



Översyn av undersökningen

Inrikes och utrikes trafik med svenska lastbilar

**Johan Eriksson
Per Anders Paulson
Bengt Rosén**

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 2003:1. Översyn av undersökningen Inrikes och utrikes trafik med svenska lastbilar / Johan Eriksson m.fl.

Digitalt skapad fil, anpassad efter de digitaliserade delarna i serien. Statistiska centralbyrån (SCB) 2016.

urn:nbn:se:scb-2003-X101OP0301

Översyn av undersökningen

**Inrikes och utrikes trafik med
svenska lastbilar**

Johan Eriksson

Per Anders Paulson

Bengt Rosén

R&D Report 2003:1

Research - Methods - Development

Evaluation of the survey Swedish national and international road goods transport

Från trycket

Maj 2003

Producent

Statistiska centralbyrån, *Statistics Sweden*, metodenheten
Box 24300, SE-104 51 STOCKHOLM

Förfrågningar

Bengt Rosén

bengt.rosen@math.uu.se

Telefon 018 - 471 57 32

Johan Eriksson

johan.eriksson@scb.se

Telefon 019 - 17 67 38

Per Anders Paulson

per.paulson@scb.se

Telefon 019 - 17 61 28

Abstract

Commissioned by the Swedish Institute for Transport and Communications Analysis (SIKA), the government agency responsible for official transport statistics, Statistics Sweden earlier carried out two separate sample surveys of goods transport with Swedish lorries, one for national and one for international transport. In year 2000 the two surveys were merged into one, "Swedish national and international road goods transports" (*Inrikes och utrikes trafik med svenska lastbilar*). Now, after a couple of rounds for the new survey, its design has been scrutinized for possible re-design. This report presents results from a study based on data and experiences from the 2000 and 2001 surveys.

A main goal was to enhance the accuracy of the statistics on international transports. The chief contribution to that effect was a so called "Blue card study". It concerned improvement of the information about "lorries likely to carry out international transports". The study comprised enterprises with more than 15 permits for international lorry transports, so called "blue cards". They were asked to give, for each one of their lorries, an estimate of how much the lorry will be used in international haulages during 2003. Chief findings from the study were: (i) This kind of information is expected to yield substantial accuracy gains for the statistics on international transports. (ii) The enterprises cooperated willingly to provide this non-compulsory extra information.

It was, of course, also desirable to improve the accuracy of the statistics on national transports. An issue of particular interest in that context was if Statistics Sweden's newly (= early in the year 2002) completed "Mileage database" (*Körsträckedatabasen*) can provide valuable auxiliary information. This turned out to be the case.

Based on findings from the Blue card study and the Mileage database a new design for stratification and sample allocation was worked out, and set into operation in the 2003 survey. Evaluations indicate that in an over-all perspective the new design will lead to considerable accuracy improvements of the statistics on international as well as national transports.

Innehållsförteckning

	Sidnr
0 Bakgrund för översynen	1
1 Undersökningens syften och uppläggning	1
1.1 Beskrivning i stort	1
1.2 Undersökningsvariabler	1
1.2.1 Sändning respektive körning	1
1.2.2 Observationsvariabler	2
1.2.3 Härledda variabler	2
1.3 Urvalsförfarande	2
1.3.1 Allmänt	2
1.3.2 Stratumindelning av urvalsramen under åren 2000 - 02	3
1.3.3 Urval och urvalsallokering	4
1.4 Skattningar	5
1.4.1 Skattning av kvartalsaggregat	5
1.4.2 Skattning av årsaggregat	6
2 Premisser för och metod aspekter på översynen	7
2.1 Allmänt om bakgrundspremisser	7
2.2 Mått på en urvalsutformnings effektivitet	8
2.3 Beräkning av skattningsprecision vid en hypotetisk urvalsutformning	8
3 Överväganden och insatser vid översynen av urvalsutformningen	10
3.1 Allmänt om undersökningens stratifiering	10
3.2 Blåkortundersökningen	11
3.3 Hjälpinformation från Körsträckedatabasen	11
3.3.1 Körsträckedatabasens täckning	12
3.3.2 Samvariation mellan årskörsträcka och undersökningsvariablerna	12
3.4 Om variabeln maxlastvikt	12
3.5 Om stratifiering med avseende på yrkesmässig / firma	13
4 Resultat av prövning av olika utformningsalternativ	13
4.1 Urvalsallokering	13
4.1.1 Urvalets fördelning på Inrikes - respektive Utrikesdel	13
4.1.2 Allokering av urvalet inom Inkesdelen	13
4.1.3 Allokering av urvalet inom Utrikesdelen	14
4.2 Stratifiering i Inrikesdelen	14
4.3 Stratifiering i Utrikesdelen	16
4.4 Förväntad skattningsprecision med den nya utformningen	17
4.4.1 Förväntade relativa felmarginaler	17
4.4.2 Förväntade relativa variansförändringar	17
Referenser	19
Bilaga 1. Enkät i blåkortundersökningen	20

0 Bakgrund för översynen

Med SIKÄ som statistikansvarig myndighet gjorde SCB:s Transportprogram tidigare två separata undersökningar av transporter med svenska lastbilar, de s.k. UVÄV - och SLUT - undersökningarna. Den förra mätte inrikes och den senare utrikes transporter. Från och med år 2000 har de två undersökningarna lagts samman till en, benämnd *Inrikes och utrikes trafik med svenska lastbilar*, vilken skall följa EU-förordningen 1172/98. En förekommande arbetsakronym för undersökningen är SLIT.

De noggrannhetskrav Eurostat ställer på årsstatistik avseende transporterad godsmängd (Ton) och transportarbete (TonKm) är $\pm 5\%$ för 95 - procentiga konfidensintervall för inrikestransporterna. För utrikestransporterna är målet satt till $\pm 10\%$. Den nuvarande redovisningen för åren 2000 och 2001 uppfyller kraven för inrikestransporterna, men noggrannheten i statistiken för utrikestrafiken ligger en bit utanför målgränserna. Det senare är en huvudledning till denna översyn, vars uppgift är att, på basis av erfarenheter och data från hittillsvarande undersökningar ge förslag till möjliga förbättringar av undersökningen.

1 Undersökningens syften och uppläggning

1.1 Beskrivning i stort

Undersökningen Inrikes och utrikes trafik med svenska lastbilar redovisar statistik om transporter som utförs av svenskregistrerade dragfordon med maxlastvikt om minst 3,5 ton. Statistiska storheter av huvudintresse är över kvartal och år aggregerade värden för transporterad godsmängd, transportarbete (mätt i tonkm) och transportsträckor. Statistiken redovisas med ett flertal indelningar (inklusive korsindelningar), varvid de viktigaste är efter inrikes resp. utrikes körning, varuslag (enligt NST/R, EU:s varuklassificering för transporter), på- och avlastställe, transportsträcka, fraktarens företagstyp (yrkesmässig eller firma), farligt gods (enligt ADR-klassificeringen) samt lasttyp (enligt FN:s rekommendationer).

Undersökningen utförs på urvalsbasis med separata undersökningar för varje kvartal. Urvalsobjekten är lastbilar / dragfordon, fortsättningsvis kallade *lastbilar*. Urval dras 1 à 2 månader före ett kvartal. En utvald lastbil tillordnas en mätvecka, och mätveckorna för utvalda lastbilar sprids jämnt över undersökningskvartalet. Lastbilar som ingår i undersökningen lämnar uppgifter på en blankett, vars viktigaste del utgörs av *sändningsjournalen*, där uppgifter lämnas om alla sändningar vars transport lastbilen påbörjar under sin mätvecka.

1.2 Undersökningsvariabler

Undersökningens observationsvariabler gäller egenskaper hos sändningar och körningar. Värden inhämtas via sändningsjournalerna. Dessutom används variabler med uppgifter om lastbilarna, vars värden hämtas från Vägverkets Bilregistret samt det s.k. Blåkort - registret. Det senare är ger sammanställning av företag med (av Länsstyrelserna) utfärdade tillstånd för yrkesmässiga utrikes transporter, s.k. *blåkort-tillstånd*.

1.2.1 Sändning respektive körning

Observationsvariablerna hänför sig till sändningar och körningar. En *sändning* är ett varuparti som transporteras från ett *pålastställe* till ett *avlastställe*. En *körning* avser en lastbils aktiviteter från ett första pålastställe till ett sista avlastställe där bilen blir tom på last, eller tvärtom dvs. från ett ställe med tom bil till ett första pålastställe. För en körning är transporterat gods oftast detsamma under hela körningen men viss variation kan förekomma genom att sändningar lastas på eller av i orter som passeras under körningen.

De uppgifter som inkommer på blanketterna gäller primärt variabelvärden för sändningar. Från dessa härleder SCB värden på variabler för körningar.

1.2.2 Observationsvariabler

De flesta variabler vars värden inhämtas via blanketten avser sändningar. De viktigaste är:

- Körsträcka mellan sändningens på- och avlastställen (KmKorda),
- Sändningens vikt (KgKvant, räknas ofta om till Ton = KgKvant/1000),
- Sändningens varuslag (VaruKod),
- Län/land för pålast (PaOmr),
- Län/land för avlast (AvOmr),
- Ort där pålast sker (PaOrt),
- Ort där avlast sker (AvOrt),
- Typ av farligt gods enligt ADR-klassning (ADRklass),
- Hur godset är lastat (LastTyp).

1.2.3 Härledda variabler

Transportarbete (TonKm) för en sändning beräknas som $\text{TonKm} = \text{Ton} \times \text{KmKorda}$.

