

Gymnasieungdomars studieintresse Lsåret 2011/12

UF0513

I denna beskrivning redovisas först administrativa och legala uppgifter om undersökningen samt dess syfte och historik. Därefter redovisas undersökningens innehåll och tillförlitlighet samt hur den genomförs och hur man kan ta del av resultaten. Genom att klicka på en rubrik i innehållsförteckningen kommer man direkt till aktuellt avsnitt.

Innehållsförteckning

A	Allmänna uppgifter	2
A.1	Ämnesområde	2
A.2	Statistikområde.....	2
A.3	SOS-klassificering.....	2
A.4	Statistikansvarig.....	2
A.5	Statistikproducent.....	2
A.6	Uppgiftsskyldighet.....	3
A.7	Sekretess och regler för behandling av personuppgifter	3
A.8	Gallringsföreskrifter	3
A.9	EU-reglering	3
A.10	Syfte och historik.....	3
A.11	Statistikanvändning.....	3
A.12	Uppläggning och genomförande.....	4
A.13	Internationell rapportering	5
A.14	Planerade förändringar i kommande undersökningar	5
B	Kvalitetsdeklaration.....	5
B.0	Inledning	5
B.1	Innehåll	5
1.1	Statistiska målstorheter	5
1.1.1	Objekt och population	5
1.1.2	Variabler.....	6
1.1.3	Statistiska mått	7
1.1.4	Redovisningsgrupper.....	7
1.1.5	Referenstider	8
1.2.	Fullständighet.....	8
B.2	Tillförlitlighet.....	8
2.1	Tillförlitlighet totalt.....	8
2.2	Osäkerhetskällor.....	8
2.2.1	Urval.....	8
2.2.2	Ramtäckning	10
2.2.3	Mätning	11
2.2.4	Svarsbortfall	12
2.2.5	Bearbetning	13
2.2.6	Modellantaganden	13
2.3	Redovisning av osäkerhetsmått.....	14

<i>B.3</i>	<i>Aktualitet</i>	<i>14</i>
3.1	Frekvens	14
3.2	Framställningstid	14
3.3	Punktlighet	14
<i>B.4</i>	<i>Jämförbarhet och sam användbarhet</i>	<i>14</i>
4.1	Jämförbarhet över tiden.....	14
4.2	Jämförbarhet mellan grupper	15
4.3	Sam användbarhet med annan statistik	15
<i>B.5</i>	<i>Tillgänglighet och förståelighet</i>	<i>16</i>
5.1	Spridningsformer.....	16
5.2	Presentation	16
5.3	Dokumentation	16
5.4	Tillgång till primärmaterial	16
5.5	Upplysningstjänster.....	16

A Allmänna uppgifter

A.1 Ämnesområde

Ämnesområde: Utbildning och forskning

A.2 Statistikområde

Statistikområde: Befolkningens utbildning

A.3 SOS-klassificering

Tillhör (SOS) Ja



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).

A.4 Statistikansvarig

Myndighet/organisation: Statistiska centralbyrån
Postadress: 701 89 Örebro
Besöksadress: Klostergatan 23
Kontaktperson: Therese Karlsson
Telefon: 019 – 17 66 21
Telefax: --
E-post: fornamn.efternamn@scb.se

A.5 Statistikproducent

Myndighet/organisation: Statistiska centralbyrån
Postadress: 701 89 Örebro
Besöksadress: Klostergatan 23
Kontaktperson: Therese Karlsson
Telefon: 019 – 17 66 21
Telefax: --
E-post: fornamn.efternamn@scb.se

A.6 Uppgiftsskyldighet

Uppgiftsskyldighet föreligger inte enligt lagen om den officiella statistiken (SFS 2001:99).

A.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.

A.8 Gallringsföreskrifter

Bevarandebehovet är under utredning. Under tiden sparas registret på SCB.

A.9 EU-reglering

Undersökningen omfattas inte av någon EU-reglering.

A.10 Syfte och historik

Syfte

Gymnasieungdomars studieintresse skall belysa intresset för att börja på högskolan bland gymnasieelever och hur intresset förändras över tiden.

De viktigaste målstorheterna är:

- Planer på högskolestudier inom tre år
- Planerad utbildningsinriktning på högskolestudierna

De viktigaste redovisningsgrupperna är gymnasieprogram och kön.

