

## Utländska doktorander Vårterminen 2013

UF0544

*I denna beskrivning redovisas först allmänna 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

<b>A</b>	<b>Allmänna uppgifter .....</b>	<b>2</b>
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	Statistik användning .....	4
A.12	Uppläggning och genomförande .....	4
A.13	Internationell rapportering .....	5
A.14	Planerade förändringar i kommande undersökningar .....	5
<b>B</b>	<b>Kvalitetsdeklaration .....</b>	<b>6</b>
B.0	Inledning .....	6
B.1	Innehåll .....	6
1.1	Statistiska målstorheter .....	6
1.1.1	Objekt och population .....	6
1.1.2	Variabler .....	7
1.1.3	Statistiska mått .....	9
1.1.4	Redovisningsgrupper .....	9
1.1.5	Referenstider .....	10
1.2	Fullständighet .....	10
B.2	Tillförlitlighet .....	10
2.1	Tillförlitlighet totalt .....	10
2.2	Osäkerhetskällor .....	11
2.2.1	Urval .....	11
2.2.2	Ramtäckning .....	12
2.2.3	Mätning .....	13
2.2.4	Svarsbortfall .....	14
2.2.5	Bearbetning .....	15
2.3	Redovisning av osäkerhetsmått .....	16

<i>B.3</i>	<i>Aktualitet</i>	16
3.1	Frekvens	16
3.2	Framställningstid	16
3.3	Punktlighet	16
<i>B.4</i>	<i>Jämförbarhet och sam användbarhet</i>	16
4.1	Jämförbarhet över tiden	16
4.2	Jämförbarhet mellan grupper	16
4.3	Sam användbarhet med annan statistik	17
<i>B.5</i>	<i>Tillgänglighet och förståelighet</i>	17
5.1	Spridningsformer	17
5.2	Presentation	17
5.3	Dokumentation	17
5.4	Tillgång till primärmaterial	17
5.5	Upplysningstjänster	17

## **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:* Klostersgatan 23  
*Kontaktperson:* Mattias Fritz  
*Telefon:* 019-17 62 67  
*Telefax:* --  
*E-post:* fornamn.efternamn@scb.se

### **A.5 Statistikproducent**

*Myndighet/organisation:* Statistiska Centralbyrån  
*Postadress:* 701 89 Örebro  
*Besöksadress:* Klostersgatan 23  
*Kontaktperson:* Mattias Fritz  
*Telefon:* 019-17 62 67  
*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**

Begreppet utländsk doktorand avser i denna undersökning personer som för första gången har påbörjat forskarstudier inom två år efter inresedatum i Sverige (uppehållstillstånd eller invandringsdatum i *Registret över totalbefolkningen*). Begreppet ska inte förväxlas med svensk/utländsk bakgrund.

Andelen doktorander som är inresande (benämns utländska doktorander fortsättningsvis) har ökat under de senaste åren. Ungefär hälften av de utländska doktoranderna stannar kvar i Sverige efter avslutad forskarexamen. Det finns därför behov av att få veta mer om de utländska doktoranderna.

Dels saknas information om deras motiv till att studera i svensk forskarutbildning, hur de rekryterats till svensk forskarutbildning, vad de har för planer efter avslutad utbildning och hur de upplever forskarstudier i Sverige. Dessutom saknas information om deras tidigare utbildning i annat land, social bakgrund etc.

Syftet är att få information om hur de utländska doktoranderna rekryteras till forskarutbildningen, motiv till utbildning i Sverige, vad de har för bakgrund, hur de upplever svensk forskarutbildning och vad de har för framtidsplaner.

De viktigaste målstorheterna är:

- Föräldrarnas högsta utbildning
- Försörjningsform
- Orsak till forskarstudier i Sverige
- Grad av nöjdhet med handledarstöd
- Framtidsplaner

Genomförandet av undersökningen är beslutat i samråd med "Användarrådet för befolkning, demografi och utbildning", där det finns representanter från Universitetskanslersämbetet (f.d. Högskoleverket), Skolverket, Arbetsförmedlingen, Stockholms universitet, Utbildningsdepartementet, Sveriges kommuner och landsting, Finansdepartementet, Socialdepartementet, Institutet för arbetsmarknadspolitisk utvärdering (IFAU) och Institutet för social forskning (SOFI).

