



Stockholms
universitet

Studie av uppgiftslämnarbördan i In- trastat 2011

Konstantin Kalinichenko

Masteruppsats 2011:11
Matematisk statistik
Oktober 2011

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Studie av uppgiftslämnarbördan i Intrastat 2011

Konstantin Kalinichenko*

Oktober 2011

Abstract

En övervägande del av företag och organisationer som handlar med varor med länder inom EU är enligt lag skyldiga att rapportera handelsstatistik till Intrastat, som hanteras i Sverige av Statistiska centralbyrån. Rapporteringen innebär en viss belastning för dessa företag och organisationer, och både statistikproducenten och uppgiftslämnarna är intresserade av att minska denna börda. I samband med detta gjordes en statistisk undersökning av uppgiftslämnarbördan i Intrastat, som utgör huvuddelen av detta examensarbete. Syftet med undersökningen är att mäta den nuvarande uppgiftslämnarbördan i Sverige, undersöka vilka delkomponenter den består av, samt analysera hur bördan kan minskas. Enkäten består av 11 frågor och undersökningen genomfördes via webblankett. En annan del av arbetet utgörs av en statistisk modell, vars syfte är att undersöka bakomliggande orsaker till långa rapporteringstider hos vissa företag och organisationer. Slutligen, har även sambandet mellan några av variabler i undersökningen analyserats med hjälp av parvisa korrelationstest. Huvudresultatet av undersökningen är att den genomsnittliga tiden som behövs för att rapportera till Intrastat per månad har minskat med drygt 30 % jämfört med år 2006, från 169 till 119 minuter. Andel uppgiftslämnare som anser att det är lätt att rapportera till Intrastat har samtidigt minskat från 81,4% till 70,8%.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden.
E-mail:kk1986@gmail.com. Supervisor: Gudrun Brattström.

ABSTRACT

STUDY OF THE REPORTING BURDEN IN INTRASTAT 2011

The majority of companies and organizations in Sweden that are involved in trade with goods and commodities with countries within the EU are legally obliged to report trade data to Intrastat, which is handled by Statistics Sweden. The reporting causes some load on these companies and organizations, and both Statistics Sweden and data providers are interested in reducing it.

In this context, a survey of the reporting burden in Intrastat has been conducted, which forms the main part of this thesis. The purpose of the survey is to measure the burden, to examine what subcomponents it consists of, and to analyze how it can be reduced. The questionnaire consists of 11 questions and the survey was conducted via the web form. Another part of this work consists of a statistical model, whose purpose is to investigate the underlying causes to long reporting times of some companies and organizations. Finally, a number of pairwise correlation tests have been conducted in order to investigate eventual links between some of the variables in the survey.

Main result of the study is that the average time needed to report to the Intrastat per month has decreased by over 30% compared to 2006, from 169 to 119 minutes. Percentage of respondents who believe that it is easy to report to the Intrastat system has at the same time decreased from 81.4% to 70.8%.

FÖRORD

Denna rapport utgör mitt examensarbete för masterexamen i Matematisk Statistik vid Stockholms Universitet. Arbetet är utfört på Statistiska centralbyrån, avdelning Ekonomisk Statistik.

Först och främst, vill jag rikta ett stort tack till min handledare på SCB, Frank Weideskog på PCA/MFFM, för all värdefull hjälp och stöd under arbetets gång.

Jag vill även tacka Gudrun Brattström, min handledare på Matematiska Institutionen på SU, för hjälp med metoddelen av arbetet och värdefulla råd.

Ett stort tack riktas till följande personer på SCB som har medverkat vid förberedelser, enkätutformningen och datainsamlingen avseende den statistiska undersökningen:

Marie George, DFO/IF,
Pekka Koski, ES/UI,
Petter Samuelsson, DFO/IF,
Runo Samuelsson, ES/UI,
Jörgen Sandberg, DFO/IF,
Fredrik Scheffer, PCA/MÅT.

Sist, men inte minst, vill jag tacka Anders Holmberg, U/ARK, för möjligheten att göra detta examensarbete.

INNEHÅLLSFÖRTECKNING

| | | |
|-----|--|----|
| 1 | Introduktion..... | 1 |
| 1.1 | Inledning..... | 1 |
| 1.2 | Syfte och metod..... | 1 |
| 2 | Uppgiftslämnarbördan och Intrastatundersökningen | 2 |
| 2.1 | Definition av uppgiftslämnarbördan..... | 2 |
| 2.2 | Val av mätmetod..... | 3 |
| 2.3 | Beskrivning av Intrastatundersökningen | 4 |
| 2.4 | Uppgiftslämnarbördan i Intrastatundersökningen | 5 |
| 3 | Mätinstrument och urval | 6 |
| 3.1 | Utformande av mätinstrument..... | 6 |
| 3.2 | Undersökningens rampopulation | 7 |
| 3.3 | Beskrivning av rampopulationen..... | 8 |
| 3.4 | Företag med flera tilläggsnummer | 10 |
| 3.5 | Urvalsdesign..... | 10 |
| 3.6 | Utformande av strata..... | 11 |
| 3.7 | Urvalsstorlek..... | 12 |
| 3.8 | Fördelning av urvalet med avseende på strata..... | 14 |
| 4 | Resultat | 15 |
| 4.1 | Genomförande av undersökningen | 15 |
| 4.2 | Analys av enhetsbortfall..... | 16 |
| 4.3 | Analys av partiellt bortfall..... | 18 |
| 4.4 | Resultat av undersökningen: Tidsåtgång..... | 19 |
| | Tidsåtgång och strata..... | 19 |
| | Antal medverkande personer och tidsåtgång..... | 21 |
| | Erfarenhet av rapportering och tidsåtgång..... | 21 |
| | Sätt att hitta varukoder och tidsåtgång..... | 22 |
| | Tillgänglighet av uppgifterna och tidsåtgång | 23 |

| | |
|--|----|
| 4.5 Resultat av undersökningen: Allmän upplevelse..... | 24 |
| Samband mellan allmän upplevelse och tidsåtgång | 24 |
| 4.6 Resultat: övriga variabler | 25 |
| Säkerhet i valet av varukoder..... | 26 |
| Möjligheten att rapportera i Euro | 26 |
| Införande av förenklad varukod för låga fakturavärden | 27 |
| Användbarhet av statistiken..... | 27 |
| Kommentarer och förslag..... | 28 |
| 4.7 Jämförelse med Undersökningen av Uppgiftslämnarbördan 2006..... | 28 |
| 5 Bearbetning och statistisk analys..... | 32 |
| 5.1 Anmärkningar gällande variabeln tidsåtgång..... | 32 |
| 5.2 Beskrivning av statistiska metoder | 33 |
| 5.3 Omkodning av variabeln "Framtagning av uppgifter" | 35 |
| 5.4 Implementering av hypotestest | 35 |
| Tidsåtgång – Allmän upplevelse..... | 36 |
| Tidsåtgång – Framtagning av uppgifter | 36 |
| Tidsåtgång – Antal medverkande | 36 |
| Allmän upplevelse – Antal medverkande..... | 37 |
| Tidsåtgång – Säkerhet i valet av varukoder | 37 |
| Allmän upplevelse – Säkerhet i valet av varukoder..... | 37 |
| 6 Undersökning av orsaker till exceptionellt lång rapporteringstid för vissa uppgiftslämnare: en statistisk modell | 39 |
| 6.1 Inledande anmärkningar | 39 |
| 6.2 Beskrivning av modellen..... | 40 |
| 6.3 Modellval samt resultat | 43 |
| Modell A: Samtliga effekter och samspel..... | 43 |
| Modell B: flöde; medium; antal varuposter; antal varuposter*medium | 44 |
| Modell C: medium; antal varuposter; antal varuposter*medium | 44 |
| 7 Diskussion och slutsatser | 47 |
| 7.1 Slutsatser | 47 |

| | |
|---|----|
| 7.2 Diskussion | 49 |
| Appendix | 51 |
| A. Exakt formulering av frågor (svenska)..... | 51 |
| B. Exakt formulering av frågor (engelska) | 52 |
| C. Lista över tabeller..... | 54 |
| D. Lista över figurer | 55 |
| E. Missivbrev (svenska) | 56 |
| F. Missivbrev (engelska)..... | 57 |
| Litteraturlförteckning | 58 |

1 INTRODUKTION

1.1 INLEDNING

Majoriteten av värdemässigt stora företag och organisationer som är inblandade i varuhandel inom Europeiska Unionen är enligt lag skyldiga att rapportera uppgifter avseende in- och utförsel av varor. Detta gäller samtliga medlemsländer och styrs av Intrastat, som är ett system för insamling av statistik över handel inom EU. Intrastat hanteras i Sverige av SCB.

Rapporteringen av statistikuppgifter är ofta tidskrävande för uppgiftslämnarna. Enligt en undersökning av uppgiftslämnarbördan i det svenska Intrastatsystemet som gjordes år 2006, tog det i genomsnitt mellan 2 och 3 timmar per företag att rapportera handelsstatistik varje månad. Denna tidsåtgång kan och bör minskas, vilket är av intresse för både uppgiftslämnarna och statistikproducenten.

Det ska konstateras att SCB bör enligt lag arbeta för att underlätta uppgiftslämnararbete (1). Enligt Regleringsbrev från år 2007, bör företagens kostnader i samband med uppgiftsredovisningen minska med minst 25 % fram till år 2010.

1.2 SYFTE OCH METOD

I dagslägen är det av intresse att undersöka den nuvarande situationen i fråga om uppgiftslämnarbördan i Sverige. Mätning av denna kan genomföras genom en statistisk undersökning bland uppgiftslämnarna. Genom att formulera frågor på ett lämpligt sätt, kan man belysa de olika dimensionerna som uppgiftslämnarbördan består av. Det är dessutom av intresse att undersöka uppgiftslämnarnas attityd till möjliga förenklingsalternativ (enligt EUs föreskrift). Enkäten skulle därför kunna innehålla ett antal frågor avseende dessa.

Vissa delkomponenter av uppgiftslämnarbördan kan anses samspela med varandra. Samband mellan dessa delkomponenter bör undersökas med hjälp av korrelationstest.

Det är dessutom känt att tidsåtgång i samband med rapportering till Intrastat är betydligt högre än genomsnittet för vissa uppgiftslämnare. Bakomliggande orsaker till dessa skillnader ska undersökas genom att framställa en statistisk modell.

Vi fortsätter med att definiera begreppet uppgiftslämnarbördan och att ge en detaljerad beskrivning av denna.

2 UPPGIFTLÄMNARBÖRDAN OCH INTRASTATUNDERSÖKNINGEN

2.1 DEFINITION AV UPPGIFTLÄMNARBÖRDAN

Ett företag eller en organisation med ett unikt organisationsnummer och tilläggsnummer¹ kallar vi för en uppgiftslämnare.

Framtagning, rapportering och eventuell bearbetning av handelsstatistik till SCB innebär en viss belastning för uppgiftslämnarna. Den kallar vi för uppgiftslämnarbörda.

Uppgiftslämnarbördan är i sig ett brett och abstrakt begrepp, men den kan brytas ned i mindre komponenter. Först och främst, kan man dela upp den totala bördan i en objektiv del, och en subjektiv del (2).

Den objektiva delen av uppgiftslämnarbördan innebär främst den delen av uppgiftslämnarbördan som innebar finansiella kostnader för uppgiftslämnare. Den absolut största delen av dessa kostnader innebär personalkostnaderna, så det är rimligt att betrakta den objektiva delen av uppgiftslämnarbördan som tidskostnad som företag och organisationer får stå för i samband med statistikrapporteringen. Tidskostnaden beror i sin tur på hur lång tid det tar per undersökning (tid) samt hur ofta undersökningarna genomförs (frekvens). Det har visat sig att objektivt mätt börda inte kan fånga helheten av den bördeupplevelse som företag och organisationer går igenom vid framtagning av uppgifter.

Den subjektiva komponenten är i sin tur ett bredare begrepp. Vad som menas här är den upplevelsen som uppgiftslämnare får gå igenom då de besvarar frågorna. En mängd olika faktorer inom företag/organisationer påverkar denna, och dessa är följande (2) (3) (4):

- Den mentala ansträngningen som behövs för uppgiftsrapporteringen
- Stress man får eventuellt gå igenom
- Kvalitet på intern redovisning i ett företag/en organisation
- Redovisningssätt (elektroniskt eller via pappersblankett)
- Tidigare erfarenhet av rapportering av uppgiftslämnarbörda
- Motivation
- Antal medverkande personer
- Vana vid att arbeta med datorer, datorstöds kvalitet på företaget

Vidare, kan man även se bördan i ett ännu bredare perspektiv som ett samspel mellan producenten, uppgiftslämnarna och statistikbeställarna (5) (6). Vi väljer dock att begränsa oss till den definitionen av bördan som föreslås av (2), (3) och (4).

Slutligen konstaterar vi att i samband med undersökningen av uppgiftslämnarbördan kan det även vara intressant att ta reda på dess bakomliggande orsaker samt undersöka var behov av insatser finns.

¹ Ett företag/en organisation med ett unikt organisationsnummer kan ha flera tilläggsnummer, vilket kan vara praktiskt då t.ex. ett företag har flera filialer runt om i landet.

2.2 VAL AV MÄTMETOD

Den objektiva komponenten av uppgiftslämnarbördan kan både under- och överskattas av uppgiftslämnare beroende på en mängd faktorer, såsom den subjektiva delen av bördan (stress, ansträngning osv.) och den mänskliga faktorn. Exempelvis kan alltför negativ (positiv) inställning till statistiska undersökningar generellt, avsaknad av (överdriven) motivation eller en alltför pessimistisk (optimistisk) personlighetstyp ha effekt på uppfattningen av den objektiva bördekomponenten.

Vid mätning av uppgiftslämnarbördan ska det helst användas en mätmetod som minimerar påverkan av de ovannämnda faktorerna på den objektiva komponenten av bördan. En metod som har visat sig uppfylla detta krav är den så kallade Standardkostnadsmodellen (4).

Denna mätmetod har sitt ursprung i Holland och fått stor spridning i Europa. Implementeringen av Standardkostnadsmodellen sker genom att man med hjälp av djupgående intervjuer uppskattar den faktiska belastningen som uppgiftslämnare får stå för. Man gör en del antaganden också, bl.a. att de undersökta företagen ska vara normaleffektiva², samt att "full-compliance"³ principen gäller. Vi går inte in i detaljerna här, men noterar att metoden inte grundar sig på statistiskt framtagna mått utan på kvalitativa bedömningar utifrån expertutlåtanden samt företagsintervjuer (7).

Detta är en av anledningarna till varför Standardkostnadsmodellen inte passar in i ramen för detta examensarbete. En annan anledning är att denna avser ett kostsamt sätt att mäta uppgiftslämnarbördan.

Mätningen av uppgiftslämnarbördan i Intrastat kommer här att göras med hjälp av en statistisk undersökning. Eftersom den objektiva delen av bördan kommer att skattas utifrån uppgiftslämnarnas svar, kan vi inte eliminera effekter av de ovannämnda faktorerna helt. Emellertid kan vissa åtgärder vidtas för att minska eventuell påverkan av dessa.