Variablerna för körningar är i stort desamma som för sändningar. Deras värden framgår dock inte direkt ur körjournalerna, utan måste härledas. När en körning omfattar bara en sändning, vilket gäller i den stora majoriteten av fall, är körning och sändning ekvivalenta. Då gäller det enkla sambandet: "körningsvariabelvärde = sändningsvariabelvärde". Men när en körning omfattar flera sändningar blir härledningen av körningsvariablers värden från körjournalens sändningsvariabelvärden litet mer komplicerat. Som exempel gäller: (i) Värdet på körningsvariabeln godsmängd (KgKvant) beräknas som ett medelvärde (enligt en fastlagd "vägningsformel") av de ingående sändningarnas KgKvant. (ii) Körningens sträcka (KmKorda) beräknas från de uppgifter som lämnas i körjournalen. (iii) En körnings transportarbete (TonKm) beräknas som körningens Ton \times körningens KmKorda.

För körningar som omfattar flera sändningar kan värdena på klassificeringsvariabler som VaruKod, ADRklass, LastTyp, liksom geografiska angivelser avseende avlast- och pålastställen vara olika för de sändningar som körningen omfattar, varvid det inte är självklart hur körningsvariablers värden skall sättas utifrån motsvarande sändningsvariablers. Huvudprincipen är följande: Vid varierande sändningsvariabelvärden under en körning används principen att *sändningen med störst godsmängd styr körningsvariabelvärdet*.

Vid framtagande av statistiken är det främst körningar och deras variabelvärden som används.

1.3 Urvalsförfarande

1.3.1 Allmänt

En årsundersökning genomförs via fyra kvartalsundersökningar med separata urval. Urvalsramen omfattar lastbilar i Vägverkets bilregister, med maxlastvikt om minst 3,5 ton, som är högst 30 år gamla, och som är aktivregistrerade vid urvalstillfället. Före urvalsdragningen görs en omfattande stratumindelning av lastbilarna. Urvalstypen är stratifierat "disjunkt" OSU, där disjunkt står för att en bil bara kan komma med i en av årets fyra kvartalsundersökningar.

Det finns flera skäl till att stratifiering är fördelaktig för precisionen i statistiken från en urvalsundersökning. Ett är att ju mer homogena (med avseende på viktiga undersökningsvariabler) grupper/stratan man kan dela in populationen i, desto bättre precision i resulterande statistik. För att skapa homogena stratan behövs adekvat hjälpinformation (viken skall vara tillgänglig när urvalet dras).

En ytterligare regel för effektiv utformning av urvalsundersökningar är att objekt som ger stora bidrag till totaler och medelvärden av särskilt intresse bör "högrepresenteras", och objekt som ger mindre bidrag bör "lågrepresenteras". Med att "högrepresentera" objekt avses att de ges större urvalssannolikhet än genomsnittet, och analogt innebär "lågrepresentera" att objek-

ten ges lägre urvalssannolikhet än genomsnittet. Ett hjälpmedel för att kunna variera urvalssannolikheter på önskat sätt är adekvat stratifiering.

De förändringar som denna översyn föranlett, vilka beskrivs mer i detalj i Avsnitt 4, gäller framför allt urvalsramens stratifiering och urvalets allokering på stratan. För att ge bakgrund för förändringarna beskrivs i nästa avsnitt den undersökningsutformning som användes under åren 2000, 2001 och 2002, medan den nya undersökningsutformningen behandlas i Avsnitt 4. Beskrivningen ges i presensform, även om det till del är fråga om "historia", men mycket av "det gamla" lever vidare.

1.3.2 Stratumindelning av urvalsramen under åren 2000 - 2002

Under 2000-2002 gjordes stratumindelning baserad på följande stratifieringsvariabler:

- lastbilens maxlastvikt,
- lastbilens karosserikod,
- om lastbilens ägarföretag hade blåkort-tillstånd (för yrkesmässiga transporter utomlands),
- om lastbilen var registrerat för yrkesmässig trafik eller firmatrafik,
- länet där lastbilen är registrerad

Mer ingående motiveringar för valet av stratifieringsvariabler ges i Avsnitt 3. Specifikation av den använda stratifieringen ges av nedanstående stegvisa procedur.

Steg 1: Ramens lastbilar delas upp i en Utrikes- och en Inrikesdel, beroende på om bilens ägare har blåkort eller ej. Till Inrikesdelen förs alla bankebilar (karosserikod = 61, för lekmannen "timmerbilar") samt alla tankbilar (karosserikod = 40, 41, 42, 45) oavsett om ägarföretaget har blåkort eller ej.

Steg 2: Bilarna i Utrikesdelen vidareuppdelas efter antalet blåkort-tillstånd som lastbilens ägarföretag har. Bilar tillhörande ägare med mindre än 16 tillstånd förs till Storleksgrupp 1 (STG1) och bilar med ägare med 16 eller fler tillstånd till Storleksgrupp 2 (STG2).

Steg 3: Inrikesdelens banke- och tankbilar förs till särskilda stratan, som sedan delas upp efter om lastbilen körs i yrkesmässig trafik eller i firmatrafik.

Steg 4: I såväl Utrikesdel som Inrikesdel görs geografisk indelning med hjälp av bilens registreringslän. För lastbilar i Inrikesdelen efter län (21 stycken), och för bilar i Utrikesdelen efter 6 regioner skapade från länen på följande sätt. Region 1 – län 01, Region 2 – länen 03, 04, 05, 06, 07, 08, 09, 18 och 19, Region 3 – länen 10 och 12, Region 4 – länen 13 och 14, Region 5 – länen 17, 20 och 21, Region 6 – länen 22, 23, 24 och 25.

Steg 5: Stratumgrupperna Inrikes×geografi delas efter yrkesmässig resp. firma.

Steg 6: Stratumgrupperna Inrikes×geografi×yrkesmässig/firma delas efter maxlastvikt enligt Tabellerna 1 och 2 nedan.

Tabell 1. Indelning efter maxlastvikt (MAXLV) i storstadslänen (01, 12 och 14)

	<i>Yrkesmässiga bilar</i>	<i>Firmabilar</i>
LÄTTA lastbilar	3,5 ton ≤ MAXLV < 8,0 ton	3,5 ton ≤ MAXLV < 7,5 ton
MEDEL lastbilar	8,0 ton ≤ MAXLV < 13,0 ton	7,5 ton ≤ MAXLV < 12,5 ton
TUNGA lastbilar	13,0 ton ≤ MAXLV	12,5 ton ≤ MAXLV

Tabell 2. Indelning efter maxlastvikt (MAXLV) utanför storstads-länen (utom Gotland, där ingen indelning görs)

	<i>Yrkesmässiga bilar</i>	<i>Firmabilar</i>
LÄTTA lastbilar	3,5 ton ≤ MAXLV < 12,5 ton	3,5 ton ≤ MAXLV < 11,5 ton
TUNGA lastbilar	12,5 ton ≤ MAXLV	11,5 ton ≤ MAXLV

1.3.3 Urval och urvalsallokering

Urvalsobjekten är som sagt lastbilar. Urvalsförfarandet är, kvartalsvis sett, stratifierat OSU med följande "skyddsregler" för undvikande av stor uppgiftslämnarbörda. (i) En lastbil som utvalts ett visst kvartal kommer inte ifråga i något annat kvartal under samma år. (ii) En lastbil som utväljs till kvartal 4 kommer inte ifråga vid urvalet till kvartal 1 året efter.

SLIT:s undersökningspopulation utgörs dock inte av lastbilar, utan av sändningar och körningar. De grupperas i sin tur i lastbilsveckor som omfattar de transporter en lastbil påbörjar under en vecka. Urvalet av lastbilsveckor fås genom att utvalda lastbilar tilldelas mätveckor med slumpmässig, men (stratumvis) jämn, utspridning över undersökningskvartalet.

Vid urvalsallokeringen (= bestämning av urvalsstorlekar i stratana) är första steget att dela undersökningsårets totala urvalsstorlek på Inrikes- och Utrikesdel, InrUrv resp. UtrUrv. Detta görs med likafördelning: $\text{InrUrv} = \text{UtrUrv} = (\text{Total urvalsstorlek})/2$. Därefter fördelas urvalsstorlekarna InrUrv och UtrUrv vidare på urvalsstratana enligt vad som sägs nedan.

Urvalsallokering i Inrikesdelen under 2000- 2002

Urvalsallokeringen i Inrikesdelen görs genom att urvalsstorleken InrUrv först fördelas på följande sex stratumgrupper: 1) Yrkes - bankebilar (stratum 4410), 2) Firma - bankebilar (stratum 4400), 3) Yrkes - tankbilar (stratum 6610), 4) Firma - tankbilar (stratum 6600), 5) Övriga yrkesbilar (= lastbilar i yrkesmässig trafik, som ej är banke - eller tankbilar), 6) Övriga firmabilar (lastbilar i firmatrafik, som ej är banke - eller tankbilar). Fördelning görs enligt nedan.

Urvalsstorlek för stratum 4400:	$0,024 \times \text{InrUrv}$,
Urvalsstorlek för stratum 4410:	$0,059 \times \text{InrUrv}$,
Urvalsstorlek för stratum 6600:	$0,016 \times \text{InrUrv}$,
Urvalsstorlek för stratum 6610:	$0,045 \times \text{InrUrv}$.
Urvalsstorlek för Övriga yrkesbilar:	$0,556 \times \text{InrUrv}$,
Urvalsstorlek för Övriga firmabilar:	$0,300 \times \text{InrUrv}$,

Den för Övriga yrkesbilar bestämda urvalsstorleken fördelas på geografiska stratan med "rest-rifierad Neyman -allokering", med vissa stratumvisa variationsproportionaler som framräknats via skattningar av stratumvarianserna för variablerna TonKm och Ton i tidigare undersökningar. Med "restifierad" avses att Neyman -allokeringen görs med restriktionen att urvalsstorleken i varje stratum skall vara minst 10. Urvalsstorleken för Firmabilar fördelades analogt på de geografiska stratana. Urvalen för tank- och bankebilar allokeras till yrkesbilar respektive firmabilar proportionellt mot stratumstorlekarna.

