken om befolkningens utbildning är bortfall i den mening att det till stor del saknas uppgift om utländska utbildningar, vilket kan leda till att utrikes föddas utbildningsnivå redovisas som uppgift saknas i stället för den faktiska utbildningsnivån. Övriga felkällor bedöms i de flesta fall ha relativt liten inverkan på tillförlitligheten. Läs mer om detta under 2.2.1-2.2.5.

##### K2.2.1 Täckning

Rampopulation för Utbildningsregistret är Sveriges folkbokförda befolkning enligt RTB. I RTB förekommer i viss mån täckningsfel vilka består i att personer som enligt gällande regelverk skall vara folkbokförda inte är det (undertäckning) samt att

personer är folkbokförda trots att de inte skall vara det (övertäckning). Det förekommer även att personer är folkbokförda på fel adress.

I den senaste skattningen av täckningen som gjorts i RTB skattades övertäckningen per den 31 december 2019 till 73 400 personer. Antalet personer i undertäckningen vid samma tidpunkt skattades till 17 000.

Den skattade övertäckningen är framför allt koncentrerad till gruppen utrikes födda. I åldersgruppen 25-64 år för RTB 2019 (och därmed UREG 2019) skattades övertäckningen till 54 000 varav 47 000 utrikes födda. Övertäckningen är klart vanligast i gruppen utrikes födda med uppgift saknas om högsta utbildning, 14,1 procent. För övriga utbildningsnivåer ligger motsvarande andel mellan 2-4 procent.

**Tabell 1. Skattad övertäckning i RTB/UREG 2019 efter utbildningsnivå för utrikes födda (25-64 år)**

Utbildningsnivå	Antal	Skattad övertäckning	%-övertäckning
Förgymnasial	244 021	4 699	1,9
Gymnasial	428 404	7 209	1,7
Eftergymnasial < 3 år	190 756	5 763	3,0
Eftergymnasial minst 3 år	360 671	14 192	3,9
Uppgift saknas	108 698	15 312	14,1
Totalt	1 332 550	47 175	3,5

Sammantaget så har täckningsbrister relativt liten inverkan på tillförlitligheten. För vissa populationsavgränsningar kan dock täckning vara en betydande källa till osäkerhet. Det handlar främst om delpopulationer där man kan anta att en betydande del utvandrat utan att lämna avtryck i register om att de lämnat landet. Ett exempel på en sådan population kan vara utbytesstudenter eller nyligen invandrade. Vid redovisning på detaljerad regional nivå kan täckningsbrister få en viss betydelse.

### K2.2.2 Mätning

Utbildningsregistret inhämtar uppgifter från flera olika register plus en större enkätundersökning riktad till utrikes födda. Mätfel kan uppstå i produktionen av varje enskilt register.

Syftet är att hämta in information om samtliga avklarade utbildningar för en individ via registren. Informationen kan avse utbildningsform, längd på utbildningen och huvudsakligt innehåll. Informationen bör vara tillräcklig för att kunna klassificera utbildningen enligt SUN 2020.

Om information avseende längden på en viss utbildning saknas så kan mätfel uppstå. Ett exempel är att informationen i en uppgiftskälla avseende utbildning på gymnasial nivå är ofullständig avseende utbildningens längd. Utbildningen klassificeras då som "kort gymnasial" (nivå 31, gymnasial utbildning 1 år, enligt SUN-nomenklaturen).

Ett annat exempel på mätfel är när uppgiften om den högsta utbildningen för en person saknas, men det finns uppgifter om minst en annan utbildning. Att vissa av en persons avklarade utbildningar inte rapporterats kan medföra att personens högsta utbildning inte väljs.

I vissa delgrupper har vi inte så detaljerad information så att vi kan precisera kodningen. Det gäller främst utrikes födda och framför allt då nyligen invandrade. Exempelvis kan information avseende längden på en viss utbildning saknas och då

kan mätfel uppstå. Ett exempel är att informationen i en uppgiftskälla avseende utbildning på gymnasial nivå är ofullständig avseende utbildningens längd.

Några formella studier har inte gjorts för att närmare kvantifiera de osäkerheter som uppstår till följd av mätfel men de bedöms vara små.

### **K2.2.3 Bortfall**

Objektsbortfall förekommer i UREG i den meningen att personer helt saknar uppgift om utbildning. I årgång 2022 av Utbildningsregistret saknar 2,6 procent av personerna, 16 år och äldre, i registret helt och hållet uppgift om utbildning. Andelen för utrikes födda är 9,3 procent, och 0,5 procent för inrikes födda.