För det första, bör frågorna formuleras på ett lämpligt sätt, nämligen så neutralt som möjligt. Detta för att undvika att frågorna blir "laddade", vilket kan influera uppgiftslämnarnas svar.

För det andra, bör enkäten innehålla rimligt antal frågor. För många frågor i enkäten ökar uppgiftslämnarnas börda och kan också påverka deras svar.

För det tredje, bör urvalet vara tillräckligt stort för att minimera eventuell påverkan av uppgiftslämnare som avviker i termer av den mänskliga faktorn.

Slutligen, bör ett antal parvisa korrelationstest göras för att undersöka samband mellan den objektiva komponenten av uppgiftslämnarbördan (främst tidsåtgång) och de

² Ett normaleffektivt företag är ett företag som varken är ovanligt effektivt som ovanligt ineffektivt i att följa lagkraven.

³ "Full compliance" innebär att alla företag som berörs av ett lagkrav uppfyller sina åtaganden fullt ut.

subjektiva delkomponenterna av den. Detta skulle utöka vår förståelse angående beroendet mellan dessa faktorer.

2.3 BESKRIVNING AV INTRASTATUNDERSÖKNINGEN

Förutom den börda som uppgiftsskyldiga företag och organisationer upplever i samband med redovisning till statistikproducenten i allmänhet ska vi här studera den specifika belastningen som upplevs i samband med Intrastatundersökningen. Vidare följer en detaljerad beskrivning av denna.

Insamling av statistikuppgifter i Intrastatundersökningen sker genom att uppgiftslämnare rapporterar data avseende in- och utförsel av varor. Denna procedur sker månadsvis och avser samtliga importörer och exportörer vars handelsomsättning med andra EU länder överstiger 4.5 miljoner kronor, i antingen in- eller utförsel. Detta tröskelvärde är beräknat så att 95 % av all införsel och minst 97 % av all utförsel täcks in av undersökningen.

Implementeringsproceduren för uppgiftsrapporteringen är EU-reglerad (8) (9). Ett antal faktorer kan dock bestämmas på nationell nivå. Uppgifterna ska lämnas in tolv dagar efter respektive månads slut i fall rapporteringen sker via pappersblankett, eller arton dagar efter respektive månads slut då rapporteringen sker elektroniskt. En skriftlig påminnelse skickas till de uppgiftslämnarna som ej har rapporterat i tid. I de fall uppgifter inte lämnats trots påminnelse, kan företaget/organisationen anmälas enligt 20 § lagen (2001:99) om den officiella statistiken.

I Sverige har man beslutat att enbart kräva redovisning av de variabler som ska rapporteras enligt EU-förordningar. Dessa är följande:

- Flöde (in- eller utförsel)
- Varukod (ca 9000 åttasiffriga varukoder, se exempelvis (10) för en lista över alla varukoder)
- EU-landskod (26 medlemsländer)
- Värde (fakturerat värde)
- Vikt (gäller ej alla varukoder)
- Annan kvantitet (kvantitet, mått för den specifika varan, t.ex. m² för vävnader, avser ej alla varukoder)
- Transaktionstyp

En unik kombination av variablerna ovan kallar vi för en varupost. Det är alltså en specifik vara med specifik transaktionstyp (köp, retur osv.), och ursprungs/ destinationsland som rapporteras vid ett tillfälle. Vi kommer att använda denna definition i fortsättningen.

Rapporteringen kan ske på fyra olika sätt. Dessa är följande:

- Genom att använda en kostnadsfri programvara IDEP (Intrastat Data Entry Package). Cirka 53% av uppgiftslämnare använder detta rapporteringssätt.
- Via webblankett. Detta sätt för att rapportera används av ca 29% av rapportörerna

- Genom att använda annan godkänd programvara (Vi kommer att kalla det sättet "övriga media" i fortsättningen). Antal uppgiftslämnare som rapporterar på detta sätt är mycket litet (ca 2%).
- Via pappersblankett. 16 % av uppgiftslämnare föredrar detta rapporteringsätt.

Det bör även noteras att handel av tjänster inte ska rapporteras till Intrastat. Inte heller reparationer och elektrisk energi ska rapporteras.

Varor som transporteras till/från Sverige som avser katastrofhjälp eller är skyddade av diplomatisk immunitet ska inte heller rapporteras. För ett antal varor (fartyg, flygplan, rymdfarkoster, sjöanläggningar och annat) gäller särskilda bestämmelser. För en detaljerad beskrivning av hur rapporteringsprocessen bör ske, se (11).

2.4 UPPGIFTSLÄMNARBÖRDAN I INTRASTATUNDERSÖKNINGEN

De tidigare undersökningarna av uppgiftslämnarbördan i Intrastat (2000, 2002, 2006) har påvisat att den delen av Intrastatundersökningen som kan anses vara mest betungande är framtagning av varukoder. Det finns totalt ca 9 000 åttasiffriga varukoder, och många är svåra att särskilja.

Exempel på varukoder:

| | |
|-----------------------------------|-----------|
| Rostfritt stål, i obearbetad form | 7218 1000 |
| Diskmaskiner av hushållstyp | 8422 1100 |
| Klänningar av bomull | 6204 4200 |

Man kan även nämna olika teoretiska möjligheter (enligt EU-förordningar) till att minska bördan i samband med Intrastatundersökningen. En del av dessa är:

- Införa förenklade varukoder, dvs. införa en "samlingskod" vid låga fakturavärden
- Endast samla in data för ett flöde (utförsel)
- Utveckla den elektroniska rapporteringen
- Rapportera det fakturerade värdet i Euro
- Inte behöva rapportera vikt om man redovisar annan kvantitet för varukoden i fråga

3 MÄTINSTRUMENT OCH URVAL

3.1 UTFORMANDE AV MÄTINSTRUMENT

Vi börjar med att konstatera att undersökningen kommer att ske via webblankett. Man kommer först att skicka ut missiv med behörigheter (användar-id och lösenord) till de företagen/organisationer som har hamnat i undersökningens urval. Personerna som är ansvariga för Intrastatrapportering hos uppgiftslämnarna kommer i sin tur att ange dessa inloggningsuppgifter på Intrastats hemsida⁴ för att komma åt webblanketten.

Missivbrevet kommer att finnas i två versioner – på svenska och på engelska. De av uppgiftslämnarna som befinner sig i Sverige kommer att få den svenska versionen, och de som befinner sig utomlands kommer att få den engelska versionen. Båda versioner av missivbrev kan hittas i Appendix.

Den främsta anledningen till att man kommer att genomföra undersökningen via webblankett är att det är mer kostsamt att implementera undersökningen via papper. Dessutom så slipper man arbete i samband med registrering av pappersblanketter.

Vi fortsätter med att bestämma frågorna som kommer att inkluderas i enkäten. Ur statistikproducentens perspektiv är det naturligt att vilja ställa så många frågor så möjligt, dock bör man även tänka på att alldeles för många frågor kommer att leda till ökad belastning för uppgiftslämnarna. Detta kan i sin tur leda dels till ökat bortfall och dels till att uppgiftslämnarnas börda ökar. Dessutom så kan en del variabler och parametrar skattas utifrån data som redan finns tillgängligt.

Nedan följer en lista över de variablerna som vi har bestämt oss att ta med:

1. Den allmänna upplevelsen i samband med rapporteringen (kategorisk variabel – B)
2. Antal personer involverade i uppgiftsrapporteringen (B)
3. Tiden det tar för att rapportera till Intrastat per redovisningsmånad (kvantitativ variabel – K)
4. Erfarenhet av rapportering till Intrastat (B)
5. Sätt att hitta varukoder (B; med flera nivåer)
6. Säkerhet i valet av varukoder (B)
7. Tillgängligheten av de efterfrågade uppgifterna för personer som är med i rapporteringsprocessen (B)
8. Förslag till förbättring: Förenkling av varukoder (B)
9. Förslag till förbättring: Möjlighet att rapportera det fakturerade värdet i Euro (B)
10. Uppfattning om relevans av den framställda statistiken för uppgiftslämnarens verksamhet (B)
11. Kommentarer (Text).

Dessa variabler antingen ingår i uppgiftslämnarbördan eller innebär de förenklingsförslagen som anses vara implementerbara i den närmaste framtiden.

⁴ www.insamling.scb.se

Variabel 3 är ett mått på den objektiva delen av bördan (tidsåtgång).

Variablerna 1,2,4 -7 samt 10-11 har som syfte att reflektera den subjektiva delen av bördan. Nämligen,

- Den mentala ansträngningen som behövs för uppgiftsrapporteringen, samt stress man får eventuellt gå igenom speglas av variabel 1.
- Kvalitet på intern redovisning hos uppgiftslämnarna, samt vana vid att arbeta med datorer och datorstöds kvalitet på företaget mäts av variablerna 5 samt 7
- Motivation reflekteras delvis av variabel 10.
- Antal personer som är involverade i rapporteringsprocessen avspeglas av variabel 2.
- Tidigare erfarenhet av rapportering av uppgiftslämnarbördan reflekteras av variabel 4.
- Variabel 6 är inkluderad för att täcka den delen av Intrastatundersökningen som anses vara svårast – framtagning av varukoder.

Variabler 8 och 9 reflekterar möjliga förslag på förenklingar som eventuellt kan implementeras i den närmaste framtiden.

När det gäller redovisningssätt (webb, papper, IDEP, eller övriga media), är denna information tillgänglig för SCB, och den kommer vi att ta hänsyn till vid stratifiering av urvalet. När det gäller förenklingsförslagen om att enbart rapportera ett flöde och inte behöva rapportera vikt, kommer inte dessa att kunna tillämpas inom den närmaste framtiden. På grund av detta inkluderas dessa inte i enkäten.

Slutligen, kan de exakta frågeformuleringarna på svenska och på engelska hittas i Appendix.

3.2 UNDERSÖKNINGENS RAMPOPULATION

Som konstaterades tidigare, är moderundersökningen (Intrastat) en cut-off⁵ undersökning, där trösklarna är beräknade så att täckningsnivån ska motsvara 95 % av all införsel och minst 97 % av all utförsel år 2010. Detta motsvarar ett tröskelvärde lika med 4.5 miljoner kronor för både införsel och utförsel av varor inom EU. Samtliga företag och organisationer vars handelsomsättning (export och/eller import) inom EU överstiger detta värde är därmed uppgiftsskyldiga och utgör populationsramen för denna undersökning. Bland dessa uppgiftslämnare, kommer vi att basera urvalet på de som har rapporterat in- eller/och utförsel under december 2010. Antalet sådana uppgiftslämnare är 8601.

Ett antal uppgiftslämnare bör exkluderas ur undersökningen. Vidare beskriver vi kriterier för att bli exkluderad samt orsaker till det.

⁵ En cut-off undersökning är en undersökning där alla enheter som är mindre än en viss storlek (i det här fallet om handelsomsättningen understiger ett visst tröskel) uteslutas ur stickprovet

- För det första, har vi bestämt oss att exkludera de företag som började rapportera till Intrastat i oktober 2009 eller senare. Detta görs för att säkerställa att urvalspopulationen enbart består av erfarna uppgiftslämnare. Antal företag som berörs av detta exklusionskriterium är 1115.
- En del företag som är uppgiftsskyldiga rapporterar via ombud. Detta innebär att uppgiftslämnare låter andra specialiserade företag ta hand om rapporteringen till Intrastat mot betalning. Vi kommer inte att inkludera sådana företag i undersökningen eftersom flera företag kunde ha anlitat ett och samma ombud, därmed inget oberoende mellan populationsenheter. Ett annat argument för exklusionen av dessa uppgiftslämnare är att rapportering via ombud innebär en viss ekonomisk kostnad, som är en del av uppgiftslämnarbördan. Om inte den kostnaden mäts så kommer börden att underskattas, i och med att man kommer då att göra ett (felaktigt) antagande om att det inte kostar något att anlita ombud. Antal företag som faller i denna kategori är 688.
- Vi utesluter även de företagen eller organisationer som enbart handlar med varor som tillhör kategorin särskilda varuförflyttningar (fartyg, flygplan mm.). De varuposter som tillhör denna kategori räknade man inte heller med. Huvudskälet till detta är att handel med varor som tillhör kategorin särskilda varuförflyttningar ska rapporteras enligt "särskilda förenklade regler"⁶ till Intrastat, dock ingår sådana uppgiftslämnare i en separat undersökning. Möjligheten att uppskatta antal sådana företag saknas, dock är det väldigt litet.
- Slutligen, i fall ett företag har använt fler än ett rapporteringsmedium för att rapportera något av flödena det har (införsel eller utförsel), kommer vi inte att analysera data som avser detta flöde. Exempelvis, om något företag har använt ett medium för att rapportera införsel men två olika media för att rapportera utförsel, kommer enbart införseldelen att vara med i urvalet. Anledningen till detta är främst att det inte är uppenbart hur man ska välja rätt stratum för dessa företag. Å ena sidan är antalet sådana företag väldigt litet (se nästa sida), vilket omöjliggör skapande av ett separat stratum för dessa. Å andra sidan är det inte korrekt att lägga dessa i något annat stratum, eftersom sådana företag kan anses skilja sig väldigt mycket från de andra med avseende på uppgiftslämnarbördan. Antal företag som har använt fler än ett rapporteringsmedium vid redovisning för åtminstone något av flödena (införsel eller utförsel) är 12. Antal företag som har använt två eller fler rapporteringsmedia för samtliga flöden är 5.

3.3 BESKRIVNING AV RAMPOPULATIONEN

Vi fortsätter med att illustrera hur de angivna exklusionskriterierna påverkar rampopulationen. Notera att vi även har tagit hänsyn till eventuella dubblettposter vid daraaggregeringen. Dubblettrapportering inträffar då försäljning eller inköp av samma vara med samma flöde, destinations/ursprungsland och transaktionstyp rapporteras två

⁶ Se www.intrastat.scb.se för mer detaljer

eller fler gånger av samma organisationsnummer/tilläggsnummer under samma tidsperiod (månad och år). En tabell med exakta siffror kan hittas nedan. Företag som enbart sysslar med särskilda förflyttningar, varuposter som avser sådana varor samt eventuella nollrapporter är inte inkluderade i tabellen.

Tabell 1: Effekt av avgränsningar på rampopulationen

| | <i>Antal företag (unika orgnr.)</i> | <i>Antal varuposter</i> |
|---|-------------------------------------|-------------------------|
| <i>Urvalspopulationen för moderundersökningen, dec 2010</i> | 8601 | 365674 |
| <i>Utesluter de uppgiftslämnarna som började rapportera under oktober 2009 eller senare</i> | 7486 | 348925 |
| <i>Utesluter de företag som rapporterar via ombud</i> | 6798 | 301971 |
| <i>Aggregerar data m.a.p. dubblettposter</i> | 6798 | 225951 |
| <i>Utesluter de företag som har använt fler än ett rapporteringsmedium för samtliga flöden de har⁷</i> | 6793 | 225234 |

Antal företag som började rapportera under oktober 2009 eller senare är alltså 1115 (16749 varuposter). Antal företag som rapporterade via ombud under december 2010 var 688 med 46954 varuposter.