Urvalsallokering i Utrikesdelen under åren 2000- 2002

De två storleksgrupperna i Utrikesdelen, STG1 och STG2, omfattar vardera sex geografiska stratan (svarande till de sex regionerna). Första steget i urvalsallokeringen vid ett visst undersökningskvartal gäller bestämning av urvalsstorlekarna för de sex stratana i STG2.

Urvalsstorlekar och urvalsdragning i STG2- stratana

Samtliga bilar i STG2 skall delta i undersökningen en gång per år. Riktmärket är att 1/4 av bilarna ingår varje kvartalsurval. Följande förfarande används, vilket dock ofta leder till en allokering som är något skev i förhållande till riktmärket "1/4 varje kvartal".

Urvalet till kvartal 1 utgörs av stratumvisa OSU som omfattar 1/3 av de lastbilar som inte var med i urvalet till kvartal 4 året innan. Urvalet till kvartal 2 utgörs av stratumvisa OSU som omfattar 1/3 av de lastbilar som inte var med i urvalet till kvartal 1. Urvalet till kvartal 3 utgörs av stratumvisa OSU som omfattar hälften av de lastbilar som inte var med i något av urvalen till kvartal 1 eller kvartal 2. Slutligen, urvalet till kvartal 4 utgörs av de lastbilar som inte var med i något av urvalen till kvartalen 1, 2 och 3.

Ovanstående allokeringregel medför att urvalsstorlekarna i vart och ett av de sex STG2-strata bestäms för varje undersökningskvartal, och därmed också den (för kvartalet) totala urvalsstorleken i STG2, $Stg2Urv = \text{summa urvalsstorlekar i de sex regionala STG2-strata}$.

Urvalsstorlekar och urvalsdragning i STG1 - stratana

Totala urvalsstorleken i STG1, $Stg1Urv$, beräknas som "efter STG2 resterande del av $UtrUrv$ ": $Stg1Urv$ (för kvartalet) = $UtrUrv$ (för året) / 4 - $Stg2Urv$ (för kvartalet). Den urvalsstorleken fördelas på de sex regionala stratana i STG1 med proportionell (mot stratumstorlek) allokering. Urvalen från STG1-strata dras sedan med OSU med användande av den tidigare nämnda "skyddsregeln". För alla bilar i Utrikesdelen gäller att de kan komma med i undersökningen högst en gång per år, och för STG2-bilar gäller att de kommer med precis en gång.

1.4 Skattningar

1.4.1 Skattning av kvartalsaggregat

Statistiska storheter

De statistiska storheter man främst önskar skatta är transportaggregat $\tau(\mathbf{x}; G)_{kv}$ enligt (1) nedan. I ord innebär $\tau(\mathbf{x}; G)_{kv}$ totalen för transportvariabeln \mathbf{x} över alla körningar under kvartal kv som ingår i redovisningsgruppen G av körningar.

$$\tau(\mathbf{x}; G)_{kv} = \sum_{\substack{j \text{ löper över alla körningar som} \\ \text{påbörjas under kvartal } kv}} \mathbf{x}_j \cdot \mathbf{1}(G)_j, \quad (1)$$

där

\mathbf{x} är någon av körningsvariablerna $KmKorda$, $KgKvant/Ton$ eller $TonKm$,

G är en redovisningsgrupp. (Exempel på G : "med visst angivet varuslag", "med yrkesmässiga lastbilar" och "med avlast i Tyskland".)

$\mathbf{1}(G)_j$ indikerar (med 1 respektive 0) huruvida körning j ingår i gruppen G eller ej.

Det urval som observeras är dock inte ett "direkt" urval av körningar, utan ett urval av lastbilsveckor, där en *lastbilsvecka* är ett *kluster av körningar*, alla körningar som lastbilen påbörjar under veckan ifråga. Genom att i (1) summera över kluster fås;

$$\tau(\mathbf{x}; G)_{kv} = \sum_{\substack{i \text{ löper över alla last-} \\ \text{bilsveckor i kvartal } kv}} \mathbf{x}(G)_i, \quad (2)$$

där

$\mathbf{x}(G) = \text{summa } \mathbf{x} \text{-värden för lastbilsveckans körningar i redovisningsgruppen } G$.

Med andra ord är $\mathbf{x}(G)$ *klustertotalen relativt gruppen } G*. (3)

En kommentar angående urvalet av lastbilsveckor

I sin helt precisa formulering är urvalsförfarandet för lastbilsveckor inte ett av urvalsteorins standardförfaranden. Det ligger nära *stratifierat OSU av lastbilsveckor* med stratifiering lastbilsstratan \times kvartalets 13 veckor, men är inte helt exakt ett sådant urval av skälet att mätveckorna för utvalda bilar inte väljs helt på måfå inom ett kvartal, utan sprids jämnt över det. Dock, förfarandet ligger så nära stratifierat OSU-urval av kluster (= lastbilsveckor) att punkt- och variansskattningar kan göras under approximationsantagandet (4) nedan. Approximationen ifråga prövades i anslutning till den UVAV-genomlysning som redovisas i Zamani & Rosén (1993), där slutsatsen är att approximationen fungerar tillfredsställande i praktiken.

Urvalet av lastbilsveckor kan med god approximation ses som ett stratifierat OSU-urval. (4)

Skattningsformler under premisen att inget svarsbortfall förekommer

Sätt;

h = index som löper över stratan, lastbilsstratan såväl som lastbilsveckostratan (eftersom dessa ett-ett-korresponderar), $h = 1, 2, \dots, H$,

N_h = storleken på lastbilsstratum h . Storleken på motsvarande stratum av lastbilsveckor är $13 \cdot N_h$, $h = 1, 2, \dots, H$,

n_h = antal utvalda lastbilar från stratum h , $h = 1, 2, \dots, H$,

σ_h^2 står för varians inom stratum h , $h = 1, 2, \dots, H$.

Under (4) ger välkända skattningsresultat nedanstående formler för punktskattningar och dessas (teoretiska) varianser. I första omgången antas uppgiftsinsamling utan svarsbortfall.

Punktskattningar

$$\hat{\tau}(\mathbf{x}; \mathbf{G}) = \sum_{h=1}^H \frac{13 \cdot N_h}{n_h} \cdot \sum_{i \text{ löper över observerade lastbilsveckor}} x(\mathbf{G})_i . \quad (5)$$

Teoretiska estimatorvarianser

$$V(\hat{\tau}(\mathbf{x}; \mathbf{G})) = \sum_{h=1}^H \frac{13^2 \cdot N_h^2}{n_h} \cdot \sigma_h^2(\mathbf{x}(\mathbf{G})) \cdot \left(1 - \frac{n_h}{13 \cdot N_h}\right). \quad (6)$$

Skattningsformler när svarsbortfall förekommer

Tyvärr förekommer svarsbortfall i de flesta undersökningar, och detta gäller även för SLIT. Då måste man på lämpligt sätt kompensera/korrigera för bortfallet, vilket i SLIT görs med s.k. *rak uppräknning*, d.v.s. man skattar som om de svarande utgjorde urvalet. (Förfarande är baserat på premisen att bortfallet sker helt slumpmässigt.) Sätt;

n'_h = antal svarande lastbilar i stratum h , $h = 1, 2, \dots, H$.

Punktskattningar

$$\hat{\tau}(\mathbf{x}; \mathbf{G}) = \sum_{h=1}^H \frac{13 \cdot N_h}{n'_h} \cdot \sum_{i \text{ löper över svarande lastbilsveckor}} x(\mathbf{G})_i . \quad (7)$$

Teoretiska estimatorvarianser

$$V(\hat{\tau}(\mathbf{x}; \mathbf{G})) = \sum_{h=1}^H \frac{13^2 \cdot N_h^2}{n'_h} \cdot \sigma_h^2(\mathbf{x}(\mathbf{G})) \cdot \left(1 - \frac{n'_h}{13 \cdot N_h}\right). \quad (8)$$

Kommentarer : (i) Ovanstående beskrivning är något förenklad. Också en korrigering för undertäckning görs, enligt beskrivning i Rosén, R&D Report 1990:3.

(ii) Formlerna (6) och (8) har rättframma pendanger för *skattning* av estimatorvarianser, men de behövs inte i detta översynssammanhang.

