

TRE ÅTERINTERVJUSTUDIER

- INKOMSTFÖRDELNINGSUNDERSÖKNINGEN
Jan Eriksson, I/UI
- HÖGSKOLENKÄTUPPFÖLJNINGEN
Anders Karlsson, AM/U
- ÅRSARBETSKRAFTEN
Majvor Karlsson och Solveig Thudin, AM/AKU



R&D Report
Statistics Sweden
Research - Methods - Development
1991:8

INLEDNING

TILL

R & D report : research, methods, development / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1988-2004. – Nr. 1988:1-2004:2.

Häri ingår Abstracts : sammanfattningar av metodrapporter från SCB med egen numrering.

Föregångare:

Metodinformation : preliminär rapport från Statistiska centralbyrån. – Stockholm : Statistiska centralbyrån. – 1984-1986. – Nr 1984:1-1986:8.

U/ADB / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1986-1987. – Nr E24-E26

R & D report : research, methods, development, U/STM / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1987. – Nr 29-41.

Efterföljare:

Research and development : methodology reports from Statistics Sweden. – Stockholm : Statistiska centralbyrån. – 2006-. – Nr 2006:1-.

R & D Report 1991:8. Kvalitetsfonden 1990. Tre återintervjustudier / Jan Eriksson, m.fl. Digitaliserad av Statistiska centralbyrån (SCB) 2016.

urn:nbn:se:scb-1991-X1010P9108

TRE ÅTERINTERVJUSTUDIER

- INKOMSTFÖRDELNINGSUUNDERSÖKNINGEN
Jan Eriksson, I/UI
- HÖGSKOLENKÄTUPPFÖLJNINGEN
Anders Karlsson, AM/U
- ÅRSARBETSKRAFTEN
Majvor Karlsson och Solveig Thudin, AM/AKU



R&D Report
Statistics Sweden
Research - Methods - Development
1991:8

Från trycket
Producent
Ansvarig utgivare
Förfrågningar

April 1991
Statistiska centralbyrån, utvecklingsavdelningen
Åke Lönnqvist
Jan Eriksson, 019-17 67 53
Anders Karlsson, 019-17 63 07
Majvor Karlsson, 08-783 50 31
Solveig Thudin, 08-783 48 32

© 1991, Statistiska centralbyrån
ISSN 0283-8680
Printed in Sweden

TRE ÅTERINTERVJUSTUDIER

INKOMSTFÖRDELNINGSUUNDERSÖKNINGEN

Jan Eriksson, I/UI

HÖGSKOLENKÄTUPPFÖLJNINGEN

Anders Karlsson, AM/U

ÅRSARBETSKRAFTEN

Majvor Karlsson och Solveig Thudin, AM/AKU

FÖRORD

Bland de projekt, som tilldelades medel från 1990 års kvalitetsfond, var de tre återintervjustudier, som presenteras i denna rapport. Studierna är ambitiöst upplagda och väl presenterade och förtjänar att uppmärksammas. De belyser olika ämnesområden och delvis olika populationer. Presenterade tillsammans ger de en mycket bra uppfattning om när intervjufrågor kan vara robusta mot svarsvariation och när de inte är det. Förhoppningsvis skall dessa tillsammans med tidigare återintervjustudier bidra till att fästa uppmärksamhet vid vikten av frågekonstruktion och formulärtest och av att hänsyn tas till svarskvaliteten vid analys.

Planeringen av studierna var icke samordnad och något olika metoder användes vid analysen. En avkortning av rapporterna var nödvändig för att de skulle kunna publiceras i serien R&D Report. Redigeringen har gjorts efter eget omdöme av varje rapportförfattare, som också svarar för synpunkter och slutsatser. Endast referenslistan är gemensam. En viss typografisk överensstämmelse har eftersträfvats. Den som vill ta del av de fullständiga rapporterna får vända sig till respektive författare.

Återintervjustudier genomförda på detta sätt är ett viktigt medel för evaluering av svarskvalitet och undersökningsresultat och borde regelbundet göras i alla produkter - även i sådana med institutionella uppgiftslämnare. Det finns också ett annat sätt att använda återintervjuer, nämligen som löpande kvalitetskontrollinslag i produktionen. Hittills har SCB icke använt sig av detta men ett program håller nu på att utarbetas i samråd mellan olika intressenter vid SCB.

Håkan L Lindström, U/STM

Referenser

- Aish, m fl. "Handout for a short course on using LISREL for analyzing measurement errors in surveys". Statistics Sweden professional training program 1990.
- Andersson, L-G "Kodningsreliabiliteten för den socio-ekonomiska (SEI) indelningen i grundskoleuppföljningen". RIS-projektet. Intern SCB-rapport 1979.
- Bergman, L R, Thorslund, M "Svarskvalitén i AKU. Resultat från två studier med återintervjuer". Metodproblem i individ- och hushållsstatistik nr 13 SCB 1980.
- Eriksson, J "Minnesanteckningar från mötet med intervjuarna 890209". Stencil 890214.
- Fleiss, J L "Statistical methods for rates and proportions". Wiley, New York 1981.
- Forsman, G "Analys av återintervjudata". Urval nr 19. Statistiska Centralbyrån 1987.
- Groves, R H "Survey errors and survey costs". Wiley 1989. Sid 324-329
- Holgersson, M "Studier av minnesfel i grundskoleuppföljningen". RISprojektet. Pm 1979.
- "Studier av reliabilitet och validitet i enkätfrågor med hjälp av återintervjuer i samband med I/UKs grundskoleuppföljning 1977". RIS-projektet. Intern SCB-rapport 1979.
- Jones, Charles m fl "Accuracy of 1970 census population and housing characteristics as measured by reinterviews". American statistical association. 1973
- Lindgren, W, B "Statistical theory", third edition. Macmillan , New York 1978.
- Mcneemar, Q "Psychological statistics". Second edition. Wiley, New York. Sid 57-58.
- SCB Kvalitén i AKUs s k Februariundersökningar, pm 1982-02-08. Rapport från DATI-projektets produktionsprov i arbetskraftsundersökningarna augusti 1989 -januari 1990, R&D Report 1990:11. Bakgrundsfakta till AKU 1990:1. Årsarbetskraften 1989, råtabeller, 1990.
- Stat. meddelanden Inkomstfördelningsundersökningen 1988. Be 21 SM 9001, SCB. Årsarbetskraften 1989. (1990a) AM 64 SM 9001, SCB.
- Wikman, A "Svarsprecisionen i surveyundersökningar om levnadsförhållanden". Metodproblem i individ- och hushållsstatistik nr 14 SCB 1980.
- Wärneryd, B "Återintervjustudier i undersökningarna av levnadsförhållanden (ULF) 81, 83 och 84". Levnadsförhållanden. Appendix 8. Sveriges officiella statistik, SCB Stockholm 1985.
- "Återintervjustudie i undersökningen av levnadsförhållanden (ULF) 1989. Sammanställning av svarsvariation". Intern rapport 901002.
- "Att fråga. Om frågekonstruktion vid intervjuundersökningar och postenkäter", SCB 1980.

ÅTERINTERVJUSTUDIE I INKOMST- FÖRDELNINGSUUNDERSÖKNINGEN (HINK) 1989.¹

Innehåll

0	Sammanfattning	4
1	Inledning	5
2	Något om teorin för återintervjustudier	5
2.1	Förutsättningar	5
2.2	Analysmått	6
2.2.1	Icke-kontinuerliga variabler	6
2.2.2	Kontinuerliga variabler	9
3	Uppläggning och genomförande	10
4	Resultatrevisning	11
4.1	Kommentarer till tabell 6 och tabell 7	11
4.2	Sammanfattning av tabell 6 och tabell 7	11
4.3	"Specialanalys" av frågorna 6, 9 och 11 (frågor om arbetstider)	13
A1	Appendix 1 (tabell 6 och 7)	14

¹ Denna rapport är en reducerad version av en tidigare rapport med samma titel daterad 901222.

0 Sammanfattning

Föreliggande studie har gjorts i syfte av att få en uppfattning om hur pass väl HINK-enkäten fungerar. Studien har gjorts med hjälp av ett s k test-retestförfarande. Samma individer har besvarat samma frågor vid två tillfällen. Om frågorna fungerar på ett tillfredsställande sätt, bör svaren vara konsistenta mellan de båda mättillfällena. Analysen är till största delen av beskrivande karaktär, d v s huvudsyftet är att "pejla in" de frågor som kan tänkas ge allvarliga mätproblem. Någon djupare diskussion om, eller analys av orsak till att dessa frågor innehåller allvarliga mätfel görs alltså inte.

I rapporten ges en kortfattad beskrivning av tankarna bakom de olika mått som används i analysen. Måtten som används är från början utvecklade inom psykologin och den experimentella statistiken. I surveysammanhang är dessa mått däremot relativt färska. Måtten mäter variationen mellan svaren. Eventuella systematiska mätfel mäts alltså inte i dessa mått.

Ett av de mått som använts i analysen är det s k inkonsistensindex. Som tumregel brukar anges att ett skattat index mellan 0 och 0.20 är att betrakta som låg inkonsistens, mellan 0.21 och 0.50 som "normal" konsistens samt värden större än 0.50 som hög inkonsistens. Då svarsalternativen på frågorna ej är klassificerade, används korrelationen mellan de observerade värdena som ett mått på överensstämmelse.

Resultaten tyder på att de enkätfrågor som berör arbetstider, av någon anledning fungerar "mindre bra". Dessa frågor uppvisar genomgående något högre värden på inkonsistensindex (0.40-0.49). Detta är dock inte att betrakta som allvarlig inkonsistens enligt tumregeln ovan. Ha dock i åtanke att tumregler inte är att betrakta som något facit!! Av frågorna om arbetstid, är det framför allt frågan som berör antal arbetade övertidstimmar (index = 0.63) samt frågan om huruvida arbetstiden förändrats (index = 0.53) som ger hög inkonsistens. Andra frågor som ger höga värden på inkonsistensindex är antal uppskattade körda mil (0.49) samt hyreskostnad för kvarboende hos föräldrarna (0.42). Övriga frågor uppvisar låga inkonsistensindex. Det verkar alltså som om svarsreliabiliteten i större delen av HINK-enkäten trots allt är god.

Orsaken till att just frågorna om arbetstider uppvisar lägre svarsreliabilitet kan man naturligtvis bara spekulera om. En orsak kan vara att man i HINK överrepresenterar företagare och jordbrukare. Dessa har förmodligen mer "instabila" förhållanden beträffande arbetstider jämfört med anställda. I avsnitt 4.3 beräknas inkonsistensindex för frågorna om arbetstider för fyra olika kategorier av sysselsatta, nämligen anställda, företagare, jordbrukare och studerande. Företagare och jordbrukare uppvisar då något högre inkonsistens än anställda. Man bör dock vara något försiktig vid denna jämförelse, eftersom inga signifikantest avseende skillnader mellan grupperna utförs.

Studien är mycket värdefull i den meningen att man för HINK-undersökningen nu har fått en bättre uppfattning om ett av mätinstrumenten i undersökningen. Vissa brister har kunnat belysas. Nästa steg borde bli att försöka genomföra förbättringar i enkäten.

Kostnaderna för "fältarbetet" i projektet fördelar sig enligt följande:

- Intervjuer	70 000 kr
- Databearbetning	37 000 kr
Summa	107 000 kr

Projektet har finansierats med 100 000 kr ur kvalitetsmiljonen. Dessa medel var avsedda att täcka kostnaderna för "fältarbetet". Övriga kostnader (datorkörningar och analys) har sektionen stått för.

1 Inledning

Inkomstfördelningsundersökningen (HINK) har som syfte att belysa inkomstförhållanden för hela Sveriges befolkning (dock med vissa avgränsningar, se nedan) och delgrupper därav. Redovisning avser såväl hushåll (familjeenheter) som individer. Populationen omfattar den del av Sveriges befolkning som ingår i kyrkobokföringen samt varit bosatta i Sverige minst sex månader. Personer som bor på institutioner ingår inte i populationen. Som institution räknas sjukhem inom långvården, fängelser och ålderdomshem.

Som urvalsram används SCB:s register över totalbefolkningen (RTB) med aktualitet vecka 30. Ramen är uppdelad i fyra strata: jordbrukare, företagare, ålderspensionärer och övriga. Ur varje strata dras ett obundet slumpmässigt urval utan återläggning (OSU).

Uppgiftsinsamling sker dels direkt från familjeenheterna genom en telefonintervju, dels genom användande av olika myndigheters register. Från intervjun erhåller man klassificeringsvariabler som t ex familjetyp och socioekonomisk grupp samt information om boendekostnad, antal arbetade timmar under undersökningsåret etc. Från registren inhämtas inkomstuppgifter som t ex deklarationsvariabler.

Föreliggande studie har gjorts i syfte av att få en uppfattning om hur pass väl HINK-enkäten fungerar. För närvarande finns en ganska god uppfattning om urvalsfelen i skattningarna. Mätproblemen i intervjuvariablerna behöver däremot analyseras närmare. Inför den kommande utvärderingen av skattereformens effekter på hushållens ekonomi, är det särskilt viktigt att man har en god uppfattning om kvaliteten i undersökningen. Intervjuvariabler som är speciellt viktiga att analysera är bl a frågor om arbetstid, boendeform, boendekostnad samt klassificeringsvariabler som t ex familjetyp och socioekonomisk grupp.

Någon återintervjustudie för HINK-undersökningen har aldrig gjorts tidigare. I anslutning till att intervjuerna för HINK-88 var avslutade, hölls ett möte med sex av de intervjuare som deltagit i fältarbetet. Syftet med detta var att få intervjuarnas synpunkter på hur frågeformuläret fungerar i praktiken. Man diskuterade formulärets layout, informationen till urvalshushållen, kontakten mellan intervjuare och intervjudperson samt frågorna i formuläret. Intervjuarnas synpunkter på ovanstående var i stort sett positiva. Intervjuarna upplevde framför allt frågan om uppskattat värde av bostadsrätt som besvärlig (besvärlig i den mening att intervjudpersonerna hade en benägenhet att bli irriterade över frågan). Detta möte finns dokumenterat i form av minnesanteckningar (se referenser).

Studien har gjorts med hjälp av ett s k test-retestförfarande. Samma individer har besvarat samma frågor vid två tillfällen. Om frågorna fungerar på ett tillfredsställande sätt, bör svaren vara konsistenta mellan de båda mättillfällena. Analysen är till största delen av beskrivande karaktär, d v s huvudsyftet är att "pejla in" de frågor som kan tänkas ha allvarliga mätproblem. Någon djupare diskussion om, eller analys av orsak till att dessa frågor innehåller allvarliga mätfel görs alltså inte.

2 Något om teorin för återintervjustudier

2.1 Förutsättningar

Vid analys av återintervjudata finns för varje individ två observerade värden för varje variabel. Man vill bl a studera skillnaderna mellan dessa observerade värden, de s k svarsavvikelserna. Om ett urval av individer besvarar samma frågor vid två olika tillfällen men under lika omständigheter, bör deras svar vara konsistenta.

Mätfelen i de två observerade värdena på variablerna antas vara oberoende av varandra. Detta antagande innebär bl a att återintervjun ej bör göras för tidigt i förhållande till den ordinarie intervjun. Om så är fallet finns det stor risk att intervjudpersonen kommer ihåg

hur han/hon svarade vid det ordinarie intervjutillfället. Man erhåller då en bild av överdriven tillförlitlighet i mätproceduren. Å andra sidan får det inte gå för lång tid mellan ordinarie- och återintervju, ty man riskerar då att intervjupersonen har glömt bort de förhållanden man frågar om. Huruvida dessa antaganden om oberoende mellan mätfelen är uppfyllda i praktiken, finns f n inga belägg för. Trots detta används (i brist på bättre) dessa mått i denna analys.