Rampopulationen består av 6793 företag med 225234 unika varuposter. Bland dessa finns det:

- 74 företag som har använt fler än ett tilläggsnummer för att rapportera under referensperioden. Antal tilläggsnummer som dessa företag disponerar är 253.
- 126 företag som har använt media klassade som övriga för att rapportera till Intrastat.

De 6593 företagen som inte tillhör till någon av de två grupperna fördelar sig på följande sätt med avseende på flöde (se tabell nedan):

Tabell 2: Fördelning av uppgiftslämnare med avseende på flöde (ej flera tilläggsnummer, övriga media)

| <i>Flöde</i> | <i>Antal företag</i> | <i>Procentuell fördelning</i> |
|-------------------------|----------------------|-------------------------------|
| <i>Införsel</i> | 3358 | 50.93 % |
| <i>Utförsel</i> | 1486 | 22.54 % |
| <i>In- och utförsel</i> | 1749 | 26.53 % |
| <i>Totalt</i> | 6593 | 100 % |

För varje flöde, kan man dela in urvalspopulationen ytterligare med avseende på rapporteringsmedium samt antal varuposter per företag.

⁷ Se beskrivning på s.10

3.4 FÖRETAG MED FLERA TILLÄGGSNUMMER

Ett antal uppgiftslämnare använder fler än ett tilläggsnummer för att rapportera till Intrastat. Att ha flera tilläggsnummer kan vara fördelaktigt då ett företag har filialer runt om i landet till exempel. Det som är viktigt här dock är att varje tilläggsnummer bör rapportera till Intrastat självständigt om moderföretaget faller in i ramarna för moderundersökningen. Det orsakar vissa komplikationer eftersom tilläggsnummer inom ett företag inte är oberoende. Därför väljer vi att dela upp rampopulationen i två grupper där en av grupperna kommer att innehålla de företag som har använt ett tilläggsnummer för att rapportera (varje uppgiftslämnare tilldelas alltid *ett* tilläggsnummer), och en annan grupp kommer att bestå av de företag som har använt två tilläggsnummer eller fler under referensperioden.

De senare ska behandlas genom att för varje organisationsnummer enbart välja ut det tilläggsnummer som har störst antal varuposter. Fördelen med detta är att man slipper beroendet mellan tilläggsnummer inom företag, samtidigt som man inte riskerar underskatta uppgiftslämnarbördan. Ett ytterligare argument för detta är att en större del av företagen med flera tilläggsnummer rapporterar majoriteten av sin handel med EU via enbart ett av sina tilläggsnummer. För att vara mer precis, 60 av 74 företag, se tabell nedan.

Tabell 3: Procentuellt antal varuposter som rapporteras via tilläggsnummer med störst antal varuposter; avser uppgiftslämnare som använde flera tilläggsnummer för att rapportera under dec 2010

| <i>% av varuposter</i> | <i>Antal företag (orgnr.)</i> | <i>Procentuell fördelning</i> |
|------------------------|-------------------------------|-------------------------------|
| <i>Mindre än 50 %</i> | 14 | 18.9 % |
| <i>50 – 75 %</i> | 35 | 47.3 % |
| <i>Mer än 75 %</i> | 25 | 33.8 % |

3.5 URVALSDESIGN

Vid genomförande av en statistisk undersökning kan man välja mellan att genomföra det som en totalundersökning eller som en urvalsundersökning. Fördelar med en totalundersökning är att man får bättre skattningar samt att arbete i samband med urvalsdesign inte behöver göras. Dock innebär totalundersökning ökad belastning för uppgiftslämnare, vilket är ett huvudskäl till att vi väljer att genomföra denna studie som en urvalsundersökning.

Vid implementering av en urvalsundersökning bör man bestämma en metod för stickprovsdragning. De vanligaste metoderna är:

- 1) Enkelt slumpmässigt urval
- 2) Proportionellt stratifierat urval
- 3) Optimalt stratifierat urval

Vi konstaterar först att om det finns variabler som kan anses vara korrelerade med den studerade variabeln, borde man stratifiera med avseende på dessa eftersom detta resulterar i ökad precision av skattningarna (12). Det är känt ifrån tidigare

undersökningar av uppgiftslämnarbördan att sådana variabler finns, och vi förkastar därför alternativet enkelt slumpmässigt urval.

Det återstår nu att välja en allokeringmetod för strata. Det bör noteras att generellt så är optimalt urval att föredra framför exempelvis proportionellt urval. Tillämpning av denna kräver dock att stratavariansernas inbördes förhållande är känt (12). I fall s^2 (variansen) kan bedömas ur tidigare undersökningar eller provundersökningar är optimalt urval att föredra. Det skall emellertid observeras att om man helt felbedömer stratavariansernas inbördes förhållande, kan man få skattningar med sämre precision än vid enkelt slumpmässigt urval ur totalpopulationen (12).

I vårt specifika fall är kunskap om stratavarianserna väldigt liten. I och med att strata kommer att konstrueras på nytt i den här undersökningen, är det inte möjligt att utnyttja information angående stratavarianser från de tidigare undersökningarna av uppgiftslämnarbördan. Utifrån detta väljer vi att använda proportionellt stratifierat urval i vår undersökning.

3.6 UTFORMANDE AV STRATA

Vi noterar först att rampopulationen är indelad i två grupper, beroende på om ett företag har använt fler än ett tilläggsnummer för att rapportera eller inte.

När det gäller företag med flera tilläggsnummer (74 st.), kommer vi inte att göra någon ytterligare indelning av denna grupp. I stället kommer vi att enbart välja ut de tilläggsnumren som har störst antal varuposter per flöde. Vidare kommer man att göra ett enkelt slumpmässigt urval på kombinationer organisationsnummer och tilläggsnummer.

Företag som har använt övriga media för att rapportera (126 st.) kommer att utgöra ett separat stratum också. I fall ett företag använder ett rapporteringsmedium klassat som övrigt för ett av flödena, men ett annat medium för det andra flödet, kommer vi enbart att ta hänsyn till det flödet som rapporterades via övriga media. Det andra flödet kommer vi att exkludera ur urvalet. Anledningen till att vi prioriterar övriga media på det här sättet är att vi inte vill utesluta några företag ur urvalet i denna grupp, som redan är ganska liten, samtidigt som att vi inte kan inkludera båda flödena i urvalet eftersom dessa inte är ömsesidigt oberoende.

I fall ett företag rapporterar både in- och utförsel av varor med hjälp av övriga media, kommer vi att slumpmässigt välja ett av flöden och bortse från det andra.

Den resterande urvalspopulationen (6593 st.) kommer att stratifieras i tre dimensioner. Dessa är:

- Flöde: införsel och utförsel. För de företag som rapporterar både in- och utförsel, kommer vi slumpmässigt välja ett av flöden som ska inkluderas i urvalet, för att utesluta möjlighet att ett företag är med i undersökningen två gånger.
- Rapporteringsmedium: IDEP, pappersblankett, och webblankett.
- Antal varuposter (mångfald av rapporterade varor): Små, Medelstora, och Stora företag. Dessa klasser väljs med avseende på antal varuposter som ett företag rapporterade under december 2010. Kategoriramen är 0-5 varuposter/år för små

företag, 6-24 för medelstora företag, samt 25+ för stora företag. Notera att denna kategorisering inte tar hänsyn till transaktionsvärden, så ett företag med hög omsättning men med liten mångfald av varor och transaktionstyper skulle förmodligen hamna i den första kategorin.

Antal företag som faller i de olika strata varierar eftersom det beror på vilka flöden som väljs ut vid bearbetning av företag som rapporterar både in- och utförelse. Fördelning av uppgiftslämnare med avseende på flöde i aggregerad form kan hittas i tabell 3. De olika strata ser ut på följande sätt (Se Tabell 4).

Tillsammans med företag som rapporterar via övriga media (126 st.) samt företag som använde fler än ett tilläggsnummer för att rapportera under referensperioden (74 st.) blir rampopulationen 6793 st. företag.

Tabell 4: Fördelning av rampopulationen med avseende på antal varuposter och rapporteringsmedium (ej unika strata)

| Företag/Medium | IDEP | Papper | Webb | Totalt |
|------------------|------|--------|------|--------|
| Små (0-5 poster) | 1133 | 494 | 1069 | 2696 |
| Medel (6-24) | 1525 | 396 | 808 | 2729 |
| Stora (25+) | 900 | 182 | 86 | 1168 |
| Totalt | 3558 | 1072 | 1963 | 6593 |

3.7 URVALSSTORLEK

Det totala urvalet består av tre delar. Vi kallar dem för basdelen, utökade delen samt "förstärkningsdelen".

Basdelens storlek bestämmer vi enligt följande. Antag att vi vill estimerar väntevärde (populationsandelen som svarar positivt) av en binär variabel och framställa ett konfidensintervall för detta. I vår studie har vi inga binära variabler dock har vi ett antal kategoriska variabler, t.ex. Allmän Upplevelse, som kan enkelt transformeras till en binär. Genom att i förhand bestämma den önskade längden av konfidensintervallet samt signifikansnivå kan man räkna ut hur många observationer man behöver.

Vi bestämmer att signifikansnivån ska vara $\alpha=0.05$, samt att halva längden av konfidensintervallet ska vara $d=0.03$. Alltså om vi estimerar väntevärdet av andelen till t.ex. 0.6, får inte konfidensintervallet vara bredare än $[0.57 ; 0.63]$. Vi antar även att skattning på väntevärdet kommer att ligga runt 0.5, därmed antar vi att variansen blir störst av samtliga möjliga (se nedan).

Det bör tilläggas att framställning av konfidensintervall för populationsandelar som svarar positivt på någon av frågorna inte är någonting som vi kommer att göra vid analys av svarsdata senare. Dock anser vi att den här metoden är lämplig för att bestämma urvalsstorleken så att det kommer att finnas tillräckligt med observationer för hypotesprövningar.

Vi definierar \bar{x} som det okända medelvärdet av den populationen vi vill undersöka. Låt vidare N stå för det totala antalet enheter i urvalspopulationen, n för antal enheter som kommer att väljas ut, samt m för det aritmetiska medelvärdet av den parameter som man

vill undersöka. Då rampopulationen är tillräckligt stor och är betydligt större än urvalet (12) kan vi utnyttja centrala gränsvärdeessatsen och konstatera att

$$\bar{x} \xrightarrow{d} N\left(m, \frac{S^2}{n} \left(1 - \frac{n}{N}\right)\right) \quad [1]$$

Här betecknar S populationsvariansen, som kan i sin tur estimeras som \sqrt{pq} , där $p =$ populationsandelen som svarar positivt, och $q = 1 - p$. Då $p = q = 0.5$, kommer även S att vara lika med 0.5 (det största möjliga värdet på S).

Vidare följer det ur [1] att ett konfidensintervall för den andelen som vi vil skatta kan konstrueras som

$$\bar{x} \pm \lambda_{\alpha} \frac{S}{\sqrt{2}} \sqrt{\frac{N-n}{Nn}} \quad [2]$$

Ur detta följer sambandet

$$\lambda_{\alpha} \frac{S}{\sqrt{2}} \sqrt{\frac{N-n}{Nn}} = d \quad [3]$$

Detta kan skrivas om ytterligare. Man får

$$n = \frac{n_0 N}{n_0 + N}; \quad n_0 = \left(\frac{\lambda_{\alpha} S}{d}\right)^2$$

I vårt fall har vi att $N = 6793$, $S = 0.5$, $d = 0.03$ samt $\lambda_{\alpha} = 1.96$. Det leder till att n_0 blir lika med 1 067,11, som i sin tur medför att den nödvändiga urvalsstorleken n blir 922.23, eller 923 efter avrundning uppåt.

Basdelen av urvalet består alltså av 923 enheter. Ett problem man inte får bortse från är eventuellt bortfall. Vid undersökningen år 2006 var bortfallet så litet som 15 %. Vid en ännu tidigare undersökning av uppgiftslämnarbördan, som skedde år 2000, var dock bortfallet betydligt högre – svarsfrekvensen uppgick då enbart till 68 % i införseln och till 54 % i utförseln. Vi väljer därför att utöka urvalet till 1500 företag. Givet att eventuellt bortfall ska fördela sig jämt mellan de olika strata, kommer vi med ett bortfall så stort som 38,5 %. $(1 - 922/1500)$ fortfarande få det önskade antalet svar.

De två delarna av urvalet som beskrevs ovan kommer att fördelas mellan strata enligt den proportionella urvalsstrategin. Detta innebär att man väljer urvalsstorlekar för de olika strata proportionellt mot stratans storlekar, under bivillkoret att summan av samtliga delurvalsstorlekar bör vara lika med den totala urvalsstorleken. För en bredare beskrivning av proportionellt stratifierat urval hänvisar vi till (12) eller (13).

Det finns dock ett par ytterligare moment man bör ta hänsyn till. För det första, är en del strata väldigt små, vilket innebär att det finns risk att man får oacceptabelt litet antal observationer i dessa. För det andra, kan vi förvänta oss att variationen av tidsåtgång för

att rapportera kommer att vara större bland företag med många varuposter än bland företag med få varuposter.

Med tanke på detta har vi bestämt oss att utöka urvalsstorleken ytterligare med 510 enheter. Dessa kommer att fördelas främst över de mindre strata samt på de strata som avser företag med många varuposter. Det är alltså den förstärkande delen av urvalet.

Sammanfattningsvis illustrerar vi struktur av urvalet i Tabell 5 nedan.

Tabell 5: Fördelning av urvalet

| <i>Del av urvalet</i> | <i>Antal enheter</i> | <i>Fördelning bland strata</i> |
|---------------------------|----------------------|---|
| <i>Basdelen</i> | 923 | Proportionellt mot stratastorlekar |
| <i>Utökade delen</i> | 587 | Proportionellt mot stratastorlekar |
| <i>Förstärkningsdelen</i> | 510 | Manuellt tillägg (se nedan). När det gäller fördelningen m.a.p. flöde, kommer urvalsenheterna att fördela sig enligt proportionellt stratifierat urval. |
| <i>Totalt i urvalet</i> | 2012 | |

3.8 FÖRDELNING AV URVALET MED AVSEENDE PÅ STRATA

Vi beskriver hur vårt urval (2012 uppgiftslämnare) fördelar sig med avseende på de olika strata. Notera att vi ej gör fördelningen med avseende på flöde ännu eftersom fördelningen av företag m.a.p. flöde är delvis slumpmässig. Se Tabell 6.