1.4.2 Skattning av årsaggregat

Årsaggregat definieras i analogi med kvartalsaggregat, med enda skillnaden att man summerar över år istället för kvartal. Det ger;

$$\tau(\mathbf{x}; \mathbf{G})_{\text{år}} = \sum_{j \text{ löper över alla körningar som påbörjas under året}} x_j \cdot 1(\mathbf{G})_j , \quad (9)$$

Mellan års- och kvartalsaggregat råder det rättframma sambandet;

$$\tau(\mathbf{x}; \mathbf{G})_{\text{år}} = \tau(\mathbf{x}; \mathbf{G})_{\text{kv1}} + \tau(\mathbf{x}; \mathbf{G})_{\text{kv2}} + \tau(\mathbf{x}; \mathbf{G})_{\text{kv3}} + \tau(\mathbf{x}; \mathbf{G})_{\text{kv4}} , \quad (10)$$

vilket leder till nedanstående naturliga punktskattning av ett årsaggregat;

$$\hat{\tau}(\mathbf{x}; \mathbf{G})_{\text{år}} = \hat{\tau}(\mathbf{x}; \mathbf{G})_{\text{kv1}} + \hat{\tau}(\mathbf{x}; \mathbf{G})_{\text{kv2}} + \hat{\tau}(\mathbf{x}; \mathbf{G})_{\text{kv3}} + \hat{\tau}(\mathbf{x}; \mathbf{G})_{\text{kv4}}. \quad (11)$$

Om det vore så att de fyra kvartalens urval dras helt oberoende av varandra, leder (11) till följande teoretiska varians för en årsskattning:

$$V[\hat{\tau}(\mathbf{x}; \mathbf{G})_{\text{år}}] = V[\hat{\tau}(\mathbf{x}; \mathbf{G})_{\text{kv1}}] + V[\hat{\tau}(\mathbf{x}; \mathbf{G})_{\text{kv2}}] + V[\hat{\tau}(\mathbf{x}; \mathbf{G})_{\text{kv3}}] + V[\hat{\tau}(\mathbf{x}; \mathbf{G})_{\text{kv4}}]. \quad (12)$$

Kvartalsurvalen under ett år är dock inte helt oberoende, av skälet att de dras som disjunkta OSU, för att skydda lastbilar från att komma med i undersökningen för ofta och för nära efter tidigare deltagande. Bedömningen görs dock att detta beroende spelar så liten roll att variansformeln (12) fungerar fullt tillfredsställande ändå.

2 Premisser för och metodaspekter på översynen

2.1 Allmänt om bakgrundspremisser

Syftet med översynen är att uppnå en så kostnadseffektiv undersökning av Inrikes och utrikes trafik med svenska lastbilar som möjligt, d.v.s. att kunna ta fram så tillförlitlig statistik som möjligt för insatta undersökningsresurser. Nedan formuleras några huvudfrågor som måste besvaras för att ge utgångspunkt för sökande efter förbättrad undersökningsutformning.

- 1) Vad vill man få så bra som möjligt?
- 2) Vad kan man påverka/styra?
- 3) Vilka krav och/eller restriktioner finns att ta hänsyn till?
- 4) Finns hjälpinformation tillgänglig nu som inte fanns tidigare?
- 5) Hur kan man påvisa förhoppningsvisa förbättringar av statistiken?

Som nämnts tidigare skall undersökningen följa EU-förordningen 1172/98, vilket medför att osäkerheten i årsstatistik för total inrikes transporterad godsmängd (Ton) respektive för totalt inrikes transportarbete (TonKm) skall vara högst $\pm 5\%$ för 95-procentiga konfidensintervall. För motsvarande utrikestransporter är målet satt till $\pm 10\%$. Statistiken för 2000 och 2001 möter kraven för inrikes transporter, medan statistiken om utrikestransporter inte riktigt klarar målgränsen. Ett huvudönskemål är därför att förbättra precisionen i statistiken om utrikes-transporter. Man vill naturligtvis också, om möjligt, förbättra noggrannheten i statistiken avseende inrikes transporter. Dessutom finns av SIKA fastställda krav på vad slags statistik som ska redovisas i Statistiska meddelanden från undersökningen, bl.a. krav på finheten i den regionala redovisningen samt att statistiken skall delas upp efter yrkesmässig resp. firmatrafik. Vidare är det angeläget att hålla ner uppgiftslämnarbyråerna.

Efter ett upphandlingsförfarande har SIKA och SCB träffat överenskommelse om att SCB skall vara undersökningsproducent åtminstone till och med år 2005, till nuvarande årliga kostnad. Förändringar i undersökningen skall alltså ske inom nuvarande kostnadsram, vilket medför att man inte kan tillgripa den allra enklaste förbättringsmetoden "utökad urvalsstorlek". Det man kan överväga är om undersökningens uppläggning kan förbättras.

För att uppnå god statistik om utrikestransporter är huvudproblemet att endast en mycket liten andel av transporterarna med svenskregistrerade lastbilar är utrikestransporter. För att förbättra utrikesstatistiken behöver man kunna styra urvalet så att fler utrikestransporter kommer att observeras. För detta behövs förbättrad information om vilka bilar som på goda grunder kan tros bli använda i utrikestrafik. Hittills har prognosvariabeln för "bilen används troligen i utrikes trafik" varit att företaget som äger bilen innehar blått kort. Förbättrad prognos bör kunna uppnås om företagen för var och en av sina lastbilar ger en bedömning av om bilen kommer att bli använd för utrikes eller enbart inrikes transporter. Sådan informationsinhämtning, kallad *blåkort-undersökningen*, genomfördes. Den beskrivs i Avsnitt 3.1.

Vad gäller förbättring av statistiken om inrikestransporter är en första fråga om det finns ny hjälpinformation som kan vara behjälplig för att skapa homogena urvalsstratan. En tidigare oprövad hjälpinformation är de körsträckor som SCB beräknar ur avläsningar av mätarställningar vid kontrollbesiktning av lastbilar hos Bilprovningen AB. Resultaten av dessa beräkningar samlas i den s.k. *Körsträckedatabasen*. Mer om detta i Avsnitt 3.2.

2.2 Mått på en urvalsutformnings effektivitet

Syftet med förändrad urvalsutformning är att uppnå förbättrad precision i skattningar av transportaggregat $\tau(\mathbf{x}; G)$ för kvartal och år. Den centrala storheten för att mäta skattningsprecision är estimatorvariansen $V[\hat{\tau}(\mathbf{x}; G)]$. Den beror dels av urvalsutformningen (stratifiering och urvalsstorlekar), dels av hur mycket variabeln \mathbf{x} varierar över populationen, och dels av vilken redovisningsgrupp G det handlar om.

Relativ felmarginal

Konkret tolkning av innebörden av en estimatorvariens fås av skattningens *relativa felmarginal* (RFM), vilken definieras enligt nedan;

$$\text{RFM}[\hat{\tau}(\mathbf{x}; G)] = \frac{2 \cdot \sqrt{V[\hat{\tau}(\mathbf{x}; G)]}}{\tau(\mathbf{x}; G)} = \frac{2 \cdot D[\hat{\tau}(\mathbf{x}; G)]}{\tau(\mathbf{x}; G)}. \quad (13)$$

Gränserna $\pm \text{RFM}[\hat{\tau}(\mathbf{x}; G)]$ ger (approx.) 95 % konfidensgränser för storheten $\tau(\mathbf{x}; G)$, mätt med värdet på $\tau(\mathbf{x}; G)$ som enhet. De nämnda EU-kraven är att inrikes årsaggregat skall ha en relativ felmarginal om högst 5 %, och att målet för utrikes årsaggregat är en relativ felmarginal om högst 10 %.

Relativ variansförändring

Ett enkelt mått på effekten av en förändrad urvalsutformning är relativ (estimator)variensförändring. Betrakta en tidigare och en ny urvalsutformning och sätt;

V_T = estimatorvariens vid tidigare urvalsutformning,

V_N = estimatorvariens vid ny urvalsutformning.

Den *relativa variansförändringen* (RVF) definieras;

$$\text{RVF} = 100 \cdot (V_T - V_N) / V_T \%. \quad (14)$$

2.3 Beräkning av skattningsprecision vid en hypotetisk urvalsutformning

Formeln (8) för estimatorvariens gäller oavsett hur ingående stratan skapats. Den kan alltså användas för beräkna statistiks noggrannhet vid en hypotetisk stratifiering och urvalsallokering, med användning av måtten RFM och / eller RVF. En förutsättning är dock att man kan ange realistiska "parametervärden" för den hypotetiska utformningen. I den fortsatta diskussionen görs beteckningsändringen att istället för att använda ett index h som löper över stratan som är numrerade $1, 2, \dots, H$ används ett index k som löper över den hypotetiska stratifieringens stratan, vilka är numrerade $1, 2, \dots, K$. De "parametrar" det handlar om är N_k , n'_k , $\mu_k(\mathbf{x}(G))$ och $\sigma_k^2(\mathbf{x}(G))$, varmed åsyftas stratumstorlekar, antal svarande i urvalen från stratana samt medelvärden och varianser i de stratan som ingår i den hypotetiska stratifieringen. Främsta svårigheten härvid är att uppnå realistiska uppskattningar av stratumvarienserna $\sigma_k^2(\mathbf{x}(G))$. Tabell 3 sammanfattar den information som behövs för beräkning av RFM för en skattning av $\tau(\mathbf{x}; G)$ vid en hypotetisk urvalsutformning.