Måtten som presenteras i avsnitt 2.2 är från början utvecklade inom psykologin och den experimentella statistiken. I surveysammanhang är dessa mått relativt färska. Litteraturen på detta område är därför något oklar om vilka egenskaper dessa mått har vid olika urvalssdesigner. Förutsättningarna som brukar anges (se Forsman 1987) är att stickprovet är ett obundet slumpmässigt urval (OSU) med återläggning. Sådana enkla designers används ganska sällan vid praktisk tillämpning av stickprovsteori.

I Forsman (1987) tillämpas inkonsistensindex på arbetskraftsundersökningen (AKU), som är ett stratifierat urval utan återläggning. Analysen görs på ovägt material, d v s hänsyn tas ej till att individernas urvalssannolikheter varierar. Forsman motiverar detta med att det är själva mätproceduren som skall utvärderas. Denna kan antas påverka individerna på likartat sätt oberoende av urvalssannolikheter. Detta antagande kan ej anses vara uppfyllt i HINK. Urvalsdesignen i HINK är stratifierat OSU, där populationen är indelad i fyra strata, nämligen jordbrukare, företagare, ålderspensionärer samt övriga. I avsnitt 4.3 "specialstuderar" frågorna om arbetstider. När inkonsistensmått beräknas för olika kategorier av sysselsatta (jordbrukare, företagare, anställda samt studerande), erhålls väldigt olika värden på måttet för respektive kategori (se avsnitt 4.3). Detta tyder på att mätproceduren mycket väl kan vara beroende av urvalsdesignen. Trots detta görs även analysen av HINKens återintervjudata på ovägt material. Resultaten skall därför icke tolkas som "bra" och väntevärdesriktiga skattningar av populationsparametrar, utan som indikatorer på var fördjupad analys krävs.

2.2 Analyismått

2.2.1 Icke-kontinuerliga variabler

Låt p_{ij} vara proportionen element som klassificerats i kategori (i) enligt den ordinarie intervjun och kategori (j) enligt återintervjun. Data kan då representeras i en kvadratisk matris enligt tabell 1, där antalet kategorier = k.

T 1

		j					
		Klassifikation enligt återintervju					
		1	2	...	...	k	Total
i	1	p_{11}	p_{12}	...	...	p_{1k}	$p_{1.}$
Klassifikation enligt ordinarie intervju	2	p_{21}	p_{22}	...	...	p_{2k}	$p_{2.}$
	...	...	...	...	...	...	...
	...	...	...	...	...	...	...
	k	p_{k1}	p_{k2}	...	...	p_{kk}	$p_{k.}$
Total		$p_{.1}$	$p_{.2}$	...	...	$p_{.k}$	1

Summan av andelen observationer som klassificerats identiskt vid båda intervjutillfällena blir

$$p_o = \sum_{i=1}^k p_{ii} \quad (2.1),$$

d v s den observerade överensstämmelsen. Detta mått tar dock inte hänsyn till att en viss överensstämmelse erhålles p g a slumpen, även om mätresultaten är slumpmässigt utspridda över kategorierna.

Förväntad överensstämmelse blir (givet marginalfördelningarna samt vid antagande om oberoende mellan mätningarna)

$$p_e = \sum_{i=1}^k p_{i.} p_{.i} \quad (2.2).$$

p_e tolkas här som summan av andelen observationer som man, rent slumpmässigt, förväntar sig ska klassificeras identiskt vid båda intervjutillfällena.

En skattning av konsistens, som tar hänsyn till den slumpmässiga överensstämmelsen, är kappa, som definieras enligt (se Fleiss 1981)

$$\hat{\kappa} = \frac{p_o - p_e}{1 - p_e} = 1 - \frac{1 - p_o}{1 - p_e} = 1 - \frac{\text{observerad missöverensstämmelse}}{\text{förväntad missöverensstämmelse}} \quad (2.3).$$

Där p_o och p_e ges av (2.1) och (2.2) ovan.

Om kappa antar värdet noll är observerad överensstämmelse identisk med förväntad överensstämmelse. Vi tolkar detta som att den observerade överensstämmelsen är "bara" vad man slumpmässigt kan förvänta sig. Om kappa antar värden nära noll, betraktas alltså detta som låg konsistens.

I surveysammanhang brukar man, istället för konsistens, mäta inkonsistens (missöverensstämmelse). Ett mått på missöverensstämmelse blir alltså

$$\hat{I} = 1 - \hat{\kappa} \quad (2.4).$$

$\hat{I}$ brukar benämnas inkonsistensindex. Som tumregel brukar anges att ett skattat index mellan 0 och 0.20 är att betrakta som låg inkonsistens, mellan 0.21 och 0.50 som "normal" inkonsistens samt värden större än 0.50 som hög inkonsistens.

Medelfelet för $\hat{I}$ skattas enligt (se Fleiss 1981)

$$Se(\hat{I}) = \frac{\sqrt{A + B - C}}{(1 - p_e) \sqrt{n}} \quad (2.5).$$

$$\text{Där } A = \sum_{i=1}^k p_{ii} [1 - (p_{i.} + p_{.i})(1 - \hat{\kappa})]^2 \quad (2.6),$$

$$B = (1 - \hat{k})^2 \sum_{i \neq j} \sum p_{ij} (p_{.i} + p_{.j})^2 \quad (2.7),$$

$$C = [\hat{k} - p_e(1 - \hat{k})]^2 \quad (2.8).$$

Det enklaste fallet av ovanstående är när frågorna endast består av två svarsalternativ. Tabell 1 reduceras då till en korstabell enligt tabell 2 nedan.

T 2

		Återintervju		
		0	1	Total
Ordinarie intervju	0	a	b	a+b
	1	c	d	c+d
Total		a+c	b+d	n

Där a,b,c och d är antalet individer i de olika kategorierna och n är antalet individer som svarat på frågan. Det enklaste måttet på missöverensstämmelse är här det s k "gross-difference-rate", som definieras enligt

$$g = \frac{b + c}{n} \quad (2.9).$$

Inkonsistensindex skattas på sedvanligt sätt (enligt (2.4)). Medelfelet för indexet skattas enligt (2.5).

Tabell 1 ovan kan delas in i "deltabeller" enligt tabell 2. Man kan då beräkna ett inkonsistensindex för varje svarskategori för sig. Inkonsistensindex enligt (2.4) är ett vägt medelvärde av dessa "individuella" index.

För fyrfältstabeller som tabell 2 finns även möjlighet att statistiskt testa om marginalfördelningarna är lika vid de båda mättillfällena, d v s om $(a+b) = (a+c)$ för hela populationen (detta är ju faktiskt vad vi antar när vi beräknar inkonsistensindexet). Först noterar vi att detta är detsamma som att testa om $b = c$ eller, efter omskrivning, att $b - c = 0$. Innebörden av detta är att under noll-hypotesen ($b = c$), så förväntar vi oss att mängden $(b+c)$ fördelar sig 50-50 på kategorierna "0,1" respektive "1,0" (se tabell 2 ovan). Om mängden $(b+c)$ är "stor", säg större än 10, kan vi approximera binomialfördelningen med normalfördelningen. Testvariabeln blir då (se McNemar)

$$T = \frac{b - c}{\sqrt{b+c}} \quad (2.10),$$

Om $T > |1.96|$ säger vi att T är statistiskt signifikant på 5% - nivån, d v s testet förkastar nollhypotesen om att marginalfördelningarna är lika vid de båda mättillfällena. McNemar ger vissa rekommendationer vid tillämpning av testet. Då $(b+c)$ är litet, säg mellan 11 och 20, förbättras normalapproximationen genom kontinuitetskorrektion med 0.5. Då $(b+c)$ är mindre än 11 bedöms testet som alltför osäkert att tillämpa. Dessa rekommendationer följs vid tillämpning av testet på HINK-data.

2.2.2 Kontinuerliga variabler

Då svarsalternativen inte är klassificerade i förväg, som för t ex frågor om arbetstid i HINK-formuläret, är måttet (2.4) ej tillämpligt. Ett sätt att göra inkonsistensindex tillämpligt i detta fall, är att i efterhand klassindela svaren. Detta är inte alltid så lämpligt, eftersom måttet är känsligt för hur man bildar dessa klasser. Om man t ex väljer att klassindela svaren i stora och få klasser, erhåller man ett lägre värde på indexet jämfört med om man skulle bilda små och fler klasser.

Ett annat mått på överensstämmelse är korrelationen mellan värdena från de båda mättilfällena. Tanken bakom detta mått är, i stora drag, följande (resonemanget nedan är hämtat från Groves 1989).

Vi mäter en variabel, X_i , för en individ vid två tillfällen. Vi erhåller då mätresultaten y_{i1} och y_{i2} . Dessa kan uttryckas som

$$y_{i1} = X_i + \varepsilon_{i1} \quad (2.11),$$

$$y_{i2} = X_i + \varepsilon_{i2} \quad (2.12),$$

där ε_{i1} och ε_{i2} är mätfelen i y_{i1} respektive y_{i2} . De observerade värdena på variabeln kan alltså uttryckas som det sanna värdet plus ett mätfel. Mätfelen antas ha följande egenskaper:

$$E(\varepsilon_{i1}) = E(\varepsilon_{i2}) = 0 \quad (2.13),$$

$$E(\varepsilon_{i1}\varepsilon_{i2}) = 0 \quad (2.14),$$

$$E(\varepsilon_{i1}X_i) = E(\varepsilon_{i2}X_i) = 0 \quad (2.15),$$

$$\sigma_{\varepsilon_1}^2 = \sigma_{\varepsilon_2}^2 = \sigma_{\varepsilon}^2 \quad (2.16),$$

d v s mätfelen antas ha väntevärdet noll, lika varians, vara okorrelerade med varandra samt vara okorrelerade med X_i . Dessa antaganden implicerar bl a följande

$$E(y_{i1}) = E(y_{i2}) = X_i \quad (2.17),$$

$$\sigma_{y_1}^2 = \sigma_{y_2}^2 = \sigma_y^2 \quad (2.18),$$

$$\sigma_y^2 = \sigma_X^2 + \sigma_{\varepsilon}^2 \quad (2.19),$$

$$\text{cov}(y_1, y_2) = \sigma_X^2 \quad (2.20).$$

Vi kan alltså (om ovanstående antaganden kan anses vara uppfyllda) använda covariansen mellan y_1 och y_2 för att mäta variansen för X . Genom att jämföra variansen för X med variansen för y_1 respektive y_2 kan man således få en uppfattning om mätfelens inverkan i svaren. Ett mått på överensstämmelse skulle i detta fall följaktligen vara

$$\rho_{y_1 y_2} = \frac{\text{cov}(y_1, y_2)}{\sqrt{\sigma_{y_1}^2 \sigma_{y_2}^2}} = \frac{\sigma_X^2}{\sigma_y^2} \quad (2.21),$$

vilket är detsamma som korrelationskoefficienten mellan y_1 och y_2 . Om denna antar värdet 1 råder ett fullständigt linjärt samband mellan y_1 och y_2 , eller annorlunda uttryckt, variationen i X (den sanna variabeln) är identisk med variationen i de observerade värdena.¹ Vi tolkar alltså detta som att X mäts utan svarsvariation. Ett inkonsistensindex blir nu

$$1 - \rho_{y_1 y_2} \quad (2.22).$$

Om indexet antar värdet noll, är korrelationskoefficienten lika med 1 och vi tolkar detta som att X mäts utan svarsvariation.

En skattning av (2.21) är (se Aish m fl)

$$\hat{\rho}_{y_1 y_2} = \frac{\widehat{\text{cov}}(y_1, y_2)}{\frac{1}{2}(\hat{\sigma}_{y_1}^2 + \hat{\sigma}_{y_2}^2)} \quad (2.23),$$

d v s man använder covariansen mellan y_1 och y_2 från stickprovet för att skatta variansen för X , samt medelvärdet av stickprovsvarianserna för y_1 och y_2 för att skatta variansen för y . En skattning av (2.22) blir således

$$1 - \hat{\rho}_{y_1 y_2} \quad (2.24).$$

3 Uppläggning och genomförande

Återintervjuerna till HINK-89 genomfördes under våren 1990. Tiden mellan ordinarie intervju och återintervju var c:a 6 veckor. Instruktionerna till intervjuerna var samma som vid den ordinarie intervjun. Vid återintervjun ställdes samma frågor till hushållet som vid den ordinarie intervjun. Intervjuerna var ej samma personer vid de båda intervjutillfällena. Några veckor innan återintervjuerna startade utsändes ett informationsbrev till urvalspersonerna och ett annat informationsbrev till intervjuerna.

Analysen har endast gjorts på urvalspersonens enkätvariabler. I intervjun svarar intervjupersonen även på samma frågor för övriga hushållsmedlemmar.

På grund av att HINK är en hushållsundersökning har man på förhand ingen kontroll över vem i hushållet som kommer att delta i intervjun. Intervjuerna uppmanas att, om möjligt, ta kontakt med urvalspersonen i hushållet. Detta är dock inte alltid möjligt (urvalspersonen är t ex ej anträffbar under intervjuperioden etc). Intervjuaren ställer då frågorna till någon annan i hushållet. Intervjupersonen vid de båda intervjutillfällena är alltså inte alltid samma person. Vid återintervjustudier är detta en störande faktor, ty om intervjupersonen ej är densamma vid de båda intervjutillfällena så mäter vi inte längre reliabiliteten mellan svar från samma person, utan reliabiliteten mellan två personer. Dessutom är förutsättningen om likvärdiga mätningar vid de båda mättillfällena ej längre uppfylld. Uppgift om vem/vilka i hushållet som deltagit i intervjun vid de två intervjutillfällena, finns i skrivande stund ej tillgänglig (uppgiften är ännu ej "stansad").

¹ Ett fullständigt linjärt samband råder även då rho antar värdet -1. Sambandet är i det fallet negativt, vilket innebär att ett högt värde för y_1 motsvaras av ett lågt värde för y_2 och vice versa. Ett negativt värde för rho är alltså ett dåligt tecken i återintervjusammanhang.

Urvalet till återintervjustudien drogs som ett suburval från de hushåll som svarat "utan problem" i det ordinarie HINK-urvalet. HINK-urvalet är ett stratifierat osu med fyra strata. Urvalet består av två paneler, där den "äldre" panelen (drygt hälften av det totala urvalet) är återkommande från HINK ett år tidigare och den "nya" panelen drages samma år som undersökningsåret. De två panelerna kan betraktas som två oberoende urval. Urvalsdesignen var samma som i det ordinarie urvalet. Urvalet till återintervjun är alltså ett tvåfasurval med stratifierat OSU i båda faserna, där de svarande hushållen från ordinarie intervjun utgör den första fasen. Urvalsfractionen i respektive strata är densamma i båda faserna.

Efter genomförd återintervju bestod urvalet av 546 svarande hushåll. Bortfallet bestod till största delen av ej anträffade hushåll (5.5%).

4 Resultatredovisning

4.1 Kommentarer till tabell 6 och tabell 7

Frågorna i tabellerna är ej exakt ordagrant återgivna från frågeformuläret. Analysen inkluderar ej samtliga i enkäten ingående frågor. I tabell 6 presenteras inkonsistensindex för frågor med kategoriserade svarsalternativ. För frågor med enbart två svarsalternativ (fyrfältstabeller) presenteras också "g-måttet" (se (2.9)) samt "T" (se (2.10)). Ingen av de testade frågorna har uppvisat någon signifikans vid test om lika marginalfördelningar. I tabell 7 presenteras inkonsistensindex för frågor med kontinuerliga svarsalternativ.