Tabell 6: Fördelning av urvalet med avseende på strata (ej flöde)

| <i>Stratum</i> | <i>Antal enheter i stratumet</i> | <i>Antal enheter i urvalet före förstärkningen</i> | <i>Antal enheter tilldelade ur förstärkningsdelen av urvalet</i> | <i>Antal enheter i urvalet efter förstärkningen</i> |
|-----------------------------|----------------------------------|--|--|---|
| <i>Flera tilläggsnummer</i> | 74 | 17 | 57 | 74 |
| <i>Övriga media</i> | 126 | 29 | 87 | 116 |
| <i>Stora / Webb</i> | 86 | 19 | 67 | 86 |
| <i>Stora / Papper</i> | 182 | 40 | 120 | 160 |
| <i>Stora / IDEP</i> | 900 | 199 | 100 | 299 |
| <i>Medel / Webb</i> | 808 | 178 | 0 | 178 |
| <i>Medel / Papper</i> | 396 | 88 | 59 ⁸ | 147 |
| <i>Medel / IDEP</i> | 1525 | 337 | 0 | 337 |
| <i>Små / Webb</i> | 1069 | 236 | 0 | 236 |
| <i>Små / Papper</i> | 494 | 109 | 20 ⁹ | 129 |
| <i>Små / IDEP</i> | 1133 | 250 | 0 | 250 |
| <i>Totalt</i> | 6793 | 1502 | 500 | 2012 |

⁸ Där 49 av 59 fördelas jämt mellan in- och utförsel, och de 10 resterande förstärker utförseln

⁹ Förstärker utförseln

4 RESULTAT

4.1 GENOMFÖRANDE AV UNDERSÖKNINGEN

Den statistiska undersökningen startade den 2:a maj 2011. Detta datum skickades 2012 missivbrev ut till uppgiftslämnarna. Efter cirka en vecka skickades påminnelser ut. Sista svarsdatumet sattes till den 22:a maj.

Totalt 1265 enkäter har besvarats av uppgiftslämnare, så den relativa svarsfrekvensen blev lika med 62,9%. Av dessa hade 858 svar (42,6% av urvalet) kommit in innan påminnelserna skickades ut, och 407 svar (20,2% av urvalet) kom in efter påminnelserna hade skickats ut.

Vid frågor eller oklarheter hade uppgiftslämnarna möjlighet att kontakta SCB¹⁰. Kontakter kunde hållas via e-post eller telefon. Under undersökningens gång hade man fått ca 30 samtal och ca 10 mejl. Majoriteten av kontakter reflekterade uppgiftslämnarnas förvåning i samband med mottagning av påminnelser. Uppgiftslämnarna hade nämligen inte förväntat sig att få en påminnelse om en frivillig undersökning. En stor del kontakter gällde vissa praktiska sidor av undersökningen. Exempelvis fick vi en del samtal och mejl där uppgiftslämnarna informerade om att de inte kan delta i undersökningen på grund av arbetsbörda. Ett par uppgiftslämnare har bytt organisationsnummer eller avbrutit verksamhet. Slutligen, har ett par kontakter ägt rum i samband med att uppgiftslämnarna har förväxlat denna undersökning med själva Intrastatundersökningen.

Vi har även gjort ett par återkontakter till uppgiftslämnarna. Totalt 5 återkontakter togs, varav 2 till respondenter med störst tidsåtgång (fråga 2 i enkäten) och 3 uppgiftslämnare med minst tidsåtgång.

När det gäller de två kontakterna med respondenter som har angett störst tidsåtgång (ca en arbetsvecka), har respektive Intrastatansvariga personer bekräftat att rapporteringen verkligen tar den tiden som de hade angett. En av personerna har konstaterat att deras företag har en väldigt komplicerad rapporteringsrutin, och den andra personen har påpekat att de rapporterar till Intrastat dagligen och att varje dag tar rapporteringen några timmar i snitt.

Av de tre företag som angav minst tidsåtgång (mindre än 5 minuter), har respektive kontaktpersoner för två av dem konstaterat att det verkligen tar så kort tid. Enligt den förklaringen som har fåttts sker rapporteringen hos dessa företag genom att utföra ett par elementära kommandon i företagens datorsystem/affärssystem.

Vid kontakt av den tredje uppgiftslämnaren har dock respektive kontaktperson poängterat att hon hade trott att man borde ange svaret *i timmar* i stället för minuter. Kontaktpersonen konstaterade vidare att hon inte läste frågan noggrant, och respektive svar för denna uppgiftslämnare korrigerades omedelbart.

¹⁰ Konstantin Kalinichenko och Petter Samuelsson på DFO/IF hade ansvar för detta.

4.2 ANALYS AV ENHETSORTFALL

I detta avsnitt kommer vi att undersöka fördelningen av det totala bortfallet med avseende på följande faktorer:

- Strata (flöde, rapporteringsmedium, antal varuposter samt två unika strata)
- Ursprung (uppgiftslämnare med svenska/utländska postadresser)

Vi börjar med att undersöka om det föreligger någon skillnad mellan uppgiftslämnarnas ursprung med avseende på bortfallet. Se Tabell 7 nedan.

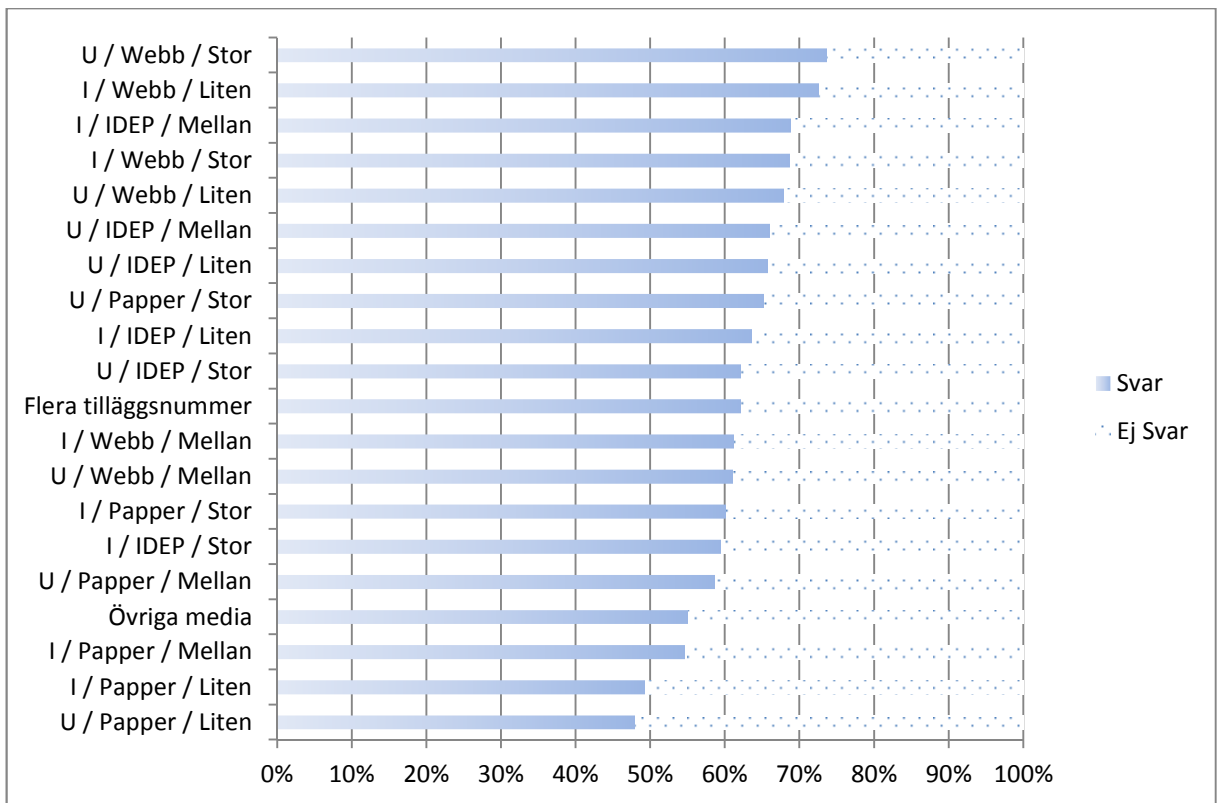
Tabell 7: Bortfall m.a.p. ursprung

| <i>Placering av företag</i> | <i>Antal mottagna svar</i> | <i>Antal företag i urvalet</i> | <i>Svarsfrekvens</i> |
|-----------------------------|----------------------------|--------------------------------|----------------------|
| <i>Sverige</i> | 1227 | 1936 | 63,4 % |
| <i>Utomlands</i> | 38 | 76 | 50 % |
| <i>Totalt</i> | 1265 | 2012 | 62,9 % |

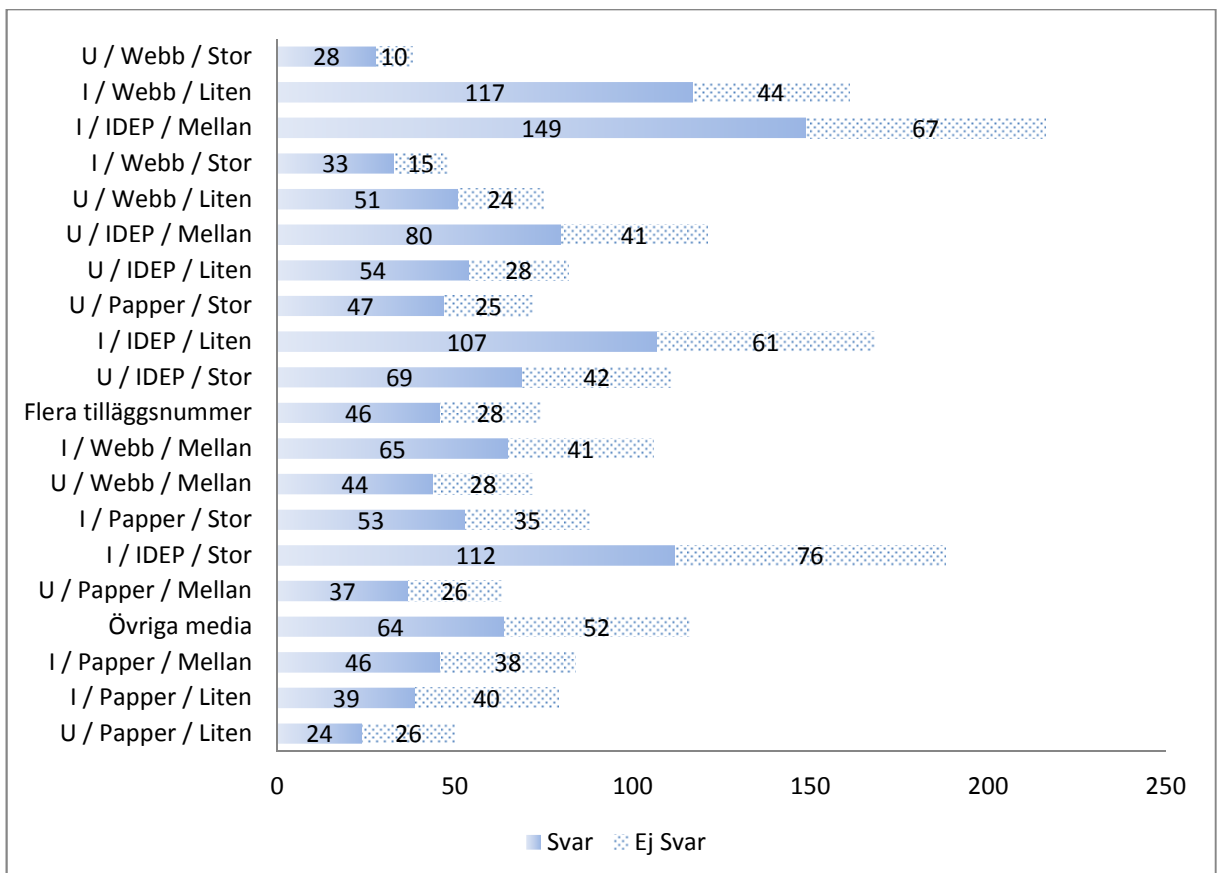
Som vi kan se, är bortfallet betydligt lägre för företag som befinner sig i Sverige jämfört med företag som befinner sig utomlands. Observera dock att antal företag som befinner sig utomlands är avsevärt mindre än de som är belägna i Sverige.

Fortsättningsvis undersöker vi hur bortfallet fördelar sig mellan olika strata. Det bör poängteras att vi redan vid urvalsplaneringsfasen hade vissa förväntningar angående fördelningen av bortfallet m.a.p. strata. Nämligen, visste man ur tidigare erfarenhet att mindre företag som rapporterar via pappersblankett tenderar att vara mindre intresserade i att delta i frivilliga undersökningar än samtliga andra uppgiftslämnare. En åtgärd man vidtog för att motverka detta var att höja samlingsfrekvensen för strata utförelse/små/papper samt införelse/små/papper.

Vid granskning av svarsdata framkom det att förväntningarna angående större bortfall i ovannämnda strata verkligen infriades. På nästa sida illustrerar vi figurer som avspeglar fördelningen av det oviktade bortfallet mellan olika strata, i absoluta och i relativa tal.



Figur 1: Bortfall per stratum, % (Beteckningar: I - införsel; U - utförsel; liten - 0-5 varuposter/mån; mellan - 6-25 varuposter/mån; stor - 26+ varuposter/mån)



Figur 2: Bortfall per stratum, absoluta tal (Beteckningar: I - införsel; U - utförsel; liten - 0-5 varuposter/mån; mellan - 6-25 varuposter/mån; stor - 26+ varuposter/mån)

Ovan kan vi se att det finns ett tydligt samband mellan rapporteringsätt (via papper eller elektroniskt) för det relativa bortfallet. Som förväntat, är bortfallet störst för utförelse/små/papper samt införelse/små/papper. Det borde även anmärkas att bortfallet var stort för uppgiftslämnare som rapporterar via övriga media.

4.3 ANALYS AV PARTIELLT BORTFALL

Förutom bortfallet på enhetsnivå, kommer vi även att titta på det partiella bortfallet, d.v.s. de fallen då respondenterna deltog i undersökningen, men inte besvarade alla frågor.