Historik

SCB genomförde den första undersökningen om gymnasieungdomars studieintresse under våren 1993 bland de elever som gick i gymnasieskolans avgångsklasser läsåret 1992/93. Nästa undersökning gjordes på hösten samma år, bland elever i gymnasieskolans avgångsklasser läsåret 1993/94. Syftet med att ändra referensperiod från vårtermin till hösttermin var att kunna presentera statistiken tidigare. Mellan läsåren 1993/94 och 2003/04 genomfördes undersökningen varje år. Därefter har den endast genomförts vartannat år.

A.11 Statistikanvändning

De viktigaste användarna av statistiken är universitet och högskolor, Högskoleverket och Utbildningsdepartementet. Statistiken används som underlag för information om samt planering av utbildningsprogram och kurser.

Innehållet i undersökningen beslutades i samråd med Användarrådet för befolkning, demografi och utbildning. Deltagande organisationer i detta råd är

Utbildningsdepartementet, Finansdepartementet, Socialdepartementet, Skolverket, Högskoleverket, Arbetsförmedlingen, Sveriges Kommuner och Landsting, Institutet för arbetsmarknadspolitisk utvärdering (IFAU), Institutet för social forskning (SOFI) samt Stockholms universitet.

A.12 Uppläggnings och genomförande

Formulering av frågor

SCB har formulerat frågorna i undersökningen i samråd med statistikanvändarna. Merparten av frågorna är återanvända från tidigare år. Blanketten var firsidig och innehöll 20 numrerade frågor, några av dem hade delfrågor vilket genererade totalt 31 frågor.

Frågan som mäter den viktigaste målstorheten har följande formulering:

1. Har du planer på att börja läsa på universitet eller högskola inom de närmaste tre åren?

Ja → *Gå till fråga 4*

Nej → *Gå till fråga 2*

Vet inte → *Gå till fråga 20*

Population och urval

Undersökningens rampopulation bestod av personer som gick årskurs 2 läsåret 2010/11. Rampopulationen hämtades från Skolverkets elevregister. Eftersom registret över elever i årskurs 3 läsåret 2011/12 inte var färdigt när undersökningen genomfördes på hösten, användes registret över elever i årskurs 2 under läsåret 2010/11 som urvalsram. Årskurs 3-registret var dock klart i samband med statistikframställningen. Detta register utnyttjades för att korrigera för täckningsbristerna. Årskurs 2-registret omfattade 121 842 personer. Årskurs 3-registret omfattade 114 846 personer. Rampopulationen stratifierades på region (län), programgrupp (naturvetenskapsprogrammet, samhällsvetenskapsprogrammet och teknikprogrammet och övriga program) samt kön.

Urvalsstorleken i denna undersökning var 21 977, vilket bedömdes ge tillräcklig precision i skattningarna av de viktigaste målstorheterna. Urvalet i denna undersökning brukar normalt ligga runt 5 000 individer

Datinsamling

Gymnasieungdomars studieintresse genomfördes som en post- och webbenkät. Den första omgången med frågeblanketter skickades ut den 4 oktober 2011. Tack- och påminnelsekort skickades den 13 oktober, ett påminnelsebrev med ny blankett den 28 oktober, 10 november samt den 16 december. Insamlingen avslutades i slutet av januari 2012. I och med den sista påminnelsen förkortades enkäten för att på så sätt få in fler svar. De frågor som fanns med i den korta enkäten var (men numrering från den långa enkäten) följande frågor: 1, 2, 3, 7, 11, 12, 15 och 16.

Att skicka fyra påminnelser och vid sista påminnelsen en förkortad enkät har inte gjorts tidigare. Resultatet av detta kommer att utvärderas inför nästa undersökningsomgång.

A.13 Internationell rapportering

Ingen internationell rapportering sker.

A.14 Planerade förändringar i kommande undersökningar

B Kvalitetsdeklaration

B.0 Inledning

SCB deklarerar kvaliteten i sina undersökningar enligt ett kvalitetsbegrepp som består av fem huvudkomponenter:

- (1) **Innehåll**, som framför allt gäller statistikens målstorheter.
- (2) **Tillförlitlighet**, som avser osäkerhetskällor och hur dessa påverkar statistiken.
- (3) **Aktualitet**, som omfattar tidsaspekter som spelar roll för hur väl statistiken beskriver nuläget.
- (4) **Jämförbarhet och sammanvändbarhet**, som avser möjligheter till jämförelser, över tiden och mellan grupper, samt till att använda statistiken tillsammans med annan statistik.
- (5) **Tillgänglighet och förståelighet**, som avser statistikens fysiska tillgänglighet och dess förståelighet.