Observera att antalet svarande (n) skiljer sig mellan frågorna. Detta har flera orsaker. Dels skiljer sig det partiella bortfallet mellan frågorna. Dessutom skall vissa frågor endast ställas till en viss kategori av hushåll (t ex sysselsatta). Kriteriet vid de selektioner som gjorts, har varit att man skall ha svarat på frågan vid båda intervjutillfällena. För frågor som endast skall ställas till en viss kategori av hushåll, har kriteriet varit att hushållet har egenskapen vid båda mättillfällena.

För frågorna om arbetstider (frågorna 6 t o m 11), har indexet beräknats på två olika antal svarande. Först beräknades indexen för samtliga svarande hushåll vid båda intervjutillfällena (exkluderat ej sysselsatta). För studerandehushåll är antal förvärsarbetade veckor på ett år få, i många fall kanske t o m noll till antalet. Det är förmodligen inte helt orimligt att anta att dessa hushåll svarar exakt lika vid båda intervjutillfällena. Man erhåller i sådana fall en bild av överdriven tillförlitlighet i svaren. När även studerandehushåll exkluderas ur materialet ökar indexskattningen drastiskt för fråga 6, som avser antal förvärsarbetade veckor. För övriga frågor om arbetstider är förändringarna däremot minimala. I avsnitt 4.3 presenteras inkonsistensindex för frågorna 6, 9 och 11 för olika kategorier av sysselsatta.

Nedan följer en sammanfattning av resultaten i tabellerna 6 och 7 (som återfinnes i appendix 1).

4.2 Sammanfattning av tabell 6 och tabell 7

Tabell 3 och 4 visar frekvenser för olika värden på det skattade inkonsistensindexet. Tabell 3 visar frekvenserna av skattade index för frågor där svaren är kategoriserade, medan tabell 4 visar frekvenser av skattade index för frågor där svarsalternativen är kontinuerliga. Tabell 4 innehåller skattade index för frågorna 6, 9 och 11, där studenter är exkluderade. Frågenummer anges endast då indexet överstiger 0.19.

T3 Inkonsistensindex för frågor med kategoriserade svarsalternativ.

Inkonsistens- index	Frekvens	Frågenr
0.00-0.09	3	
0.10-0.19	12	
0.20-0.29	1	33
0.30-0.39	1	32
0.40-0.49	0	
0.50-0.59	2	7, 20
0.60-0.69	0	
0.70-	0	
	19	

De flesta frågorna uppvisar låg inkonsistens. Frågorna 7 och 20 har ganska höga värden, vilket tyder på att dessa frågor, av någon anledning, fungerar mindre bra. Frågorna 32 och 33 uppvisar "mellanhöga" värden. Fråga 7 berör förändring av arbetstid, medan fråga 20 handlar om huruvida intervjupersonen erhållit ej skattepliktiga stipendier eller ej. Det bör här tilläggas att fråga 20 är extremt snedfördelad. Av de 526 som svarat på frågan (se tabell 5) har 520 st svarat "nej" vid båda intervjutillfällena. Av de resterande 6 svarande har 2 st svarat "ja" vid båda intervjuerna och 4 st svarat "ja" vid första intervjun men "nej" vid den andra intervjun. En sådan snedfördelning av svaren gör att kappa tenderar att bli litet (inkonsistensindex blir stort). Denna snedfördelning tycks också påverka medelfelet för indexet, som i detta fall blir väldigt stort trots att antalet svarande är många. I frågorna 32 och 33 vill man veta om hushållsel respektive värme ingår i hyran.

T4 Inkonsistensindex för frågor med kontinuerliga svarsalternativ (xx(f) anger att frågan är följdfråga till fråga xx).

Inkonsistens- index	Frekvens	Frågenr
0.00-0.09	1	
0.10-0.19	5	
0.20-0.29	2	
0.30-0.39	1	36(f)
0.40-0.49	4	6, 9 , 25(f), 37(f)
0.50-0.59	0	
0.60-0.69	1	11
0.70-	0	
	14	

Frågorna 6, 9 25(f), 37(f) samt 36(f) uppvisar "mellanhöga" värden på indexet, medan fråga 11 har ett högt värde. Frågorna 6, 9 och 11 behandlar arbetstider. Fråga 25 ställes till intervjupersoner som under 1989 var kvarboende hos föräldrarna. Man vill där veta om intervjupersonen betalade någon form av hyra till föräldrarna. I fråga 25(f) vill man veta storleken på denna eventuella hyra. I fråga 36(f) skall intervjupersonen uppskatta marknadsvärdet av eventuell bostadsrätt. I fråga 37(f) skall han/hon uppskatta antal mil körda privat under 1989, där hushållet haft utgift för drivmedel.

4.3 "Specialanalys" av frågorna 6, 9 och 11 (frågor om arbetstider).

Frågorna om arbetstider är speciellt viktiga vid analys av skattereformens effekter. Det kan därför finnas skäl att granska svaren på dessa frågor något mer ingående. Tabell 5 visar inkonsistensindex för olika kategorier av sysselsatta för frågorna 6, 9 och 11. Fråga 6 berör antalet arbetade veckor, fråga 9 normalt antal arbetade tim/vecka under 1989 och fråga 11 antal timmar övertid under 1989.

T5 Inkonsistensindex för olika kategorier av sysselsatta

Frågenr	Sysselsättningskategori	Inkonsistensindex	n
6	Anställda	0.40	342
	Företagare	0.59	50
	Jordbrukare	0.50	29
	Studerande	0.04	16
9	Anställda	0.37	342
	Företagare	0.62	50
	Jordbrukare	0.72	30
	Studerande	0.68	17
11	Anställda	0.61	343
	Företagare	1.01	50
	Jordbrukare	1.00	30
	Studerande	0.20	17

Vi kan notera att gruppen anställda genomgående uppvisar lägre värden på inkonsistensindex än företagare och jordbrukare. För frågorna 6 och 11 är inkonsistensindexen betydligt lägre för studenter än övriga. Notera dock att antal i gruppen (n) är få och att vi inte utför några signifikanstest avseende skillnader mellan grupperna.

Resultaten för företagare och jordbrukare i fråga 11 bör kommenteras. Bland de 50 företagarna har 46 st uppgett noll timmar övertid vid båda mättillfällena. ständigt. De fyra som har uppgett olika antal övertidstimmar vid de båda intervjuerna, uppvisar alltså väldigt stor svarsvariation. För jordbrukare gäller att 29 av de 30 har uppgett noll antal övertidstimmar vid båda mättillfällena. Det är alltså endast en intervjuperson som gett olika svar vid de båda intervjufallen. Indexen för företagare och jordbrukare bör alltså tagas med en nypa salt i detta fall.

Resultaten i tabell 5 tyder på att urvalsdesignen kan ha stor betydelse för skattning av inkonsistensindex i populationen.

A1 Appendix 1 (tabell 6 och 7)

T6 Inkonsistensindex för frågor med kategoriserade svarsalternativ

Fråge- nr		Inkonsistens- index	1.96*medel- felet	T ¹	g	n	k ²
1	Antal hushållsmedlemmar	0.03	0.02	-	-	546	4
2	Gifta/sambo	0.12	0.04	1.89	0.05	546	2
5	Huvudsysselsättning	0.11	0.04	-	-	546	4
7	Förändrades förvärsarbets- tiden under 1989?	0.53	0.14	0.00	0.10	450	2
7	Samma som ovan men studerande är exkluderade	0.57	0.14	0.00	0.10	437	2
10	Avsåg veckoarbetstiden heltid?	0.16	0.07	1.38	0.05	382	2
10	Samma som ovan men studerande är exkluderade	0.17	0.07	1.38	0.05	375	2
13	Socioekonomisk grupp	0.16	0.04	-	-	450	6
14	Yrkesställning	0.12	0.03	-	-	544	5
15	Var ... fackligt ansluten?	0.12	0.04	1.13	0.06	508	2
16a	Vilken facklig organisation tillhörde ...?	0.09	0.04	-	-	292	4
19	Har ... betalt underhålls- bidrag för barn under 1989?	0.19	0.15	.	0.01	523	2
20	Har ... mottagit stipendier som inte är skattepliktiga under 1989?	0.50	0.42	.	0.01	526	2
24	Var ... kvarboende hos någon av/båda föräldrarna under mer än halva 1989?	0.16	0.13	.	0.08	65	2
25	Betalade ... något för att bo kvar?	0.00	0.00	.	0.00	35	2
27	Fastighetstyp	0.11	0.04	-	-	545	3

1 Om (b+c) < 11 markeras detta med punkt (.).
2 k = antal svarsalternativ.

T6 (forts från föregående sida)

Fråge- nr		Inkonsistens- index	1.96*medel- felet	T	g	n	k
28	Boendeform	0.09	0.03	-	-	479	5
32	Ingår hushållsel i hyran?	0.31	0.16	0.00	0.07	193	2
33	Ingår värme i hyran?	0.20	0.16	.	0.03	193	2
36	Ägde hushållet bostadsrätt den 31/12 1989?	0.16	0.07	1.18	0.03	535	2
37	Hade hushållet tillgång till bil under 1989?	0.14	0.06	0.46	0.03	544	2

T7 Inkonsistensindex för kontinuerliga variabler

Fråge- nr		Inkonsistens- index	n
6	Hur många veckor förvärs- arbetade ... sammanlagt under 1989?	0.11	456
6	Samma som ovan men studerande är exkluderade	0.41	435
9	Hur många tim/vecka för- värsarbetade ... under 1989?	0.43	458
9	Samma som ovan men studerande är exkluderade	0.42	433
11	Hur många timmar arbetade ... övertid under 1989?	0.63	459
11	Samma som ovan men studerande är exkluderade	0.63	437
Om ... har betalt underhållsbidrag:			
19(f)	Hur stort var beloppet per månad?	0.27	13

T7 (forts från föregående sida)

Frågenr	Inkonsistens- index	n
Om ... har mottagit skattefria stipendier:		
20(f) Hur mycket under 1989?	0.00	2
23 Hur stor var hushållets sammanlagda avgift för barnomsorgen per månad?	0.23	66
Om kvarboende betalade någon hyra:		
25(f) Hur mycket per månad?	0.42	13
29 Hur stor var hyran/avgiften per månad 1989?	0.10	193
30a Ränteutgift för lån på bostadsrätt (hela 1989)?	0.11	55
30b Skuldbelopp 31/12 på lån till bostadsrätt?	0.19	55
34 Hur stor är bostadens yta i kvm?	0.14	511
35 Hur många rum har bostaden förutom kök/kokvrå?	0.15	539
Om hushållet ägde bostadsrätt under 1989:		
36(f) Vilket är det uppskattade marknadsvärdet?	0.32	50
Om hushållet hade tillgång till bil under 1989:		
37(f) Hur många mil har Ni uppskattningsvis kört privat under 1989 där ni själv haft utgift för drivmedel?	0.49	404

Återintervjustudie av högskoleenkätuppföljningen 1990 ¹

AM/U
Anders Karlsson

Innehållsförteckning

	Sida
0. Sammanfattning	18
1. Inledning	19
1.1 SCBs elevuppföljningar	19
1.2 Högskoleuppföljningen 1990	19
1.3 Bakgrund och syfte med återintervjustudien	19
2. Undersökningens uppläggning	20
3. Resultatredovisning	21
3.1 Jämförelse av andelen lika svar	21
3.2 Undersökning av metodskillnader	24
3.3 Specialstudie av enkätfråga 1	26
Bilaga	
1 Frågeblankett	27

1 Detta är en förkortad version av en rapport med samma titel daterad 1991-01-28.

0. Sammanfattning

Föreliggande studie har gjorts med syfte att få en uppfattning om hur väl frågorna fungerar i våra enkäter, som riktas till olika elevgrupper inom utbildningsväsendet. I detta fall var undersökningsgruppen ett urval av examinerade från högskolan 1985/86, som våren 1990 erhöll en postenkät med frågor kring verksamhet, arbetsuppgifter, vidareutbildning m m. Studien har gjorts m h a ett s k test-retest-förfarande. Samma individer har fått besvara samma frågor vid två olika tillfällen. Om frågorna fungerar på ett tillfredsställande sätt, bör svaren överensstämma (vara konsistenta) mellan de båda mät-tillfällena.

Vid utvärderingen av resultaten har valts att utgå från tre aspekter. För det första har vi jämfört andelen lika svar mellan den ordinarie undersökningen och återintervjun för samtliga enkätfrågor, och i de fall markanta skillnader föreligger försökt förklara dessa. Ett andra angreppssätt har varit att studera skillnader beroende på intervjumetod (postenkät eller telefonintervju). Slutligen har en specialstudie gjorts av variabeln som ska mäta huvudsaklig verksamhet, m h a erhållna svar på ett antal tilläggsfrågor som ställdes i återintervjun.

M h a det enkla måttet index of crude agreement studerades andelen lika svar vid de båda mättillfällena. Rena faktafrågor, som verksamhet under mätveckan, arbetsledande befattning, arbetsmarknadssektor, anställningsform, arbetstimmar, arbetslöshet och deltagande i utbildning, uppvisade en stor överensstämmelse i svaren mellan den ordinarie undersökningen och återintervjun. Frågor kring orsaker, attityder och värderingar, som t ex varför man valt ett visst arbete och vad man tyckte om högskoleutbildningen ur olika aspekter, fungerade betydligt sämre. Även frågan om yrke och arbetsuppgifter gav en lägre överensstämmelse än väntat vid jämförelse av de kodade uppgifterna.

För att få en uppfattning på överensstämmelsen för de kontinuerliga variablerna beräknades korrelationskoefficienten. Att exakt ange varaktigheten uttryckt i tid (månader, veckor) av någon företeelse, som t ex arbetets eller utbildningens längd, visade sig inte vara helt lätt. Däremot visade svaren på hur mycket man tjänade mycket god överensstämmelse mellan den ordinarie intervjun och återintervjun.

För att undersöka om valet av insamlingsmetod i någon mån påverkade svarsresultatet, bestämdes redan vid undersökningens uppläggning att den ena hälften av postenkätsvarens i den ordinarie undersökningen skulle återintervjuas via postenkät och den andra hälften per telefon. Ett p*q*-test användes som mått för att approximativt testa om signifikans förelåg mellan insamlingsmetoderna. Sex av variablerna visade sig vara signifikanta på 5 %-nivån. De största metodskillnaderna förelåg för frågan om man hade en arbetsledande befattning eller ej samt för frågan om man deltagit i någon utbildning efter examen 1985/86.

Eftersom bortfallet var avsevärt högre bland de postenkätutvalda än de telefonintervjuutvalda i återintervjufasen, standardiserades på försök värdena för de variabler som uppvisat signifikanta resultat i metodskillnadstesten för att justera för eventuella bortfalls- och selekteringseffekter.

Frågan om huvudsaklig verksamhet under mätveckan är av central betydelse för alla enkätuppföljningar inom sektionen, och vid återintervjutillfället ställdes fyra tilläggsfrågor för att få en uppfattning om validiteten för denna fråga. Det visade sig att validiteten verkade vara god, d v s frågan mäter det vi vill att den ska mäta.

Rapporten har sammanställts av Anders Karlsson (AM/U), och i analys- och metodarbetet har även Claes Andersson (I/LEDN) samt Karin Arvemo-Notstrand (AM/U) deltagit. Kostnaderna för studien har totalt uppgått till ca 165 tkr, varav 26 tkr på D/DB, 76 tkr på D/IE och ca 63 tkr till utredare och metodstatistiker. Den har till stor del, 125 tkr, finansierats m h a medel ur "kvalitetsmiljonen".