Tabell 3. Behövd information för beräkning av RFM vid en hypotetisk stratifiering med stratan A_1, A_2, \dots, A_K

<i>Stratum</i>	<i>Stratum-storlek</i>	<i>Medelvärde i stratumet</i>	<i>Standardavvikelse i stratumet</i>	<i>Antal svarande</i>
A_1	N_1	$\mu(\mathbf{x}(G))_1$	$\sigma(\mathbf{x}(G))_1$	n'_1
A_2	N_2	$\mu(\mathbf{x}(G))_2$	$\sigma(\mathbf{x}(G))_2$	n'_2
\cdot	\cdot	\cdot	\cdot	\cdot
A_K	N_K	$\mu(\mathbf{x}(G))_K$	$\sigma(\mathbf{x}(G))_K$	n'_K

Med information enligt Tabell 3 kan estimatorvarianser $V[\hat{\tau}(\mathbf{x};G)]$ beräknas via formeln (8). I det steget behövs inte stratummedelvärdena $\mu(\mathbf{x}(G))_k$, men de behövs för att beräkna aggregaten som ingår i relativa felmarginaler. Dessa beräknas;

$$\tau(\mathbf{x};G) = \sum_{k=1}^K N_k \cdot \mu(\mathbf{x}(G))_k. \quad (15)$$

Värdena på $\mu(\mathbf{x}(G))_k$ och $\sigma(\mathbf{x}(G))_k$ känner man normalt inte, men de kan uppskattas. För att pröva effekten av alternativa urvalsutformningar i framtida undersökningar antas framtida populationer se ut ungefär som den senast tillgängliga, d.v.s. 2001 års population. Under den premissen skattas $\mu(\mathbf{x}(G))_k$ och $\sigma(\mathbf{x}(G))_k$ i framtida undersökningar enligt nedan.

$$\mu(\mathbf{x}(G))_k = \frac{1}{N_k} \cdot \sum_{j \in A_k} x(G)_j, \quad (16)$$

$$\sigma^2(\mathbf{x}(G))_k = \frac{1}{N_k - 1} \cdot \left(\sum_{j \in A_k} x(G)_j^2 - \frac{\left(\sum_{j \in A_k} x(G)_j \right)^2}{N_k} \right), \quad (17)$$

där komponenterna i (16) och (17) skattas utifrån 2001 års material enligt nedan, där h löper över de stratan som användes i 2001 års undersökning;

$$N_k = \sum_{A_k} 1 \quad \text{skattas med} \quad \sum_{h=1}^H \frac{13 \cdot N_h}{n'_h} \cdot \sum_{i=1}^{n'_h} \mathbf{1}(A_k)_i, \quad (18)$$

$$\sum_{j \in A_k} x(G)_j \quad \text{skattas med} \quad \sum_{h=1}^H \frac{13 \cdot N_h}{n'_h} \cdot \sum_{i=1}^{n'_h} x(G)_i \cdot \mathbf{1}(A_k)_i, \quad (19)$$

$$\sum_{j \in A_k} x(G)_j^2 \quad \text{skattas med} \quad \sum_{h=1}^H \frac{13 \cdot N_h}{n'_h} \cdot \sum_{i=1}^{n'_h} x(G)_i^2 \cdot \mathbf{1}(A_k)_i, \quad (20)$$

där $\mathbf{1}(A_k)$ är indikatorn för att en 2001-körning ingår i den hypotetiska stratifieringens stratum A_k .

Insättning av (18), (19) och (20) i (16) och (17) ger skattningar av N_k , $\mu_k(\mathbf{x}(G))$ och $\sigma_k^2(\mathbf{x}(G))$. Dessa används sedan i (8) tillsammans med lämpligt valda värden på n'_k , och ger då uppskattade värden för estimatorvarianser för kvartalsaggregat. Vid specifikation av n'_1, n'_2, \dots, n'_K har antagits att framtida svarsfrekvenser kommer att vara densamma som "svarsprocenten overall" under 2001, nämligen 73%. Via relationen (12) erhålls sedan uppskattade estimatorvarianser för årsaggregat.

3 Överväganden och insatser vid översynen av urvalsutformningen

3.1 Allmänt om undersökningens stratifiering

Vid stratifierat urval indelas urvalsramen i disjunkta grupper (= stratan), och oberoende separata urval dras sedan från stratana. Stratifieringen görs före urvalsdragningen med efterföljande observationer, och måste därför baseras på information från källor utanför själva undersökningen, s.k. *hjälpvariabler / stratifieringsvariabler*. Huvudprinciper för effektiv stratifiering är som följer. (i) Det är fördelaktigt med stratan som medför *homogenisering* av populationen, varmed avses att de viktiga undersökningsvariablerna varierar mindre inom stratana än över hela populationen. (ii) Det är ofta fördelaktigt att vid urvalet *hög - eller lågrepresentera vissa objekt* (se Avsnitt 1.3.1). Sådana objekt måste då avgränsas i särskilda stratan för att urvalssannolikheterna skall kunna varieras på önskat sätt.

I SLIT är populationsobjekten "körningar" grupperade i "lastbilsveckor", och de viktigaste variablerna gäller lastbilars körningar under en mätvecka; transportarbete (TonKm), godsmängd (Ton) och körsträcka (KmKorda). Urvalsramen utgörs dock inte av lastbilsveckor, utan av "lastbilar". Urvalet av lastbilar genererar ett urval av lastbilsveckor genom att utvalda lastbilar tillordnas en slumpmässig mätvecka (inom kvartalet), och det urvalet ger sedan (ett klustrat) urval av körningar. Stratifieringen av lastbilarna medför en stratifiering av lastbilsveckorna, som i sin tur medför en stratifiering av körningarna.

Som sägs i Avsnitt 1.3.2 baserades stratifieringen under åren 2000-2002 på följande variabler för lastbilar: (i) Maxlastvikt, (ii) Karosserikod, (iii) Yrkesmässig/firma, (iv) Om ägarföretaget har blåkorttillstånd, (v) Registreringslän. Skäl för valet av dessa variabler ges nedan.

- Samvariation föreligger mellan en lastbils maxlastvikt och dess transportinsats. Ju större bil, desto större genomsnittlig insats. Indelning efter maxlastvikt medför homogenisering av populationen av lastbilsveckor.
- Karosserikoden används för att avgränsa banke- och tankbilar, vilka är transportmässigt högpresterande, och ger huvudbidrag till de särskilt intressanta varugrupperna rundvirke och oljeprodukter. Det är önskvärt att högrepresentera dessa bilar, och där för förs de till särskilda stratan.
- För precisionen i utrikesstatistiken är det angeläget att många utrikes transporter kan observeras. Man vill högrepresentera "troliga utrikeskörare". Bilar i blåkortföretag är sådana, även om merparten av deras transporter faktiskt går inrikes. Vidare finns skäl att tro att ju fler blåkort ett företag har, desto mer "utrikesbenägna" är deras bilar.
- Indelningen yrkesmässig/firma görs för att statistik redovisas med den indelningen.
- Den regionala stratumindelningen görs för att statistik redovisas med geografisk indelning. Av särskilt intresse är den s.k. flödesstatistiken.

Med så många indelningsvariabler blir antalet stratan rätt stort. I och för sig kan man använda många stratan, men alltför många stratan kan få menlig inverkan på statistikens noggrannhet, framför allt genom det restrifierade urvalet med minst 10 bilar från varje stratum. Onödigt många stratan medför att ett antal små, och rätt ointressanta, stratan överrepresenteras.

Centrala utgångsfrågor vid översynen av stratifieringen var följande:

- a) Kan antalet observerade utrikeskörningar ökas (utan att totala urvalstorleken ökas)?
- b) Finns ny hjälpinformation som kan vara behjälplig för homogeniserande stratifiering?
- c) Är det fördelaktigt att minska antalet stratan något?

Fråga a) ledde till den s.k. blåkortundersökningen som beskrivs i nästa Avsnitt 3.2. Fråga b) ledde till studium av informationen i Körsträckedatabasen, vilket beskrivs i Avsnitt 3.3.

3.2 Blåkortundersökningen

Den helt övervägande andelen lastbilstransporter med svenska lastbilar sker inrikes. Men, för att uppnå bra statistik för utrikestransporter fordras många observationer av sådana. Det gäller då att försöka avgränsa "utrikesbenägna" bilar, för att kunna högrepresentera dessa i urvalet. De återfinns hos blåkortföretag, vars bilar utgör Utrikesdelen i urvalsramen. Den delas sedan i Storleksgrupperna 1 och 2 (STG1 resp. STG2) beroende på om företaget innehar färre än 16 blåkort eller fler. STG2 omfattar bilar i förhållandevis stora speditörsföretag, många med profil "utrikesspeditörer". Dock, majoriteten av blåkortföretagens körningar är ändå inrikes körningar. Så även om förhållandevis många utrikeskörningar kunnat observeras har antalet varit för lågt för att uppnå riktigt bra utrikesstatistik.

En möjlighet att öka antalet observerade utrikeskörningar ges om man ur STG2 kan utesluta bilar som så gott som säkert gör enbart inrikes körningar. Den för detta nödvändiga informationen finns dock bara hos företagen, och måste inhämtas från dem om man vill kunna använda den. Syftet med blåkortundersökningen var dels att utröna om STG2 - företagen kunde dela upp sin lastbilspark i en "inrikesflotta" resp. "troliga/möjliga utrikeskörare", dels att få veta hur villiga företagen är att utföra det aktuella uppgiftslämnandet när det är frivilligt.

En undersökning genomfördes under september 2002. Den omfattade de 70 företag som hade 16+ blåkort. Undersökningsblanketten återfinns i Bilaga 1. Av 2 765 berörda fordon inkom svar avseende 2 537 fordons utnyttjande under år 2003. Bedömningen var att 672 lastbilar kommer att användas för utrikeskörningar till minst 80 % av körtiden, 304 bilar för utrikeskörningar till 2 - 80 % av körtiden samt 1 562 lastbilar för i stort sett bara inrikes körningar (mindre än 2 % av körtiden till utrikes körningar).

Den åtgärd som vidtagits på basis av resultaten är att de 1562 fordon som bedömts användas för i stort sett bara inrikes körningar förts från STG2 till STG1. Lastbilar ägda av företag som ej svarade på enkäten ligger dock kvar i STG2. Detta innebär att definitionerna av STG1 och STG2 ändrats, men trots detta används alltjämt benämningarna STG1 och STG2 (med utrikes körgrupper 1 och 2 som synonymer).

Vidare skall bilarna i nya STG2 delta i undersökningen två gånger per år, medan de i tidigare STG 2 ingick en gång per år under 2000 - 2002. Därigenom kommer antalet observerade utrikeskörningar öka. Det är visserligen sant att bilarna i nya STG2 får ökat uppgiftslämnande, men sett med företagets ögon kommer uppgiftslämnandet att minska för i stort sett alla företag med bilar i tidigare STG2, eftersom deras "inrikesbilar" kommer att slippa avsevärt lindrigare undan ifråga om att ingå i undersökningen.