Önskemål från deltagare vid ett konsumentmöte i februari 2013 har också tagits i beaktning vid planering av undersökningen. Vid konsumentmötet fanns representanter från Karolinska institutet, Kungl. Tekniska högskolan, Stockholms universitets studentkår, Sveriges förenade studentkårer, Utbildningsdepartementet, Universitetskanslerämbetet, Universitets- och högskolerådet och Statistiska centralbyrån.

## **A.11 Statistikanvändning**

De viktigaste användarna av statistiken är universitet och högskolor, Universitetskanslerämbetet, Utbildningsdepartementet och andra myndigheter. Statistiken används som ett underlag för att få information om hur utländska doktorander upplever att forskarutbildningen fungerar samt om deras sociala bakgrund. Resultaten går att jämföra mot en kontrollgrupp av svenska doktorander (ej inresta doktorander) som fick ett förkortat frågeformulär.

## **A.12 Uppläggning och genomförande**

### **Population och urval**

Undersökningen består av två populationer. Urvalsramen avgränsas i *Universitets- och högskoleregistret* med aktualitet höstterminen 2012 till de som var studieaktiva under höstterminen 2012, inte avbrutit forskarutbildningen eller tagit forskarexamen under höstterminen 2012 samt för första gången registrerats i svensk forskarutbildning för högst 10 år sedan. Ofullständiga personnummer och personer som inte var folkbokförda i Sverige exkluderades ur ramen.

I huvudstudien ingick alla doktorander som för första gången hade aktivitetsgrad större än noll på forskarstudier inom två år efter inresedatum i Sverige (uppehållstillstånd eller invandringsdatum i RTB) och som var aktiva doktorander under vårterminen 2013 (avgränsas med selektionsfråga i enkät) och folkbokförda i Sverige. Med aktivitetsgrad större än noll menas att man inte bara var registrerad på forskarutbildningen, utan också i någon grad var studieaktiv. I huvudstudiens rampopulation ingick 4 657 personer.

I kontrollstudien ingick alla övriga doktorander efter att huvudstudiens population selekterats fram ur urvalsramen och som var aktiva doktorander under vårterminen 2013 (avgränsas med selektionsfråga i enkät) och folkbokförda i Sverige. I kontrollstudiens rampopulation ingick 11 366 personer. Syftet med kontrollstudien var att kunna jämföra svarmönster mellan utländska doktorander och övriga doktorander.

I huvudstudien undersöktes alla 4 657 personer i urvalsramen. Urvalet i kontrollstudien drogs som stratifierat obundet slumpmässigt urval. Urvalsstorleken för kontrollstudien var 3 003 personer.

### **Formulering av frågor**

SCB har formulerat frågorna i undersökningen i samråd med statistikanvändarna och mättekniker på SCB, som är specialister på frågekonstruktion. Blanketten till de utländska doktoranderna omfattade åtta A4-sidor med frågor och innehöll 30 numrerade frågor, några av dem hade delfrågor vilket genererade totalt 58 frågor. Blanketten till kontrollgruppen innehöll 22 numrerade frågor fördelade på sju A4-sidor. Några av dem hade delfrågor vilket genererade totalt 41 frågor. Samtliga 22 frågor i kontrollgruppens blankett fanns med i blanketten till de utländska doktoranderna. De övriga frågorna i blanketten till de utländska doktoranderna var av sådan karaktär att de inte passade för campusstudenter.

### **Datainsamling**

Undersökningen genomfördes som en post- och webbenkät. Pappersblanketten till de utländska doktoranderna var på engelska, men på webben fanns möjlighet att välja mellan svensk- och engelskspråkig blankett. Till kontrollgruppen fanns både pappersblankett och webblankett endast på svenska.