1. Inledning

1.1 SCBs elevuppföljningar

SCB producerar löpande statistik över utbildningars omfattning och inriktning i Sverige. Huvudsakligen utgår denna statistik från individbaserade register, som finns för grundskolan (årskurs 9), gymnasieskolan, komvux och högskolan. Vid sidan av detta genomförs också elevuppföljningar, vilka är enkätundersökningar bland urval av elever som lämnat en viss skolform. Som urvalsram för dessa undersökningar används oftast de ovan nämnda individregistren för resp skolform. Resultat från elevuppföljningarna används i utredningar och analyser som görs såväl vid andra enheter inom SCB, främst Prognosinstitutet vid avdelningen för arbetsmarknadsstatistik (AM/PI), som utanför SCB, t ex av offentliga utredningar, skolöverstyrelsen (SÖ) och universitets- och högskoleämbetet (UHÄ).

Genomförandet av en undersökning följer i stort sett samma mönster oavsett vilken undersökningsgrupp det gäller. Oftast är det fråga om urvalsundersökningar, med urval stratifierade efter t ex utbildningslinjer, kurser, regioner och kön. Inom varje stratum tillämpas en utvidgning av Hansen-Hurwitz-planen, med urvalsdragning i två faser (se Bengt Swenson: Två optimala urvalsplaner för kombinerad användning av enkät och telefonintervju vid bortfall i telefonfasen. SCB, Metodinformation 79:3). I första fasen dras ett obundet slumpmässigt urval utan återläggning (OSU-UÅ), vars samtliga ingående personer tillsänds en enkät. Bland dem som efter en viss tid och ett antal påminnelser inte besvarat enkäten, dras i andra fasen ett OSU-UÅ. Personerna i detta urval söks sedan per telefon för en intervju.

1.2 Högskoleuppföljningen 1990

Enligt programmet för ovan nämnda elevuppföljningar samt efter diskussioner med intressenter för högskolestatistik - bland vilka kan nämnas UHÄ, enskilda högskolor/ universitet och arbetsmarknadens parter - genomfördes under april till juni 1990 en undersökning bland dem som examinerats från en utbildningslinje i högskolan läsåret 1985/86. Syftet med uppföljningen var att få en beskrivning av verksamheten ett antal år efter examen (om t ex förvärvsarbete, var detta inom ett yrke/yrkesområde som utbildningen avsåg), hur utbildningen svarade mot de krav arbetsuppgifterna ställde samt hur man ur olika aspekter bedömde den utbildning man examinerats från. Enkätens utformning framgår av bilaga 1.

Undersökningen inriktades mot endast vissa grupper, nämligen examinerade inom datautbildningar, samhällsvetarutbildningar, vissa utbildningar inom kultur- och informationssektorn, vårdsektorn samt undervisningssektorn, ekonomexaminerade samt examinerade på enstaka kurs, fil kand och individuella linjer. Totalt omfattade denna population 14 200 individer, varav 7 800 utvaldes till att vara med i undersökningen. En stratifiering gjordes m a p utbildningslinje och kön, varvid många grupper som ansågs särskilt intressanta att studera kom att totalundersökas.

1.3 Bakgrund och syfte med återintervjustudien

Vid statistiska undersökningar av denna typ, finns risken att fel uppstår under olika faser som påverkar resultatens tillförlitlighet. Storleken av felen och deras bidrag till den totala felvariansen är oftast mycket svåra att uppskatta. De flesta tidigare studier i kvalitetshöjande syfte som gjorts med anknytning till elevuppföljningarna har gällt frågor om urvalsallokering, insamlingsplanering och bortfallets betydelse. Däremot saknas kunskap om enkätfrågornas reliabilitet och validitet. Den senaste studien genomfördes 1979 (RIS-projektet: Studier av reliabilitet och validitet i enkätfrågor med hjälp

av återintervjuer i samband med I/UKs grundskoleuppföljning 1977), men eftersom enkätfrågornas utformning i viss mån förändrats och frågor av värderande och attitydmässig karaktär tillkommit, är det värdefullt att ånyo göra en studie av kvaliteten. Frågor som mäter variabeln verksamhet (arbete, studier, arbetslöshet etc) är gemensamma för alla enkätuppföljningar inom sektionen. De spelar en central roll oavsett utbildningsnivå hos gruppen de riktar sig till, medan frågor om motiv för studier m m främst berör högskoleuppföljningarna.

Mot denna bakgrund söktes medel ur "kvalitetsmiljonen" till att genomföra en återintervjustudie i anslutning till högskoleuppföljningen 1990.

2. Undersökningens uppläggning

I samband med valet av examensgrupper till den ordinarie undersökningen, utvaldes även de som skulle ingå i återintervjustudien. För att få en spridning av examina inom olika utbildningssektorer medtogs följande: examinerade på civilingenjörslinjen-elektroteknik, ekonomlinjen-nya, rehabiliteringslinjen och fritidspedagoglinjen. Det bör nämnas att urvalssannolikheten varierade mellan olika strata. Av de totalt 7 800 utvalda i undersökningen omfattade ovan nämnda grupper 1 050 personer. För att senare ha möjlighet att studera olika metoders inverkan på undersökningens resultat, bestämdes att hälften av de postenkätsvarande skulle återintervjuas via postenkät och hälften genom telefonintervju. I bortfallsuppföljningen skedde insamlingen via telefonintervjuer i både den ordinarie undersökningen och i återintervjustudien. I den ordinarie undersökningen genomfördes telefonintervjuerna av intervjuare vid telefongruppen i Örebro, medan samtliga återintervjuer per telefon gjordes av en utvald trupp regionala intervjuare runt om i landet.

Ett antagande vid analys av återintervjudata är att de två observerade värdena för varje individ är oberoende av varandra. Görs återintervjun för tidigt i förhållande till den ordinarie intervjun är det möjligt att intervjupersonen kommer ihåg hur han/hon svarade vid det ordinarie tillfället, och om för lång tid föreligger mellan intervjutillfällena finns risken att intervjupersonen har glömt bort förhållandena om det efterfrågade. Efter att ha konsulterat olika källor bestämdes därför att tiden mellan ordinarie intervju och återintervju skulle vara 2-4 veckor. Instruktionerna var samma vid båda tillfällena. Vid återintervjun ställdes samma frågor till undersökningsspersonen som vid den ordinarie intervjun samt 4 tilläggsfrågor för att utvärdera validiteten för enkätfråga 1, som är av stor betydelse i samtliga elevuppföljningar.

En sammanfattning av svarsfrekvenser, bortfall m m i de olika faserna framgår av tabell 1 nedan.

Tabell 1 Svarsfrekvenser, bortfall m m i återintervjustudien av högskoleuppföljningen 1990.

			Återintervjustudien		Antal svar		Bortfall	
			"Utsända"		postenk		postenk	
			postenk	tfnint	postenk	tfnint	postenk	tfnint
Ordinarie undersökningen	Antal svar	764	346	418	165	354	181	64
	varav postenkät	706	346	360	165	320	181	40
	varav telefonint	58	0	58	0	34	0	24

3. Resultatredovisning

Vid utvärderingen av resultaten från föreliggande studie har valts att utgå från följande tre aspekter:

- 1 Jämförelse av andelen lika svar från den ordinarie undersökningen till återintervjun för samtliga enkätfrågor, och i de fall markanta skillnader föreligger försöka förklara dessa.
- 2 Jämförelse av om det finns skillnader i svaren beroende på vilken intervjumetod (postenkät eller telefonintervju) som valts.
- 3 En specialstudie av enkätfråga 1 (huvudsaklig verksamhet) m h a erhållna svar på tillägsfrågorna.

3.1 Jämförelse av andelen lika svar

I litteraturen förekommer en mängd olika, mer eller mindre sofistikerade analysmått när det gäller att mäta och tolka svarsvariabilitet (en del tips finns i källförteckningen för den som är intresserad). Varje mått är dock behäftat med sina svagheter beträffande förutsättningar som måste föreligga, känslighet för sneda svarsfördelningar o s v och i denna studie har vi därför och p g a att begränsat med tid stått till förfogande valt att använda oss utav relativt enkla metoder.

Ett enkelt sätt när man handskas med kategoriska data är följande. Vid jämförelse av resultat för en speciell variabel (fråga) kan data representeras av den kvadratiske matrisen $[p_{ij}]$, där p_{ij} är proportionen element som klassificerats i kategori (svarsalternativ) "i" enligt den ordinarie intervjun och i kategori "j" enligt återintervjun.

		Klassifikation enligt ordinarie intervju				
		1	2	. . .	L	Totalt
Klassifikation enligt återintervju	1	p_{11}	p_{12}	. . .	p_{1L}	$p_{1.}$
	2	p_{21}	p_{22}	. . .	p_{2L}	$p_{2.}$
	.	.	.	.	.	.
	.	.	.	.	.	.
	.	.	.	.	.	.
	L	p_{L1}	p_{L2}	. . .	p_{LL}	$p_{L.}$
Totalt		$p_{.1}$	$p_{.2}$	. . .	$p_{.L}$	1

Diagonalelementen, p_{ii} , $i = 1, 2, \dots, L$, i denna matris utgörs av de fall där resultaten mellan intervjuomgångarna överensstämmer. Proportionen av fall som klassificerats identiskt vid de två observationerna är då

$$p_0 = \sum_i p_{ii},$$

och brukar kallas index of crude agreement. En nackdel med detta mått är att det inte tar hänsyn till att en viss överensstämmelse erhålls p g a slumpen, även om mätresultaten är slumpmässigt utspridda över kategorierna.

I tabell 2 redovisas index of crude agreement för de icke-kontinuerliga variablerna i undersökningen. Yrkesvariabeln (fråga 2) består av en femsiffrig kod, och har här transformerats om till fem nya variabler, där
 variabel 2.1=1 om yrkeskoden överensstämmer på 1-siffernivå, 0 annars ,
 variabel 2.2=1 om yrkeskoden överensstämmer på 2-siffernivå, 0 annars ,
 variabel 2.3=1 om yrkeskoden överensstämmer på 3-siffernivå, 0 annars ,
 variabel 2.4=1 om yrkeskoden överensstämmer på 4-siffernivå, 0 annars samt
 variabel 2.5=1 om yrkeskoden överensstämmer på 5-siffernivå, 0 annars .

Allmänt kan konstateras att rena faktafrågor som verksamhet under mätveckan (fråga 1), arbetsledande befattning (fråga 3), arbetsmarknadssektor (fråga 4), anställningsform (fråga 5), arbetstimmar (fråga 6), arbetslöshet (fråga 17) och deltagande i utbildning (fråga 19) uppvisade stor överensstämmelse i svaren mellan ordinarie undersökningen och återintervjun. Däremot visade sig frågor kring orsaker, attityder och värderingar fungera betydligt sämre, särskilt frågorna 13 och 18 (hur högskoleutbildningen svarade mot arbetsuppgifternas krav resp bedömning av högskoleutbildningen ur olika aspekter). Även frågan om yrket visade upp kanske förvånansvärt stora skillnader i svaren mellan de båda tillfällena.

Tabell 2 Index of crude agreement för de icke-kontinuerliga variablerna i återintervjustudien av högskoleuppföljningen 1990.

Fråga	Index of crude agreement (proc)	Bastal** (antal svarande)	Fråga	Index of crude agreement (proc)	Bastal** (antal svarande)
1	96	513	16:antal inom yrk	80	512
2.1*	88	491	16:antal andra yrk	89	512
2.2*	80	491	17	96	505
2.3*	71	491	18:delfråga 1	78	496
2.4*	65	491	18:delfråga 2	74	496
2.5*	62	491	18:delfråga 3	71	464
3	92	483	18:delfråga 4	65	498
4	99	489	18:delfråga 5	77	493
5	97	493	18:delfråga 6	74	474
6	98	492	18:delfråga 7	67	492
8	73	487	19:svarsalt 1	88	519
10	72	490	19:svarsalt 2	84	519
11	79	492	19:svarsalt 3	85	519
12	65	26	19:svarsalt 4	97	519
13:delfråga 1	64	453	19:svarsalt 5	77	519
13:delfråga 2	55	446	20:persutb	75	519
13:delfråga 3	65	452	20:högskutb	89	519
13:delfråga 4	62	450	20:komvux	98	519
13:delfråga 5	51	423	20:AMU	99	519
14	82	487	20:ann utb	84	519
15	71	48			

* Yrkesvariabeln är omtransformerad, se kommentar i texten sid 21.

** Endast de som svarat vid båda tillfällena är medräknade.

Nedan följer några kommentarer till de största avvikelserna, fråga för fråga.

Fråga 1

Fråga 1 fungerade i sin helhet bra, men eftersom den huvudsakliga verksamheten är av så central betydelse i våra undersökningar studerades även om det förelåg någon missstämmelse m a p svarsalternativ. Det visade sig att överensstämmelsen var mycket god för alla de enskilda svarsalternativen.

Fråga 2

Yrke kodades med en femsiffrig kod enligt yrkesregistret som används i Folk- och bostadsräkningen 1985. Underlag har främst varit svaret på fråga 2, men även svaren på frågorna 3 och 4 samt uppgift om examen har utnyttjats som hjälpinformation. Som framgår av tabell 2 var det rätt stora skillnader i yrkeskoden mellan ordinarie intervju och återintervju; på 3-siffernivå var t ex överensstämmelsen endast 71 procent. Yrke används ofta som redovisningsvariabel och indelningen i redovisningsgrupper görs på 2- eller 3-siffernivå. En djupare analys av skillnaderna var därför angelägen, vilket inkluderade en studie av primärmaterialet.

Resultatet av denna studie visade att cirka 70 procent av skillnaderna kunde förklaras av att svaren var tillräckligt avvikande mellan de två tillfällena för att rendera olika koder. Yrkesregistret har i viss mån bidragit till problemet, genom att det i hög grad bygger på uppgifter om näringsgren och arbetsmarknadssektor. Det innebär t ex att om yrket "badbiträde" kompletterats med upplysningen att arbetsmarknadssektorn är landstinget erhålls en kod som redan på 1-siffernivå skiljer sig från den kod som sätts om "badbiträde" angetts ospecificerat. Särskilt frekventa visade sig avvikelserna vara när det gällde grupperna med tekniskt, ekonomiskt eller administrativt arbete. I de återstående fallen (på 3-siffernivå 10 procent av samtliga), där nästan eller exakt samma upplysningar erhöles för yrkesvariabeln vid båda mättillfällena ändå lett till olika kodning, beror skillnaderna huvudsakligen på att yrket i sig självt kan vara "diffust" och inte självklart kan kopplas samman med en enda unik kod i det utnyttjade yrkesregistret (t ex "konsult - konstruktion") eller på olika utnyttjandegrad av hjälpinformationen vid kodningstillfällena.

En teori om att det systematiskt vid något av intervjutillfällena skulle ha avgetts ett utförligare svar (t ex att man vid återintervjun tog sig tid till att precisera sina arbetsuppgifter mer "fylligt") kunde inte bekräftas.

Fråga 13

Frågan om den högskoleutbildning man examinerats från 1985/86 svarade mot arbetsuppgifternas krav i det arbete man hade under mätveckan visade sig vara den som fungerat sämst i undersökningen; endast 51 till 65 procents överensstämmelse för de olika delfrågorna. En analys gjordes med svarsalternativ 2 (ja, något) och 3 (ja, i hög grad) sammanslagna för delfrågorna 1-4. För delfråga 5 sammanslogs alternativ 1 och 2 (mycket/något mer) samt 4 och 5 (något/mycket mindre). Index of crude agreement ökade visserligen till mellan 63 och 79 procent, men är fortfarande inte tillfredsställande. Ett förhållandevis stort partiellt bortfall för frågan tyder också på att svårigheter förelagat vid besvarandet, och att både utformning av alternativen och formulering måste tänkas igenom.