Av 1265 besvarade enkäter har samtliga frågor besvarats i 1172 (92,7 %). Den resterande delen (93 enkäter; 7,3 %) innehöll minst en obesvarad fråga. Vid en närmare undersökning av dessa har man fått följande bild av det partiella bortfallet (Vi hänvisar till listan över frågor på s.6 eller till Appendix för exakt formulering av frågorna):

- Av 93 enkäter där minst en fråga lämnades obesvarad, fanns det 2 enkäter där inga frågor har besvarats, 1 enkät där enbart frågorna 1 och 2 har besvarats, samt 11 enkäter där enbart frågorna 1,2, och 3 har besvarats.
- När det gäller de resterande 79 enkäterna, har frågorna 1, 2, 4, 6 och 7 besvarats av samtliga respondenter i denna grupp. Strukturen för variabelbortfallet gällande frågorna 8, 9 och 10 är följande:
 - 3 uppgiftslämnare har inte besvarat fråga 8
 - 6 uppgiftslämnare har inte besvarat fråga 9 (där 4 rapporterar via webb)
 - 5 uppgiftslämnare har inte besvarat fråga 10 (där 4 rapporterar via webb, och samtliga inte har besvarat fråga 9)
- Det partiella bortfallet avseende frågorna 3 och 5 för de 79 enkäterna ser ut på följande sätt:

Tabell 8: Variabelbortfall avseende frågor 3 och 5

| <i>Fråga 3 \ Fråga 5</i> | <i>Obesvarad helt</i> | <i>Obesvarad delvis</i> | <i>Besvarad helt</i> | <i>Totalt</i> |
|--------------------------|-----------------------|-------------------------|----------------------|---------------|
| <i>Besvarad</i> | 1 | 36 | 5 | 42 |
| <i>Obesvarad</i> | 1 | 1 | 35 | 37 |
| <i>Totalt</i> | 2 | 37 | 40 | 79 |

Det bör konstateras att en större del av det partiella bortfallet faller på frågorna 3 och 5. När det gäller fråga 3, kan en möjlig förklaring till det relativt stora bortfallet vara att det i vissa fall kan vara svårt att ta reda på tidsåtgången för uppgiftsrapporteringen. Dessutom kan beräkningen av tidsåtgången innebära en viss mental och/eller fysisk ansträngning (behov av aritmetiska beräkningar, behov av kontakt med kollegor mm.), vilket kan leda till ökat bortfall också.

När det gäller fråga 5, noterar vi att ett relativt stort antal respondenter inte svarade på den helt. Vi antar att detta beror på att frågan var mer tidskrävande och mer komplicerad än andra.

Slutligen, tror vi att de 12 respondenterna som enbart besvarade de 2(3) första frågorna kunde ha gjort detta av misstag. Det är nämligen så att vid ifyllande av webbenkäten innehöll första sidan enbart de tre första frågorna. För att komma vidare borde man trycka på en knapp "Gå vidare", som kunde lätt förväxlas med knappen "Skicka uppgifterna till SCB". Om man inte tryckte på den rätta knappen utan på den andra, skickades uppgifterna till SCB utan någon möjlighet att ändra inskickade uppgifter. Ett argument som stödjer denna teori är att vi har blivit kontaktade av en person som informerade oss om att hon råkade trycka på fel knapp och bad oss att lägga in svaren på de resterande frågorna i svarsdatabasen.

4.4 RESULTAT AV UNDERSÖKNINGEN: TIDSÅTGÅNG

Det viktade medelvärdet för tidsåtgången för att rapportera till Intrastat per månad (fråga 3 i enkäten) visade sig vara 119 minuter om data från samtliga strata ingår. Räknar man bort stratumet "flera tilläggsnummer" blir medelvärdet lika med 117 minuter.

De viktade medelvärdena fås på följande sätt. Som ett exempel, betrakta den gruppen av observationer som tillhör kategorin införsel/IDEP. Denna grupp består av tre strata där varje stratum motsvarar en storleksklass. För att få fram det viktade medelvärdet för dessa strata tillsammans, kan man multiplicera varje stratums medelvärde med respektive populationsstorleken. Därefter summerar man dessa och dividerar summan med det totala antalet uppgiftslämnare i populationen avseende dessa tre strata. Alltså får man

$$\frac{77 \times 761 + 126 \times 978 + 177 \times 567}{761 + 978 + 567} = 122.36 \approx 122$$

Det generaliserade uttrycket för uträkning av de viktade medelvärdena är alltså

$$\bar{x}_{1\dots J} = \frac{\sum_{i=1}^J \bar{x}_i n_i}{\sum_{i=1}^J n_i}$$

Här innebär $\bar{x}_{1\dots J}$ det viktade medelvärdet över strata 1 ... J, \bar{x}_i – medelvärdet i stratum i , och n_i – antal element i populationen i stratum i .

Vidare illustreras genomsnittliga rapporteringstider med avseende på strata, antal involverade personer (fråga 2), erfarenhet av rapporteringen (fråga 4), sätt att hitta varukoder (fråga 5) samt tillgänglighet av uppgifterna (fråga 7).

TIDSÅTGÅNG OCH STRATA

Nedanstående tabell illustrerar fördelning av tidsåtgången med avseende på strata.

Tabell 9: Tidsåtgång med avseende på strata

| Flöde | Rapporterings- medium | Storleks- klass | Antal i urvalet | Antal i populatione n | Viktad snittid (min) | Mediantid (min) |
|--|--------------------------|--------------------|--------------------|-----------------------------|-------------------------|--------------------|
| Införsel | IDEP | Små | 168 | 761 | 77 | 45 |
| | | Medel | 216 | 978 | 126 | 60 |
| | | Stora | 188 | 567 | 177 | 90 |
| | | Totalt | 572 | 2306 | 122 | 60 |
| | Papper | Små | 79 | 360 | 131 | 30 |
| | | Medel | 84 | 243 | 158 | 80 |
| | | Stora | 88 | 100 | 162 | 60 |
| | | Totalt | 251 | 703 | 144 | 60 |
| | Webb | Små | 161 | 730 | 58 | 30 |
| | | Medel | 106 | 483 | 165 | 60 |
| | | Stora | 48 | 48 | 241 | 180 |
| | | Totalt | 315 | 1261 | 105 | 60 |
| | Totalt | | 1138 | 4270 | 121 | 60 |
| Utförsel | IDEP | Små | 82 | 372 | 68 | 48 |
| | | Medel | 121 | 547 | 111 | 80 |
| | | Stora | 111 | 333 | 181 | 90 |
| | | Totalt | 314 | 1252 | 116 | 60 |
| | Papper | Små | 50 | 134 | 83 | 60 |
| | | Medel | 63 | 153 | 67 | 45 |
| | | Stora | 72 | 82 | 129 | 60 |
| | | Totalt | 185 | 369 | 87 | 60 |
| | Webb | Små | 75 | 339 | 53 | 40 |
| | | Medel | 72 | 325 | 76 | 38 |
| | | Stora | 38 | 38 | 272 | 90 |
| | | Totalt | 185 | 702 | 76 | 45 |
| | Totalt | | 684 | 2323 | 100 | 60 |
| Unikt stratum: Övriga media | | | 116 | 126 | 302 | 90 |
| Unikt stratum: Flera tilläggsnummer | | | 74 | 74 | 269 | 120 |
| Totalt (exkl. Flera tilläggsnummer) | | | 1938 | 6719 | 117 | 60 |
| Totalt | | | 2012 | 6793 | 119 | 60 |

Som förväntat, ser vi att rapporteringstiden varierar kraftigt mellan de olika strata. Eftersom de två unika strata inte är fördelade med avseende på antal varuposter, kan vi konstatera att i dessa strata är tidsåtgången störst. Beträffande företag och organisationer som rapporterar via övriga media, konstaterar vi att detta stratum innehåller en större del av Sveriges största företag. Det är naturligt att förvänta sig att sådana uppgiftslämnare för in/ut stora mängder av varierande varor.

När det gäller uppgiftslämnare som tillhör stratumet "flera tilläggsnummer", antar vi att den relativt höga snitttiden för rapportering följer ur det urvalsförfarandet som implementerades med avseende på detta stratum. Som konstaterades tidigare, har man för varje uppgiftslämnare som hamnade i stratum "flera tilläggsnummer" valt ut tilläggsnummer med maximalt antal varuposter. Detta innebär att tidsåtgång per tilläggsnummer uppenbarligen har överskattats.

Vidare konstaterar vi att antal varuposter har ett tydligt samband med tidsåtgången (avser ej kategorin utförsel/papper, se tabellen). Även flödet verkar ha påverkan på

tidsförbrukningen i samband med rapportering till Intrastat. Dessa samband är välkända ur tidigare undersökningar av uppgiftslämnarbördan och förvånar oss inte. Dock förväntades det inte att tidsåtgången för stratumet utförelse/papper/mellan kommer att vara så ovanligt låg.

Slutligen, noterar vi att de uppgiftslämnare som rapporterar via webb och som har litet antal varuposter behöver betydligt mindre tid för att rapportera än uppgiftslämnare i motsvarande storleksklass men med papper eller IDEP som rapporteringsmedium. Dock gäller denna företeelse inte för uppgiftslämnare med medel eller stort antal varuposter. Sådana företag/organisationer behöver i regel längre tid för att rapportera än deras motparter som rapporterar via IDEP eller papper. Vi konstaterar därför att rapportering via webb passar bäst för rapportörer med litet antal varuposter, dock ej för uppgiftslämnare med medelstort eller stort antal varuposter.

ANTAL MEDVERKANDE PERSONER OCH TIDSÅTGÅNG

Tabell 10 nedan visar den procentuella fördelningen av uppgiftslämnare med avseende på antal personer som medverkar i rapporteringen till Intrastat (fråga 2), samt sambandet mellan denna och tidsåtgången.

Tabell 10: Antal medverkande personer

| <i>Antal personer som medverkar i rapporteringsprocessen</i> | <i>Antal</i> | <i>Relativt antal,</i> | <i>Snittid (min)</i> |
|--|--------------|------------------------|----------------------|
| <i>1 person</i> | 879 | 71,8 % | 116 |
| <i>2 personer</i> | 267 | 21,8 % | 162 |
| <i>3 personer eller fler</i> | 73 | 6,0 % | 311 |
| <i>Vet ej/Ingen åsikt</i> | 6 | 0,5 % | 211 |
| <i>Totalt</i> | 1225 | | |

Ur tabellen framkommer ett tydligt samband mellan antal personer som medverkar i uppgiftsrapporteringen och tidsåtgången. Vidare noterar vi att rapporteringen till Intrastat utförs av en person hos majoriteten av uppgiftslämnare. Hos mer än 90 % av uppgiftslämnare utförs rapporteringen av 1 eller 2 personer.

Det är inte klart vad sambandet mellan antal involverade personer och tidsåtgången orsakas av. Å ena sidan kan fler medverkande personer leda till längre rapporteringstid p.g.a. behov av kommunikation, eventuellt dubbelarbete osv. Å andra sidan kan lång rapporteringstid leda till att rapporteringen till Intrastat förstärks av ytterligare personal.

ERFARENHET AV RAPPORTERING OCH TIDSÅTGÅNG

En annan viktig bakgrundsvariabel är erfarenhet av Intrastatrapportering hos personer som besvarade undersökningen (fråga 4). I tabellen nedan visas hur tidsåtgången fördelar sig med avseende på denna variabel.

Tabell 11: Erfarenhet av rapportering

| <i>Tiden för start av rapporteringen¹¹</i> | <i>Antal</i> | <i>Relativt antal,</i> | <i>Snittid (min)</i> |
|---|--------------|------------------------|----------------------|
| <i>Jan 2011 eller senare</i> | 71 | 5,9 % | 138 |
| <i>Sep 2010 – Dec 2010</i> | 62 | 5,1 % | 102 |
| <i>Aug 2010 eller tidigare</i> | 1023 | 84,3 % | 143 |
| <i>Vet ej/Ingen åsikt</i> | 58 | 4,8 % | 94 |
| <i>Totalt</i> | 1214 | | |

Vi kan konstatera att större erfarenhet med att rapportera till Intrastat inte behöver leda till kortare rapporteringstider. Till och med tar det en aning längre tid för "erfarna" uppgiftslämnare i jämförelse med de mindre erfarna.

Vi antar att avsaknad av samband mellan erfarenhet och tidsförbrukning kan bero på att Intrastatrapportering inte kräver mycket förberedelser och speciell kunskap utan innebär oftast ett monotont arbete. Därför är uppgiftslämnare relativt snabba på att lära sig att rapportera.

SÄTT ATT HITTA VARUKODER OCH TIDSÅTGÅNG

Tabellerna nedan innehåller information om hur ofta de olika sätten att hitta varukoder används (fråga 5) samt hur användning av dessa samspelar med tidsåtgången.

Tabell 12 visar hur ofta de olika sätten för framtagning av varukoder används av uppgiftslämnare. Cellerna i tabellen innehåller absolut och relativt antal uppgiftslämnare som har hamnat i respektive kategori. Variationen i den sista kolumnen (totalerna) beror på att några uppgiftslämnare inte lämnade svar avseende samtliga framtagningssätt.

Tabell 12: Sätt att hitta varukoder

| <i>Framtagningssätt / Svarsalternativ</i> | <i>Antal uppgiftslämnare (absolut och relativt)</i> | | | | |
|---|---|----------------------|---------------|---------------|---------------|
| | <i>Alltid/Ofta</i> | <i>Ibland/Sällan</i> | <i>Aldrig</i> | <i>Vet ej</i> | <i>Totalt</i> |
| <i>På fakturan</i> | 340 27,6 % | 385 31,3 % | 406 33 % | 99 8,1 % | 1230 100 % |
| <i>Via SCB</i> | 247 20,2 % | 477 38,9 % | 411 33,6 % | 90 7,4 % | 1225 100 % |
| <i>Via Tullverket</i> | 249 20,2 % | 475 38,5 % | 406 32,9 % | 103 8,4 % | 1233 100 % |
| <i>Egen kunskap /bedömning</i> | 387 31,4 % | 461 37,5 % | 285 23,2 % | 97 7,9 % | 1230 100 % |
| <i>Från leverantören (partnerföretaget)</i> | 245 19,9 % | 468 37,9 % | 402 32,6 % | 119 9,6 % | 1234 100 % |
| <i>Övrigt</i> | 88 7,2 % | 208 17 % | 414 33,9 % | 512 41,9 % | 1222 100 % |

¹¹ Tiden då personen som har besvarat enkäten hade börjat medverka i rapporteringsprocessen

I Tabell 13 illustreras samband mellan tidsåtgången och användning av ett eller annat framtagningssätt. Cellerna i denna tabell innehåller snitttider för rapportering beräknade för uppgiftslämnare som använder motsvarande framtagningssätt och som har kryssat i respektive svarsalternativ.

Tabell 13: Sätt att hitta varukoder och tidsåtgång

| Framtagningssätt / Svarsalternativ | Snittid (minuter) | | | |
|---|-------------------|---------------|--------|--------|
| | Alltid/Ofta | Ibland/Sällan | Aldrig | Vet ej |
| <i>På fakturan</i> | 191 | 162 | 87 | 72 |
| <i>Via SCB</i> | 136 | 166 | 116 | 80 |
| <i>Via Tullverket</i> | 157 | 155 | 118 | 93 |
| <i>Egen kunskap/bedömning</i> | 140 | 171 | 106 | 83 |
| <i>Från leverantören (partnerföretaget)</i> | 136 | 173 | 115 | 83 |
| <i>Övrigt</i> | 84 | 193 | 116 | 142 |

Det framtagningssätt som verkar användas oftast för att ta fram varukoder är att utnyttja egen kunskap och/eller bedömning. Det är även ett av de framtagningssätten som resulterar i snabbast rapportering. Vi noterar även att de uppgiftslämnare som svarade "Vet ej" avseende något av ovannämnda rapporteringssätten tenderar att ha mycket lägre tidsförbrukning i relation till andra rapportörer. En möjlig förklaring till detta kan vara att sådana företag/ organisationer alltid rapporterar handel av samma varor och behöver därför inte ta fram några nya varukoder då de rapporterar. Detta kan innebära dels att ansvarig personal inte vet hur varukoderna har tagits fram, dels att rapporteringstiden är mycket kortare jämfört med andra uppgiftslämnare. En annan möjlig förklaring kan vara att sådana rapportörer anlitar externa konsulter som tar hand om framtagningen av varukoder.