Den första omgången med frågeblanketter skickades ut den 6 maj 2013. Sedan skickades 3 påminnelser till dem som inte besvarat frågeblanketten. Ett tack- och påminnelsekort skickades den 15 maj. En första påminnelse med ny enkät skickades den 29 maj och en andra påminnelse med ny enkät den 13 juni. Insamlingen avslutades den 21 augusti 2013.

## **A.13 Internationell rapportering**

Ingen internationell rapportering sker.

## **A.14 Planerade förändringar i kommande undersökningar**

Ingen kommande undersökning är planerad.

## 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 sam användbarhet**, 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](http://www.scb.se).

### B.1 Innehåll

#### 1.1 Statistiska målstorheter

##### 1.1.1 Objekt och population

#### Objekt

Statistikens objekt är personer.

#### Målpopulation

Målpopulationen var studieaktiva doktorander i svensk forskarutbildning under vårterminen 2013, uppdelade på utländska doktorander för huvudstudien och övriga för kontrollstudien. För att tillhöra målpopulationen skulle de också ha blivit registrerade i svensk forskarutbildning för högst 10 år sedan samt vara folkbokförda i Sverige.

Målpopulationen för huvudstudien var alla doktorander som för första gången hade aktivitetsgrad större än noll på forskarutbildningen inom två år efter inresedatum i Sverige (uppehållstillstånd eller invandringsdatum i RTB) och som var aktiva doktorander under vårterminen 2013 (avgränsades med selektionsfråga i enkät) och folkbokförda i Sverige. Med aktivitetsgrad större än noll menas att man inte bara var registrerad på forskarutbildningen, utan också i någon grad var studieaktiv.

Med utländska doktorander avses personer som hos Migrationsverket angivit studier som grund för bosättning och som invandrat mindre än två år innan studierna påbörjades. Eftersom inte alla doktorander ansöker om uppehållstillstånd hos Migrationsverket så räknas även utrikes födda doktorander som invandrat mindre än två år innan doktorandstudierna påbörjades. Upphållstill-

stånden kan också förnyas, varpå en person kan ha vistats i Sverige längre än två år och ändå klassas som inresande. Doktorander med okänd bakgrund ingår i redovisningen av utländska doktorander. Okänd bakgrund har doktorander dels med okänt födelseland, dels med ofullständiga personnummer. Doktorander med ofullständigt personnummer ingår dock inte i målpopulationen för denna undersökning.

I kontrollstudien ingick alla övriga doktorander efter att huvudstudiens population selekterats fram ur urvalsramen.

### **Rampopulation**

Urvalsramen avgränsades i *Universitets- och högskoleregistret* med aktualitet höstterminen 2012 till de personer som var studieaktiva i svensk forskarutbildning under höstterminen 2012, inte avbrutit forskarutbildningen eller tagit forskarexamen under höstterminen 2012 samt för första gången registrerats i svensk forskarutbildning för högst 10 år sedan. Uppgiften om aktivitetsgrad rapporteras av lärosätena till LADOK och därifrån hämtas uppgiften till *Universitets- och högskoleregistret*. Ofullständiga personnummer och personer som inte var folkbokförda i Sverige exkluderades ur urvalsramen. De personer som kontaktat SCB efter utskick av frågeblanketter och sagt sig inte tillhöra rampopulationen, eller svarat på enkätens kontrollfråga att de inte tillhör rampopulationen har i efterhand exkluderats. I huvudstudiens nettoram ingick 4 657 personer och i kontrollstudiens nettoram ingick 11 366 personer.

#### *1.1.2 Variabler*

Variablerna insamlades dels via frågor i post- och webbenkät, dels från *Universitets- och högskoleregistret* och *Registret över totalbefolkningen*.

### **Registervariabler**

Från *Universitets- och högskoleregistret* hämtades uppgifter om universitet/högskola, forskningsområde (grov nivå) och tidpunkt för studiestart på forskarutbildningen.