I SCB:s MIS 2001:1 Kvalitetsbegrepp och riktlinjer för kvalitetsdeklaration av officiell statistik ges generella beskrivningar och definitioner av SCB:s kvalitetsbegrepp och de olika kvalitetskomponenterna, se www.scb.se.

Denna statistik redovisar intresset för högskolestudier inom tre år bland elever på gymnasieskolans avgångsklasser läsåret 2011/12. Statistiken baserades på uppgifter som samlades in via postenkäter till ett urval bestående av 21 977 elever. Undersökningen har genomförts sedan läsåret 1993/94. Fram tills läsåret 2003/04 genomfördes undersökningen varje år, därefter har den genomförts vartannat år.

B.1 Innehåll

1.1 Statistiska målstorheter

1.1.1 Objekt och population

Objekt

Statistikens objekt är individer.

Målpopulation

Målpopulationen utgjordes av de elever som gick sista året på gymnasieskolans nationella program, specialutformat program och international baccalaureate läsåret 2011/12.

1.1.2 Variabler

Variablerna insamlades dels via frågor i post- och webbenkät dels från Skolverkets elevregister läsåret 2010/11 och Registret över totalbefolkningen med aktualitet september 2011.

Registervariabler

Från Skolverkets elevregister läsåret 2010/11 hämtas gymnasieprogram. Gymnasieprogram utgörs av samtliga program inom gymnasieskolan förutom det individuella programmet. Det görs en gruppering av gymnasieprogram. De delas upp i naturvetenskapsprogrammet, samhällsvetenskapsprogrammet, teknikprogrammet samt övriga program.

I övriga program ingår: Barn- och fritidsprogrammet, Byggprogrammet, Elprogrammet, Energiprogrammet, Estetiska programmet, Fordonsprogrammet, Handels- och administrationsprogrammet, Hantverksprogrammet, Hotell- och restaurangprogrammet, Industriprogrammet, International baccalaureate, Livsmedelsprogrammet, Medieprogrammet, Naturbruksprogrammet samt Omvårdnadsprogrammet.

Specialutformade program ligger under närliggande nationellt program.

Från registret över totalbefolkningen med aktualitet september 2011 hämtas personnummer, kön och hemkommun. Indelningen av region (län) görs med hjälp av registervariabeln hemkommun.

Enkätvariabler

Nedan definieras de viktigaste enkätvariablerna. Övriga variabler samt samtliga svarsalternativ i varje fråga finns redovisade i enkäten som återfinns i dokumentationen *SCBDOK* under Statistikens framtagning på SCB:s webbplats, www.scb.se/UF0513.

Planer på högskolestudier inom tre år: Undersökningsspersonen gör själv en bedömning utifrån fasta svarsalternativ om han/hon planerar att börja på högskolan. Endast ett alternativ får anges. Svarsalternativen är;

- Ja
- Nej
- Vet inte

Planerad utbildningsinriktning på högskolestudierna: Här ingår de personer som svarar att de har planer på att börja studera på högskolan inom tre år. Undersökningsspersonen gör själv en bedömning utifrån fasta svarsalternativ vilket ämnesområde han/hon planerar att börja studera. Endast ett alternativ får anges.

Planerat universitet/högskola där studierna skall bedrivas: Här ingår de personer som svarar att de har planer på att börja studera på högskolan inom tre år. Undersökningsspersonen gör själv en bedömning utifrån fasta svarsalternativ vilket universitet/högskola han/hon planerar att studera vid. Endast ett alternativ får anges.

Intresse för utlandsstudier

Här ingår de personer som svarar att de har planer på att börja studera på högskolan inom tre år. Undersökningsspersonen gör själv en bedömning om han eller hon är intresserad av att genomföra en del av studierna utomlands. Endast ett alternativ får anges.

Avsedd längd på studierna

Här ingår de personer som svarar att de har planer på att börja studera på högskolan inom tre år. Undersökningsspersonen gör själv en bedömning hur länge han/hon planerar att läsa på universitet/högskola utifrån fasta svarsalternativ. Endast ett alternativ får anges.

Studieupplägg

Här ingår de personer som svarar att de har planer på att börja studera på högskolan inom tre år. Undersökningsspersonen gör själv en bedömning utifrån fasta svarsalternativ om hur han/hon önskar lägga upp studierna i form av fristående kurser eller program. Endast ett alternativ får anges.

