

Hushållens icke-vinstdrivande organisationer 2005

1 Inledning

Emma-projektet, eller paraplyprojektet för förbättring av den ekonomiska statistiken, omfattar i huvudsak förbättringsförslagen i åtgärdsplanen för den ekonomiska statistiken. Åtgärdsplanen har sin utgångspunkt i utredningen om översyn av den ekonomiska statistiken (SOU 2002:118).

Inom ramen för Emma-projektet startades ett delprojekt för att förbättra kvaliteten i undersökningen "Hushållens icke-vinstdrivande organisationer" (HIO) avseende undersökningsår 2005. HIO är en årlig undersökning sedan år 2004 men provundersökningar har gjorts avseende åren 1999 och 2001. HIO utgör ett viktigt underlag för NationalRäkenskaperna (NR) vad gäller beräkning av sektorn "Hushåll, inklusive ideella organisationer".

Ett antal områden pekades ut som prioriterade inom delprojektet:

- att öka inflödet av besvarade blanketter
- att uppskatta bortfallets inverkan på skevhet (bias) i skattningarna och även (om möjligt) föreslå biasreducerande åtgärder
- att undersöka om ytterligare källor kan användas för att förbättra avgränsning/ stratifiering av undersökningspopulationen

Arbete har tidigare gjorts för att hitta effektiva källor för avgränsning/ stratifiering av HIO varför detta område, av resursskäl, prioriterades ned.

2 Bakgrund

HIO är utformad så att den främst ska uppfylla de behov av uppgifter NR har på denna sektor. Att undersöka ideella organisationer jämföras juridiskt med att undersöka individer vilket innebär att det är *frivilligt* att delta i undersökningen. Undersökningen vänder sig till ett urval av organisationer och huvudvariablerna som mäts är intäkter och kostnader specificerade på olika intäkts- och kostnadsslag samt investeringar och tillgångar med specifikationer. Även uppgifter om organisationens verksamhet, medlemsantal och medelantal sysselsatta samlas in. Undersökningen innehåller också uppgifter som SCB:s finansräkenskaper har behov av, i huvudsak variablerna finansiella tillgångar och skulder samt en specifikation av dessa på olika tillgångs- och skudslag.

Undersökningen omfattar i princip två olika typer av blanketter, en som vänder sig till de minsta organisationerna och en som vänder sig till de större organisationerna. En större blankett utökad med uppgifter om finansiella tillgångar och skulder skickas ut till organisationer som förväntas ha stora finansiella tillgångar.

Den minsta blanketten innehåller uppgifter om antal medlemmar, antal sysselsatta och organisationens huvudsakliga verksamhet. Tanken är att

organisationens resultat- och balansräkning ska bifogas blanketten så att personer vid SCB kan fylla i de ekonomiska uppgifterna. Syftet med denna enkla blankett är att hålla nere uppgiftslämnarbördan så långt det går och därigenom öka möjligheten att många organisationer besvarar enkäten.

Den större blanketten innehåller, förutom uppgifterna i den minsta, också uppgifter om organisationens verksamhet fördelad på olika områden samt en specifikation av totala intäkter, kostnader, tillgångar och investeringar.

3 Undersökningens utformning år 2005

3.1 Avgränsning av rampopulationen

Samtliga Juridiska Enheter (JE) i SCB:s företagsregister (FDB) i november år 2005 klassade som tillhörande sektorn "Hushållens icke-vinstdrivande organisationer utom registrerade trossamfund" enligt standarden för institutionell sektorindelning 2000 (INSEKT 2000) avgränsas till rampopulationen.

Antal organisationer var totalt 168 044 inom sektor 710 men de flesta var mycket små vad gäller ekonomisk verksamhet. För att ringa in organisationer med betydande ekonomisk aktivitet ska organisationen uppfylla något av nedanstående kriterium:

- Organisationer med anställda
- Organisationer med lönesumma
- Organisationer med minst en inlämnad kontrolluppgift (KU)
- Organisationer med tillgångar över 10 miljoner kronor

Efter att ovanstående kriterium tillämpats tillhör 28 692 organisationer den så kallade undersökningspopulationen.