Fråga 18

Att göra en bedömning av den högskoleutbildning man avslutat 1985/86 ur olika aspekter tillhörde också de frågor som fungerade mindre bra. Utformning av alternativen och formulering av delfrågorna bör tänkas igenom. En tendens var att svarsalternativ 2 (ganska bra) föredrogs. En vanlig kommentar var att skillnaden var stor mellan olika kurser som ingick på utbildningslinjen och att det därför var svårt att ange ett svarsalternativ.

Fråga 20

Bland svaren på frågan om vilka utbildningar man deltagit i efter 1985/86 var det framför allt relativt dålig överensstämmelse beträffande personalutbildning. Detta indikerar att definitionen som ges i enkäten inte är lyckad utan bör ses över.

Ett enkelt mått på överensstämmelse när det gäller kontinuerliga variabler är att studera korrelationen mellan värdena från de båda mätillfällena. Analysen bjöd inte på några större överraskningar. Att exakt ange varaktigheten uttryckt i tid (månader, veckor) av någon företeelse visade sig inte vara helt lätt, och i fråga 20 komplicerades det också av att utbildningens längd skulle räknas om till heltidsstudier. Svaren på fråga 7 om hur mycket man tjänade visade god överensstämmelse mellan den ordinarie intervjun och återintervjun.

3.2 Undersökning av metodskillnader

För att undersöka om valet av insamlingsmetod i någon mån påverkade svarsresultatet, bestämdes redan vid planeringen av urvalet att den ena hälften av postenkätsvararna i den ordinarie undersökningen skulle återintervjuas via postenkät och den andra hälften per telefon.

$$\text{Låt } p^* = \frac{n_{ee}p_{ee} + n_{et}p_{et}}{n_{ee} + n_{et}}$$

där n_{ee} = antal som svarat i den ordinarie undersökningen med postenkät och i återintervjun med postenkät

p_{ee} = andel lika svar bland dem som svarat i den ordinarie undersökningen med postenkät och i återintervjun med postenkät

n_{et} = antal som svarat i den ordinarie undersökningen med postenkät och i återintervjun per telefon

p_{et} = andel lika svar bland dem som svarat i den ordinarie undersökningen med postenkät och i återintervjun per telefon.

Ett approximativt mått för att testa om signifikans föreligger mellan insamlingsmetoderna är då följande:

$$\frac{p_{ee} - p_{et}}{\sqrt{p^*(1 - p^*)\left(\frac{1}{n_{ee}} + \frac{1}{n_{et}}\right)}}, \text{ ett s k } p^*q^*\text{-test.}$$

I tabell 3 på nästa sida redovisas resultatet av p^*q^* -testen. Sex av variablerna visade sig vara signifikanta på 5 %-nivån. De största metodskillnaderna förelåg för fråga 3 (arbetsledande befattning) och svarsalternativ 1 på fråga 19. På den sistnämnda frågan svarade en betydligt högre andel av postenkätsvararna än de telefonintervjuade att man inte deltagit i någon utbildning efter examen 1985/86, 21 mot 15 procent.

Som framgått tidigare (bl a tabell 1) var bortfallet avsevärt högre bland de postenkätutvalda än de telefonintervjuutvalda i återintervjufasen. För att justera för eventuella bortfallseffekter samt selekteringseffekter vid urvalet standardiserades därför värdena för de variabler som uppvisat signifikanta resultat i metodskillnadstesten. För en beskrivning av hur standardiseringen gjordes, hänvisas till originalversionen av denna rapport. En test av nollhypotesen $H_0: p_{ee} - p_{et} = 0$ med en p^*q^* -test visade på signifikanta skillnader för fråga 11, 13 och 14.

Tabell 3 Index of crude agreement samt p*q*-test av intervjumetod för de icke-kontinuerliga variablerna i återintervjustudien av högskoleuppföljningen 1990. Postenkätsvarande i den ordinarie undersökningen fördelade efter intervjumetod i återintervjufasen.

Fråga	Återintervjumetod: postenkät		Återintervjumetod: telefon		p*q*-test
	Index of crude agreement	Bastal** (antal svarande)	Index of crude agreement	Bastal** (antal svarande)	
1	98	160	96	319	1,15
2.1*	86	154	88	305	- 0,60
2.2*	79	154	79	305	0,00
2.3*	71	154	68	305	0,66
2.4*	67	154	62	305	1,05
2.5*	64	154	58	305	1,24
3	97	153	89	299	2,92***
4	99	154	99	304	0,00
5	96	156	98	305	- 1,26
6	99	156	97	304	1,35
8	80	155	68	301	2,71***
10	77	155	69	303	1,80
11	85	156	76	304	2,25***
12	55	11	67	12	- 0,59
13:delfråga 1	57	143	67	281	- 2,02***
13:delfråga 2	59	139	53	278	1,16
13:delfråga 3	69	142	63	281	1,22
13:delfråga 4	62	141	62	280	0,00
13:delfråga 5	54	134	50	262	0,75
14	88	153	79	303	2,36***
15	73	22	67	21	0,43
16:antal inom yrk	84	161	79	318	1,31
16:antal andra yrk	88	161	90	318	- 0,70
17	97	157	97	315	0,00
18:delfråga 1	79	158	76	306	0,73
18:delfråga 2	75	155	74	308	0,23
18:delfråga 3	72	148	71	287	0,22
18:delfråga 4	65	156	65	309	0,00
18:delfråga 5	78	156	74	305	0,94
18:delfråga 6	80	148	71	296	2,04***
18:delfråga 7	71	154	64	307	1,50
19:svarsalt 1	95	165	85	320	3,26***
19:svarsalt 2	85	165	83	320	0,56
19:svarsalt 3	88	165	84	320	1,18
19:svarsalt 4	98	165	96	320	1,16
19:svarsalt 5	78	165	76	320	0,49

* Yrkesvariabeln är omtransformerad, se kommentar i texten sid 21.

** Endast de som svarat vid både det ordinarie tillfället och i återintervjun är medräknade.

*** Signifikant resultat på 5 %-nivån.

3.3 Specialstudie av enkätfråga 1

Frågan om huvudsaklig verksamhet under mätveckan är som tidigare nämnts av mycket stor betydelse i alla elevuppföljningar, eftersom denna fråga utgör referens i tiden för hela undersökningen. Vid återintervjutillfället ställdes därför, förutom samma frågor som vid det första tillfället, fyra tilläggsfrågor för att få en uppfattning om validiteten för enkätfråga 1, d v s försöka ringa in "den absoluta sanningen" om vad man i huvudsak sysslade med under den aktuella perioden.

I den första tilläggsfrågan efterfrågades om förvärvsarbete över huvudtaget utförts under mätveckan. Fem svarsalternativ fanns att välja emellan, ett nej-svar samt fyra alternativ med olika intervall för arbetstidens längd. Mycket god överensstämmelse rådde mellan svaren på enkätfråga 1 och tilläggsfråga 1. Tilläggsfrågan handlade om vad som inträffat under mätveckan, medan man i fråga 1 skulle bortse från tillfällig frånvaro (t ex sjukdom, föräldraledighet, semester, tjänstledighet). Att 42 personer i enkäten svarat att de förvärvsarbetade men angett ett nej-svar på tilläggsfrågan behövde alltså inte innebära någon motsägelser. Se vidare kommentarerna till tilläggsfråga 3.

I tilläggsfråga 2 efterfrågades om man hade någon anställning under mätveckan. Även här överensstämde erhållna svar väl med svaren för enkätfråga 1. Text sade sig de arbetslösa enligt enkätfrågan ej heller inneha någon anställning. Att man huvudsakligen förvärvsarbetade men saknade anställning (angavs av 4 personer) är riktigt för t ex egna företagare.

Tilläggsfråga 3 besvarades endast av dem som uppgett sig ha någon anställning i tilläggsfråga 2. Återigen skulle den huvudsakliga verksamheten under mätveckan anges, men man hade fler svarsalternativ att välja emellan än i enkätfråga 1 (bl a semester, sjuk, föräldraledig och tjänstledig). En av farhågorna när det gällde validiteten för enkätfrågan om huvudsaklig verksamhet var att man trots instruktionen vid förvärvsarbete om att bortse från tillfällig frånvaro i dylikt fall placerade sig under annan verksamhet. Detta visade sig dessbättre inte vara fallet i någon större utsträckning. I de flesta fall angav man förvärvsarbete som huvudsaklig aktivitet i fråga 1, även om man den aktuella mätveckan var frånvarande p g a sjukdom, föräldraledighet, semester, tjänstledighet m m.

Tilläggsfråga 4 besvarades endast av dem som i tilläggsfråga 2 uppgett att de inte hade någon anställning. Liksom i tilläggsfråga 3 efterfrågades den huvudsakliga verksamheten under mätveckan, men med fler svarsalternativ att välja emellan än i enkätfråga 1 (bl a militärtjänst, sökte arbete under veckan men fick inget, ville ha arbete men sökte inget, väntade på svar på arbete jag sökt tidigare, slutat en anställning och inte hunnit börja en ny). Inga klara felaktigheter kunde påvisas, och de föreskrivna alternativen visade sig fånga upp större delen av svaren.

Det är viktigt att frågorna besvaras av Dig som brevet är adresserat till.

Examensuppföljningen 1990

Denna enkät vänder sig till Dig som slutförde en högskoleutbildning läsåret 1985/86.

Om Du har en högskoleexamen på fil kand, enstaka kurs eller individuell linje, ange inriktningen (huvudämnet)

(11)

Fråga 1 Besvaras av alla.

1 Vad gjorde Du **huvudsakligen** under veckan 2-8 april 1990?

(14)

- Ange endast ett alternativ.
- Bortse från tillfällig frånvaro (t ex sjukdom, föräldraledighet, semester, tjänstledighet).

- ☐ 1 Förvärvsarbetade (även arbete i eget företag)
- ☐ 2 Studerade
- ☐ 3 Arbetslös (dvs sökte, men fick ingen anställning under veckan)
- ☐ 4 Annat, ange vad (t ex arbetade i eget hushåll, militärtjänst)

(15)

Fråga 2-15 Besvaras av alla som hade arbete (även deltid) under veckan 2-8 april 1990.

Frågorna gäller även Dig som var tillfälligt frånvarande p g a föräldraledighet, semester eller tjänstledighet t ex för studier.

2 Vilket arbete hade Du under veckan 2-8 april 1990?

(16) Yrke och arbetsuppgifter

- Om yrkesbenämningen är allmän, ange Dina arbetsuppgifter, t ex sjuksköterska, mentalvård, konsulent-försäljning.

3 Hade Du en arbetsledande befattning?

(46)

- ☐ 1 Ja ☐ 2 Nej

4 Inom vilken arbetsmarknadssektor hade Du arbete under veckan 2-8 april 1990?

(47)

- Ange även företagets namn.

- ☐ 1 Statlig
- ☐ 2 Kommunal, landsting
- ☐ 3 Enskild, annan

Företagets namn

5 Vilken anställningsform hade Du under veckan 2-8 april 1990?

(48)

- ☐ 1 Fast anställning (tillsvidareanställning)
- ☐ 2 Beredskapsarbete
- ☐ 3 Tillfällig/tidsbegränsad anställning (vikariat, objektsanställning, forskarförordnande)
- ☐ 4 Egen företagare

6 Hur många timmar arbetar Du en normal arbetsvecka?	(49)	☐ 1 35 timmar eller mer per vecka (heltid) ☐ 2 Mellan 20-34 timmar per vecka ☐ 3 Mindre än 20 timmar per vecka
7 Hur länge har Du haft det arbete Du hade under veckan 2-8 april 1990?	(50)	årmån
8 Hur fick Du kännedom om det arbete Du hade under veckan 2-8 april 1990? - Ange endast ett alternativ.	(54)	☐ 1 Läste annons i dagspressen ☐ 2 Information genom arbetsförmedlingen ☐ 3 Information via yrkeshandledning, öppet hus eller dylikt på högskolan ☐ 4 Information under praktiktjänstgöring/feriearbete på företaget ☐ 5 "Värvning" av företaget ☐ 6 Bytte arbete inom företaget ☐ 7 Tips från bekant/släkting ☐ 8 På annat sätt, ange hur.....⁽⁵⁵⁾
9 Hur mycket tjänade Du i månaden före skatt på det arbete Du hade under veckan 2-8 april 1990?	(56)	kronor
10 Vad var det viktigaste skälet till att Du tog det arbete Du hade under veckan 2-8 april 1990? - Ange endast ett alternativ.	(62)	☐ 1 Ville arbeta inom just det yrkesområdet ☐ 2 Bra lön ☐ 3 Bra möjligheter till utveckling och avancemang i arbetet ☐ 4 Bra möjligheter att anpassa arbetet till barn och familjeliv ☐ 5 För att få stanna på den ort där jag bodde ☐ 6 Annat skäl, ange vad.....⁽⁶³⁾
11 Hur överensstämde det arbete Du hade under veckan 2-8 april 1990 med den högskoleutbildning Du avslutade 1985/86?	(64)	☐ 1 Arbetet var helt inom det yrke/yrkesområde utbildningen avsåg (målyrket) ☐ 2 Arbetet var delvis inom det yrke/yrkesområde utbildningen avsåg ☐ 3 Arbetet var inom ett annat yrke/yrkesområde än det utbildningen avsåg } → Gå till fråga 13 → Besvara fråga 12
12 Vad var det viktigaste skälet till att Du arbetade inom ett yrke som Inte svarade mot den högskoleutbildning Du avslutade 1985/86? - Ange endast ett alternativ.	(65)	☐ 1 Kunde inte få något arbete inom yrkesområdet på den ort jag bor ☐ 2 Dålig lön inom målyrket ☐ 3 Är inte längre intresserad av arbetsuppgifterna inom målyrket ☐ 4 Dåliga arbetsförhållanden/för stressigt inom målyrket ☐ 5 Annat skäl, ange vad.....⁽⁶⁶⁾
—————→ Gå vidare till fråga 14		

13 Hur tycker Du att den högskoleutbildning Du avslutade 1985/86 svarade mot de krav arbetsuppgifterna ställde i det arbete Du hade under veckan 2-8 april 1990?

- Svara med ett alternativ på varje fråga och ange gärna, i förekommande fall, ämne/n (t ex engelska, statistik, farmakologi).

	Nej (67)	Ja, något	Ja, i hög grad	Ange ämne(n) (68)
Krävde arbetsuppgifterna kunskaper i ämnen som utbildningen ej omfattade.....	☐ 1	☐ 2	☐ 3	
Krävde arbetsuppgifterna mer/fördjupade kunskaper i ämnen som utbildningen inte omfattade.....	☐ 1 (87)	☐ 2	☐ 3	 (78) (88)
Omfattade utbildningen ämnen som arbetsuppgifterna inte krävde kunskap i.....	☐ 1	☐ 2	☐ 3	
Omfattade utbildningen mer/fördjupade kunskaper i vissa ämnen än vad arbetsuppgifterna krävde.....	☐ 1 (97)	☐ 2	☐ 3	 (98)

	Mycket mer (107)	Något mer	Lika mycket	Något mindre	Mycket mindre	Ange vad (108)
Hur mycket krävde arbetsuppgifterna av de "allmänna högskolekunskaper" som Du fick i utbildningen (t ex kritiskt/vetenskapligt förhållningssätt, självständigt arbete)?.....	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	

14 Planerar Du att byta arbete eller övergå till annan verksamhet under det närmaste året?

- Ange endast ett alternativ.