Redovisningsvariabler

Nedan redovisas de viktigaste redovisningsvariablerna i undersökningen

- Intresse för högskolestudier inom tre år
- Planerad utbildningsinriktning på högskolestudierna
- Planerat universitet/högskola där studierna skall bedrivas
- Intresse för utlandsstudier
- Avsedd längd på studierna
- Studieupplägg för de fortsatta studierna (fristående kurser eller program)

Av sekretesskäl redovisas inte resultat som bygger på 3 observationer eller färre. Skattningar med för stor urvalsosäkerhet redovisas inte.

1.1.3 Statistiska mått

De statistiska storheterna (parametrarna) utgörs av **totaler** och **kvoter**. De viktigaste målstorheterna i undersökningen är:

- Andel av eleverna i gymnasieskolans avgångsklasser läsåret 2011/12 som planerade att studera på högskolan inom tre år.
- Andel av de studieintresserade eleverna som planerade att läsa inom ett visst ämnesområde.
- Andel av de studieintresserade eleverna som planerade att läsa vid ett visst universitet eller en viss högskola.

De totaler som redovisas är skattat antal personer med en viss egenskap. Storheter av typen kvot redovisas främst i form av skattad andel personer med en viss egenskap. Skattningarna redovisas med ett flertal gruppindelningar, vilka beskrivs i avsnitt *1.1.4 Redovisningsgrupper*.

1.1.4 Redovisningsgrupper

Statistiken redovisas på kön, gymnasieprogram (grupperat) samt för enstaka frågor region (län). Indelningen för gymnasieprogram är beskriven under registervariabler.

1.1.5 Referenstider

Referenstid för objekt är elever i gymnasieskolans avgångsklasser läsåret 2011/12. Referensperioden för variablerna är hösten 2011.

1.2. Fullständighet

Användare har visat intresse för redovisning av resultaten på regional nivå (kommunnivå). Resultaten bygger på ett för litet urval för att detta ska kunna tillgodoses.

B.2 Tillförlitlighet

2.1 Tillförlitlighet totalt

Ett statistikvärdes tillförlitlighet/osäkerhet beror på den totala avvikelserna (totala felet) mellan skattningen och målstorheten. Det totala felet kan beskrivas enligt följande:

Totala felet = Urvalsfel + Täckningsfel + Bortfallsfel + Mätfel + Bearbetningsfel

I denna undersökning förekommer alla fel. Den osäkerhet som orsakas av att skattningarna baseras på ett urval uppskattas med hjälp av 95 % -iga konfidensintervall. Det har dock inte genomförts några metodstudier för att kunna redovisa kvantitativa mått över storleken på andra osäkerhetskällor än urvalsfelet.

Under respektive avsnitt kan endast bedömningar om de olika felens effekt på skattningarna ges. Det är alltså viktigt att tänka på att de redovisade konfidensintervallen endast innefattar storleken på urvalsfelet.

2.2 Osäkerhetskällor

2.2.1 Urval

Urval

Undersökningens rampopulation består av personer som gick årskurs 2 på gymnasieskolan läsåret 2010/11. Rampopulationen hämtas från Skolverkets elevregister 2010/11. Den omfattade 121 842 personer.

Rampopulationen stratifierades efter gymnasieprogram, län och kön. Sammanlagt bildades 168 strata. Storleken på urvalet var 21 977 personer. Vid bestämningen av urvalsstorleken togs hänsyn till det förväntade bortfallet.

I varje stratum drogs ett obundet slumpmässigt urval. Urvalet fördelades så att felmarginalen för huvudfråga 1 (andelen studieintresserade) var lika stor i alla stratum, baserat på kännedom om skattningarnas ungefärliga storlek från tidigare undersökningar.

Estimation

För att skatta målstorheterna används ett estimationsförfarande som baseras på s.k. kalibrerade uppräkningsvikter. Detta förfarande, som utnyttjar hjälpinforma-

tion i form av registervariabler, används främst för att kompensera för den skevhet som kan förmodas uppstå till följd av de bortfalls- och täckningsproblem som föreligger men även för att reducera den slumpmässiga osäkerhet, den varians, som estimationsförfarandet är behäftat med.

I kalibreringen används följande registervariabler (hjälpvariabler):
Kön, riksområde, län, program, vårdnadshavarens födelse-land, vårdnadshavarens utbildningsnivå, vårdnadshavarens civilstånd, slutbetyg i årskurs 9. Särskilt vårdnadshavarens utbildningsnivå och slutbetyg i årskurs 9 har visat sig vara starka variabler.