3.2 Stratifiering

Undersökningspopulationen avseende HIO stratifieras i två dimensioner för att möjliggöra effektiva urval, dels efter krav på redovisning och dels efter organisationernas storlek.

Den branschstratifiering som görs är:

Branschstratum	Ingående branschkode (NACE Rev 1.1)
65-67	65-67 (Finansiell verksamhet)
70000	70 (Fastighetsverksamhet)
80000	80 (Utbildning)
85000	85 (Hälsa- och sjukvård mm.)
91000	91 (Intresseorg. samt religiösa samf.)
92000	92 (Rekreativ verksamhet)
Bland	Resten, däribland okänd bransch

I undersökningspopulationen varierar bidraget till det som ska mätas (intäkter, kostnader mm.) mycket mellan olika organisationer. För att dra ett

effektivt urval måste undersökningspopulationen homogeniseras i termer av det som ska mätas. Ett bra storleksmått vad gäller HIO är svårt att hitta men för att ändå göra en storleksindelning av organisationerna har lönesumma valts som storleksmått.

3.3 Typ av blankett

Av de 28 692 företagen undersökningspopulationen tilldelades 92 företag den större blanketten utökad med uppgifter om finansiella tillgångar och skulder (typ C), 7 347 företag den ordinarie större blanketten (typ B) och 21 253 företag den mindre blanketten (typ A).

3.4 Urval

Den totala urvalsstorleken om 2 000 organisationer i HIO har bestämts i huvudsak efter ekonomiska ramar. De 2 000 organisationerna har sedan fördelats över stratum efter några huvudprinciper:

- Resultat från de viktigaste redovisningsgrupperna ska redovisas med samma precision
- En urvalsstorlek om minst 30 organisationer i varje stratum
- Inom en redovisningsgrupp fördelas urvalet så effektivt som möjligt över storleksgrupper (med avseende på minimerad varians)
- Urvalsvikten i ett enskilt stratum får inte vara orimligt hög

Ett obundet slumpmässigt urval utan återläggning drogs inom varje stratum.

4 Ökat inflöde

4.1 Inflöde i tidigare undersökningar

Att undersöka ideella organisationer jämföras, som tidigare nämnts, juridiskt med att undersöka individer vilket innebär att det är *frivilligt* att delta i undersökningen. Det är ofta svårt att motivera uppgiftslämnare till att fylla i blanketten och följaktligen har inflödet i HIO historiskt varit lågt. Andelen utvalda organisationer som besvarade blanketten i 2004:års undersökning uppgick till 47,2%.

Ett bortfall på drygt 50 % är högt, dels med avseende på precision i de skattningar som ska beräknas men också med avseende på bortfallsbias. Som kompensation för de organisationer som inte besvarar enkäten används uppgifter från dem som besvarat enkäten. Fördelar sig bortfallet slumpmässigt inom ett stratum genererar denna metod ingen bias (skevhet) bara en variansökning. Fördelar sig bortfallet däremot inte slumpmässigt kommer skattningarna att vara behäftade med bias. Med bias (skevhet) i en skattning menas här att ett 95%-igt konfidensintervall för en skattad total (baserad på de svarande i urvalet) inte täcker värdet på den sanna totalen.

De år som HIO gjorts har bortfallsandelen legat konstant på drygt 50 %. Med en återkommande bortfallsandel på drygt 50 % är frågan om det är

meningsfullt att göra reguljär statistik över HIO berättigad. Därför görs denna satsning avseende år 2005 som syftar till att öka inflödet av enkäter (resultat- och balansräkningar i förekommande fall).

4.1 Inflöde HIO 2005

Satsningen på ökat inflöde har framförallt inneburit fler påminnelser, både via post och via telefon. *Nytt i årets undersökning är påminnelser via telefon.*

I april år 2006 sändes blanketter ut till de 2 000 utvalda organisationerna och sista insändningsdatum var satt till 19 maj. Tabell 1 visar att knappt 30 % då hade kommit in och påminnelser (enbart via post) gick ut den 24 maj till resterande drygt 70 % av organisationerna.