Vi konstaterar vidare att de uppgiftslämnarna som är vana vid att använda övriga framtagningssätt behöver i regel mycket mindre tid för att rapportera än andra. Eftersom "övriga framtagningssätt" är ett brett begrepp som kan inkludera olika framtagningssätt kan vi inte veta exakt vad det beror på. Emellertid tror vi att orsaker till detta fenomen är samma som orsaker till kortare rapporteringstider hos uppgiftslämnare som svarade "Vet ej" avseende något av rapporteringssätten (se ovan).

Vidare bör det anmärkas att uppgiftslämnare som svarade "Aldrig" gällande användning av något av rapporteringssätten har i snitt lägre tidsförbrukning än andra. Detta kan bero på att sådana uppgiftslämnare använder färre framtagningssätt än de som inte svarade "Aldrig" för något av framtagningssätten, vilket gör att tidsåtgången blir lägre.

Slutligen, ser vi att uppgiftslämnare som brukar ta fram varukoder på fakturan behöver oftast längre tid för att rapportera än rapportörer som tar fram varukoder på andra sätt.

TILLGÄNGLIGHET AV UPPGIFTERNA OCH TIDSÅTGÅNG

En annan faktor som kan influera tidsåtgången är tillgängligheten av de efterfrågade uppgifterna (fråga 7). Nedan reflekteras sambandet mellan denna faktor och tidsåtgången. I den nedanstående tabellen innebär den andra kolumnen antal uppgiftslämnare som angav motsvarande svarsalternativ. Kolumnen "Snittid" innehåller i sin tur medelvärden för rapporteringstider för dessa uppgiftslämnare. Notera att vid besvarande av frågan kunde respondenterna ange flera svarsalternativ.

Tabell 14: Tillgänglighet av uppgifterna

| <i>Uppgifterna tas fram ...</i> | <i>Antal positiva svar</i> | <i>Totalt antal</i> | <i>Snittid (minuter)</i> |
|--|--------------------------------|-------------------------|------------------------------|
| <i>Direkt ur företagets datorsystem</i> | 369 29,5 % | 1249 | 90 |
| <i>Direkt ur företagets datorsystem, men behöver bearbetas</i> | 699 56,0 % | 1249 | 147 |
| <i>Hämtas från andra källor</i> | 233 18,7 % | 1249 | 224 |
| <i>Uppskattas utifrån annan kunskap</i> | 20 1,6 % | 1249 | 273 |

Ur tabellen ovan framkommer ett tydligt beroende mellan tillgängligheten av uppgifterna och tidsåtgång. Uppgiftslämnare som har uppgifterna direkt tillgängliga i sina datorsystem/affärssystem behöver i snitt betydligt mindre tid för att rapportera till Intrastat än de som behöver bearbeta uppgifterna innan dessa skickas till SCB. Tidsåtgången är ännu högre för de företag och organisationer som inte har uppgifterna tillgängliga i sina datorsystem. Dock utgör sådana företag enbart drygt 20 % av alla uppgiftslämnare.

4.5 RESULTAT AV UNDERSÖKNINGEN: ALLMÄN UPPLEVELSE

En av basfrågorna i undersökningen är frågan om allmän upplevelse i samband med Intrastatrapporteringen. I Tabell 16 visas fördelningen av svar på denna fråga med avseende på strata.

Vi ser att mer än 70 % av uppgiftslämnare tycker att det är lätt att rapportera till Intrastat. Denna andel är störst för utförelseföretag som använder IDEP och papper och som har litet antal varuposter (75-80% av dem svarade positivt på frågan om allmän upplevelse). Uppgiftslämnare med stort antal varuposter och som använder webblankett är dock minst tillfredsställda med uppgiftsrapporteringen. En anmärkning som bör göras är att de företagen som rapporterar införsel via IDEP eller papper tenderar att ha bättre upplevelse i samband med rapporteringen om antal rapporterade varuposter är stort.

SAMBAND MELLAN ALLMÄN UPPLEVELSE OCH TIDSÅTGÅNG

Vi fortsätter med att undersöka samband mellan upplevelsen i samband med uppgiftsrapporteringen och tidsåtgången. Se nedanstående tabell.

Tabell 15: Allmän upplevelse vs. Tidsåtgång

| | <i>Antal</i> | <i>Snittid (minuter)</i> |
|--------------------------------------|--------------|--------------------------|
| <i>Mycket lätt och Ganska lätt</i> | 890 | 90 |
| <i>Ganska svårt och Mycket svårt</i> | 354 | 258 |
| <i>Ingen åsikt</i> | 19 | 188 |
| <i>Summa</i> | 1263 | |

Som förväntat, är sambandet uppenbart mellan upplevelsen av rapportering till Intrastat och tiden det tar för att rapportera.

Tabell 16: Allmän upplevelse i samband med uppgiftsrapporteringen

| Flöde | Rapporte rings-medium | Storleksklass | Svarsalternativ | | | Antal svar |
|-------------------------------------|--------------------------|----------------------------------|-----------------------------|-------------------------------|---------------|--------------|
| | | | Mycket lätt och Ganska lätt | Ganska svårt och Mycket svårt | Ingen åsikt | |
| Införsel | IDEP | Små | 67,3 % | 30,0 % | 2,8 % | 107 |
| | | Medel | 71,1 % | 27,5 % | 1,3 % | 149 |
| | | Stora | 75,0 % | 22,3 % | 2,7 % | 112 |
| | | Införsel /IDEP (viktat) | 70,8 % | 27,0 % | 0,2 % | 368 |
| | Papper | Små | 64,1 % | 33,3 % | 2,6 % | 39 |
| | | Medel | 67,4 % | 30,4 % | 2,2 % | 46 |
| | | Stora | 79,3 % | 18,9 % | 1,9 % | 53 |
| | | Införsel /Papper(viktat) | 67,4 % | 30,3 % | 2,3 % | 138 |
| | Webb | Små | 74,4 % | 24,8 % | 0,9 % | 117 |
| | | Medel | 61,5 % | 38,5 % | 0 % | 65 |
| | | Stora | 50,0 % | 50,0 % | 0 % | 32 |
| | | Införsel /Webb (viktat) | 68,5 % | 31,0 % | 0,5 % | 214 |
| | Införsel (viktat) | | | 69,6 % | 28,7 % | 1,7 % |
| Utförsel | IDEP | Små | 81,5 % | 16,7 % | 1,9 % | 54 |
| | | Medel | 70,0 % | 28,8 % | 1,2 % | 80 |
| | | Stora | 69,6 % | 30,4 % | 0 % | 69 |
| | | Utförsel /IDEP (viktat) | 73,3 % | 25,6 % | 1,1 % | 203 |
| | Papper | Små | 83,3 % | 16,7 % | 0 % | 24 |
| | | Medel | 80,6 % | 19,4 % | 0 % | 36 |
| | | Stora | 72,3 % | 27,7 % | 0 % | 47 |
| | | Utförsel /Papper (viktat) | 79,7 % | 20,3 % | 0 % | 107 |
| | Webb | Små | 74,5 % | 25,5 % | 0 % | 51 |
| | | Medel | 68,2 % | 27,3 % | 4,6 % | 44 |
| | | Stora | 57,1 % | 42,9 % | 0 % | 28 |
| | | Utförsel /Webb (viktat) | 70,6 % | 27,2 % | 2,1 % | 123 |
| | Utförsel (viktat) | | | 73,5 % | 25,2 % | 1,2 % |
| Unikt stratum: Övriga media | | | 68,8 % | 28,1 % | 3,1 % | 64 |
| Unikt stratum: Flera tilläggsnummer | | | 60,9 % | 37,0 % | 2,2 % | 46 |
| Samtliga strata (viktat) | | | 70,8 % | 27,6 % | 1,6 % | 1263 |

4.6 RESULTAT: ÖVRIGA VARIABLER

De andra frågorna som ingick i enkäten avsåg säkerhet i valet av varukoder (fråga 6), införande av förenklad varukod för låga fakturavärden (fråga 8), möjlighet att rapportera i Euro (fråga 9), användbarhet av statistiken som framställs av Intrastatundersökningen (fråga 10) samt kommentarer och förslag (fråga 11). Vi fortsätter med att undersöka

uppgiftslämnarnas respons avseende dessa variabler. Samtliga data som illustreras nedan är viktade med avseende på antal element i populationen inom respektive strata.

SÄKERHET I VALET AV VARUKODER

När det gäller säkerhet i valet av varukoder, ser uppgiftslämnarnas svar ut på följande sätt (se tabell nedan)

Tabell 17: Säkerhet i valet av varukoder

| Säkerhet i valet av varukoder | Mycket säker | Ganska säker | Ganska osäker | Mycket osäker | Ingen åsikt | Antal svar |
|-------------------------------|--------------|--------------|---------------|---------------|-------------|------------|
| Införsel (%) | 22,4 | 63,3 | 11,1 | 1,3 | 1,9 | 713 |
| Utförsel (%) | 33,2 | 57,5 | 6,0 | 1,2 | 2,1 | 429 |
| Unika strata (%) | 40,1 | 48,8 | 0,8 | 0,8 | 9,5 | 109 |
| Totalt (%) | 26,6 | 60,9 | 9,0 | 1,3 | 2,2 | 1251 |

Vi ser att absoluta majoriteten av uppgiftslämnare känner sig säkra i valet av varukoder. Nästan 90 % av uppgiftslämnare som tillhör de unika strata är antingen mycket säkra eller ganska säkra när de väljer varukoder.

MÖJLIGHETEN ATT RAPPORTERA I EURO

Vi fortsätter med att undersöka hur respondenterna svarade på frågan om möjligheten att rapportera i Euro. Tabell 18 innehåller svar avseende samtliga strata. I Tabell 19, som befinner sig nedanför den första, är samtliga svar från IDEP-användare exkluderade, eftersom dessa redan har möjlighet att rapportera i Euro.

Tabell 18: Möjlighet att rapportera i Euro, samtliga rapporteringsmedia

| Inställning till möjlighet att rapportera i Euro | Positiv | Varken positiv eller negativ | Negativ | Möjligheten finns redan eftersom vi använder IDEP | Antal svar |
|--|---------|------------------------------|---------|---|------------|
| Införsel (%) | 20,2 | 49,2 | 3,6 | 27,0 | 710 |
| Utförsel (%) | 22,6 | 49,8 | 5,1 | 22,4 | 428 |
| Unika strata (%) | 17,7 | 57,2 | 6,5 | 18,6 | 107 |
| Totalt (%) | 21,0 | 49,7 | 4,2 | 25,2 | 1245 |

Tabell 19: Möjlighet att rapportera i Euro, svar från IDEP-användare exkluderade

| Inställning till möjlighet att rapportera i Euro | Positiv | Varken positiv eller negativ | Negativ | Möjligheten finns redan eftersom vi använder IDEP | Antal svar |
|--|---------|------------------------------|---------|---|------------|
| Införsel (%) | 28,8 | 64,1 | 6,0 | 1,1 | 344 |
| Utförsel (%) | 26,4 | 64,8 | 8,1 | 0,7 | 227 |
| Unika strata (%) | 17,7 | 57,2 | 6,5 | 18,6 | 107 |
| Totalt (%) | 27,3 | 63,7 | 6,9 | 2,1 | 678 |

Som vi kan se ovan (Tabell 19), är mer än 25 % av icke-IDEP användare intresserade i att kunna välja rapporteringsvaluta mellan Euro och Svenska Kronor.

Vi ser även att 18 % av uppgiftslämnare som tillhör de unika strata angav att de rapporterar via IDEP. En möjlig förklaring till detta är att dessa uppgiftslämnare härstammar från stratumet "flera tilläggsnummer" och verkligen använder IDEP för att rapportera.

INFÖRANDE AV FÖRENKLAD VARUKOD FÖR LÅGA FAKTURAVÄRDEN

Tabell 20 illustrerar uppgiftslämnarnas svar på frågan om låga fakturavärden.

Tabell 20: Införande av förenklad varukod för låga fakturavärden

| <i>Inställning till införande av förenklad varukod för låga fakturavärden</i> | <i>Positiv</i> | <i>Varken positiv eller negativ</i> | <i>Negativ</i> | <i>Antal svar</i> |
|---|----------------|-------------------------------------|----------------|-------------------|
| <i>Införsel (%)</i> | 52,5 | 46,2 | 1,2 | 711 |
| <i>Utförsel (%)</i> | 46,8 | 48,9 | 4,4 | 428 |
| <i>Unika strata (%)</i> | 45,5 | 44,2 | 10,3 | 109 |
| <i>Totalt (%)</i> | 50,3 | 47,1 | 2,6 | 1248 |

Vi ser att runt hälften av uppgiftslämnare är positivt inställda till förslaget, och att runt 45 – 50 % av uppgiftslämnarna är neutrala till det. En anmärkning som bör göras att 10 % (11 st.) uppgiftslämnare som tillhör de unika strata har en negativ inställning till ovannämnda förslaget. En möjlig förklaring till detta kan vara att dessa rapportörer har en rapporteringsrutin som kommer att behöva ändras i fall förslaget implementeras, vilket dessa uppgiftslämnare möjligtvis ser som extra belastning.

ANVÄNDBARHET AV STATISTIKEN

När det gäller den näst sista frågan i enkäten, användbarhet av statistiken, har uppgiftslämnarna svarat på följande sätt (Se tabell nedan).

Tabell 21: Användbarhet av statistiken

| <i>Svarsalternativ avseende frågan om användbarhet av statistiken</i> | <i>Mycket intresserade</i> | <i>Ganska intresserade</i> | <i>Inte särskilt intresserade</i> | <i>Helt ointresserade</i> | <i>Antal svar</i> |
|---|----------------------------|----------------------------|-----------------------------------|---------------------------|-------------------|
| <i>Införsel (%)</i> | 4,6 | 28,2 | 49,0 | 18,2 | 710 |
| <i>Utförsel (%)</i> | 6,3 | 29,3 | 49,0 | 15,4 | 428 |
| <i>Unika strata (%)</i> | 9,8 | 38,5 | 41,0 | 10,7 | 108 |
| <i>Totalt (%)</i> | 5,3 | 28,9 | 48,7 | 17,1 | 1246 |

Vi kan se att majoriteten av uppgiftslämnare som tillhör de "vanliga" (ej unika) strata inte är särskilt intresserade i att ta del av resultatet av Intrastatundersökningen. Dock är bilden annorlunda när det gäller de unika strata: runt hälften av företag i dessa svarade positivt på frågan.