Var och en av de hjälpvariabler som används i kalibreringen har valts utifrån minst ett av följande tre kriterier:

- (i) de samvarierar med svarsbenägenheten
- (ii) samvarierar med (viktiga) undersökningsvariabler
- (iii) avgränsar (viktiga) redovisningsgrupper

I synnerhet är det punkterna (i) och (iii) som varit vägledande. Om punkt (i) är uppfylld kan man förvänta sig att användandet av kalibrerade vikter leder till en reduktion av bortfallsskevheden. Är punkt (ii) uppfylld bör såväl variansen (samplingvarians + bortfallsvarians) som bortfallsskevheden reduceras för estimatorer baserade på de (viktiga) undersökningsvariablerna. Slutligen, om punkt (iii) är uppfylld bör variansen (samplingvarians + bortfallsvarians) minska för estimatorer av storheter definierade på redovisningsgruppsnivå.

För att reducera för den täckningsskevhet som uppstår till följd av den täckningsproblematik som föreligger, har de totaler som används för att fastställa de kalibrerade vikterna hämtats från det uppdaterade registret för årskurs 3. Detta förfarande bygger på ett implicit, icke kontrollerbart, modellantagande om hur undersökningsvariablerna är relaterade till kalibreringsvariablerna.

De kalibrerade vikterna har den egenskapen att de vid skattning av populations-totalen, för var och en av de registervariabler som används för att fastställa de kalibrerade vikterna, resulterar i en skattning som överensstämmer exakt med den registertotal som används för att fastställa de kalibrerade vikterna. Detta innebär bl.a. att alla antalsskattningar för kalibreringsvariablerna som baseras på de kalibrerade vikterna överensstämmer med de storheter som finns i Skolverkets Elevregister läsåret 2011/12.

Osäkerhetsintervall beräknas enligt formeln för nominellt 95-procentiga konfidensintervall:

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

där såväl punkt- som medelfelsestimaten beräknas med hjälp av ett av SCB egenutvecklat estimationsprogram, CLAN97. Under förutsättning att stickprovstorleken är stor samt

- (i) att den estimator som används är approximativt väntevärdesriktig,

- dvs. icke-skev
- (ii) att estimatorns varians endast beror av samplingvariansen och bortfallsvariansen

och

- (iii) att den hjälpinformation som används i kalibreringen är starkt korrelerad med de okända svarssannolikheterna

gäller att intervallet har en faktisk konfidensgrad på ca 95 procent, dvs. under förutsättningarna är sannolikheten ca 95 procent att man med det använda urvals- och estimationsförfarandet erhåller ett intervall som täcker in den sökta storheten.

Naturligtvis är inte samtliga ovanstående förutsättningar perfekt uppfyllda i praktiken. Även om hjälpvariablerna är klokt valda kommer estimatorn att vara behäftad med viss skevhet, delvis till följd av kvarstående bortfalls- och täckningsfel men också till följd av andra fel av systematisk karaktär. Vidare är det troligt att estimatorns varians underskattas. Utöver samplingvariansen och bortfallsvariansen tillkommer även ett variansbidrag från andra osäkerhetskällor av slumpmässig natur, men vanligtvis kan endast en del av detta variansbidrag "fångas upp" av den varians estimator som används. Tyvärr är det inte möjligt att avgöra i vilken utsträckning som avvikelser från förutsättningarna ovan föreligger och vilka effekter dessa avvikelser får för tolkningen av de osäkerhetsintervall som presenteras i termer av empirisk konfidensgrad. En rekommendation är dock att man bör vara försiktig med att tolka de presenterade intervallen som regelrätta konfidensintervall.

2.2.2 *Ramtäckning*

Målpopulationen var de elever som gick i årskurs 3 läsåret 2011/12. I rampopulationen ingick de elever som gick i årskurs 2 på gymnasieskolan läsåret 2010/11. Anledningen till att ett år gammalt register används för att skapa ramen är att det aktuella inte är klart när undersökningen genomförs.

I och med att rampopulationen skilde sig från målpopulationen förekommer täckningsbrister i denna undersökning. Övertäckningen i statistiken, dvs. enheter i rampopulationen som inte tillhör målpopulationen, utgörs av elever som avbrutit eller gjort uppehåll i sina studier under eller efter årskurs 2. Undertäckningen i statistiken, dvs. enheter i målpopulationen som inte finns med i rampopulationen, utgörs av elever som hade studieuppehåll läsåret 2010/11, men som hade återupptagit sina gymnasiestudier ett år senare.

För att minimera de täckningsbrister som finns i undersökningen används årskurs 3-registret vid statistikframställningen. Det går därmed att identifiera och ta hänsyn till övertäckningen med hjälp av det nya registret. Även undertäckningen identifieras med årskurs 3-registret.