Från *Registret över totalbefolkningen* (RTB) hämtades uppgifter om kön, ålder, medborgarskapsland (grupperat) och födelseland (grupperat).

### **Variabler i statistiken**

Nedan definieras de viktigaste variablerna i statistiken. 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/uf0544](http://www.scb.se/uf0544).

Av sekretesskäl redovisas inte resultat som bygger på 3 observationer eller färre. Skattningar med för stor urvalsosäkerhet redovisas inte, där konfidensintervallet täcker värdet noll för parameterskattningen.

*Föräldrarnas högsta utbildning:* Urvalspersonen gjorde själv en bedömning utifrån fasta svarsalternativ om vilken högsta utbildningsnivå dess mor respektive far har/hade. Endast ett alternativ fick anges. Svarsalternativen var;

- Förgymnasial utbildning eller ingen utbildning alls
- Gymnasial utbildning
- Eftergymnasial utbildning kortare än 3 år
- Eftergymnasial utbildning 3 år eller mer
- Forskarutbildning
- Vet inte

*Försörjningsform:* Urvalspersonen fick svara utifrån fasta svarsalternativ hur han/hon försörjde sig under vårterminen 2013. Flera alternativ fick anges. Svarsalternativen var;

- Doktorandanställning
- Annan anställning inom högskolan
- Utbildningsbidrag
- Studiemedel CSN
- Stipendium från Sverige (t.ex. SI)
- Stipendium från annat land än Sverige
- Anställning utanför högskolan (med tid för egen forskning)
- Anställning utanför högskolan (utan tid för egen forskning)
- Egen finansiering
- Annan finansiering

*Orsak till forskarstudier i Sverige:* Denna fråga ställdes endast till utländska doktorander (ej kontrollgruppen). Urvalspersonen fick svara på delfrågor med fasta svarsalternativ om hur viktiga olika faktorer var i deras beslut att välja Sverige för studier på forskarnivå. Svarsalternativen hade samma skala i alla delfrågorna; Mycket viktigt, Ganska viktigt, Inte alls viktigt, Inte aktuellt. Huvudfrågan var; "Hur viktigt var följande för dig i ditt beslut att välja just Sverige för dina studier på forskarnivå?" Delfrågorna var;

- Att utbildningen är på engelska?
- Att utbildningen är avgiftsfri?
- Möjligheten att få utbildningsbidrag/ doktorandtjänst under utbildningstiden?
- Att utbildningens kvalitet har ett gott rykte?
- Att utbildningens studiemiljö har ett gott rykte?
- Att den svenska arbetsmarknaden lockar/det finns karriärmöjligheter?
- Att Sverige är ett tryggt land?
- Annat



*Grad av nöjdhet med handledarstöd:* Urvalspersonen fick svara på delfrågor med fasta svarsalternativ i vilken grad de upplevt att dess handledare uppfyllt vissa handledarkriterier. Svarsalternativen hade samma skala i alla delfrågorna; I mycket stor grad, I ganska stor grad, I ganska liten grad, I mycket liten grad/Inte alls. Huvudfrågan var; ” I vilken grad har din/dina handledare... ” Delfrågorna var;

- ...låtit dina egna intressen styra valet av avhandlingsämne?
- ...visat intresse för dina studier på forskarnivå?
- ...diskuterat metodfrågor med dig?
- ...diskuterat teori med dig?
- ...gett dig råd kring vilken tidigare forskning du behöver läsa in dig på?
- ...gett konstruktiv kritik på ditt forskningsarbete?
- ...diskuterat dina framtidsplaner efter forskarutbildningen med dig?
- ...gett dig handledning i den omfattning du önskat?