Personer som avslutat en utbildning inom särskolan, grundskolan och gymnasiet, ingår inte i den årliga uppdateringen av UREG. Det beror på att SCB inte får spara uppgifter om elever i särskolan i mer än två år samt att det saknas uppgift om personen avslutat utbildningen med ett avgångsbetyg eller motsvarande. Vår bedömning är att omkring hälften av bortfallet i gruppen inrikes födda i åldern 17–20 år beror på att personerna gått i särskolan.

Personer som aldrig har gått i skola redovisas också som bortfall/uppgift saknas. Ingen möjlighet finns dock att kontrollera vilka som saknar uppgift om utbildning och vilka som helt saknar utbildning.

Vissa utbildningar har inte rapporterats till Utbildningsregistret under vissa perioder. Exempel är folkhögskoleutbildning på gymnasial nivå under perioden 1990–2000 och många äldre kortare yrkesutbildningar på gymnasial nivå. Ett annat exempel är Försvarsmakten som inte levererat uppgifter om specialistofficersutbildning fr.o.m. årgång 2013.

Bortfallet (uppgift om högsta utbildning saknas) påverkar framför allt redovisningen av utrikes föddas utbildningsbakgrund. De redovisningsgrupper där det har störst betydelse är födelseland och invandringstidpunkt. Bortfallet är störst bland nyinvandrade och varierar dessutom mycket mellan personer från olika födelseländer.

### **K2.2.4 Bearbetningar**

Uppgifterna om utbildning från källregistren kodas enligt Svensk utbildningsnomenklatur (SUN 2020). Därefter läggs alla utbildningar i Utbildningsregistrets ackumulerade s.k. ”bruttofil”. Med hjälp av prioriteringsregler väljs därefter den högsta utbildningen för varje person. Det är från denna ”personfil” med varje persons för närvarande högsta utbildning de statistiska uppgifterna huvudsakligen hämtas.

Prioriteringsreglerna bygger i första hand på utbildningens nivå och antagande om kvaliteten i olika källregister. Detta kan leda till att en annan utbildning än den högsta utbildningen väljs, om den riktiga högsta utbildningen finns i ett källregister som är lågt prioriterat.

Några formella studier har inte gjorts för att närmare kvantifiera de osäkerheter som uppstår till följd av prioriteringsreglerna, men med den kännedom som finns om källorna till Utbildningsregistret bedöms de vara små. För mer detaljerad information om hur valet av personens högsta utbildning görs, se dokumentation av Statistikens framställning (StaF) på SCB:s webbplats, <http://www.scb.se/uf0506>.

### **K2.2.5 Modellantaganden**

Det förekommer att utbildningsnivån approximeras utifrån andra variabler, exempelvis utbildningens längd eller avklarade poäng (se F.2.5.5). Dessa approximationer bedöms fungera väl och därmed ha en liten betydelse för statistikens kvalitet.

### **K2.3 Preliminärt register jämfört med slutligt**

Endast slutligt observationsregister tas fram.

## **K3 Aktualitet och punktlighet**

### **K3.1 Framställningstid**

Den totala framställningstiden för registret är cirka 4 månader.

### **K3.2 Frekvens**

Insamlingen av data från olika källor och framställningen av registret sker en gång per år. Registret tillgängliggörs under våren varje år.

### **K3.3 Punktlighet**

Tillgänglighetsförändringen har varit punktligt för samtliga varianter och versioner.

## **K4 Tillgänglighet och tydlighet**

### **K4.1 Tillgång till registret**

Uppgifter ur UREG kan lämnas ut för forskning och statistikändamål. Varje förfrågan till SCB om utlämnande av mikrodata sekretessprövas. SCB lyder under statistiksekretessen som återfinns i 24 kap 8 § Offentlighets- och sekretesslagen ([2009:400](#)). Vid automatiserad behandling av personuppgifter gäller även reglerna i dataskyddsförordningen.

Efter godkänd prövning kan forskare och utredare få tillgång till avidentifierade mikrodata för egna bearbetningar i SCB:s system för bearbetningar av avidentifierade personuppgifter, [Microdata Online Access \(MONA\)](#).

En skriftlig begäran om de uppgifter som önskas skickas till SCB. Mer information om hur man begär uppgifter samt beställningsblankett finns på SCB:s webbplats. När uppgifterna ska användas i forskning ser SCB gärna att en projektbeskrivning bifogas. Om forskningsprojektet har prövats av en regional etikprövningsnämnd bör även ansökan och beslut bifogas. I de fall som uppgifterna kan lämnas ut lämnas en offert som innehåller en beskrivning av SCB:s åtagande, leveranstidpunkt och pris.