KOMMENTARER OCH FÖRSLAG

Efter besvarande av samtliga frågor hade uppgiftslämnarna möjlighet att lämna kommentarer angående Intrastatundersökningen. Totalt hade man fått in 110 kommentarer från uppgiftslämnarna, som vi delade in i följande kategorier (se nedan). Notera att kategorierna inte är disjunkta, dvs. vissa kommentarer kunde ha hamnat i fler än en kategori

- Varukoder (30 kommentarer). Kommentarererna i denna grupp avsåg framtagning av varukoder och problematiken i samband med detta. En del uppgiftslämnare poängterade att det är för betungande att leta efter varukoder, och att ofta kan man inte garantera att varukoder väljs rätt.
- Vikt och Kvantitet (19 kommentarer). I de flesta kommentarerna i denna kategori har uppgiftslämnarna ifrågasatt krav på rapportering av vikt eller kvantitet.
- Rapporteringmedium (20 kommentarer). Kommentarererna gällde i regel problem och svårigheter som uppstår vid rapportering via IDEP, webblankett eller pappersblankett. Bland annat konstaterade uppgiftslämnare att IDEP inte är tillräckligt användarvänligt och att webblankettsystemet¹² blir långsamt vid inmatning av många varuposter.
- Fakturor (15 kommentarer). Uppgiftslämnarna tyckte att fakturor borde innehålla information om varukoder, vikt, annan kvantitet och värde enligt EU-standard.
- Dåligt eller bra (14 kommentarer). En uppgiftslämnare tycker att " ... det är bra som det är". De tretton andra uppgiftslämnare ifrågasatte relevansen av rapportering till Intrastat i allmänhet eller konstaterade att den är för tidskrävande.
- Oväsentligt för undersökningen (11 kommentarer). Uppgiftslämnarna informerade om byte av organisationsnummer, ändring av företagsnamn och liknande.
- Övrigt (15 kommentarer). Varierande kommentarer och förslag angående Intrastatrapporteringen som inte ingick i någon av ovanstående kategorier.

4.7 JÄMFÖRELSE MED UNDERSÖKNINGEN AV UPPGIFTSLÄMNARBÖRDAN 2006

I detta avsnitt görs en analys av hur uppgiftslämnarbördan har förändrats under de senaste åren. Denna kommer att gå ut på att jämföra resultatet av denna undersökning med resultat av den senaste undersökningen av uppgiftslämnarbördan. Denna undersökning gjordes år 2005/2006 och var en delundersökning i ett projekt om Intrastat (14). Enkäten i undersökningen år 2006 var inte identisk med den här undersökningen, dock var 3 frågor identiska med motsvarande frågor i denna undersökning. Dessa var frågan om tidsåtgång (fråga 3 i denna undersökning), frågan om allmän upplevelse (fråga 1), samt frågan om säkerhet i valet av varukoder (fråga 5). Ytterligare två frågor - frågan om rapportering i Euro (fråga 9), samt frågan om införande av förenklad varukod för låga transaktionsvärden (fråga 8), - var samma i båda undersökningarna i termer av

¹² Även känd som SIV-systemet

frågeställning men olika i termer av svarsskala. Vi illustrerar diskrepanser mellan svarsskalorna avseende dessa två frågor i Tabell 22.

Problemet med diskrepanserna mellan de två skalorna löses på följande sätt:

- a) Slå ihop a) och b) samt d) och e) för undersökningen 2006
- b) Räkna om fördelningen av svar så att svarsalternativet "Möjligheten finns redan eftersom vi använder IDEP" inte är med gällande den nuvarande undersökningen (endast fråga 9)

Tabell 22: Jämförelse av svarsskalor avseende denna undersökning och undersökning år 2006: Frågor 8 och 9

| Svarsalternativ | Undersökning (År) | |
|-----------------|-------------------------------------|--|
| | 2006 | 2011 |
| a) | Stor förbättring | Positiv |
| b) | Liten förbättring | |
| c) | Varken förbättring eller försämring | Varken positiv eller negativ |
| d) | Liten försämring | Negativ |
| e) | Stor försämring | |
| f) | -- | Möjligheten finns redan eftersom vi använder IDEP (endast fråga 9) |

Vi fortsätter med att jämföra svaren mellan undersökningarna. Vi konstaterar först att svarsfrekvensen avseende undersökningen 2006 var **85,4 %** (854 av 1000) jämfört med 62,9 % i den nuvarande undersökningen. När det gäller förändringen av den genomsnittliga tidsåtgången för att rapportera (i minuter) under dessa år, är situationen följande:

Tabell 23: Förändring av genomsnittliga rapporteringstiden m.a.p. flöde, unika strata ej inkluderade

| Undersökning (År) | Snittid (införsel) | Snittid (utförsel) |
|----------------------------|--------------------|--------------------|
| 2006 | 177 | 145 |
| 2011 | 121 | 100 |
| Förändring i % mot År 2006 | - 31,6 % | - 31,0 % |

Notera att stratamet "alternativa media" i undersökningen 2006 samt strata "flera tilläggsnummer" och "övriga media" i denna undersökning inte ingår i tabellen. Anledningen till detta är att dessa strata inte delades in med avseende på flöde. En viktig anmärkning här är att tidsåtgången i de unika strata är (302 minuter för "övriga media" och 269 minuter för "flera tilläggsnummer") är betydligt större än tidsåtgången i stratamet "alternativa media" år 2006 (150 minuter). Därför är minskningen i tidsförbrukning för samtliga strata lite mindre än minskningen för införsel och utförsel (se Tabell 24).

När det gäller stratamet "flera tilläggsnummer", är det stora värdet på tidsåtgången en indirekt följd av urvalsförfarandet, eftersom enbart tilläggsnummer med maximalt antal varuposter togs med i urvalet för varje sådan uppgiftslämnare. Dock är det inte uppenbart varför tidsåtgången för stratamet "övriga media" är dubbelt så stor som tidsåtgången för stratamet "alternativa media" i den tidigare undersökningen.

Tabell 24: Förändring av den genomsnittliga rapporteringstiden: ordinära strata, unika strata, och totalt

| Undersökning (År) | Snittid (unika strata) | | Snittid (alla strata) |
|----------------------------|-------------------------|----------------------------|-----------------------|
| 2006 | 150 (alternativa media) | | 169 |
| 2011 | 302 (övriga media) | 269 (flera tilläggsnummer) | 119 |
| Förändring i % mot År 2006 | - | | - 29,6 % |

Vi fortsätter med att titta på undersökningarnas andra basvariabel – den allmänna upplevelsen i samband med rapporteringen.

Tabell 25: Jämförelse mellan undersökningarna: Allmän upplevelse

| Rapporteringen upplevs som ... | Införsel (%) | | Utförsel (%) | | Totalt (%) | |
|--------------------------------|--------------|------|--------------|------|------------|------|
| | 2006 | 2011 | 2006 | 2011 | 2006 | 2011 |
| Mycket lätt/ Ganska lätt | 79,7 | 69,6 | 86,5 | 73,5 | 81,4 | 70,8 |
| Ganska svår/ Mycket svår | 18,4 | 28,7 | 13,1 | 25,2 | 17,2 | 27,6 |
| Ingen åsikt/vet inte | 1,8 | 1,7 | 0,5 | 1,2 | 1,5 | 1,6 |
| Antal svar | 491 | 711 | 207 | 428 | 838 | 1248 |

Vi ser att minskningen i tidsåtgången under de senaste 5 åren inte ledde till positiv förändring av rapporteringsupplevelsen, utan tvärtom, kan vi se att attityden till rapporteringen har blivit mindre positiv för både in- och utförsel.

En annan fråga som var med i båda undersökningarna är säkerhet i valet av varukoder. I tabellen nedan illustreras det hur denna har förändrats under åren 2006-2011.

Tabell 26: Jämförelse mellan undersökningarna: Säkerhet i valet av varukoder

| Hur säker är man vid valet av varukoder | Införsel (%) | | Utförsel (%) | | Totalt (%) | |
|---|--------------|------|--------------|------|------------|------|
| | 2006 | 2011 | 2006 | 2011 | 2006 | 2011 |
| Mycket säker | 14,1 | 22,4 | 22,8 | 33,2 | 16,0 | 26,6 |
| Ganska säker | 60,2 | 63,3 | 55,9 | 57,5 | 59,6 | 60,9 |
| Ganska osäker | 16,9 | 11,1 | 12,3 | 6,0 | 15,7 | 9,0 |
| Mycket osäker | 4,4 | 1,3 | 3,3 | 1,2 | 3,9 | 1,3 |
| Ingen åsikt/vet inte | 3,8 | 1,9 | 5,7 | 2,1 | 4,2 | 2,2 |
| Antal svar | 497 | 713 | 211 | 429 | 850 | 1251 |

Vi ser att uppgiftslämnarna har blivit betydligt säkrare i valet av varukoder, vilket gäller både in- och utförsel. Vi fortsätter med att titta på frågan om införande av en förenklad varukod för låga fakturavärden. Se tabell 27.

Populariteten av förslaget om införande av förenklad varukod för låga fakturavärden har minskat med ca 10 % för införsel och har ökat obetydligt (ca 3 %) för utförsel.

Tabell 27: Jämförelse mellan undersökningarna: Införande av förenklad varukod för låga fakturavärden

| <i>Inställning till införande av förenklad varukod för låga fakturavärden</i> | <i>Införsel (%)</i> | | <i>Utförsel (%)</i> | | <i>Totalt (%)</i> | |
|---|---------------------|-------------|---------------------|-------------|-------------------|-------------|
| | <i>2006</i> | <i>2011</i> | <i>2006</i> | <i>2011</i> | <i>2006</i> | <i>2011</i> |
| <i>Positiv</i> | 62,8 | 52,5 | 43,8 | 46,8 | 57,8 | 50,3 |
| <i>Varken positiv eller negativ</i> | 35,7 | 46,2 | 53,8 | 48,9 | 39,6 | 47,1 |
| <i>Negativ</i> | 1,4 | 1,2 | 2,4 | 4,4 | 2,6 | 2,6 |
| <i>Antal svar</i> | 487 | 711 | 208 | 428 | 836 | 1248 |

Ett annat förbättringsförslag som ingick i båda undersökningarna är möjlighet till att rapportera det fakturerade värdet i Euro. Se Tabell 28 för resultatet av jämförelse mellan undersökningarna avseende denna variabel. Notera att samtliga svar som kom från uppgiftslämnare som rapporterar via IDEP är exkluderade från tabellen. Anledningen till detta är att IDEP-systemet har en inbyggd funktion som gör det möjligt att rapportera det fakturerade värdet i Euro och andra EU-valutor.

En annan anmärkning som bör göras är att de procentuella andelarna avseende den nuvarande undersökningen är beräknade med avseende på de tre första svarsalternativen (men inte den sista som avser rapportering via IDEP). Detta är för att underlätta jämförelsen mellan undersökningarna.

Tabell 28: Jämförelse mellan undersökningarna: Möjlighet att rapportera det fakturerade värdet i Euro

| <i>Inställning till möjlighet att rapportera det fakturerade värdet i Euro</i> | <i>Införsel (%)</i> | | <i>Utförsel (%)</i> | | <i>Totalt (%)</i> | |
|--|---------------------|-------------|---------------------|-------------|-------------------|-------------|
| | <i>2006</i> | <i>2011</i> | <i>2006</i> | <i>2011</i> | <i>2006</i> | <i>2011</i> |
| <i>Positiv</i> | 21,0 | 28,8 | 10,1 | 26,4 | 16,7 | 27,3 |
| <i>Varken positiv eller negativ</i> | 44,1 | 64,1 | 45,4 | 64,8 | 45,4 | 63,7 |
| <i>Negativ</i> | 34,8 | 6,0 | 37,8 | 8,1 | 37,7 | 6,9 |
| <i>Möjlighet finns redan eftersom vi använder IDEP</i> | - | 1,1 | - | 0,7 | - | 2,1 |
| <i>Antal svar</i> | 488 | 344 | 208 | 227 | 837 | 678 |

Jämfört med år 2006 är betydligt fler uppgiftslämnare positivt inställda till möjlighet att rapportera i Euro. Andel av uppgiftslämnare som är negativt inställda till förslaget har minskat drastiskt.

5 BEARBETNING OCH STATISTISK ANALYS

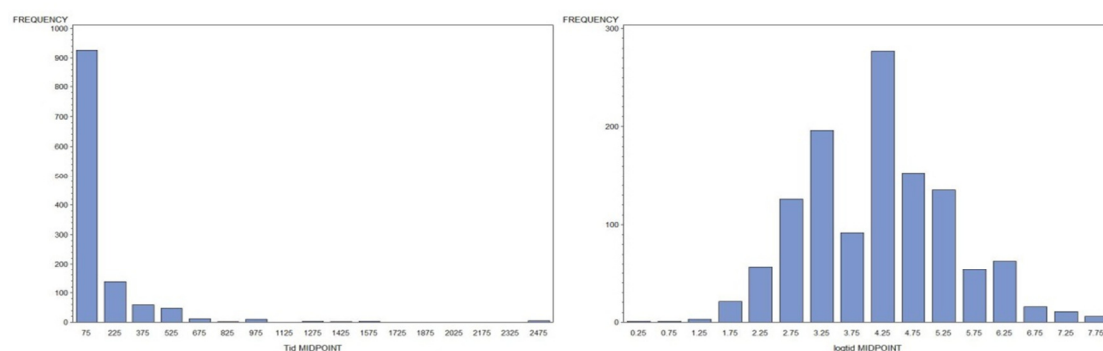
5.1 ANMÄRKNINGAR GÄLLANDE VARIABELN TIDSÅTGÅNG

I denna undersökning hade vi 11 variabler, varav 9 kategoriska variabler, 1 textvariabel (kommentarfältet), och en kvantitativ variabel (tidsåtgång). I denna avsnitt tittar vi på datastrukturen avseende den kvantitativa variabeln tidsåtgång samt undersöker vilka egenskaper den har.

Tiden det tog att rapportera varierar kraftigt från uppgiftslämnare till uppgiftslämnare. Det kan ta från några minuter till några dagar att rapportera till Intrastat. De allra flesta rapportörer behöver dock inte mer än några timmar för detta. För att vara mer precis, ligger rapporteringstiden inom intervall 0 - 4 timmar för 87 % av uppgiftslämnare. När de gäller de resterande 13 %, kan det ta upp till en vecka för dessa att rapportera, vilket gör att fördelning för tidsåtgång blir mycket skev (se Figur 3, vänster).

Ett naturligt tillvägagångssätt för skeva data är att ta en naturlig logaritm av observationerna för att få en mer normaliserad¹³ fördelning av data. Histogrammet av den logaritmerade data kan hittas i Figur 3 (till höger).

En annan strukturell egenskap av data är att vissa tidsvärden observeras mycket oftare än andra. Till exempel har hela 190 uppgiftslämnare rapporterat tidsåtgång lika med 30 minuter, och 209 uppgiftslämnare har angett 60 minuter. En naturlig förklaring till varför vissa tidsvärden förekommer betydligt oftare än andra är att människor är vana vid att mäta tiden i 5, 15, 30 eller 60 minuters intervall. Därför "avrundar" många uppgiftslämnare tidsåtgång till värden som de är vana vid att använda.