☐ 1	Ja, byta arbete till ett annat yrkesområde	} → Gå till fråga 16
☐ 2	Ja, byta arbete inom det yrkesområde jag nu arbetar	
☐ 3	Ja, övergå till studier	
☐ 4	Ja, övergå till annan verksamhet (t ex arbete i eget hushåll)	
☐ 5	Nej	

15 Vilket är det viktigaste skälet till att Du planerar att byta till ett annat yrkesområde?

- Ange endast ett alternativ.

☐ 1	Vill ha mer kvalificerade arbetsuppgifter
☐ 2	Vill ha ett arbete där jag kan utnyttja min högskoleutbildning bättre
☐ 3	Vill ha ett arbete med bättre lön och/eller anställningsförhållanden
☐ 4	Annat skäl, ange vad..... (111)

Fråga 16-20 Besvaras av alla.

16 Hur många arbeten och hur länge har Du arbetat sedan Du avslutade Din högskoleutbildning 1985/86?

- Räkna endast arbeten där Du förvärvat arbetat minst halvtid under minst 3 månader.
- Räkna samman hur lång tid Du arbetat inom målyrket resp inom andra yrkesområden.

	Inget arbete (112)	Ett arbete	Två arbeten	Tre eller flera arbeten	Sammanlagd tid Du arbetat (113)
Arbete inom yrkesområden som svarar mot utbildningen (målyrken).....	☐ 1	☐ 2	☐ 3	☐ 4	 år..... mån
andra yrkesområden.....	☐ 1 (116)	☐ 2	☐ 3	☐ 4	 år..... mån (117)

- 17 Hur lång tid har Du sammanlagt varit arbetslös sedan Du avslutade Din högskoleutbildning 1985/86? (120)
- ☐ 1 Jag har inte varit arbetslös (121)
- ☐ 2 Arbetslös under.....månader
- Med arbetslös menas att Du inte haft något arbete och samtidigt sökt arbete.

18 Vad tycker Du om den högskoleutbildning Du avslutade 1985/86?

- Gör en översiktlig bedömning för hela utbildningen, för varje aspekt nedan och skriv gärna kommentarer.
- Om Du inte kan ta ställning till någon aspekt t ex för att den inte förekom i Din utbildning, ange det i kommentarerna.

	Mycket bra (123)	Ganska bra	Dåligt	Kommentarer/förtydliganden (t ex för lite för mycket, ge exempel) (124)
Fördelningen av totala poäng på olika ämnesområden.....	☐ 1 (125)	☐ 2	☐ 3	 (126)
Teorinnehållet i de olika ämnesområdena.....	☐ 1 (127)	☐ 2	☐ 3	 (128)
Praktiken, arbetsträningen.....	☐ 1 (129)	☐ 2	☐ 3	 (130)
Fördelningen mellan schemabunden undervisning och självständigt arbete.....	☐ 1 (131)	☐ 2	☐ 3	 (132)
Kvaliteten på undervisningen.....	☐ 1 (133)	☐ 2	☐ 3	 (134)
Forskningsanknytningen, kontakter med forskare inom olika ämnesområden.....	☐ 1 (135)	☐ 2	☐ 3	 (136)
Träning i att utveckla förmågan kritiskt/vetenskapligt tänkande.....	☐ 1	☐ 2	☐ 3	

- 19 Har Du deltagit i någon utbildning/kurs sedan Du slutade högskoleutbildningen 1985/86? (137)
- ☐ 1 Nej (138)
- ☐ 2 Ja, studerat som en naturlig fortsättning på min examen (139)
- ☐ 3 Ja, studerat för nöjes skull (140)
- ☐ 4 Ja, studerat för att kunna byta yrkesområde (141)
- ☐ 5 Ja, studerat för att tillgodose de krav som arbetsuppgifterna ställde

Besvara fråga 20

20 Vad var det för kurser, hur många och hur långa var de?

- Som **personalutbildning** räknas all utbildning som helt eller delvis bekostas av arbetsgivaren i form av avgifter eller som betald arbetstid, oberoende om den arrangeras av arbetsgivaren själv eller i samarbete med t ex högskola, KOMVUX.
- Räkna kursens/kursernas **smmanlagda längd** som summan av antalet veckors heltidsstudier. En högskolepoäng motsvarar en veckas heltidsstudier.

	Antal kurser	Sammanlagd längd (antal veckor)		Antal kurser	Sammanlagd längd (antal veckor)
(142)	☐ 1	Personalutbildning....	(160)	☐ 4	Arbetsmarknadsutbildning..
(148)	☐ 2	Högskoleutbildning...	(166)	☐ 5	Annan utbildning, ange vad
(154)	☐ 3	KOMVUX.....			

Eventuella kommentarer..... (172) Jag önskar en sammanställning av resultaten ☐ (173)

Återintervjustudie av Årsarbetskraften 1989

Am/AKU
**Majvor Karlsson
Solveig Thudin**

Innehåll	Sida
0 Sammanfattning	32
1 Inledning	33
2 Metod	33
3 Genomförande	34
4 Resultat	35

Bilaga

- | | |
|---|--|
| 1 | Dikotomiserade tabeller
med avvikelsemått |
|---|--|

Detta är en reviderad version av en tidigare rapport med samma titel daterad 910130

0 Sammanfattning

Svarsvariabiliteten redovisas i rapporten för nära nog alla frågor och kan användas som ett underlag vid bedömningen av den totala kvalitén i undersökningen av årsarbetskraften. Flera idéer om hur mätinstrumentet skulle kunna förbättras har framkommit. Detta arbete är resurskrävande och ligger utanför ramen för detta projekt.

Svarsavvikelseerna varierar kraftigt mellan frågorna i undersökningen om årsarbetskraften. Några frågor visar en svarsvariabilitet som är acceptabel enligt vår enkla analysmetod (T3, T5, T11), medan några frågor ger ett nedslående resultat (T25, T31, T32 och T34). För övriga frågor är svarsavvikelseerna något för stora för att man helt ska anse variabiliteten godtagbar.

Den inledande frågan (T1), som kartlägger antalet sysselsatta veckor, har inte alarmerande svarsavvikelser, men leder till fel i andra frågor. Därför bör man omarbета denna fråga tillsammans med fråga T34, som kartlägger antalet arbetslöshetsveckor. Summan av svaren i dessa två frågor ger nämligen antal veckor i årsarbetskraften, som är en bakgrundsvariabel i undersökningen. Därefter kan frågorna som avser huvudsakligt arbete under året, ombytessökandet och byten av arbete ställas. Här bör statistikanvändarnas åsikter inhämtas. En ändring av några frågors formuleringar kan ge besvärande tidsseriebrott. Detta har inte tidigare skett, utan undersökningen om årsarbetskraften har i stort sett varit oförändrad sedan den startade 1967.

Detta leder oss till funderingar om indelningen i helårs- och delårsarbetskraften är väsentlig i undersökningen, eller om det är byten och sökandet av (nytt) arbete som är viktigt, oavsett om man tillhör arbetskraften hela eller del av året. I våra kontakter med användarna har det framkommit att intresset är mycket stort för uppgifter som belyser både den externa och den interna rörligheten. Speciellt när det gäller ombytessökandet finns det egentligen ingen annan statistikkälla än Årsarbetskraften. Detta är något att arbeta vidare med och beakta vid eventuella förändringar av frågorna T1, T33A/T33B och T34.

Svarsavvikelseerna tycks till en del bero på glömska, men det finns flera indikationer på att man inte förstår frågan. För flertalet frågor har vi angett några kvalitetshöjande rekommendationer, som sammanfattat är:

- Formuleringen av frågan som kartlägger antalet sysselsatta veckor (T1) bör ändras. För att underlätta för intervjupersonen att minnas skulle intervjuaren kunna använda en teknik där hela året kartläggs i exempelvis tremånadersintervall.
- Frågan som kartlägger antalet arbetslösa veckor (T34) bör förtydligas och ställas i nära anslutning till T1.
- Frågan som kartlägger genomsnittlig arbetstid (T2) bör förtydligas.
- Kombinationen byten av arbetsgivare (företag) och externt ombytessökande bör utredas.
- Vi bör utreda hur man kan förbättra och lyfta fram frågorna om ombytessökandet och byten av arbete, eventuellt då i kombination med en nedtoning av kartläggningen av hel- och delårsarbetskraften.
- Instruktionerna till intervjuarna bör förbättras.

Projektkostnaderna har uppgått till sammanlagt 54 000 kronor, varav 46 000 har finansierats med medel ur 1990 års "kvalitetsmiljon". Kostnaderna har fördelat sig på intervjuer (14 000 kr), dataregistrering (2 000 kr) och analys (38 000 kr). Övriga kostnader har belastat AKU (DATI-projektets produktionsprov).

1 Inledning

Undersökningen av årsarbetskraften har genomförts varje år sedan 1967 som en tilläggsundersökning till AKU i februari året efter undersökningsåret. Fr o m år 1991 genomförs undersökningen i januari månad.

Årsarbetskraften belyser individernas arbetsmarknadssituation under hela det föregående året. Undersökningen beskriver bl a hur många personer som varit sysselsatta hela året eller del av året, hur många personer som varit arbetslösa och i förekommande fall antal arbetslöshetsveckor. Vidare kartläggs hur många som varit ombytessökande och hur många som bytt arbete under föregående år.

I alla intervjuundersökningar förekommer svarsfel både systematiska och slumpmässiga, som uppstår genom att intervjupersonen inte minns eller har missförstått frågan eller att intervjuaren inte har helt klart för sig vad frågan ska ge svar på. Det finns en misstanke att svarsfelen är stora i undersökningen om årsarbetskraften.

Det är naturligtvis problematiskt att ställa frågor där referensperioden är så lång och ligger så långt tillbaka i tiden, då risken för svarsfel p g a glömskeeffekter hos intervjupersonerna ökar.

Några kvalitetsstudier av Årsarbetskraften (tidigare benämnd Årssysselsättningen) utförda 1979-1981 av SCB visade bl a att man kommer ihåg vad man *huvudsakligen* gjort under året, medan *tillfälliga* händelser tenderar att glömmas bort. Svarsfel beroende på glömska tycks ha den systematiska effekten att antalet personer sysselsatta under hela året överskattas och antalet som varit arbetslösa någon gång under året underskattas.

Glömskeeffekter leder troligen även till slumpmässiga svarsfel. Det torde exempelvis vara svårt för intervjupersonen att ange antalet veckor man haft ett arbete om arbetet är ett oregelbundet arbete och om man dessutom haft flera arbetsgivare under året.

Föreliggande rapport avser att belysa svarsvariabiliteten vid två oberoende intervjuer för ett urval av intervjupersoner som deltog i undersökningen av årsarbetskraften i februari 1990.

En frågeställning som behandlas är om svarsavvikelserna kan anses vara acceptabla eller alltför stora. Vid alltför stora avvikelser blir nästa steg en diskussion kring blankettkonstruktionen. Kan frågor förbättras? Bör vissa frågor uteslutas eller läggas till? Är frågornas inbördes ordning motiverad?

Efter en översiktlig metodgenomgång och beskrivning av hur studien har genomförts redovisas resultaten med kommentarer.

2 Metod

Återintervjuer är ett effektivt verktyg när man ska studera svarsfelen i en undersökning och utvärdera ett mätinstrument.

Teorin bakom *oberoende återintervjuer* är att om ett urval av individer besvarar samma frågor vid två olika tillfällen, men under i övrigt lika omständigheter, så bör svaren vara konsistenta.

Tidsavståndet mellan de båda mätningarna kan påverka svaren på så vis att ligger de för nära varandra kan intervjupersonen minnas sina svar från det första tillfället och svara likadant. Detta skulle i så fall ge ett alltför positivt resultat för återintervjustudien. Men det skulle även kunna gälla att intervjupersonen minns själva frågan och kan bemöta den bättre vid det andra intervjutillfället än vid den första intervjun, då svaren getts både oförberett och utan särskilt lång betänketid.

Tidsperioden mellan de två mätningarna bör alltså väljas omsorgsfullt så att å ena sidan återintervjun inte ligger för långt ifrån referensperioden (för att motverka glömskeffekter) och å andra sidan på lagom avstånd från föregående intervju för att behålla oberoendet mellan intervjuerna.

Svarsvariabiliteten är den slumpmässiga delen av svarsfelet, som kan hänföras till intervjuperson eller intervjuare, under förutsättning att övriga felkällor kan hållas oförändrade mellan de två tillfällena. Med detta menas att urvalsfel, ramfel, bortfallsfel, kodningsfel, stansfel och övriga bearbetningsfel visserligen finns men har teoretiskt sett samma variation och riktning vid båda tillfällena.

Den systematiska delen av svarsfelet, den s k svarsbiasen, kan också hänföras till både intervjuperson och intervjuare med ovan angivna förutsättning.

Vanligen presenteras grundmaterialet från en återintervjustudie i korstabeller där huvuddiagonalen utgörs av intervjupersoner som svarat lika vid de två tillfällena. Svarsavvikelse kan utläsas runt omkring huvuddiagonalen. Procentuell avvikelse framräknas på ett enkelt sätt, totalt och marginellt.

Två avvikelsemått som kan användas vid bedömning av storleken på svarsavvikelse finns beskrivna i Wärnerud m fl, 1990, sid 49ff. (Se även Forsman, 1987 samt Wikman, 1980). Vid framtagning av dessa mått måste de korstabeller vars frågor har fler än två alternativ dikotomiseras eller tudelas till fyrfältstabeller. Nedan visas en schematisk beskrivning av hur en sådan tudelad tabell ser ut samt hur avvikelsemått framräknas:

Ordinarie intervju	Återintervju		
	Ja	Nej (eller motsv)	
Ja	a	b	a+b
Nej	c	d	c+d
	a+c	b+d	n

$$g = \frac{b+c}{n} * 100 = \text{procentuell avvikelse}$$

$$M = (b+c) / \left[\frac{(a+b)(b+d)}{n} + \frac{(c+d)(a+c)}{n} \right] = \frac{\text{Observerad avvikelse}}{\text{Förväntad avvikelse givet marginalfördelningarna och oberoende}}$$

Höga värden på g och M visar att svarsvariabiliteten är hög och kan vara en indikation på att frågan inte helt fungerar utan bör granskas mer ingående. Index M är, under vissa förutsättningar om marginalfördelningarna, identiskt med det amerikanska Index of inconsistency, som bl a används för att utvärdera kvalitén i de amerikanska arbetskraftsundersökningarna. Värden på detta index som är mindre än 0,2 betraktas som låga och värden större än 0,5 betraktas som höga. (Se Forsman 1987 och Wärnerud, 1990). Vi använder dessa gränser som riktlinjer i vår analys. I bilaga 1 återfinns dikotomiserade tabeller över centrala variabler med beräkningar av avvikelsemått g och M.

3 Genomförande

Under perioden augusti 1989 t o m februari 1990 genomfördes en återintervjustudie av Arbetskraftsundersökningarna (AKU) med syftet att se om skattningskillnader förelåg mellan resultat baserade på datorstödda intervjuer (DATI) och resultat baserade på intervjuer utförda med pappersblankett (PAPI). Urvalet till återintervjuerna AKU i februari uppgick till 897 personer när bortfallet vid det ordinarie tillfället frånräknats.

Urvalet stratifierades efter arbetskraftsstatus och anknytningsvariabelns tre huvudkategorier FAST, LÖS respektive EJ ANKNYTNING till arbetsmarknaden. Urvalsdesignen syftade till att uppnå maximal precision vid skattningen av mätfelsbiasen för ordinarie AKUs sysselsättnings- och arbetslöshetsvariabler samt anknytningsvariablerna.