En matchning mellan årskurs 2-registret och årskurs 3-registret visar följande:

- 8,9 procent av de elever som fanns i rampopulationen (årskurs 2-registret) fanns inte med i årskurs 3-registret. Dessa elever har definierats som övertäckning.
- 3,3 procent av de elever som ingick i årskurs 3-registret fanns inte med i rampopulationen (årskurs 2-registret). Dessa elever har definierats som undertäckning.
- 111 011 elever ingick i både årskurs 2- och årskurs 3-registret

Genom att kalibrera mot hjälptotaler från årskurs 3-registret i estimationen görs ett försök att justera för de täckningsbrister som uppstår i och med att urvalet dras ifrån årskurs 2-registret. Kvarstående täckningsbrister och deras effekt på andelsskattningarna är svåra att bedöma då ingen särskild studie gjorts av täckningsfel.

2.2.3 Mätning

Mätförfarandet

Undersökningen har genomförts som en **post- och webbenkät** till personer som går sista året på gymnasieskolan. Det finns både frågor med fasta och öppna svarsalternativ. För mer information om utseende och innehåll i enkäten, se dokumentationen *SCBDOK* under Statistikens framtagning på SCB:s webbplats, www.scb.se/UF05013.

Den första undersökningen av liknande karaktär gjordes 1993. Erfarenheter från hur frågorna fungerat i de tidigare undersökningarna har tagits till vara och justeringar/förbättringar har gjorts vid behov. Inför årets undersökning har några frågor lagts till och i vissa frågor har fler svarsalternativ lagts till. Detta har gjorts i samråd med Högskoleverket och Utbildningsdepartementet.

Indikationer på mätproblem

Det har inte gjorts några studier över eventuella mätfel. Det förekommer att en del uppgifter som lämnats på de öppna svarsalternativen är för knapphändiga för att kunna kodas.

Det finns ingen fråga som har fått en nämnvärt stor andel svarande för alternativet "annat" eller "övrigt". Detta alternativ finns med för att inte tvinga in respondenterna på svarsalternativ. Skulle någon fråga ha en hög andel svar för dessa alternativ skulle det kunna innebära problem med att tolka resultaten.

Enkäten innehåller inga frågor som kan bedömas som särskilt känsliga.

Effekter på statistiken

Ingen mätfelsstudie har gjorts, därför är det svårt att göra en bedömning av effekterna på resultaten. Det finns dock inga indikationer på att det finns speciellt stora problem i någon grupp av populationen, exempelvis så verkar det partiella bortfallet fördela sig slumpmässigt över redovisningsgrupperna.

2.2.4 Svartsbortfall

Svartsbortfall, vanligen förkortat till bortfall, uppstår när värden på en eller flera observationsvariabler (t.ex. registervariabler eller enkätfrågor) i en undersökning inte kan inhämtas.

- Saknas alla värden för en individ är det fråga om **objektsbortfall**.
- Saknas enbart vissa värden, handlar det om **partiellt bortfall**.

Objektsbortfall

Följande vägda bortfallsmått har använts vid beräkning av objektbortfallets storlek:

$$\text{Bortfallsandel} = 1 - \frac{\sum_S d_k}{\sum_S d_k + \sum_B d_k + u \sum_O d_k},$$

där d_k är designvikten, dvs. en vikt som tar hänsyn till urvalsdesignen och u är andelen med okänd status som tillhör målpopulationen. Här har u satts lika med 0,876, vilket innebär att bortfallsmåttet har beräknats med antagandet att 87,6 procent ej svarande i bortfallet tillhör målpopulationen.

Summeringarna i detta bortfallsmått har gjorts över följande delmängder:

- S: Element tillhörande målpopulationen för vilka fullständiga eller partiella svar erhållits.
- B: Element tillhörande målpopulationen som utgör objektbortfall.
- O: Element vars målpopulationsstatus är okänd (i detta fall ej avhörda).

Objektbortfallet var 44,4 procent.

Åtgärder för att reducera bortfallet

SCB har använt en uppföljningsstrategi där målet varit att reducera bortfallet så mycket som möjligt. Den första omgången med frågeblanketter skickades ut den 4 oktober 2011. Tack- och påminnelsekort skickades den 17 oktober, ett påminnelsebrev med ny blankett den 28 oktober, 10 november och en tredje påminnelse med ny blankett den 16 december.