*Framtidsplaner:* Urvalspersonen fick svara utifrån fasta svarsalternativ om han/hon har för avsikt att stanna i Sverige efter avslutad forskarutbildning. För huvudstudien följde sedan följdfrågor om orsaker beroende på svaret i denna fråga. Ett svarsalternativ fick anges. Svarsalternativen var;

- Ja
- Nej
- Kanske/Vet ej

### 1.1.3 Statistiska mått

Statistiken presenteras som summor och kvoter. De summor som redovisas är skattat antal personer med en viss egenskap. Kvoter 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 nedanstående redovisningsgrupper eftersom de bedöms kunna förklara skillnader i svarsmönster mellan utländska doktorander och övriga doktorander.

Redovisningsgrupper för utländska doktorander och övriga doktorander:

- Kön
- Ålder (3 grupper)
- Antal terminer sedan första registrering i svensk forskarutbildning (4 grupper)
- Forskningsämnesområde
- Lärosäte (krav på minst 200 aktiva doktorander våren 2013 för att särredovisas)

Redovisningsgrupper endast för utländska doktorander:

- Medborgarskapsland (2 grupper, EU/EES + Schweiz eller övriga)
- Födelseland (3 grupper)

### 1.1.5 Referenstider

Referensperioden för de flesta variablerna är vårterminen 2013. Några variabler avser tiden för hela forskarutbildningen och den varierar mellan personer.

## 1.2. Fullständighet

Statistiken över utländska doktorander ger en god bild över bl.a. motiv till studier i Sverige, social bakgrund, försörjningsformer, handledarstöd och framtidsplaner men redovisar t.ex. inte något kring doktorandernas verksamhet eller studieresultat.

## B.2 Tillförlitlighet

### 2.1 Tillförlitlighet totalt

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

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

Huvudgruppen, de utländska doktoranderna, totalundersöktes. Det innebär att alla personer i urvalsramen tillhörande huvudgruppen har fått möjlighet att besvara frågeformuläret. I och med det är skattningarna avseende de utländska doktoranderna inte påverkade av något urvalsfel. Skattningar för kontrollgruppen bygger dock på ett urval och lider således av urvalsfel. Övriga fel som bidrar till det totala felet förekommer i båda undersökningsgrupperna.

Den osäkerhet som orsakas av att skattningarna baseras på ett urval uppskattas med hjälp av 95-procentgiga konfidensintervall. Det har inte genomförts några metodstudier för att kunna redovisa kvantitativa mått över storleken på andra osäkerhetskällor än urvalsfelet. En mindre bortfallsanalys gjordes dock i samband med viktberäkningen. Detta beskrivs i Kalibreringsrapporten som finns på SCB:s webbplats, [www.scb.se/uf0544](http://www.scb.se/uf0544).

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

Urvalsramen avgränsades i *Universitets- och högskoleregistret* med aktualitet höstterminen 2012 till de personer som var studieaktiva i svensk forskarutbildning under höstterminen 2012, inte avbrutit forskarutbildningen eller tagit forskarexamen under höstterminen 2012 samt för första gången registrerats i svensk forskarutbildning för högst 10 år sedan. Uppgiften om aktivitetsgrad rapporteras av lärosätena till LADOK och därifrån hämtas uppgiften till Universitets- och högskoleregistret. Ofullständiga personnummer och personer som inte var folkbokförda i Sverige exkluderades ur urvalsramen. De personer som kontaktat SCB efter utskick av frågeblanketter och sagt sig inte tillhöra rampopulationen, eller svarat på enkätens kontrollfråga att de inte tillhör rampopulationen har i efterhand exkluderats. I huvudstudiens nettoram ingick 4 657 personer och i kontrollstudiens nettoram ingick 11 366 personer.

I huvudstudien undersöktes alla 4 657 personer i nettourvalsramen och därmed är huvudstudien inte påverkad av något urvalsfel. Urvalet i kontrollstudien drogs som stratifierat obundet slumpmässigt urval. Urvalsstorleken för kontrollstudien var 3 003 personer.

#### Estimation

För att skatta målstorheterna används ett estimationsförfarande som baseras på s.k. kalibrerade uppräkningsvikter. Detta förfarande, som utnyttjar hjälpinformatiön 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 användes följande registervariabler (hjälpvariabler):

Kön (2 grupper), tid sedan första registrering i svensk forskarutbildning (6 grupper), lärosäte (3 grupper), födelseland (3 grupper) och storlek på folkbokföringskommun (4 grupper).