Återintervjuerna gjordes av 15 intervjuare som genomgick en särskild utbildning med tonvikt på kunskap om de definitioner i AKU som är avgörande för anknytningsgrad och status. Utförlig beskrivning av denna studie finns i Rapport från DATI-projektets produktionsprov, 1990. Definitioner och begrepp som används i AKU finns beskrivna i Bakgrundsfakta till AKU 1990:1

Detta tillfälle utnyttjades till att genomföra återintervjuer på undersökningen av årsarbetskraften, trots att urvalsdesignen inte passade helt.

Urvalet till återintervjustudien Årsarbetskraften utgjordes av de 897 personerna. Bortfallet bestod av personer som var bortfall i minst en av ordinarie Årsarbetskraften och återintervjun av Årsarbetskraften. Det uppgick till 75 personer. För personer som är långvarigt sjuka, intagna för anstaltsvård eller arbetar utomlands görs ingen intervju i Årsarbetskraften. Denna grupp (grad EX) approximeras till att inte tillhöra arbetskraften under hela det föregående året (= mätperioden). Antalet som tillhörde kategorin grad EX var i februari 22 personer. Det var alltså 800 personer som intervjuades både i ordinarie Årsarbetskraften och i återintervjun av densamma.

Tidsavståndet mellan ordinarie intervjun och återintervjun i februari var i genomsnitt 16 dagar.

SCB har sedan många år tillbaka, på uppdragsbasis, ett antal tilläggsfrågor av varierande innehåll och omfattning, som integreras i undersökningen av årsarbetskraften. Dessa har inte ingått i återintervjun, vilket är anledningen till att frågenumereringen inte är kontinuerlig.

Den särskilda utbildningen intervjuarna genomgick före återintervjuerna gällde inte undersökningen av årsarbetskraften, utan intervjuarna hade samma instruktioner till tilläggsfrågorna vid bägge tillfällena. Kodning och övrig databearbetning genomfördes så långt det var möjligt på samma sätt vid återintervjun som vid den ordinarie intervjun.

Variabiliteten av näringsgren och yrke för de sysselsattas huvudsakliga arbete behandlas inte i rapporten, eftersom variationerna för dessa SNI- och NYK-klassificeringar endast till en del kan hänföras till intervjuperson eller intervjuare. Vår bedömning är att variabiliteten för dessa SNI- och NYK-klassificeringar till större delen i stället kan hänföras dels till den manuella kodningen och dels till den maskinella överföringen från ordinarie AKU.

Ingen uppräknings till befolkningstotaler har gjorts på återintervjuerna av Årsarbetskraften utan endast frekvenserna behandlas.

4 Resultat

Frågorna T1, T33A/T33B, T34 samt summa T1+T34 är centrala frågor i undersökningen och bestämmer om up

- varit sysselsatt någon gång under året ($T1 > 0$),
- varit arbetslös någon gång under året ($T34 > 0$),
- tillhört helårsarbetskraften ($50 \text{ veckor} \leq T1 + T34 \leq 52 \text{ veckor}$)
- tillhört delårsarbetskraften ($0 < T1 + T34 \leq 49 \text{ veckor}$) eller
- varit utanför arbetskraften under hela det föregående året. ($T1 = 0$ och $T33A = 2$)

I bilaga 1 (övre delen av sid 1) finns en sammanställning av dessa centrala variabler. Av avvikelsemått att döma kan svarsvariabiliteten anses godtagbar för alla utom för variabeln "arbetslös någon gång under året", där måttet M är något för högt.

Definitioner av de variabler och begrepp som behandlas, återfinns i "Årsarbetskraften 1989" i serien Statistiska meddelanden.

Frågorna T1, T24 och T33A/T33B är s k filterfrågor, dvs svaret på dessa styr vilka frågor i blanketten intervjupersonen fortsättningsvis ska ha. Det är därför motiverat med en fördjupad analys av dessa frågor.

Fråga T1 Hur många veckor hade du jobbförvärvsarbete 1989 (på hel eller deltid)? Du ska även inräkna semester, sjukledighet, tjänstledighet samt eventuell permittering.

Antal arbetade veckor enligt ordinarie intervjun	Antal arbetade veckor enligt återintervjun							Summa	Avvikelse procent
	0	1-4	5-8	9-13	14-26	27-51	52		
0	29	3	-	-	1	2	1	36	19
1-4	3	12	-	-	2	-	-	17	29
5-8	1	2	14	6	-	2	-	25	44
9-13	-	1	-	8	2	1	-	12	36
14-26	1	-	-	-	12	11	3	27	56
27-51	-	-	-	-	6	27	8	41	34
52	1	-	1	1	1	20	618	642	4
Summa	35	18	15	15	24	63	630	800	
Avvikelse procent		14	33	7	50	50	57	2	

Vi har valt den intervallindelning, som AKU använder i sin resultatredovisning. I 90 procent av intervjuerna, dvs för 720 av 800 intervjuer överensstämmer svaren i återintervjun med den ordinarie intervjun. 10 procent av svaren finns alltså utanför huvuddiagonalen.

Avvikelsprocenten är överlag stor bland dem som är sysselsatta under en del av året. Största avvikelserna, mätt i antalet veckor, har personer i de skuggade fälten. Här avviker svaren från de båda tillfällena med minst ett halvår, vilket kan vara svårt att förklara med glömska. Snarare är frågekonstruktionen och situationen vid intervjutillfället anledningen till den stora spridningen.

618 personer uppgav i bägge intervjuerna att de arbetade hela året, dvs 52 veckor. Avvikelsen från detta i återintervjun var endast 4 procent. 29 personer uppgav vid bägge tillfällena att de inte hade arbetat något alls.

7 personer ändrade sitt svar vid återintervjun från att de inte arbetat alls till att de hade arbetat en eller flera veckor under året. Detta kan bero på att man har varit frånvarande under året och därför spontant svarat att man inte alls förvärvsarbetat. Att 3 personer av de 7 vid återintervjun svarade att de arbetade 1-4 veckor, kan bero på exempelvis feriearbete som glömdes bort vid den första intervjun. De återstående 4 personernas avvikelser kan troligen hänföras till att de inte förstått frågan, eller att intervjuaren varit otydlig.

6 personer ändrade sitt svar från att de hade arbetat till att de inte hade arbetat under året. Med ett liknande resonemang som ovan skulle 4 eller 5 av dessa kunna förklaras av glömska medan 2 eller 1 skulle kunna förklaras av missförstånd eller otydlighet.

För personer som enligt ordinarie intervjun arbetade 1-51 veckor kan avvikande svar i återintervjustudien till stor del troligtvis förklaras av glömskeeffekter. Det är betydligt svårare att kunna ge ett korrekt svar på vad som har hänt del av året. Ser man till på den gruppen som i ordinarie intervjun uppgav att de arbetade 14-26 veckor är avvikelseprocenten så hög som 56 procent.

Det tycks vara så att man överhuvudtaget är osäker på antalet arbetade veckor om man arbetar mindre än 51 veckor. Eftersom vi inte accepterar partiellt bortfall på den här frågan vet vi, bl a genom kontakter med intervjuerna, att man mer eller mindre tvingar fram ett svar i de fall som uppgav att man inte minns.

Beräkning av index M i bilaga 1 (överst på sid 1) ger ett lågt värde, vilket skulle kunna tolkas som att svarsavvikelserna för T1 = 0 är acceptabla. Vi antar dock att denna enkla analys inte är tillräcklig för de sk filterfrågorna, eftersom hänsyn inte tas till vad ett eventuellt felsvar leder till för efterföljande frågor.

I bilaga 1 (sid 1 nedre delen) har korstabellen för de sysselsatta dikotomiserats efter de olika intervallen och den procentuella svarsavvikelsen g och index M beräknats för varje sådan tabell. T1=1-4 och T1=52 har låga värden på M, vilket kan tolkas så att svarsavvikelserna för dessa svar är acceptabla. I intervallen 14-26 veckor och 27-51 veckor har höga värden på M, där alltså intervjupersonen är vag i sin angivelse av antal arbetade veckor.

Rekommendation:

Ett sätt att dämpa svarsvariationerna på grund av glömskeeffekter på fråga T1 skulle kunna vara att vi använder oss av AKUs tablåmodell när man undersöker hur up arbetat under senaste året. 3 månader åt gången kartläggs hur up arbetat och man försöker få up att minnas hur det var under just dessa månader. Denna metod blir fullt möjlig att använda sig av när DATI är helt genomförd. I PAPI är tilläggsfrågorna med alla hopp och härledningar redan mycket komplicerade när man skall kartlägga ett förflutet år.

Det finns indikationer på att frågan missförstås. Eftersom svaret i T1 är avgörande för vilka ytterligare frågor som kommer att ställas, är det därför väsentligt att frågeformuleringen noggrant genomgås. Frågan är lång och jobbig att läsa upp och att lyssna till. Samtliga begrepp i den andra meningen skulle kunna misstolkas utifrån AKUs definitioner. Ordet "permittering" kan i vardagsspråk betyda "arbetslös" vilket i så fall ger ett helt missvisande svar för både fråga T1 och T34.

Det skulle emellertid leda alltför långt att i denna rapport utreda hur frågekonstruktionen i stället skulle kunna se ut. En idé är att i ett sammanhang utreda både antalet sysselsatta veckor och antalet arbetslösa veckor innan frågor som leder till näringsgrens- och yrkesklassificering samt om man bytt /sökt (annat) arbete ställs.

Fråga T2 Hur många timmar arbetade du i genomsnitt per vecka när du hade förvärvsarbete 1989?

Antal arbetade timmar i genomsnitt enligt ordinarie intervjun	Antal arbetade timmar i genomsnitt enligt återintervjun					Procentuell avvikelse
	1-19	20-29	30-34	35-	Totalt	
1-19	19	4	1	2	26	27
20-29	9	75	12	12	108	31
30-34	1	9	55	19	84	35
35-	6	4	9	521	540	4
Summa	35	92	77	554	758	
Procentuell avvikelse	46	18	29	6		

12 procent av svaren finns utanför huvuddiagonalen.

Frågan är lätt att besvara för personer som arbetat heltid hela året. Även vid deltid med samma antal veckotimmar under hela året är frågan antagligen enkel att besvara. För dem som haft varierande arbetstider är det troligen svårt att komma ihåg hur man faktiskt arbetat och det kan dessutom vara svårt för intervjuaren att beräkna rätt genomsnitt.

Ett problem med den här frågan är när up har haft en bisyssla under året. Vi vet av intervjuarnas frågor att det finns en stor osäkerhet hur man skall räkna arbetstiden genomsnittligt när den avviker från de mer lätthanterliga mönstren. Är det genomsnittlig arbetstid i det huvudsakliga arbetet som vi noterat i fråga T3 som avses? Eller är det genomsnittligt arbete totalt (inklusive bisyssla) som ska anges? Även inom AKU-funktionen har diskussioner förts om vad vi skall knyta den genomsnittliga arbetstiden till, eftersom Årsarbetskraften kartlägger det längsta (huvudsakliga) arbetet i frågorna T3–T7.

Enligt bilaga 1 (sid 1) är svarsavvikelserna "acceptabla" för alternativet 35 timmar eller mer. Övriga alternativ har inte helt acceptabla svarsavvikelser men dessa är dock inte alarmerande.

Rekommendation:

Förtydliga frågan och klargör bättre för intervjuarna vad frågan skall ge svar på. Ge exempel på hur genomsnittlig arbetstid ska beräknas vid varierande arbetstid och flera arbeten.

Fråga T3 Hos vem hade du ditt längsta (huvudsakliga) arbete 1989?

Här är svarsavvikelsen relativt liten. Endast 4 procent av svaren ligger utanför huvuddiagonalen. Det vi vill ha svar på i denna fråga är den arbetsgivare hos vilken man arbetat flest timmar under året. Frågan avgör om näringsgrenskodning skall göras eller om denna uppgift kan överföras maskinellt från ordinarie AKU.

Fråga T6 Vilka var dina huvudsakliga arbetsuppgifter?

6 procent av svaren ligger utanför huvuddiagonalen. Frågan om arbetsuppgifter är till hjälp vid yrkeskodningen. Att definiera vad som är arbetsuppgift är svårt och lämnar därmed utrymme för subjektiva bedömningar.

Fråga T11 Bytte du arbetsgivare (företag) någon gång under 1989?

Drygt 4 procent av svaren finns utanför huvuddiagonalen och index M är lågt dvs svarsavvikelseerna ligger på en acceptabel nivå. Återintervjuerna fångar upp flera bytare.

Ett problem med svaren på frågan är att när man bytt arbetsgivare inom delårsarbetskraften, vet man inte om intervjupersonen gått direkt mellan arbetena eller om en period förflutit däremellan. Ett annat problem som framkommit vid kontakter med intervjuarna är när up har haft en bisyssla som övergått i huvudsyssla samtidigt som up har slutat sin ordinarie huvudsyssla. Här är det svårt att svara om "byte" skett.

Fråga T21 Har du någon gång under 1989, samtidigt som du haft arbete, sökt arbete i annat företag?

Nära 10 procent av svaren avviker från huvuddiagonalen. Avvikelsemåttet i bilaga 1 är större än för frågan om man bytt arbete och ligger utanför gränsen för vad som kan anses acceptabelt enligt huvudregeln. De är dock inte alarmerande höga. Även i denna fråga fångas flera ombytessökande upp vid återintervjun än vid ordinarie intervjun.

Tänkbara orsaker till svarsavvikelseerna är glömska eller att man inte förstår innebörden av att "söka arbete".

Rekommendation:

Att göra en djupare analys i kombination med frågan om man bytt arbetsgivare. Ofta förekommer det att man svarat Ja på T11 och samtidigt Nej på T21.

Fråga T23 På vilket /vilka sätt sökte du annat arbete?

Enligt bilaga 1 är avvikelsemåttet för höga för att man skall anse svarsavvikelseerna acceptabla. Intervjupersonen kan ange ett eller flera av alternativen arbetsförmedling, arbetsgivare, annons och övrigt som svar. Frågans avsikt tycks vara oklar.

Rekommendation:

Syftet med frågan bör utredas och förtydligas.

Fråga T24 Arbetade du någon gång under 1989 som anställd?

T24 är en filterfråga för att sortera anställda vidare till frågor som klargör om de bytt plats/befattning eller sökt annan plats/befattning inom företaget de arbetat hos.

Enligt bilaga 1 ligger 2 procent av svaren utanför huvuddiagonalen och index M är inte anmärkningsvärt högt.

Fråga T25 Bytte du någon gång under 1989, för minst 6 månader, plats eller befattning inom det företag du arbetade i?

6 procent av svaren ligger utanför huvuddiagonalen och index M är högt. Anledningen till de stora svarsavvikelseerna är troligen avgränsningssvårigheter av begreppet plats/befattning och av företagsbegreppet.

Fråga T31 Hade du sökt den platsen eller befattningen som du senast bytte till?

Fråga T32 Sökte du någon gång annan plats eller befattning inom det företag du arbetade i?

Det interna ombytessökandet bestäms av svar Ja i T31 eller T32. 6 procent av svaren ligger utanför huvuddiagonalen och index M enligt bilaga 1 (sid 2) är högt, dock något lägre än i fråga T25. Anledningen till de stora svarsavvikelseerna är troligen desamma som i T25.

Rekommendation:

Våra statistikanvändare har visat stort intresse för uppgifter om byten av plats/befattning och ombytessökandet inom företagen. Vi bör därför utreda hur frågorna T25, T31 och T32 skulle kunna förbättras.