3.3 Hjälpinformation från Körsträckedatabasen

Lastbilar besiktigas hos Bilprovningen en gång om året, varvid deras vägmätarställning numera registreras. Från dessa registreringar uppskattar SCB en årlig körsträcka för varje enskild lastbil. Uppskattade årskörsträckor sammanställs i den s.k. Körsträckedatabasen, vilken omfattar nästan alla lastbilar. Undantagen är framför allt bilar som är mindre än ett år gamla, och därmed ännu inte hunnit besiktas för första gången. Årskörsträckan förväntades samvariera med en lastbils transportinsats, bilar med stor respektive liten årlig körsträcka utför (genomsnittligt sett) större respektive mindre insatser. Därför förmodades årskörsträckor vara god, och tidigare icke utnyttjad, hjälpinformation för homogeniserande stratifiering. För att klarlägga saker gällde att utröna dels hur god täckningen i Körsträckedatabasen är, dels hur starkt årskörsträckevariabeln samvarierar med undersökningsvariablerna. Resultat redovisas nedan.

3.3.1 Körsträckedatabasens täckning

Tabell 4 nedan beskriver täckningsförhållanden i databasen.

Tabell 4. Antal lastbilar med respektive utan årskörsträcka för 2001

<i>Kvartal</i>	<i>Antal lastbilar med körsträcka</i>	<i>Antal lastbilar utan körsträcka</i>	<i>Därav i nya STG2</i>
1	49 827	6 511	440
2	47 477	7 564	565
3	50 697	5 430	449
4	51 338	5 600	463

3.3.2 Samvariation mellan årskörsträcka och undersökningsvariablerna

Mått på styrkan i samvariationen mellan årliga körsträckor och de tre undersökningsvariablerna Km, TonKm och Ton erhöles via skattade korrelationskoefficienter. Eftersom stratifiering med hjälp av körsträckor främst var tänkt att användas i Inrikesdelen beräknades korrelationskoefficienterna dels för "alla lastbilar", dels för bilar i Inrikesdelen.

Tabell 5. Korrelationskoefficienter för årskörsträcka och undersökningsvariabler för alla lastbilar år 2001

<i>Kvartal</i>	<i>Km</i>	<i>TonKm</i>	<i>Ton</i>
1	0,69	0,68	0,26
2	0,68	0,64	0,13
3	0,69	0,66	0,16
4	0,73	0,69	0,15

Tabell 6. Korrelationskoefficienter för årskörsträcka och undersökningsvariabler för lastbilar i Inrikesdelen år 2001

<i>Kvartal</i>	<i>Km</i>	<i>TonKm</i>	<i>Ton</i>
1	0,70	0,77	0,45
2	0,71	0,71	0,30
3	0,72	0,72	0,24
4	0,69	0,73	0,28

Korrelationskoefficienterna är som synes höga för körsträcka (Km) och transportarbete (TonKm). Slutsatsen är att körsträckor bör vara en bra homogeniserande stratifieringsvariabel.

3.4 Om variabeln maxlastvikt

Tidigare var maxlastvikt den homogeniserande stratifieringsvariabeln, men den överspelas av årskörsträcka åtminstone vad gäller samvariation med Km och TonKm. Man kan undra om både årskörsträcka och maxlastvikt skall användas vid stratifieringen, eller om det räcker med den ena. Dock, maxlastvikt kan förmodas samvariera starkast med Ton. Som komplement till ovannämnda studie av körsträckors samvariation med körningsvariablerna gjordes en motsvarande studie för maxlastvikt. Resultatet ges i Tabell 7. Där ses att maxlastvikt är korrelerad med alla körningsvariablerna, men för Km och TonKm inte tillnärmelsevis så starkt som årskörsträcka. För Ton har maxlastvikt dock något högre korrelation än årskörsträcka.

Tabell 7. Korrelationer mellan maxlastvikt och körningsvariablerna år 2001

<i>Kvartal</i>	<i>Km</i>	<i>TonKm</i>	<i>Ton</i>
1	0,19	0,35	0,34
2	0,20	0,31	0,31
3	0,20	0,35	0,37
4	0,18	0,32	0,29

3.5 Om stratifiering med avseende på yrkesmässig/firma

Under 2000 - 2002 gjordes stratumindelning efter yrkesbilar/firmabilar, med efterföljande urvalsallokering som var ett arv från den tidigare UVAV - undersökningen. Eftersom firmabils-trafiken över åren minskat avsevärt, är denna stratifiering inte längre av samma dignitet som tidigare. År 1993 fanns 22 130 firmabilar som utförde ett transportarbete om 3 256 miljoner TonKm medan det år 2001 fanns endast 17 641 firmabilar som utförde ett transportarbete om 2 278 miljoner TonKm. Antalet yrkesbilar däremot har ökat från 31 690 år 1993 till 39 066 år 2001 och deras transportarbete har ökat från 22 652 miljoner TonKm till 27 692 miljoner TonKm. Mot den bakgrunden var det naturligt att ifrågasätta om stratifiering efter yrkesmässig/firma snarare är till förfång än till fördel för statistikens precision. (Se Avsnitt 3.1 om "många stratan".) Det skall framhållas att statistik med uppdelning efter yrkesmässig/firma utmärkt väl kan tas fram även om man inte stratifierar efter denna uppdelning.

4 Resultat av prövning av olika utformningsalternativ

Som redan indikerats prövades ett antal hypotetiska urvalsutformningar (stratifieringar och urvalsallokeringar) via beräknade relativa felmarginaler och relativa variansförändringar för aggregat-skattningar, med metodik enligt Avsnitt 2.2. I det följande beskrivs den utformning som framstod som bäst. Den har, med SIKAs godkännande, tagits i bruk från och med innevarande undersökningsår 2003. Även om det vore mest logiskt att börja med beskrivning av stratifieringen, behandlas allokeringen först. Stratifieringen behandlas i Avsnitt 4.2.

4.1 Urvalsallokering

För att uppnå god undersökningsutformning gäller inte bara att dela in urvalsramen i adekvata stratan, utan också att allokera urvalet på ett bra sätt på urvalsstratan. Nedan redovisas de allokeringar som föll ut som varande bäst av de prövade allokeringarna.

4.1.1 Urvalets fördelning på Inrikes- respektive Utrikesdel

Totala årsurvalet av lastbilar kommer även fortsättningsvis att vara 3 000. Första steget vid fördelningen av den urvalsstorleken gäller tilldelningen till Inrikes- respektive Utrikesdel. Vid översynen anammades den tidigare fördelningen, hälften av urvalet, dvs. 1 500 bilar, till Inrikesdelen och hälften till Utrikesdelen. Denna fördelningsproblematik gavs dock inte ett ingående studium. Valet är till stor del baserat på att urvalet i Utrikesdelen inte bör minskas, eftersom utrikesstatistiken åtminstone hittills haft sämst noggrannhet.

4.1.2 Allokering av urvalet inom Inrikesdelen

Som tidigare allokeras urvalet i Inrikesdelen med fasta proportioner på grupperna bankebilar, som får 8,3 % , tankbilar, som får 6,1 % och "övriga bilar" som får 85,6 %. Inom grupperna tank- och bankebilar uppdelade i årskörsträckeklasser görs sedan vidareallokering proportionellt mot stratumstorlek. Allokering med hjälp av stratumvarianser för körsträckor testades, men förbättrade inte precisionen i skattningarna.

Inom gruppen "övriga lastbilar" testades följande allokeringsvariabler: årskörsträckor, transportarbete (TonKm) samt ett 50-50% -medelvärde av Neyman-allokeringar med avseende på TonKm respektive Ton, beräknat på basis av data från tidigare undersökningar. Den variant som framstod som mest robust för samtliga undersökningsvariabler, och som används framöver, är 50-50% -allokeringen.

4.1.3 Allokering av urvalet inom Utrikesdelen

Även för lastbilar i nya STG1 (dvs. bilar hos företag med färre än 16 blåkort, samt bilar som klassats som "inrikeskörare" hos företag med 16+ blåkort) testades olika allokeringar. Även här framstod 50 - 50 % - varianten som den lämpligaste. För STG2 föreligger inget egentligt allokeringsproblem, samtliga bilar skall delta i undersökningen två gånger per år.

4.2 Stratifiering i Inrikesdelen

Det framkom att tank- och bankebilar ger så substantiella bidrag till viktiga transportaggregat att de även fortsatt får ligga i separata stratumgrupper. Steg 1 i stratifieringen är, som tidigare, uppdelning i grupperna bankebilar, tankbilar samt "övriga bilar".

För att hålla ned antalet stratan i gruppen "övriga bilar" byttes i Steg 2 tidigare geografiska indelning efter län till indelning efter Sveriges åtta NUTS2 - regioner SE01, SE02, SE03, SE04, SE0A, SE06, ... , SE08. Via simuleringar framkom dock att Gotland fortsatt bör vara en egen region, för att inte riskera att bli nästan helt utan utvalda lastbilar. I detta Steg 2 görs inte geografisk indelning för banke- och tankbilar.