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 (samplingvariens + bortfallsvariens) som bortfallsskevheden reduceras för estimatorer baserade på de (viktiga) undersökningsvariablerna. Slutligen, om punkt (iii) är uppfylld bör variansen (samplingvariens + bortfallsvariens) minska för estimatorer av storheter definierade på redovisningsgruppsnivå.

För att reducera för den täcknings skevhet som uppstår till följd av den täckningsproblematik som föreligger, har de totaler som används för att fastställa de kalibrerade vikterna hämtats från Registret över totalbefolkningen (RTB) och Universitets- och högskoleregistret. 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 populationstotalen, 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.

Osäkerhetsintervall beräknas enligt formeln för nominellt 95-procentiga konfidensintervall:

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

där såväl punkt- som medelsfelsestimaten 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 estimatören att ha en 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.

### 2.2.2 Ramtäckning

Det förekommer täckningsbrister i denna undersökning. Omfattningen av täckningsbristerna var vid skapandet av rampopulationen helt okänd, men innan utskick av enkäter framkom information om betydande övertäckning i huvudstudien

då 12 procent av urvalet inte var folkbokförda i Sverige och 2 procent lämnat Sverige. Motsvarande andel för kontrollstudien var 0,1 procent som inte var folkbokförda i Sverige och 0,7 procent som lämnat Sverige. Även under insamlingen framkom information om viss övertäckning. En del personer som enligt registeruppgifter gällande höstterminen 2012 fortfarande var aktiva i svensk forskarutbildning och folkbokförda i Sverige har meddelat SCB via frågeformulärets kontrollfråga eller på annat sätt att de inte tillhör målpopulationen.

Till övertäckning räknas också personer som tagit ut examen under vårterminen 2013 före mättillfället som påbörjades i maj. Denna del av övertäckningen är skattad till 6 procent av hela rampopulationen genom att i efterhand ta fram registerantal över antalet doktorander som examinerats före mättillfället. Tillsammans med tidigare känd övertäckning skattas övertäckningen till 11 procent av hela rampopulationen.

Undertäckningen, dvs. personer som inte fanns med i rampopulationen fast de borde, utgörs främst av personer som för första gången var aktiva i svensk forskarutbildning under våren 2013. Undertäckningen är skattad till 8 procent av hela rampopulationen genom att i efterhand ta fram registerantal över antalet doktorander som påbörjade forskarstudier under vårterminen 2013. Undertäckningen bedöms vara av ungefär lika stor grad som övertäckningen. Täckningsbristernas effekt på skattningarna är svåra att bedöma då ingen särskild studie gjorts av täckningsfel. Effekterna skiljer sig troligen åt mellan olika frågor och är sannolikt större för gruppen utländska doktorander jämfört med övriga doktorander.

### 2.2.3 Mätning

#### **Mätförfarandet**

Undersökningen har genomförts som en post- och webbenkät till personer som var aktiva i svensk forskarutbildning vårterminen 2013. I en fråga ombads urvalspersonerna att fylla i vilket år de var födda. Detta tillsammans med en fråga om kön var kontrollfrågor för att bedöma rimligheten i att rätt person hade svarat. Samtliga övriga frågor hade fasta svarsalternativ men vissa frågor kompletterades med öppna svarsrutor för att fånga upp svar som inte passade in bland de fasta alternativen. För mer information om utseende och innehåll i enkäten, se dokumentationen *SCBDOK* under Statistikens framtagning på SCB:s webbplats, [www.scb.se/uf0544](http://www.scb.se/uf0544).

Detta är den första undersökningen med liknande karaktär riktat mot utländska doktorander i svensk forskarutbildning, så några erfarenheter från hur liknande frågorna i tidigare undersökningar fungerat finns inte.