Fråga T33A/T33B

Frågorna T33A och T33B är filterfrågor, vilka ska ge svar på om intervjupersonen var arbetssökande under föregående år.

Om intervjupersonen, enligt fråga T1, har arbetat 52 veckor avslutas intervjun före dessa frågor. Om intervjupersonen svarat 0 veckor i fråga T1 ställdes frågan *Sökte du arbete under 1989?*

Till de intervjupersoner som svarat att de arbetat 1-51 veckor i fråga T1 ställdes frågan *Du hade alltså jobb ca (antal enligt fråga T1) veckor. Var du arbetssökande någon av de övriga veckorna (dvs 52 minus antal veckor enligt T1)?*

Om intervjupersonen svarar Nej, avslutas intervjun. Om svaret är Ja i antingen T33A eller T33B, ställs frågan hur många veckor man var arbetssökande i fråga T34.

Alla kombinationer av svar vid de båda intervjutillfällena ges av nedanstående tablå:

Ordinarie intervju	Återintervju				T1 = 52	Summa
	T1 = 0		T1 = 1-51			
	1/T33A	2/T33A	1/T33B	2/T33B		
T1=0						
1/T33A	5	1	1	1	-	8
2/T33A	-	23	-	4	1	28
T1=1-51						
1/T33B	1	-	22	10	5	38
2/T33B	-	4	7	67	6	84
T1 = 52	-	1	10	13	618	642
Summa	6	29	40	95	630	800

De svarsavvikelser som finns kan till stor del härledas från svaret i T1. Därför är metoden att titta enbart på korstabeller inte fullt tillräcklig. Vi kan ändå konstatera att gruppen T1 = 1-51 vid båda intervjutillfällena har större svarsavvikelse än gruppen som svarat T1 = 0. (Se bilaga 1, sid 2.)

Rekommendation:

Utred om det är bättre att kartlägga arbetssökande i samma sammanhang som antalet sysselsatta veckor i T1.

*Fråga T34**Hur många veckor var du arbetssökande 1989?*

Antal arbets sökande veckor enligt ordinarie intervjun	Antal arbetssökandeveckor enligt återintervjun					Avvikelse- procent
	1-4	5-13	14-26	27- 52	Totalt	
1-4	5	1	1	1	8	38
5-13	3	4		1	8	50
14-26			5	1	6	20
27-52			3	4	7	43
Summa	8	5	9	7	29	
Avvikelse- procent	38	20	80	43		

38 procent av svaren återfinns utanför huvuddiagonalen med ovanstående intervallindelning. Avvikelsemåttet i bilaga 1 (sid 3) är i det närmaste höga. Beräkningarna grundar sig dock på ett relativt litet material. Därför kan inte alltför långtgående slutsatser dras.

Som i fråga T1 är det svårt att förklara stora skillnader i angivna antal arbetssökande-veckor med glömska (se skuggade fälten i tabellen.) Dessa stora avvikelser ger oss anledning att kritiskt granska frågeformuleringen. Förstår folk innebörden av att söka arbete? Blandar folk ihop ombytessökande med arbetssökande?

Rekommendation:

Ordet permittering i fråga T1 kan vara en felkälla här och bör bytas ut. Man bör avväga om antal veckor man varit arbetssökande i stället ska förläggas i direkt anknytning till T1 och i så fall bearbeta frågeformuleringen.

Fråga T35

*Avser de..... (Antal enligt T34) veckorna som du var
arbetssökande en enda period?*

Ett smärre blankettfel samt att materialet är litet gör att det inte är meningsfullt att analysera svarsavvikelserna.

Fråga T36

På vilket sätt sökte du då arbete senaste gången?

Frågan innehåller samma svarsalternativ som i fråga T23. Materialet är för litet för ananlys.

”Fråga” T37 ställs inte vid intervjun utan utgör summan av frågorna T1 och T34, dvs ger svar på hur många veckor up befunnit sig i arbetskraften under året. Den används som en bakgrundsvariabel i Årsarbetskraftens resultatredovisning. Med samma indelning i veckointervall som i Årsarbetskraftens råtabeller ges summan av intervjupersonernas svar i T1 och T34 i nedanstående tablå.

”Fråga” 37 Antal veckor i årsarbetskraften

Antal veckor enligt ordinarie intervjun	Antal veckor enligt återintervjun								Avvikelse- procent
	1-4	5-13	14-26	27-39	40-49	1-49 (Delårs- arbets- kraften)	50-52 (Helårs- arbets- kraften)	Summa	
1-4	11	-	4	-	-	15	-	15	27
5-13	3	26	1	1	1	34	1	35	26
14-26	2	1	10	9	1	21	3	24	58
27-39	-	-	2	11	1	14	2	16	31
40-49	-	-	-	3	4	7	6	13	69
1-49 (Delårs- arbetskraften)	16	27	17	24	7	91	12	103	12
50-52 (Helårs- arbetskraften)	1	2	2	4	9	18	645	663	3
Summa	17	29	19	28	16	109	657	766	
Avvikelse procent	35	10	47	61	75	17	2		

Intervallindelningen av antal veckor är densamma som den vi använder i Årsarbetskraftens råtabeller. Med denna intervallindelning avviker 8 procent av svaren från huvuddiagonalen. 3 procent i helårsarbetskraften och 12 procent i delårsarbetskraften har svarat annorlunda i återintervjun.

Index M i bilaga 1 (övre delen sid 1) är lågt för både helårsarbetskraften och delårsarbetskraften, så svarsavvikelsena kan enligt vår metod anses vara acceptabla om vi använder denna grova indelning. Delårsarbetskraften ger däremot höga värden på index M för intervallen 14-26 veckor, 27-39 veckor och 40-49 veckor bilaga 1 (sid 3).

Rekommendation:

Se frågorna T1 och T34.

Variabel	Tabell	Avvikelsemått	Variabel	Tabell	Avvikelsemått
T1	Återintervju		T34	Återintervju	
Ordinarie intervju	1 0		Ordinarie intervju	1 0	
Antal syssel-	1 758 6 764	g=1,6	Antal arbets-	1 29 17 46	g=4,2
satta 1-52			sökande1-52		
veckor	0 7 29 36	M=0,19	veckor	0 17 737 754	M=0,39
	765 35 800			46 754 800	
T37	Återintervju		T37	Återintervju	
Ordinarie intervju	1 0		Ordinarie intervju	1 0	
Antal i års-	1 766 6 772	g=1,4	Antal i helårs-	1 645 19 664	g=4,0
arbetskraften			arbetskraften		
(hel+del)	0 5 23 28	M=0,20	(50-52 veckor)	0 13 123 136	M=0,14
	771 29 800			658 142 800	
T37	Återintervju		T1, T33A	Återintervju	
Ordinarie intervju	1 0		Ordinarie intervju	1 0	
Antal i delårs-	1 91 17 108	g=4,9	Antal ej i	1 23 5 28	g=1,4
arbetskraften			arbetskraften		
(1-49 veckor)	0 22 670 692	M=0,20	någon gång	0 6 766 772	M=0,20
	113 687 800		under året	29 771 800	

T1	Återintervju		T1	Återintervju	
Ordinarie intervju	1 0		Ordinarie intervju	1 0	
Antal syssel-	1 12 2 14	g=0,7	Antal syssel-	1 14 10 24	g=1,5
satta veckor			satta veckor		
1-4	0 3 741 744	M=0,18	5-8	0 1 733 734	M=0,29
	15 743 758			15 743 758	
T1	Återintervju		T1	Återintervju	
Ordinarie intervju	1 0		Ordinarie intervju	1 0	
Antal syssel-	1 8 4 12	g=1,5	Antal syssel-	1 12 14 26	g=3,3
satta veckor			satta veckor		
9-13	0 7 739 746	M=0,41	14-26	0 11 721 732	M=0,53
	15 743 758			23 735 758	
T1	Återintervju		T1	Återintervju	
Ordinarie intervju	1 0		Ordinarie intervju	1 0	
Antal syssel	1 27 14 41	g=6,3	Antal syssel-	1 618 23 641	g=4,5
satta veckor			satta veckor		
27-51	0 34 683 717	M=0,50	52	0 11 106 117	M=0,16
	61 697 758			629 129 758	
T2	Återintervju		T2	Återintervju	
Ordinarie intervju	1 0		Ordinarie intervju	1 0	
Genomsnittligt	1 19 7 26	g=3,0	Genomsnittligt	1 75 33 108	g=6,6
antal arbetade			antal arbetade		
timmar per	0 16 716 732	M=0,39	timmar per	0 17 633 650	M=0,29
vecka 1-19	35 723 758		vecka 20-29	92 666 758	
T2	Återintervju		T2	Återintervju	
Ordinarie intervju	1 0		Ordinarie intervju	1 0	
Genomsnittligt	1 55 29 84	g=6,7	Genomsnittligt	1 521 33 554	g=6,9
antal arbetade			antal arbetade		
timmar per	0 22 652 674	M=0,35	timmar per	0 19 185 204	M=0,17
vecka 30-34	77 681 758		vecka 35-	540 218 758	

T3		Återintervju			
Ordinarie intervju		1	0		
Längsta	1	614	22	636	g=4,2
(huvudsakliga)					
arbetet	0	10	112	122	M=0,15
		624	134	758	

T5		Återintervju			
Ordinarie intervju		1	0		
Huvudsakliga	1	606	31	637	g=5,9
arbetsuppgifter					
	0	14	107	121	M=0,21
		620	138	758	

T11		Återintervju			
Ordinarie intervju		1	0		
Byte av arbets-	1	82	9	91	g=4,5
givare (företag)					
	0	25	642	667	M=0,20
		107	651	758	

T21		Återintervju			
Ordinarie intervju		1	0		
Ombytes-	1	106	25	131	g=9,5
sökande					
	0	47	580	627	M=0,31
		153	605	758	

T23		Återintervju			
Ordinarie intervju		1	0		
Söksätt	1	10	3	13	g=9,4
Arbetsför-					
medling	0	7	86	93	M=0,39
		17	89	106	

T23		Återintervju			
Ordinarie intervju		1	0		
Söksätt	1	25	10	35	g=14,2
Arbets-					
givare	0	5	66	71	M=0,33
		30	76	106	

T23		Återintervju			
Ordinarie intervju		1	0		
Söksätt	1	38	6	44	g=14,1
Annonns					
	0	9	53	62	M=0,29
		47	59	106	

T24		Återintervju			
Ordinarie intervju		1	0		
Anställd	1	709	5	714	g=2,2
	0	12	32	44	M=0,22
		721	37	758	

T25		Återintervju			
Ordinarie intervju		1	0		
Byte av plats/	1	13	25	38	g=5,6
befattning inom					
företaget	0	15	656	671	M=0,63
		28	681	709	

T31, T32		Återintervju			
Ordinarie intervju		1	0		
	1	19	25	44	g=5,6
Internt omby-					
tessökande	0	15	650	665	M=0,54
		34	675	709	

T33A		Återintervju			
Ordinarie intervju		1	0		
Antal arbets-	1	5	1	6	g=3,4
sökande personer					
då T1=0	0	0	23	23	M=0,11
		5	24	29	

T33B		Återintervju			
Ordinarie intervju		1	0		
Antal arbets-	1	22	10	32	g=16,0
sökandepersoner					
då T1 = 1-51	0	7	67	74	M=0,39
		29	77	106	

T33A eller T33B		Återintervju			
Ordinarie intervju		1	0		
Antal arbets-	1	29	12	41	g=13,0
sökande personer					
	0	7	98	105	M=0,33
		36	110	146	

T34		Återintervju			
Ordinarie intervju		1	0		
Antal arbets-sökandeveckor	1	5	3	8	$g=20,7$
1-4	0	3	18	21	$M=0,52$
		8	21	29	

T34		Återintervju			
Ordinarie intervju		1	0		
Antal arbets-sökandeveckor	1	4	4	8	$g=17,2$
5-13	0	1	20	21	$M=0,49$
		5	24	29	

T34		Återintervju			
Ordinarie intervju		1	0		
Antal arbets-sökandeveckor	1	5	1	6	$g=17,2$
14-26	0	4	19	23	$M=0,44$
		9	20	29	

T34		Återintervju			
Ordinarie intervju		1	0		
Antal arbets-sökandeveckor	1	4	3	7	$g=20,7$
27-52	0	3	19	22	$M=0,56$
		7	22	29	

$0 < T37 \leq 52$		Återintervju			
Ordinarie intervju		1	0		
Antal i års-arbetskraften	1	11	4	15	$g=1,3$
1-4	0	6	745	751	$M=0,32$
		17	749	766	

$0 < T37 \leq 52$		Återintervju			
Ordinarie intervju		1	0		
Antal i delårs-arbetskraften	1	13	10	23	$g=1,4$
5-8	0	1	742	743	$M=0,30$
		14	752	766	

$0 < T37 \leq 52$		Återintervju			
Ordinarie intervju		1	0		
Antal i års-arbetskraften	1	8	4	12	$g=1,4$
9-13	0	7	747	754	$M=0,41$
		15	751	766	

$0 < T37 \leq 52$		Återintervju			
Ordinarie intervju		1	0		
Antal i års-arbetskraften	1	10	14	24	$g=3,0$
14-26	0	9	733	742	$M=0,55$
		19	747	766	

$0 < T37 \leq 52$		Återintervju			
Ordinarie intervju		1	0		
Antal i års-arbetskraften	1	11	5	16	$g=2,9$
27-39	0	17	733	750	$M=0,51$
		28	738	766	

$0 < T37 \leq 52$		Återintervju			
Ordinarie intervju		1	0		
Antal i års-arbetskraften	1	4	9	13	$g=2,7$
40-49	0	12	741	753	$M=0,74$
		16	750	766	

$0 < T37 \leq 52$		Återintervju			
Ordinarie intervju		1	0		
Antal i års-arbetskraften	1	645	18	663	$g=3,9$
50-52	0	12	91	103	$M=0,16$
		657	109	766	

R & D Reports är en för U/ADB och U/STM gemensam publikationsserie, som fr o m 1988-01-01 ersätter de tidigare "gula" och "gröna" serierna. I serien ingår även **Abstracts** (sammanfattning av metodrapporter från SCB).

R & D Reports Statistics Sweden are published by the Department of Research & Development within Statistics Sweden. Reports dealing with statistical methods have green (grön) covers. Reports dealing with EDP methods have yellow (gul) covers. In addition, abstracts are published three times a year (light brown /beige/ covers).

Reports published during 1991:

- | | |
|-------------------|--|
| 1991:1
(grön) | Computing elementary aggregates in the Swedish consumer price index
(Jörgen Dalén) |
| 1991:2
(grön) | Översikt av estimatorer för stora och små redovisningsgrupper (Sixten Lundström) |
| 1991:3
(grön) | Interacting nonresponse and response errors (Håkan L Lindström) |
| 1991:4
(grön) | Effects of nonresponse on survey estimates in the analysis of competing exponential risks (Ingrid Lyberg) |
| 1991:5
(grön) | Study of the estimation precision in the Chinese MIS surveys
(Bengt Rosén) |
| 1991:6
(beige) | Abstracts I - Sammanfattning av metodrapporter från SCB |
| 1991:7
(grön) | Kvalitetsfonden 1990 Projekt - handläggning - utvärdering
(Roland Friberg och Håkan L Lindström) |

Kvarvarande **beige** och **gröna** exemplar av ovanstående promemorior kan rekvireras från Inga-Lill Pettersson, U/LEDN, SCB, 115 81 STOCKHOLM, eller per telefon 08-783 49 56.

Kvarvarande **gula** exemplar kan rekvireras från Ingvar Andersson, U/ADB, SCB, 115 81 STOCKHOLM, eller per telefon 08-783 41 47.