Tabell 1 Antal och andel inkomna blanketter efter de olika påminnelserna

Inkomna	Antal svar	Ackumulerat	Totalt i urval	Andel svar, %	Ackumulerat, %
Inkomna	558	558	2 000	27,9	27,9
Påminnelse 0524	260	818	2 000	13,0	40,9
Påminnelse 0615	241	1 059	2 000	12,1	53,0
Påminnelse 0824	181	1 240	2 000	9,1	62,0

Ytterligare 13 % av organisationerna kom då in med svar, se tabell 1. Den påminnelse som gick ut den 15 juni innebar användning av både post och telefon. Ett slumpmässigt urval om 150 organisationer bland de största icke inkomna organisationerna gjordes och påmindes via telefon. Urvalets storlek i respektive stratum, dvs. fördelningen av de 150, gjordes med hjälp av proportionell allokering. De 150 fördelades proportionellt mot bortfallets storlek, dvs. enligt principen större bortfall fler telefonpåminnelser. Eftersom vi önskar samma svarsandel i alla totalundersökta stratum och vi tror att telefon ger högre svarsfrekvens jämfört med blankett är det rimligt att göra så. Resterande icke inkomna organisationer fick påminnelse via blankett.

Denna omgång kom ytterligare 12 % in med svar, se tabell 1. Fördelningen mellan inkomna via post respektive telefon kan studeras i nedanstående tabell 2. Materialet är litet och studien inte upplagd som ett experiment enligt alla konstens regler (inte lika många per telefon som blankett i varje stratum), men betraktar man svarsbenägenheten som lika i alla totalundersökta strata kan man summera över strata och få en hygglig bild av resultatet, se tabell 2. Resultatet pekar mot att inflödet från dem som påmindes via telefon är högre.

Tabell 2 Jämförelse av blankettinflöde efter strategi för påminnelse

Påminnelse 0615	Antal påminda	Antal svar	Andel svar, %
Telefon	150	39	26,0
Blankett	1 024	202	19,7

En sista påminnelse gick ut den 24 augusti där de totalundersökta som tidigare påmints via blankett nu telefonpåmindes. Totalt telefonpåmindes 404 organisationer och 540 blankettpåmindes. Efter att de stora organisationerna plockats ut för att telefonpåminnas drogs ett slumpmässigt urval bland de mindre. Det innebär att antalet telefonpåminda blev totalt 404. Allokeringssprincipen för att välja ut mindre organisationer för telefonpåminnelse följde den ovan beskrivna principen.

Tabell 3 Jämförelse av blankettinflöde efter strategi för påminnelse

Påminnelse 0824	Antal påminda	Antal svar	Andel svar, %
Telefon	404	117	29,0
Blankett	540	64	11,9

Resultatet pekar även denna gång mot att inflödet från dem som påmindes via telefon är högre jämfört med de som påmindes via blankett, se tabell 3. Troligen mer tillförlitligt resultat eftersom materialet är större (åtminstone i termer av antal telefonpåminnelser).

Fler påminnelser var inte rimligt, undersökningen är trots allt frivillig och därför är det orimligt att påminna oändligt många gånger. Slutsatsen måste ändå bli den att extra påminnelse (särskilt i form av telefonpåminnelser) lönar sig eftersom andel inkomna år 2005 är drygt 60 % att jämföra med år 2004 då andel inkomna var knappt 50 %.

I tabell 4 nedan presenteras antal i undersökningspopulationen (Antal i ramen), antal utvalda (Antal utvalda), antal svarande (Antal svarande) och andel svarande (Andel svarande) per blanketttyp samt totalt för hela undersökningen år 2005. Tabellen visar att andelen inkomna blanketter är störst vad gäller den enkla blanketten. Det styrker antagandet om att en enklare blankett motiverar uppgiftslämnarna att fylla i och skicka in.

Tabell 4 Antal och andel inkomna efter typ av blankett

Blankett Typ	Antal i ramen	Antal utvalda	Antal svarande	Andel svarande, %
A	21 253	754	524	69,5
B	7 347	1 154	663	57,5
C	92	92	53	57,6
Totalt	28 692	2 000	1 240	62,0

Notera att organisationer med blanketttyp B och C i större utsträckning telefonpåmindes eftersom stora organisationer i första hand var föremål för telefonpåminnelser. Trots detta är inflödet märkbart sämre vad gäller organisationer med blanketttyp B och C.