#### **Indikationer på mätproblem**

Det har inte gjorts några studier över eventuella mätfel. Ett stort partiellt bortfall, dvs. att många låter bli att besvara en viss fråga, kan indikera mätproblem. Det kan exempelvis vara att frågan är känslig, svår att förstå eller att något svarsalternativ saknas. Det partiella bortfallet varierar mellan 1 och 2 procent på alla frågor vilket anses normalt. Ingen fråga sticker ut när det gäller partiellt bortfall.

Känsliga frågor kan också få skevhet i svarsmönster om respondenter medvetet svarar fel. Enkäten innehåller inga frågor som kan bedömas som särskilt känsliga, men frågor om diskriminering samt doktorandernas handledare och handledarstöd kan upplevas känsliga, särskilt som forskarstudierna fortfarande pågick vid tiden för undersökningen.

Några av frågorna var s.k. attitydfrågor, för vilka det är ett välkänt problem att definiera och uppskatta mätfel och dess inverkan på resultaten. Exempelvis kan personer från olika delar av världen med olika kulturella bakgrunder uppleva samma grad på en attitydfråga olika. Två doktorander som fått lika mycket tid med handledare kan uppleva tillräckligheten olika.

### **Effekter på statistiken**

Ingen mätfelsstudie har gjorts, därför är det svårt att göra en bedömning av effekterna på resultaten.

#### *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*

Objektsbortfallet i huvudstudien var 42,1 procent och i kontrollstudien 31,4 procent, korriberat för känd övertäckning. Bland de utländska doktorander som genomfört forskarstudier i Sverige längst är det färre som deltagit i undersökningen. Några har meddelat att de inte anser sig vara utländska doktorander, även om definitionen i undersökningen påvisar det. Detta är också skälet till att vi avgränsat populationen till att doktoranden ska ha påbörjat forskarstudier för högst 20 terminer sedan för att tillhöra målpopulationen.

#### *Åtgärder för att reducera bortfallet*

SCB har använt en uppföljningsstrategi där målet varit att reducera bortfallet till så stor grad som möjligt med de på förhand tilldelade resurserna. Insamlingen kom igång ganska sent under vårterminen 2013, 6 maj, men insamlingen låg öppen till mitten av augusti. Ett påminnelsekort och två påminnelser med enkäter skickades ut, det sista 13 juni.

I huvudstudien utgjordes största delen objektsbortfall av ej anträffade personer, men också en betydande del postreturer. Postreturer bland utländska personer indikerar att de inte längre finns kvar i landet. Nästan hela objektsbortfallet i kontrollstudien utgjordes av ej anträffade personer.

Möjligheten för telefonuppföljning i bortfallet för huvudstudien undersöktes, men en bedömning gjordes att det skulle bli svårt att få tag i dessa personer. Dessutom fanns inte engelskspråkig intervjupersonal tillgänglig förrän sent under hösten 2013.

### ***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 varierade det partiella bortfallet för enskilda frågor mellan 0,1 och 1,8 procent för huvudgruppen och mellan 0,1 och 1,1 procent för kontrollgruppen. Frågorna om språkkunskaper i svenska (endast till huvudgruppen) hade högst partiellt bortfall.

### ***Metoder för bortfallskompensation***

För att kompensera för objektsbortfallet har en *kalibreringsestimator* använts, se avsnittet om estimation under 2.2.1 *Urval*. Om bortfallet skiljer sig från de svarande med avseende på undersökningsvariablerna så kan skattningarna som grundar sig på enbart de svarande bli skeva.

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 av öppna svar har gjorts 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. Partiellt bortfall (saknade värden) redovisas i en egen kategori i publicerad statistik. Ingen imputering av saknade värden har gjorts.