### **K4.2 Informationsspridning**

På SCB:s hemsida finns publiceringsdatum för statistiken som är framtagen från registret. Det datumet kan registret börja användas externt.

<https://www.scb.se/hitta-statistik/publiceringskalendern/>

Internt finns informationen på Inblick:

<http://inblick.scb.intra/Stod/Statistikproduktion/register/individer/ureg/>

Kontakter sker kontinuerligt mellan registerproducenterna och de interna användarna.



### **K4.3 Dokumentation**

Registret dokumenteras i Statistikens detaljerade innehåll på SCB:s webbplats (MetaPlus).

Variant: Personens högsta utbildning,  
Versioner: 2008–2022

Variant: Personens högsta utbildning, 16–74 år,  
Versioner: 1997–2007

Variant: Personens samtliga utbildningar,  
Versioner: 1997–2022

Dokumentation, av registrens detaljerade innehåll, finns på SCB:s webbplats. Där beskrivs alla variabler och värdemängder m.m. Dokumentationen finns här:  
<https://www.h6.scb.se/metadata/mikrodataregister.aspx>.

## **K5 Jämförbarhet och sam användbarhet**

### **K5.1 Jämförbarhet över tid**

SCB strävar hela tiden efter att öka Utbildningsregistrets kvalitet, t.ex. genom att nya källor kan användas. Detta gör att försiktighet måste iakttas vid studier av tidsserier. Nedan följer de viktigaste förändringarna sedan 1990.

#### **Version 1990:**

- Utbildningsdata från Folk- och bostadsräkningen (FoB) 1990 används vid framställning av Utbildningsregistret fr.o.m. årgång 1990 och har förbättrat registrets kvalitet avsevärt. Vi vill uppmana till försiktighet vid jämförelser mellan årgångarna 1985–1989 av registret och årgångar från 1990 och framåt, då de senare innehåller uppgifter om utbildning från FoB 90.

Utbildningsuppgiften via FoB 90 har medfört att registret fått in en uppdaterad uppgift om personens högsta utbildning vilket har minskat andelen uppgift saknas, särskilt för gruppen utrikes födda som invandrat mellan 1970 och 1990.

#### **Version 1994:**

- Våren 1995 genomfördes en enkät till drygt 100 000 personer födda utomlands vilka saknade uppgift i Utbildningsregistret. Årgångarna fr.o.m. 1994 har framställts med dessa uppgifter, vilket höjt registrets kvalitet ytterligare, d.v.s. minskat andelen uppgift saknas bland de utrikes födda.

#### **Version 1999:**

- Hösten 1999 genomfördes en enkät till drygt 60 000 utrikes födda personer vilka saknade uppgift i Utbildningsregistret. Årgångarna fr.o.m. 1999 har framställts med dessa uppgifter, vilket höjt registrets kvalitet ytterligare. Från och med 1999 är SCB:s enkät till utrikes födda, "Utbildning i annat land än Sverige", årlig och omfattar personer i åldern 20–59 år som under året blivit folkbokförda. Enkätsvaren har medfört att andelen uppgift saknas i gruppen utrikes födda har minskat.

### Version 2000:

- Till årgång 2000 genomgick Utbildningsregistret större kvalitetshöjande förändringar vilket leder till ett tidsseriebrott och gör att jämförelser med tidigare årgångar av registret måste göras med stor försiktighet:

1) Svensk Utbildningsnomenklatur, SUN, anpassades genom nya SUN 2000 till den internationella nomenklaturen ISCED 97 (International Standard Classification of Education).

2) Ett antal nya uppgiftskällor till Utbildningsregistret tillkom.

Detta gör att utbildningsnivån i riket enligt Utbildningsregistret har ökat kraftigt i version 2000. Ökningarna, som vilket delvis är en effekt av de ovan beskrivna förändringarna, ligger främst på nivåerna "Treårig gymnasieutbildning" och "Eftergymnasial utbildning 3 år eller längre".

Mer information om kvalitetshöjningen finns på SCB:s webbplats.

<http://www.scb.se/uf0506>

För en fördjupad studie av tidsseriebrottet med kvantitativa mått på förändringar hänvisas den intresserade läsaren till en särskild analys av Cecilia Wass, "Tidsseriebrott i Utbildningsregistret" som finns tillgänglig på SCB:s webbplats.