Partiellt bortfall

Partiellt bortfall kan bero på att frågan är svår att förstå, är känslig, att uppgiftslämnaren glömmer att besvara frågan, att lämpligt alternativ saknas eller att instruktionerna vid hoppfrågor misstolkas. Till partiellt bortfall räknas även dubbelmarkeringar och svar som inte kan tydas. I denna undersökning varierar det partiella bortfallet från att vara 1–4 procent för de flesta frågor.

Metoder för bortfallskompensation

För att kompensera för objektbortfallet har en *kalibreringsestimator* använts (se avsnittet om urval och estimation).

Om bortfallet skiljer sig från de svarande med avseende på undersökningsvari-

ablerna så kan skattningarna som grundar sig på enbart de svarande bli skeva. Om exempelvis de svarande till högre andel vill påbörja högskolestudier än de som inte svarat finns risk för överskattning av andelen som planerar att påbörja högskolestudier.

För att försöka reducera eventuella bortfallsskevheter har vikter beräknats med hjälp av kalibrering. Storleken på kvarstående skevheter orsakade av bortfallet är dock svåra att mäta.

2.2.5 *Bearbetning*

Dataregistrering och kodning

Registrering av data sker genom att inkomna enkäter läses in maskinellt (skanning). Skanningprogrammet kontrollerar så att valida värden registreras vid skanningen. De värden som faller ut som fel i dessa kontrollerar rättas upp manuellt med stöd av skanningprogrammet, s.k. verifiering.

Eftersom nästan alla variabler i undersökningen är kategoriska med slutna svarsalternativ bör inte skanningen ge upphov till några stora fel vid registrering av data. Givetvis uppstår dock fel, därför görs granskning av skannade data (se avsnitt granskning och rättning).

Ingen kodning görs numera i undersökningen.

Granskning och rättning

Efter verifieringen görs ytterligare kontroller av registrerade data. Logiska kontroller görs så att svar på olika frågor inte är orimliga i förhållande till varandra. I de fall där endast ett svarsalternativ är tillåtet kontrolleras så att inte dubbelmarkeringar finns. Rättning av dubbelmarkeringar för kategoriska variabler sker efter en viss prioritetsordning. I vissa fall sker rättningen utifrån svaren på andra frågor i enkäten. Eftersom antalet upprättningar i denna undersökning har varit litet bedöms effekterna på de slutliga skattningarna som relativt små.

Partiellt bortfall (saknade värden) redovisas i en egen kategori när statistiken publiceras. Ingen imputering av saknade värden görs.

Framställning av resultat

Punktskattningar och konfidensintervall har beräknats med hjälp av ett av SCB egenutvecklat estimationsprogram.

Det är svårt att bedöma effekten av de fel som kan uppstå vid bearbetningen. Ingen studie har gjorts för att försöka bedöma storleken av denna felkälla.

2.2.6 *Modellantaganden*

Statistiken beror inte av andra modellantaganden än de som tidigare angivits under avsnittet 2.2.1 *Urval*.

2.3 Redovisning av osäkerhetsmått

95-procentiga konfidensintervall anges för samtliga skattningar i samtliga tabeller. Av sekretesskäl redovisas inte resultat som bygger på 3 observationer eller färre. Skattningar med för stor urvalsosäkerhet redovisas inte.

B.3 Aktualitet

3.1 Frekvens

Undersökningen genomförs och publiceras vartannat år.

3.2 Framställningstid

Referensperiod för huvudsakliga parametrar i undersökningen är hösten 2011. Statistiken publicerades den 16 april 2012.

3.3 Punktlighet

Statistiken var först planerad att publiceras 20 mars 2012 men flyttades sedan fram till den 16 april 2012 vilket också fullföljdes.

B.4 Jämförbarhet och sam användbarhet

4.1 Jämförbarhet över tiden

Till 2011 års undersökning gjordes en mindre ändring i frågeblankettens första och viktigaste fråga. Förändringen medförde främst att de som inte visste om de hade planer på att börja läsa på universitet eller högskola inom de närmaste tre åren ändå fick svara på fråga 20 som handlar om hur mycket information respondenten fått om högskolestudier från ett antal informationskällor.

1	Har du planer på att börja läsa på universitet eller högskola inom de närmaste tre åren?	<input type="checkbox"/> Ja	→ Gå till fråga 4
		<input type="checkbox"/> Nej	→ Gå till fråga 2
		<input type="checkbox"/> Vet inte	→ Gå till fråga 20

Till 2003 års undersökning gjordes en förändring i frågeenkäten. I den viktigaste frågan om intresset för högskolestudier ändrades bland annat ett av svarsalternativen. Nedan följer en fullständig redovisning av förändringarna i frågan om intresset för högskolestudier.