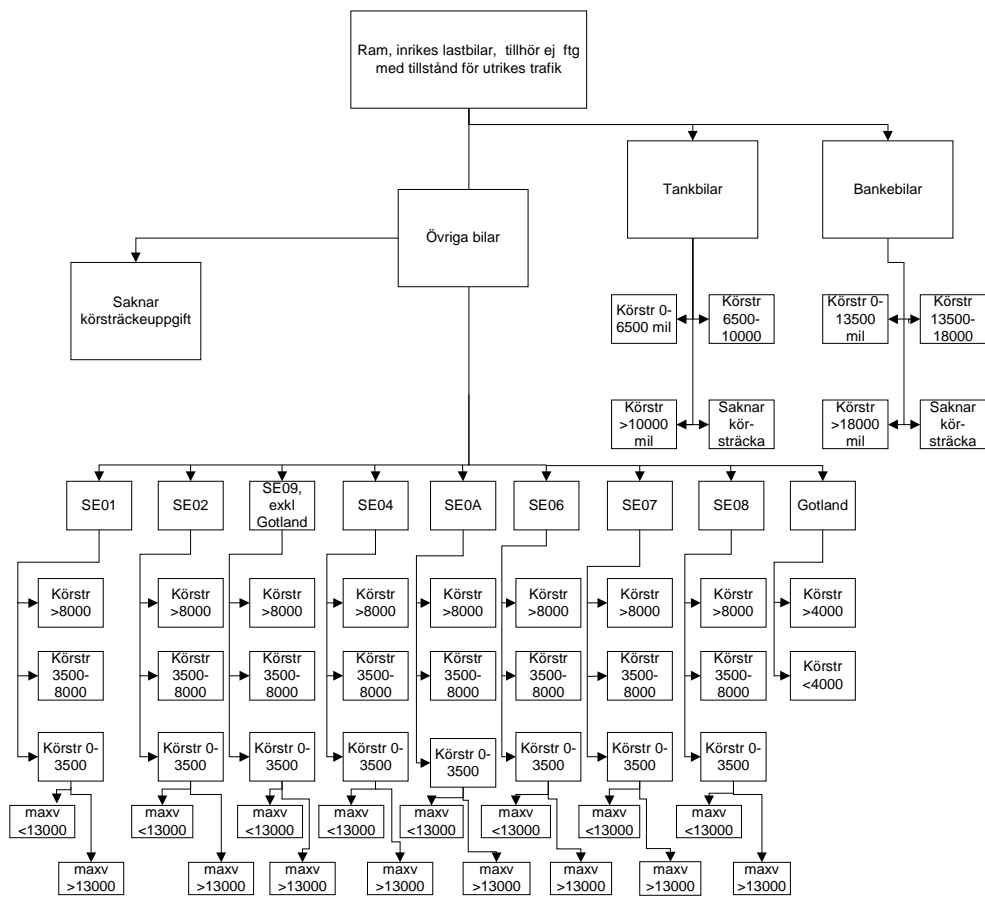
Vid testning visade det sig att stratifieringen yrkes/firma kan tas bort utan att precisionen hos statistiken för total yrkesmässig trafik och total firmabilstrafik försämras. För vissa mindre redovisningsgrupper blir dock statistiken för firmabilstrafiken något sämre. Indelningen ifråga togs dock bort för att hålla ned antalet stratan.

Årskörsträcka visade sig, som väntat, ha avsevärd homogeniseringskraft. I Steg 3 stratifieras efter den variabeln i var och en av de tre grupperna bankebilar, tankbilar och "övriga bilar". Bilar utan uppgift om körsträcka förs till separata stratan inom var och en av de tre grupperna. Med trial and error bestämdes (förhoppningsvis) optimala stratumgränser för årskörsträckor.

Som nämnts övervägdes om stratifiering efter maxlastvikt också skulle tas bort. Test gav dock vid handen att den stratifieringen bör bibehållas för bilar med korta årskörsträckor. Bilarna i den kategorin gör visserligen korta körningar, men ofta med stor godsmängd.

Gången i stratifieringen är alltså följande: Steg 1: Tankbil, bankebil, övrigbil. Steg 2: Geografisk indelning. Steg 3: Indelning efter årskörsträcka. Steg 4: Indelning efter maxlastvikt. Den precisa utformningen av stratifieringen, inklusive stratumgränser, framgår i Diagram 1 nedan.

Diagram 1. Schema över stratifieringen i Inrikesdelen



4.3 Stratifiering i Utrikesdelen

Som tidigare nämnts används information från Blåkortregistret och Blåkortundersökningen för att dela upp blåkortföretagens bilar på ett nytt sätt. STG2 (Körgrupp 2) utgörs nu av fordon som ägs av företag med 16+ blåkort och som bedömts vara "icke inrikeskörare". Fordon som ägs av företag med 16+ blåkort och som bedömts vara "inrikeskörare" förs numera till STG1 (Körgrupp 1). I övrigt utgörs STG1 av bilar i företag med mindre än 16 blåkort.

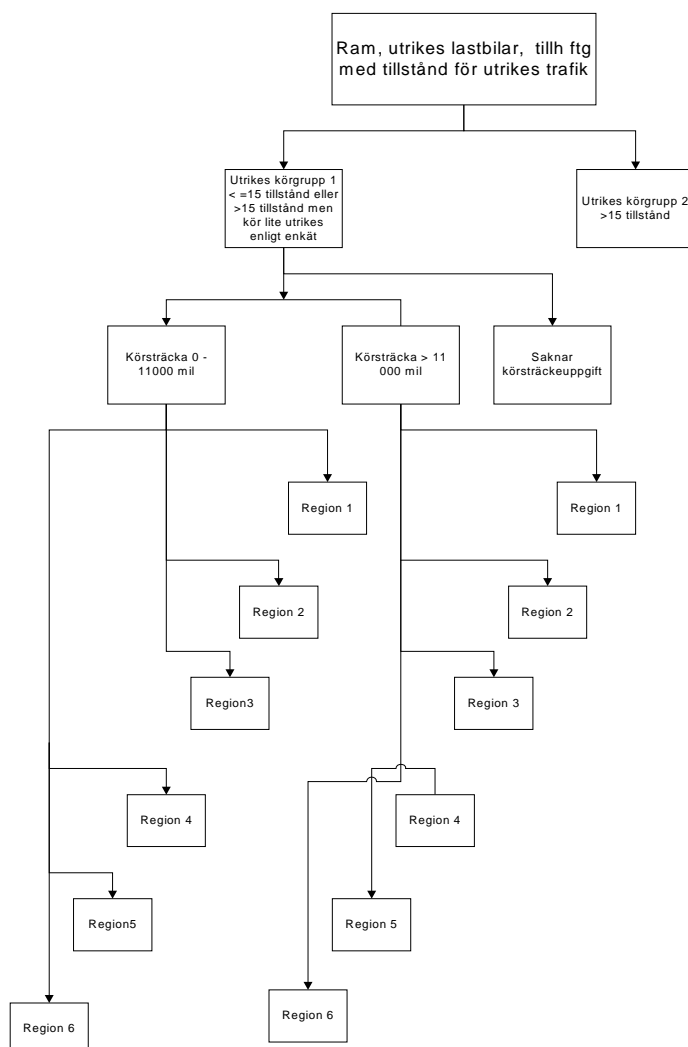
Geografisk stratifiering av STG1 görs som tidigare med stratifiering efter 6 regioner.

För STG 1 framkom vid testen att estimatorvarianserna även här minskar om årskörsträcka används för stratifiering. Lastbilar utan information om körsträcka läggs i ett eget stratum.

I STG 2 stratifieras lastbilarna inte vidare. Alla kommer under ett undersökningsår att ingå i två kvartalsundersökningar (dock ej i två på varandra följande kvartal).

Stratifieringen, inklusive stratumgränser, framgår i Diagram 2 nedan.

Diagram 2. Schema över stratifieringen i Utrikesdelen



4.4 Förväntad skattningsprecision med den nya utformningen

Det är naturligtvis av intresse att försöka uttala sig om vilka skattningsprecisions den nya uralutformningen kommer att leda till. Underlag för detta ges av de (upp)skattade estimatorvarianser som framräknades vid sökandet efter bra stratifieringar och allokeringar.

4.4.1 Förväntade relativa felmarginaler

I nedanstående tabell anges för några olika statistikuppgifter de relativa felmarginaler (se formeln (13) i Avsnitt 2.2) som förelåg vid 2001 års undersökning samt vad de förväntas bli vid 2003 års undersökning.

Tabell 8. Relativa felmarginaler för viss årsstatistik med indelning efter varuslag för undersökningen år 2001 och förväntade relativa felmarginaler år 2003

Varuslag	Relativa felmarginaler (i %) år 2001			Förväntade relativa felmarginaler (i %) år 2003		
	TonKm	Ton	Km	TonKm	Ton	Km
Totalt	3,9	5,0	2,8	3,0	4,6	2,2
Inrikes	4,3	5,2	3,0	3,4	4,8	2,5
Utrikes	10,7	14,2	9,6	9,2*	13,1*	8,1*
Jord, sten, grus och sand	15,7	14,4	12,6	14,0	13,5	11,9
Livsmedel och djurfoder	12,9	12,9	9,7	10,9	11,9	8,2
Rundvirke	10,4	10,3	10,3	8,1	8,5	8,0
Råolja o oljeprodukter	13,6	15,1	13,0	12,9	15,2	12,3
Övrigt	5,1	7,3	3,6	4,0	6,5	3,0
Kommentar: De relativa felmarginaler som redovisas för 2001 är beräknade med n' = faktiskt antal inkomna svar, medan de förväntade felmarginalerna för 2003 är beräknade under schablonantagande om 27 % svarsbortfall i varje stratum. Det antagandet medför att felmarginalerna förmodas bli något större i verkligheten.				*) Effekterna av omdefinitionen av STG2 och dess ändrade undersökningsfrekvens har inte tagits med i beräkningarna. De relativa felmarginalerna för utrikestrafiken tros bli åtskilligt mindre.		

4.4.2 Förväntade relativa variansförändringar

Tabellerna nedan redovisar förväntade relativa variansreduktioner (RVF, se formeln (14) i Avsnitt 2.2) vid jämförelse mellan undersökningarna år 2001 och år 2003.

Eftersom svarsfrekvenserna för år 2003 ej är kända, är resultaten i Tabellerna 9, 10 och 11 inte baserade på antal svarande utan på antal utvalda lastbilar (såväl för 2001 som för 2003).