<https://www.scb.se/contentassets/758c162300754eb28ba15236f63ca0a0/metodstudieureg.pdf>

### Version 2011:

- Till version 2011 av Utbildningsregistret så har det gjorts en reviderad beräkning av variabeln "YAndel" i källa Komvux. Variabeln "YAndel" används för att klassificera utbildningsinriktning (allmän eller yrkesinriktning) på Komvuxkurserna.

Den reviderade beräkningen av variabeln "YAndel" har införts i fem årgångar (2006–2010) bakåt av Utbildningsregistret

Den reviderade beräkningen av variabeln "YAndel" har medfört att de yrkesinriktade gymnasiala utbildningsnivåerna, ej fullföljd, ökat i antal medan det skett motsvarande minskning bland de studieförberedande gymnasiala utbildningsnivåerna, ej fullföljd. För utbildningsinriktning har de med allmän inriktning minskat i antal medan det skett en motsvarande ökning av de gymnasiala yrkesinriktningarna.

### Version 2017:

- Till version 2017 begränsades enkäten "Utbildning i annat land än Sverige". För de personer som hade en utbildningsuppgift registrerad vid Arbetsförmedlingen t.o.m. 30 september 2017 har uppgifterna från Arbetsförmedlingen använts och personerna fick inte årets enkät.

### **Version 2019:**

- I Utbildningsregistret 2019 infördes en ny version av utbildningsnomenklaturen, SUN 2020. Klassificeringen av utbildningsnivå förändrades inte jämfört med tidigare år utan förtydliganden har gjorts hur utbildningar ska nivåklassificeras. De förändringar som skett i den nya klassifikationen rör inriktningskodningen och då framför allt på den finaste nivån (4-siffernivån) för att på ett bättre sätt spegla de förändringar som skett i det svenska utbildningssystemet sedan SUN 2000. Det är huvudsakligen eftergymnasiala utbildningar som berörs av övergången till den nya klassifikationen.

På 1-siffernivån av inriktningskoden är det små förändringar mellan SUN 2000 och SUN 2020. Drygt 4 000 personer (16 år och äldre) byter inriktningskod på den grävsta nivån mellan klassifikationerna, 2 100 personer går från tjänster (inriktning 8 i SUN 2000) till naturvetenskap, matematik och IKT (inriktning 4 i SUN 2020) och 2 300 personer går från tjänster (inriktning 8 i SUN 2000) till teknik och tillverkning (inriktning 5 i SUN 2020).

### **Version 2020:**

- En variabel som anger om personens högsta avslutade utbildning är genomförd i Sverige eller utlandet infördes i registret. Variabeln skapas utifrån informationen i den källa personens utbildningsuppgift hämtas i från.
- Inga förändringar av målstorheter eller metoder har gjorts till följd av pandemin, som kan tänkas påverka jämförbarheten över tid.

### **Version 2021:**

- I version 2021 infördes en mer detaljerad indelning av utbildningsgrupper enligt SUN 2020, anpassade till dagens utbildningsutbud och den nya klassifikationen. Denna gruppering finns nu i registret fr.o.m. version 2019. Indelningen i grupper görs utifrån utbildningens nivå och inriktning och omfattar cirka 130 utbildningsgrupper.
- Inga förändringar av målstorheter eller metoder har gjorts till följd av pandemin, som kan tänkas påverka jämförbarheten över tid.

## **K5.2 Samanvändbarhet med andra register**

Utbildningsregistret är den enda svenska registerkällan med uppgifter om befolkningens högsta avslutade utbildningsnivå- och inriktning. Samanvändbarhet är möjlig med de register som innehåller objektet personnummer. Ett exempel är sysselsättningsregistret där möjlighet ges att få en uppfattning om utbildningsnivån bland förvärvsarbetande respektive bland ej förvärvsarbetande.

## Allmänna uppgifter

### U1 Sekretess och personuppgiftsbehandling

I myndigheternas särskilda verksamhet för framställning av statistik gäller sekretess enligt 24 kap. 8 § offentlighets- och sekretesslagen ([2009:400](#)).

Vid behandling av personuppgifter, dvs. information som direkt eller indirekt kan hänföras till en person som är i livet, gäller lagen ([2001:99](#)) och förordningen ([2001:100](#)) om den officiella statistiken samt EU:s dataskyddsförordning ([2016/679](#)).

Aidentifierade mikrodata från UREG kan lämnas ut till forskning och statistikändamål efter sekretessprövning om SCB bedömer att det finns stöd för att behandla uppgifterna.