En renodlad studie över skillnader i inflöde mellan liten och stor blankett har gjorts genom ett experiment i provundersökningen år 2001. Där fick de mindre organisationerna slumpmässigt blankett A eller B och resultatet från den studien visar också på att liten blankett medför högre inflöde.

5 Bortfallets inverkan på skevhet (bias) i skattningarna

Bortfallet kompenseras genom omviktning inom varje stratum. Den metoden bygger på att de svarande organisationerna kan ses som ett obundet slumpmässigt urval inom stratum. Ett problem är att man ofta inte vet så mycket om bortfallets struktur. Men, vad gäller undersökningen avseende år 2004 har bortfallsproblematiken studerats genom att använda variabler som är kända för hela undersökningspopulationen och kan antas samvariera med ett antal av huvudvariablerna som ska mätas.

5.1 Studie över bortfallets inverkan på skattningarna

Till varje organisation i undersökningspopulationen kan uppgift om sysselsatta från den Registerbaserade ArbetsMarknadsStatistiken (RAMS) avseende år 2004 påföras.

Tabell 5 visar en summering av antal sysselsatta enligt RAMS för hela undersökningspopulationen, dvs. ett sant värde (facit). Tabell 5 visar också en skattning av totalt antal sysselsatta (pSyss) samt medelfel (sSyss) baserad på de svarande i HIO-urvalet. Skattningen avviker signifikant (utanför ett 95%-igt konfidensintervall) från det sanna värdet. En underskattning av det sanna värdet föreligger i samtliga redovisningsgrupper.

Tabell 5 visar också en skattning av totalt antal sysselsatta (pSyss) samt medelfel (sSyss) baserad på hela urvalet i HIO (Utvalda). Den skattningen ligger i nivå (inom ett 95%-igt konfidensintervall), till och med mycket nära, det sanna värdet vilket ytterligare styrker påståendet att bortfallet introducerar bias i skattningarna.

Vad gäller branschgruppsskattningarna, baserade på hela HIO-urvalet, föreligger ingen systematik vad gäller över- eller underskattning.

Tabell 5 Sant värde samt motsvarande uppgifter skattade med svarande i HIO och med hela HIO-urvalet (PRams = skattning av antal sysselsatta, SRams = medlefel)

BGrupp	Antal sysselsatta enligt RAMS 2004				
	Sant värde	Svarande		Utvalda	
		pSyss	sSyss	pSyss	sSyss
65-67	1 140	605	229	1 159	95
70000	1 129	488	167	738	252
80000	16 189	14 753	1 117	15 573	729
85000	11 200	8 053	539	10 897	222
91000	26 049	24 781	3 339	27 734	2 115
92000	18 865	16 343	1 886	18 536	1 067
bland	6 434	5 414	745	6 406	475
<i>Totalt</i>	<i>81 006</i>	<i>70 437</i>	<i>4 108</i>	<i>81 044</i>	<i>2 548</i>

Motsvarande studie har gjorts för ett antal andra variabler (som lönesumma och antal anställda) där man också känner värdena för varje organisation i undersökningspopulationen och de har visat samma resultat vad gäller bortfallets inverkan på skattningarna.

5.2 Nytt skattningsförfarande i HIO

Det låga inflödet av ifyllda blanketter introducerar skevhet (bias) i skattningarna kan vi sluta oss till med ledning av studien i avsnitt 5.1. Ett sätt att minska skevheten i skattningarna är att använda ett skattningsförfarande som utnyttjar hjälpinformation. För att skevheten ska minskas måste den hjälpinformation som används vara korrelerad med de variabler som ska mätas i undersökningen samt gärna vara korrelerad med svarsbenägenheten bland organisationerna. En skattning av korrelationen mellan de viktigaste undersökningsvariablerna och ett antal tänkbara hjälpvariabler har gjorts. Resultatet visar att lönesumma samvarierar bra med intäkter och kostnader. Vad gäller undersökningsvariablerna investeringar och tillgångar var samvariationen sämre vilket inte är förvånande när det gäller den typen av variabler.