Som teknisk förenkling uteslöts bidrag från STG2 vid beräkning av förväntade relativa variansreduktioner. Förhoppningsvis leder de förenklade räkningarna till rätt storleksordningar för relativa förändringar för de flesta av redovisningsgrupperna. Därmed har ändringarna i undersökningutformningen på basis av Blåkortundersökningen inte beaktats. Med största trolighet kommer varianserna för skattningar av utrikestransporter att reduceras ytterligare.

Inga insatser gjordes för att parera effekterna av s.k. outliers. Numera ingår korrigering för extrema värden i den slutliga skattningsprocessen i SLIT. Sådana korrigeringar kan påverka punktskattningarna, men också variansreduktionen i positiv riktning.

Den nya utformningen ger bäst resultat för transportarbete (TonKm) och körda kilometer (KmKorda), se Tabell 9. Detta beror på att årskörsträckor har starkt samband med transportarbete och körda kilometer, men svagare med godsmängd (Ton). Som tidigare nämnts har hänsyn till detta tagits i den nya utformningen, dels genom stratifiering efter maxlastvikt för bilar med korta årskörsträckor, dels genom allokering med hjälp av såväl Ton som TonKm.

Tabell 9. Relativa variansförändringar (i %) för viss årsstatistik indelad efter varuslag vid jämförelse av åren 2001 och 2003

<i>Varuslag</i>	<i>TonKm</i>	<i>Ton</i>	<i>Km</i>
Totalt	28	6	22
Inrikes	27	6	22
Utrikes	16	10	18
Jord, sten, grus o sand	14	2	-1
Livsmedel o djurfoder	26	15	27
Rundvirke	24	21	24
Råolja o oljeprodukter	0	1	2
Övrigt	27	12	19

Stratifieringen efter yrkesmässig respektive firmabilstrafik togs bort av skäl som tidigare redogjorts för. För den yrkesmässig trafiken är varianserna mindre eller oförändrade med den nya utformningen för samtliga varugrupper i Tabell 10. Firmabilstrafiken har lite mer varierade resultat, men på totalnivå sker variansförbättring för TonKm och Km. För Ton sker dock försämring även på totalnivå för firmabilstrafiken.

Tabell 10. Relativa variansförändringar för viss årsstatistik indelad efter yrkesmässig/firma vid jämförelse av åren 2001 och 2003

<i>Varuslag</i>	<i>TonKm</i>	<i>Ton</i>	<i>Km</i>
TotalFirma	11	-33	10
TotalYrkes	26	6	17
InrikesFirma	7	-33	8
InrikesYrkes	26	6	16
UtrikesFirma	59	54	54
UtrikesYrkes	14	9	15
Jord, sten grus o sand firma	-17	-37	-59
Jord, sten grus o sand yrkes	14	4	3
Livsmedel o djurfoder firma	33	16	19
Livsmedel o djurfoder yrkes	24	14	30
Rundvirke firma	-138	-201	-148
Rundvirke yrkes	24	21	23
Råolja o oljeprodukter firma	-45	-102	17
Råolja o oljeprodukter yrkes	-1	10	-3
Övrigt firma	12	-20	13
Övrigt yrkes	27	13	16

Varianserna för skattning av transportarbete efter pålastningslän förväntas bli bättre med den nya urvalsutformningen, se Tabell 11. För körda kilometer, d.v.s. transportsträcka, beräknas resultatet också oftast bli bättre. Även om estimatorvarienserna ökar för fler län än de minskar för, minskar variansen för totalt transporterad godsmängd. De län där varianserna ökar är små ifråga om total godsmängd.

Tabell 11. Relativa variansförändringar vid jämförelse av viss årsstatistik med geografisk indelning vid jämförelse mellan åren 2001 och 2003

<i>Län för pålast</i>	<i>TonKm</i>	<i>Ton</i>	<i>Km</i>
Totalt	27	5	22
Stockholm	20	43	-14
Stockholms kommun	22	34	17
Uppsala	17	7	-2
Södermanland	29	-18	22
Östergötland	20	-8	2
Jönköping	38	14	25
Kronoberg	-6	-35	-8
Kalmar	10	1	-7
Gotland	3	37	6
Blekinge	6	5	-8
Skåne	15	-8	-7
Malmö kommun	12	-30	11
Halland	-1	-8	7
Västra Götaland	13	-18	26
Göteborgs kommun	9	22	18
Värmland	26	-6	-1
Örebro	34	-37	5
Västmanland	20	22	14
Dalarna	33	29	18
Gävleborg	16	-21	-0
Västernorrland	18	-6	13
Jämtland	23	-35	14
Västerbotten	8	-22	-6
Norrbotten	21	15	36

5 Referenser

Rosén, B. Om justering för undertäckning i undersökningar med urval i "rum och tid".
SCB R&D Report 1990:3.

Rosén, B. & Zamani, M. Översyn av Undersökningen av lastbilstransporter i Sverige (UVAV).
SCB R&D Report 1993:2.

Bilaga 1. Enkät i blåkortundersökningen



Undersökning av åkerier med tillstånd till internationell godstransport på väg i yrkesmässig trafik s.k. blå kort

Här nedan är en lista över de lastbilar/dragbilar (med maxlastvikt 3,5 ton eller mer) som enligt Vägverkets Fordonsregister 2002-08-15 är registrerade på Er företag. Kryssa i aktuell ruta för respektive bil hur Ni **uppskattar** att bilen har använts i utrikes trafik under de **senaste 12 månaderna** (nov 2001 - okt 2002) och hur Ni **bedömer** att bilen kommer att köra i utrikes trafik under den **kommande 12-månaders perioden** (nov 2002 - okt 2003).

Med **utrikeskörning** avses **helheten** i en körning där båda eller endera av orten för avlastning eller pålastning ligger utanför Sverige.

Ofta = Lastbilen/dragbilen används **mer än 80%** (i körtid under året) för utrikeskörningar.

Ibland = Lastbilen/dragbilen används **2 - 80%** (i körtid under året) för utrikeskörningar.

Sällan/aldrig = Lastbilen/dragbilen används **mindre än 2%** (i körtid under året) för utrikeskörningar.

Om lastbilar/dragbilar sålts eller skrotats ber vi Er kryssa i rutan såld/skrotad. Om företaget nyförvärvat bilar behöver ingen komplettering göras.

Skicka in blanketten i bifogat svarskuvert så snart som möjligt, dock senast den 15 oktober 2002. Var vänlig och fyll i kontaktperson nedan.

Företagets kontaktperson: _____ tfn: _____

FID	REGNR	Uppskattat utrikeskörande under de senaste 12 månaderna			Bedömning av utrikeskörande under kommande 12 månader			Såld/ skrotad
		Ofta >80%	Ibland 2-80%	Sällan/aldrig <2%	Ofta >80%	Ibland 2-80%	Sällan/aldrig <2%	
498	XXX001	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
498	XXX002	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
498	XXX003	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
498	XXX004	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
498	XXX005	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
498	XXX006	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Förteckning över utkomna R&D Reports

R&D Reports är en för IT-enheten och Metodenheten gemensam publikationsserie, som 1988-01-01 ersatte de tidigare ”gula” och ”gröna” serierna.

Reports published during 2000 and onwards:

- 2000:1 Kalibrering av vikter – beskrivning av tekniken och de SCB-fall den prövats i
(*Sixten Lundström et al*)
- 2000:2 On Inclusion Probabilities and Estimator Bias for Pareto π ps Sampling
(*Nibia Aires and Bengt Rosén*)
- 2000:3 Bortfallsbarometer nr 15 (*Per Nilsson, Ann-Louise Engstrand, Sara Tångdahl, Stefan Berg, Tomas Garås och Arne Holmqvist*)
- 2000:4 Bortfallsanalys av SCB-undersökningarna HINK och ULF (*Jan Qvist*)
- 2000:5 Generalized Regression Estimation and Pareto π ps (*Bengt Rosén*)
- 2000:6 A User's Guide to Pareto π ps Sampling (*Bengt Rosén*)
- 2001:1 Det statistiska registersystemet. Utvecklingsmöjligheter och förslag
(*SCB, Registerprojektet*)
- 2001:2 Order π ps Inclusion Probabilities Are Asymptotically correct (*Bengt Rosén*)
- 2002:1 On the Choice of Sampling Design under GREG Estimation in Multiparameter Surveys (*Anders Holmberg*)
- 2002:2 Model-based calibration for survey estimation, with an example from expenditure analysis Surveys (*Claes Cassel, Peter Lundquist and Jan Selén*)
- 2002:3 On the Choice of Sampling Design in Business Surveys with Several Important Study Variables (*Anders Holmberg, Patrik Flisberg and Mikael Rönnqvist*)
- 2002:4 The Sampling- and the Estimation Procedure in the Swedish Labour Force Survey (*Hassan Mirza and Jan Hörngren*)
- 2003:1 Översyn av undersökningen Inrikes och utrikes trafik med svenska lastbilar
(*Johan Eriksson, Per Anders Paulson och Bengt Rosén*)

ISSN 0283-8680

Tidigare utgivna **R&D Reports** kan beställas genom Katarina Klingberg, SCB, MET, Box 24 300, 104 51 STOCKHOLM (telefon 08-506 942 82, fax 08-506 945 99, e-post katarina.klingberg@scb.se). **R&D Reports** from 1988-1999 can - in case they are still in stock - be ordered from Statistics Sweden, attn. Katarina Klingberg, MET, Box 24 300, SE-104 51 STOCKHOLM (telephone +46 8 506 942 82, fax +46 8 506 945 99, e-mail katarina.klingberg@scb.se).