### U2 Bevarande och gallring

UREG omfattas av Riksarkivets bevarandebeslut RA-MS 2007:64 (ändrad RA-MS 2011:59, RA-MS 2013:55, RA-MS 2017:15).

### U3 Uppgiftsskyldighet

Uppgiftsskyldighet föreligger enligt förordning om den officiella statistiken ([2001:100](#)), vilket innebär att myndigheter ska lämna de uppgifter som behövs för framställning av statistiken.

Uppgiftsskyldighet till framställning av annan statistik än den officiella föreligger inte.

### U4 EU-reglering och internationell rapportering

Ingen EU-reglering finns för UREG.

### U5 Historik

Registrets första version avsåg förhållandena per 1985-12-31. Därefter genomförs årliga uppdateringar. Före 1985 togs detaljerad och totalräknad statistik över utbildningsnivå fram i samband med folk- och bostadsräkningarna 1930 respektive 1970.

För mer information, se de dokumentationer som finns här:

<http://www.scb.se/uf0506>

### U6 Kontaktuppgifter

<b>Information om det statistiska registret</b>	
<b>Kontaktperson</b>	Michael Karlsson
<b>Enhet</b>	D/REG/IR
<b>E-post</b>	<a href="mailto:michael.karlsson@scb.se">michael.karlsson@scb.se</a>
<b>Telefon</b>	010 – 479 6481

<b>Information om det statistiska registret</b>	
<b>Kontaktperson</b>	Lotta Voss
<b>Enhet</b>	D/REG/IR
<b>E-post</b>	<a href="mailto:lotta.voss@scb.se">lotta.voss@scb.se</a>
<b>Telefon</b>	010 – 479 6686

<b>Mikrodatautlämnande</b>	
<b>Kontaktperson</b>	Marie Kåberg
<b>Enhet</b>	SSA/UE/MD
<b>E-post</b>	<a href="mailto:Marie.Kaberg@scb.se">Marie.Kaberg@scb.se</a>
<b>Telefon</b>	010 – 479 6579

<b>Statistikuppdrag</b>	
<b>Kontaktperson</b>	Tomas Westling
<b>Enhet</b>	SSA/AU/UTB
<b>E-post</b>	<a href="mailto:tomas.westling@scb.se">tomas.westling@scb.se</a>
<b>Telefon</b>	010- 479 6178

## Bilaga 1 Frågeformulär, Utbildning i annat land än Sverige

Besvara gärna frågorna på internet så behöver du inte posta något!  
Gå in på [www.insamling.scb.se/ual](http://www.insamling.scb.se/ual) och logga in med BankID eller dina inloggningsuppgifter:

Användarnamn:

Lösenord:

Answer the questions online - then you won't have to post the paper form.  
Go to [www.insamling.scb.se/ual](http://www.insamling.scb.se/ual) and log in with your BankID or your login details:

1. Har du utbildat dig i något annat land än Sverige?

Ja

Nej

→ Tack för ditt svar. Sänd in frågeformuläret i svarskuvertet som följer med.

2. Hur många år har du sammanlagt gått i grund-, gymnasie-, yrkesskola i annat land än Sverige?  
Markera endast ett alternativ.

1–8 år

9 år

10–11 år

12 år eller mer

3. Hur många år har du studerat vid universitet/högskola eller annan eftergymnasial utbildning i annat land än Sverige?

Markera endast ett alternativ.

Ingen eftergymnasial utbildning

1 år

2 år

3 år

4 år eller längre

Doktorsexamen (motsvarande PhD)

4. Vilket år tog du din högsta examen?

Skriv årtal:

5. Vilken inriktning har din högsta utbildning i annat land än Sverige?

Markera endast ett alternativ.

Allmän

Lärarutbildning, pedagogik

Konstnärlig utbildning

Humaniora, språk, religion

Samhälls- och beteendevetenskap

Juridik

Ekonomi

Handel, administration

Naturvetenskap, data

Teknik och tillverkning (t.ex. arkitekt)

Civilingenjör (minst 4-årig högskola)

Lant- och skogsbruk (t.ex. veterinär eller agronom)

Hälso- och sjukvård, social omsorg (t.ex. apotekare eller receptarie)

Läkare

Tandläkare

Sjuksköterska

Service (t.ex. restaurang, turism, lokalvård, polis eller militär)

Har du valt att besvara pappersenkäten lägger du den i svarskuvertet som följer med. Du behöver inget frimärke när du postar enkäten. Tack!

If you chose to use this paper form, please return it in the enclosed prepaid reply envelope. Thank you!