För undersökningsvariablerna intäkter och kostnader bör ett skattningsförfarande som utnyttjar, den för hela undersökningspopulationen, kända variabeln lönesumma reducera bias. En antydning till hur detta skulle falla ut går att studera genom att använda ett sådant skattningsförfarande med hjälp av de för hela undersökningspopulationen kända variablerna antal sysselsatta (från RAMS) och lönesumma. Tabell 6 visar antal sysselsatta enligt RAMS summerade för redovisningsgrupperna och totalt (facit). Tabell 6 visar också skattningar (pSyss) av antal sysselsatta enligt RAMS producerade på traditionellt sätt (HT-estimator) samt skattningar (pSyss) producerade med hjälp av en estimator som utnyttjar hjälpinformation (GREG-estimator). Båda skattningarna är baserade på de svarande i HIO-urvalet.

Regressionsmodellen utnyttjar uppgift om årslönesumma samt uppgift om antal organisationer per branschgrupp (BGrupp) som hjälpinformation.

Tabell 6 Facit för antal sysselsatta samt motsvarande uppgifter skattade med GREG- respektive HT-estimator (pSyss = skattning av antal sysselsatta, sSyss = medlefel och relstd = relativt medlefel (variationskoefficient))

	Facit	GREG-estimator			HT-estimator		
BGrupp	Syss	pSyss	sSyss	relstd	pSyss	sSyss	relstd
65-67	1 140	3 262	274	8,4	605	229	37,9
70000	1 129	831	103	12,4	488	167	34,2
80000	16 189	15 874	581	3,7	14 753	1 117	7,6
85000	11 200	11 244	219	1,9	8 053	539	6,7
91000	26 049	25 461	2 410	9,5	24 781	3 339	13,5
92000	18 865	17 688	934	5,3	16 343	1 886	11,5
bland	6 434	6 071	381	6,3	5 414	745	13,8
<i>Total</i>	<i>81 006</i>	<i>80 431</i>	<i>2 702</i>	<i>3,4</i>	<i>70 437</i>	<i>4 108</i>	<i>5,8</i>

Skattningarna, per branschgrupp, av antal sysselsatta enligt RAMS (pSyss) ligger mycket nära det sanna värdet (Facit Syss) i de flesta branschgrupper när GREG-estimatorn används. Ett undantag är branschgrupp 65-67 där avvikelserna förklaras av ett fel i hjälpinformationen som senare rättades. I jämförelse med den traditionella HT-estimatorn fungerar GREG-estimatorn betydligt bättre, både vad gäller bias och variansreduktion. Variationskoefficienten (relstd) är lägre i alla branschgrupper för GREG-estimatorn jämfört med HT-estimatorn.

Liknande studier har gjorts för ett några andra variabler kända för hela undersökningspopulationen och de visar alla att GREG-estimatorn fungerar bättre än en HT-estimator. Från och med undersökningsår 2005 används en GREG-estimator för att skatta totala intäkter och kostnader per redovisningsgrupp. Men för variabler som investeringar och tillgångar används fortfarande en HT-estimator eftersom det är svårt (omöjligt) att hitta bra hjälpinformation.

6 Slutsats

Studier visar att den enkla blanketten fungerar bra eftersom inflödet från organisationer med den typen av blankett är högre än för organisationer med de andra typerna av blankett. Ett ökat påminnelsearbete lönar sig och särskilt lönar sig den dyrare formen med telefonpåminnelser. Ett ökat inflöde av blanketter är mycket viktigt i HIO eftersom studier visar att bortfallet introducerar bias (skevhet) i skattningarna. Studier visar också att det är möjligt att reducera bortfallsbiasen med hjälp av ett nytt skattningsförförande för några huvudvariabler, men inte för alla variabler.

Från och med undersökningsår 2005 strävar SCB efter att så många organisationer som möjligt ska få den enkla typen av blankett. Påminnelsearbetet har fortsättningsvis fått lov att dra mer resurser och varit utformat så som beskrivits ovan. Totala intäkter och kostnader, per redovisningsgrupp, skattas från och med undersökningsår 2005 med en GREG-estimator.