Ny blankett (2003 och senare)

1	Har du planer på att börja läsa på universitet eller högskola inom de närmaste tre åren?	<input type="checkbox"/> 1 Ja	→ Gå till fråga 3
		<input type="checkbox"/> 2 Nej	→ Gå till fråga 2
		<input type="checkbox"/> 3 Vet inte	→ Tack för ditt svar! Skicka in blanketten till SCB.

Gammal blankett (2002 och tidigare)

1	Har du planer på att börja läsa på universitet eller högskola inom de närmaste tre åren?	<input type="checkbox"/>	1 Ja → Gå till fråga 3
		<input type="checkbox"/>	2 Jag har inte bestämt mig än → Gå till fråga 3
		<input type="checkbox"/>	3 Nej → Gå till fråga 2

Dessa förändringar fick till följd att andelen som var intresserade att läsa på högskolan inom tre år ökade kraftigt (en kontrollgrupp som fick den gamla blanketten användes i undersökningen). Av de elever som erhöll den nya blanketten (2003 års och senare) planerade **60 procent** att läsa vidare på högskolan inom de närmaste tre åren. Motsvarande andel som erhöll den gamla blanketten (2002 års och tidigare) var endast **50 procent**. Skillnaden är statistiskt säkerställd.

Detta medför att inga jämförelser kan göras mellan undersökningar gjorda innan 2003 och efter 2003 i fråga om gymnasieelevernas studieintresse. Det visade sig att nästintill all ökning i studieintresset (2003 var studieintresset 60 procent och 2002 var det 49 procent) kunde hänföras till förändringen i frågekonstruktionen. Däremot kan jämförelser från 2003 års undersökning och senare göras.

En ändring om gjordes i 2011 års enkät och påverkar jämförbarheten är i fråga 2 (Varför är du inte intresserad av att läsa på universitet/högskola inom de närmsta tre åren?). Här var det tidigare bara möjligt att markera ett alternativ. I och med att frågan gjordes om 2011 lades fler svarsalternativ till samt att det också blev möjligt för respondenterna att svara fler än ett svarsalternativ.

En annan ändring som gjordes 2011 är att fler svarsalternativ lades till fråga 11 (Varför vill du läsa på universitet/högskola?) och det blev också möjligt att välja mer än ett svarsalternativ. En ny fråga infördes dock som löd Vilket av alternativen i fråga 11 är den viktigaste orsaken? Statistiken från tidigare år är därmed jämförbar men användare bör uppmärksamma utökningen av antalet svarsalternativ.

Att urvalet tillfälligt ökat från 5 303 till 21 977 kan påverka jämförbarheten, någon studie av detta har dock inte gjorts.

4.2 Jämförbarhet mellan grupper

Jämförelse är möjlig mellan de redovisningsgrupper som används, dvs. kön, gymnasieprogram samt för enstaka frågor även län.

4.3 Samanvändbarhet med annan statistik

Sammanvändbarhet med annan utbildningsstatistik är möjlig. Genom att utnyttja SCB:s registerstatistik om övergången från gymnasieskolan till högskolan ges möjlighet att få uppfattning om andelen gymnasieelever som inte fullföljer sina studieplaner.

B.5 Tillgänglighet och förståelighet

5.1 Spridningsformer

Statistiken publiceras i temarapporten *Gymnasieungdomars studieintresse 2011/12*. På SCB:s webbplats finns också utgivna pressmeddelanden och en länk till Statistikdatabasen. Delar av statistiken publiceras också i Utbildningsstatistisk årsbok och Fickskolan.

5.2 Presentation

Statistiken presenteras i form av tabeller, diagram och text.

5.3 Dokumentation

Framtagningen av statistikregistret och statistiken beskrivs i *SCBDOK*. Statistikens kvalitet beskrivs i föreliggande dokument, *Beskrivning av statistiken (BaS)*. Detaljerad information om mikrodata finns beskrivet i *Dokumentation av mikrodata (MetaPlus)*. Samtliga dokumentationer finns att tillgå på SCB:s webbplats, www.scb.se/uf0513.

5.4 Tillgång till primärmaterial

Primärdata för enskilda individer finns sparade för samtliga undersökningsår. Specialbearbetningar kan utföras på uppdragsbasis. Forskare, utredare m.fl. kan få tillgång till avidentifierat material efter en särskild prövning.

5.5 Upplysningstjänster

Therese Karlsson, enheten för utbildning och arbetsmarknad vid SCB.
Telefon: 019-17 66 21
e-post: fornamn.efternamn@scb.se