### ***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 avsnitten i delkapitlet 2.2 *Osäkerhetskällor*.

## 2.3 Redovisning av osäkerhetsmått

Resultaten från undersökningen presenteras i en temarapport "Utländska doktorander" där utvalda frågor redovisas som andelar utan konfidensintervall (urvalsosäkerhet). Istället redovisas skattningar med konfidensintervall i *Statistikdatabasen (SSD)* och på webbsidan [www.scb.se/uf0544](http://www.scb.se/uf0544) under *tabeller och diagram*. Av sekretesskäl redovisas inte resultat som bygger på 3 observationer eller färre. Skattningar med för stor urvalsosäkerhet redovisas inte, dvs. där konfidensintervallet täcker värdet noll för parameterskattningen.

## B.3 Aktualitet

### 3.1 Frekvens

Detta är första gången undersökningen genomförs. Det finns inga planerade framtida undersökningsomgångar.

### 3.2 Framställningstid

Referensperiod för huvudsakliga parametrar i undersökningen är vårterminen 2013. Insamlingen av statistiken skedde under maj 2013 – augusti 2013. Statistiken publicerades den 16 december 2013.

### 3.3 Punktlighet

Statistiken publicerades enligt tidsplanen.

## B.4 Jämförbarhet och sam användbarhet

### 4.1 Jämförbarhet över tiden

Ej relevant eftersom detta är första gången undersökningen genomförs. Inga fler undersökningsomgångar är planerade.

### 4.2 Jämförbarhet mellan grupper

I huvudundersökningen totalundersöktes rampopulationen för att kunna bryta ner statistiken på redovisningsgrupper i största möjliga utsträckning. Syftet med att undersöka en kontrollgrupp av doktorander som inte klassats som inresande (utländska doktorander) var att kunna jämföra resultat från de utländska doktoranderna med resultat från resterande doktorander.

I vissa enkätfrågor ombads urvalspersonen göra bedömningar där skillnader i förväntningar påverkar bedömningen. Enkätfrågor där svaren är subjektiva



bedömningar mäter upplevda förhållanden som inte nödvändigtvis speglar faktiska förhållanden. Olika personer kan bedöma samma förhållande på olika sätt beroende på förväntningar och vilken nivå som krävs för att känna sig nöjd. Enkäten innehöll t.ex. frågor om grad av läs- och skrivkunskaper i engelska och svenska, där upplevda kunskaper inte behöver vara likställt med faktiska kunskaper. Inom vissa forskningsämnesområden behöver man kanske inte lika stora språkkunskaper som i andra. Därför bör man överväga om skillnader i resultat mellan grupper för variabler som grundas på subjektiva bedömningar kan förklaras av att grupperna av någon anledning kan uppleva samma förhållande olika innan slutsatser dras.

#### **4.3 Samanvändbarhet med annan statistik**

SCB har inte publicerat någon annan statistik inom området som är lämplig att sammanvända med denna studie

### **B.5 Tillgänglighet och förståelighet**

#### **5.1 Spridningsformer**

Statistiken publiceras som en Temarapport "Utländska doktorander" där utvalda frågor redovisas och kommenteras. På SCB:s webbplats, [www.scb.se/uf0544](http://www.scb.se/uf0544), finns också pressmeddelande, tabeller och diagram samt en länk till *Statistikdatabasen* (SSD) där en utökad redovisning av undersökningens resultat finns.

#### **5.2 Presentation**

Statistiken presenteras i form av tabeller, diagram och text.

#### **5.3 Dokumentation**

Framtagningen av statistikregistret och statistiken beskrivs i *SCBDOK*. Statistikkens kvalitet beskrivs i föreliggande dokument, *Beskrivning av statistiken (BaS)*. Detaljerad information om mikrodata finns beskrivet i *Dokumentation av mikrodata (MetaPlus)*. Samtliga dokumentationer finns på SCB:s webbplats, [www.scb.se/uf0544](http://www.scb.se/uf0544).

#### **5.4 Tillgång till primärmaterial**

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

#### **5.5 Upplysningstjänster**

Mattias Fritz, enheten för utbildning och arbetsmarknad vid SCB.  
Telefon: 019-17 62 67  
E-post: [fornamn.efternamn@scb.se](mailto:fornamn.efternamn@scb.se)