Figur 3: Histogram för det ursprungliga (vänster) och det logaritmerade (höger) data avseende variabeln Tidsåtgång

Med tanke på de två ovannämnda egenskaperna av data har vi beslutat att dela in tidsvariabeln i ett antal klasser för fortsatt analys. Det väcker i sin tur en fråga angående hur många grupper ska tidsvariabeln delas in i samt vilka tidsintervall ska dessa täcka. Ett alternativ som kan anses passa bra här är att genomföra indelningen på så sätt att varje värde som kan anses förekomma ofta får sin egen grupp. Emellertid anser vi att detta tillvägagångssätt inte är implementerbart på grund av att antal sådana ofta-

¹³ En struktur som är mer lik normalfördelning

förekommande värden, eller "toppar", är stort, vilket medför risk att för få observationer kommer att finnas inom varje kategori.

Tabell 29: Gruppering av variabeln Tidsåtgång

| <i>Klassificering</i> | <i>Tidsintervall som täcks av resp. grupp (min.)</i> | <i>Antal observationer</i> | <i>Typvärde (min.)</i> | <i>Medianvärde (min.)</i> | <i>Medelvärde (ej viktat) (min.)</i> |
|-------------------------------|--|----------------------------|------------------------|---------------------------|--------------------------------------|
| <i>Mycket kort tid</i> | 0 – 15 | 163 (13 %) | 15 | 10 | 11,6 |
| <i>Kort tid</i> | 16 – 45 | 335 (27 %) | 30 | 30 | 32,0 |
| <i>Medellång tid</i> | 46 – 105 | 288 (24 %) | 60 | 60 | 66,8 |
| <i>Lång tid</i> | 106 – 240 | 288 (24 %) | 120 | 120 | 156,9 |
| <i>Mycket lång tid</i> | 241 – 480 | 100 (8 %) | 480 | 360 | 387,3 |
| <i>Exceptionellt lång tid</i> | 481 - 2500 | 51 (4 %) | 600 | 940 | 1046,7 |

Därför har vi valt att dela in tidsvariabeln på följande sätt (Se Tabell 29). För varje kategori illustreras det medel- och medianvärde, samt det värde som förekommer oftast inom varje grupp. Dessutom illustrerar tabellen det absoluta och det relativa antalet observationer inom varje grupp.

Notera att tidsintervallernas längd växer exponentiellt. Detta reflekterar den skeva fördelningen av data.

Man kan se att median-, typ- samt medelvärden för grupperna ovan stämmer i regel inte överens, vilket följer ur datas skevhet. En intressant anmärkning är att mer än 95% behöver mindre än 8 timmar för att rapportera till Intrastat.

Vi fortsätter med att undersöka om det finns ett beroende mellan olika variabler. Detta kommer att ske genom att ta fram ett mått på korrelation mellan kategoriska variabler som också är ordinala. Detta mått heter Gamma, γ . Vi beskriver detta i mer detalj i nästa avsnitt.

5.2 BESKRIVNING AV STATISTISKA METODER

Efter att vi har delat in tidsvariabeln i grupper består undersökningsdata enbart av ordinala kategoriska variabler. Att en variabel är ordinal innebär att kategorierna i variabeln är ordnade, till skillnad från nominala variabler där ingen sådan ordning finns. Exempelvis representerar indelning m.a.p. rapporteringsmedia [papper, IDEP, webb, övriga] en nominal indelning och indelning m.a.p. antal varuposter [små, medelstora, stora företag] en ordinal indelning.

Om vi tabulerar två ordinala variabler tillsammans får vi en $I \times J$ kontingenstabell, där I och J betecknar antal kategorier i variablerna. Det som man oftast vill undersöka är om variablerna är ömsesidigt oberoende, vilket vanligen testas genom att ta fram en oddskvot eller genom att tillämpa Pearsons χ^2 -test. Dessa statistiska metoder passar väl då variablerna är nominala, dock förloras det information om eventuell trend i data om

variablerna är ordinala. För att kunna fånga denna information finns det andra statistiska metoder, och den metoden som vi kommer att tillämpa här är Gamma test, som även är känd som Goodman och Kruskals Gamma. Precis som andra korrelationsmått, kan γ ligga i intervallet $[-1; 1]$, där -1 innebär 100 % negativt samband (perfekt inversion), och 1 innebär 100% positivt samband (perfekt överensstämmelse). Ett värde $\gamma=0$ innebär avsaknad av någon korrelation mellan de testade variablerna. Vi tillämpar följande tolkning av värden av γ :

Tabell 30: Tolkning av γ

| Intervall för γ | Mått på korrelation |
|-----------------------------------|---------------------|
| $[0.5 ; 1]$ och $[-0.5 ; -1]$ | Stark korrelation |
| $[0.2 ; 0.5]$ och $[-0.2 ; -0.5]$ | Måttlig korrelation |
| $[-0.2 ; 0.2]$ | Ingen korrelation |

Gamma behandlar variabler symmetriskt, alltså behöver man inte identifiera en av variablerna som respons. Följande formel används för uträkning av skattning av Gamma:

$$\hat{\gamma} = \frac{C - D}{C + D} \quad [4]$$

Här innebär C och D antal konkordanta respektive diskordanta par i en kontingenstabell. Om en kontingenstabell reflekterar samband mellan två kategoriska variabler X och Y , är ett par av observationer konkordant i fall den observationen som har högre rank med avseende på X också har högre rank med avseende på Y . På samma sätt är ett par diskordant om observationen som har högre rank med avseende på X har lägre rank med avseende på Y . I fall någon av observationerna i paret har samma rank med avseende på någon av variablerna kallas paret *tied*.

Stort antal konkordanta par och litet antal diskordanta par kan vara ett tecken på att det finns en positiv trend i data. Å andra sidan, kan stort antal diskordanta par (och litet antal konkordanta par) vara ett tecken på negativ association. Vi går inte in i detaljer på hur man räknar ut antal konkordanta respektive diskordanta par, men en intresserad läsare hänvisas till (15). Vi noterar dock att $\hat{\gamma}$ är en maximum-likelihood skattning av γ , och själva γ definieras enligt nedan:

$$\gamma = \frac{\Pi_c - \Pi_d}{\Pi_c + \Pi_d} \quad [5]$$

Här är Π_c och Π_d sannolikheter för ett slumpmässigt valt par av observationer att vara konkordant (c) eller diskordant (d). Igen, hänvisar vi till Agrestis bok för mer detaljer. Asymptotisk varians för γ kan räknas ut vilket gör det möjligt att bl.a. implementera asymptotiska, normalbaserade hypotestest. Det enda som krävs är att antal observationer i cellerna i kontingenstabellen är tillräckligt stort för centrala gränsvärdesatsen att kunna gälla. Ett exakt uttryck för variansen av γ samt en detaljerad beskrivning av detta kan en intresserad läsare hitta i (16).

5.3 OMKODNING AV VARIABELN "FRAMTAGNING AV UPPGIFTER"

När det gäller variabeln "Framtagning av uppgifter" (Fråga 7), bör omkodning av svar göras eftersom man kunde ange fler än ett svarsalternativ då man besvarade denna fråga. Efter deskriptiv undersökning av data har vi beslutat att omkoda denna variabel på följande sätt:

- Svaren kommer att delas in i 5 ordnade grupper: (Uppgifterna togs fram ...)
 - 1) direkt ur företagets datorsystem (svarsalternativ 1)
 - 2) ur företagets datorsystem, men behövde bearbetas (svarsalternativ 2)
 - 3) ur företagets datorsystem, men behövde bearbetas, samt hämtades från andra källor
 - 4) hämtades från andra källor (svarsalternativ 3)
 - 5) uppskattades utifrån annan kunskap (svarsalternativ 4)
- De uppgiftslämnare som enbart har angett ett svarsalternativ, kommer att hamna i grupp 1, 2, 4 eller 5 beroende på hur de har svarat. Nästan 95 % av respondenter tillhör denna grupp.
- De uppgiftslämnare som har kryssat i både svarsalternativ 2 och 3 kommer att hamna i grupp 3 (43 st. uppgiftslämnare). De som har kryssat i svarsalternativ 1 och 4 kommer också att hamna i grupp 3 (4 st. uppgiftslämnare). På motsvarande sätt kommer de som har angett svarsalternativ 1 och 3 hamna i grupp 2 (4 st. uppgiftslämnare); och de som har angett alternativ 2 och 4 hamnar i grupp 4 (1 uppgiftslämnare).
- Ur analysen väljer vi att exkludera de 6 st. uppgiftslämnare som har angett svarsalternativ 1 och 2, samt de uppgiftslämnare som har kryssat i alternativ 3 och 4 (2 st.). Dessutom kommer 5 st. uppgiftslämnare som har angett 3 svarsalternativ att uteslutas. Ingen har kryssat i samtliga 4 svarsrutor.

5.4 IMPLEMENTERING AV HYPOTESTEST

Vi anser att det är av intresse att genomföra parvisa korrelationstest mellan följande variabler:

Tidsåtgång – Allmän upplevelse. Eventuell korrelation mellan dessa variabler kan säga mycket om eventuell påverkan av den subjektiva delen av uppgiftslämnarbördan på dess objektiva del.

Tidsåtgång – Framtagning av uppgifter. Kan korta rapporteringstider kopplas till hög automatiseringsnivå av rapporteringsprocessen hos uppgiftslämnarna?

Tidsåtgång – Antal medverkande. Vad finns det för samband mellan antal medverkande personer och den objektiva uppgiftslämnarbördan?

Allmän upplevelse – Antal medverkande. Vad finns det för samband mellan antal medverkande personer och den subjektiva uppgiftslämnarbördan?

Tidsåtgång – Säkerhet i valet av varukoder. Orsakar osäkerhet i valet av varukoder fördröjningar i rapporteringsprocessen?

Allmän upplevelse – Säkerhet i valet av varukoder. Hur mycket bidrar säkerhet/osäkerhet i valet av varukoder till att hela rapporteringsrutinen upplevs mer positivt/negativt?

Hypotestesterna som ska genomföras kommer att basera sig på allt svarsdata som vi har tillgängligt. Observationer där det förekommer partiellt bortfall för någon av de två variabler som testas kommer att exkluderas ur analysen.

Svarsalternativen beträffande variablerna 1 (Allmän upplevelse) samt variabel 6 (Säkerhet i valet av varukoder), kommer att ordnas på följande sätt:

Mycket lätt/säker → Ganska lätt/säker → Ingen åsikt → Ganska svårt/osäker → Mycket svårt/osäker

När det gäller variabel 2 (Antal medverkande), kommer vi att bortse från de svaren som hamnade i kategorin "Kan ej uppskatta antal medverkande".

TIDSÅTGÅNG – ALLMÄN UPPELVELSE

Redan vid den deskriptiva undersökningen av data (se Tabell 16) kan man konstatera ett tydligt samband mellan dessa variabler.

Programpaketet SAS ger $\hat{\gamma} = 0,558$ (baserat på 1225 observationer). 95 % asymptotisk konfidensintervall för $\hat{\gamma}$ är $[0,502 ; 0,615]$, och testning av $H_0: \gamma = 0$ resulterar i (tvåsidigt) p-värde mindre än 0.001. Vi förkastar alltså nollhypotesen om ingen association mellan tidsåtgång och allmän upplevelse och konstaterar att det förekommer stark positiv korrelation mellan dessa variabler (Längre tid är associerad med sämre attityd till rapportering).

TIDSÅTGÅNG – FRAMTAGNING AV UPPGIFTER

Även här förväntar vi oss en tydlig korrelation också (Se Tabell 14). Antal observationer är 1200, och skattningen av γ visar sig vara lika med 0.35 med asymptotisk konfidensintervall $[0,287 ; 0,414]$. Nollhypotesen $\gamma = 0$ förkastas på grund av att p-värde är mycket liten, mindre än 0.001. Dessutom är nedre gränsen av konfidensintervallet större än 0.2.

Alltså konstaterar vi att det förekommer positiv korrelation mellan tidsåtgång och sätt att ta fram varukoder, i att de "mindre avancerade" framtagningssätt kan kopplas till längre rapporteringstider. Dock är korrelationen något svagare här än den är för det föregående variabelparet.

TIDSÅTGÅNG – ANTAL MEDVERKANDE

Se Tabell 10 för deskriptiv undersökning av data. Implementering av hypotestest via programpaketet SAS resulterar i $\hat{\gamma} = 0,312$; 95% konfidensintervall $\in [0,233 ; 0,39]$. Detta är baserad på 1219 observationer. P-värdet vid test av $H_0: \gamma = 0$ är återigen mindre än 0.001, så nollhypotesen förkastas. Alltså är tidsåtgång måttligt positivt korrelerad till antal personer som medverkar i rapporteringen till Intrastat.

ALLMÄN UPPLEVELSE – ANTAL MEDVERKANDE

Nedan illustreras kontingenstabell som reflekterar samband mellan dessa variabler (se Tabell 31).

Vi undersöker dessa data med hjälp av SAS. Skattningen av γ blir lika med 0.347, och 95 % konfidensintervall blir [0.256 ; 0.437]. H_0 om ingen association emellan variablerna förkastas på grund av ett lågt p-värde, som är igen mindre än 0.001.

Tabell 31: Allmän upplevelse – Antal medverkande

| Allmän upplevelse / Antal medverkande | 1 | 2 | 3 eller fler | Totalt |
|---------------------------------------|-----------|-----|--------------|--------|
| Mycket lätt | 125 (87%) | 17 | 2 | 144 |
| Ganska lätt | 536 (75%) | 149 | 34 | 719 |
| Ingen åsikt | 9 (69%) | 3 | 1 | 13 |
| Ganska svårt | 178 (65%) | 71 | 26 | 275 |
| Mycket svårt | 31 (46%) | 27 | 10 | 68 |
| Totalt | 879 | 267 | 73 | 1219 |

TIDSÅTGÅNG – SÄKERHET I VALET AV VARUKODER

Samband mellan variablerna tabuleras nedan.

Tabell 32: Tidsåtgång – Säkerhet i valet av varukoder

| Tidsåtgång (min.) / Säkerhet i valet av varukoder | Mycket säker | Ganska säker | Ingen åsikt | Ganska osäker | Mycket osäker | Totalt |
|---|--------------|--------------|-------------|---------------|---------------|--------|
| 0 – 15 | 63 | 90 | 4 | 5 | 0 | 162 |
| 16 – 45 | 87 | 204 | 11 | 27 | 2 | 331 |
| 46 – 105 | 60 | 188 | 3 | 25 | 7 | 283 |
| 106 – 240 | 78 | 165 | 7 | 32 | 2 | 284 |
| 241 – 480 | 23 | 61 | 5 | 7 | 2 | 98 |
| 481 - 2500 | 13 | 29 | 0 | 7 | 1 | 50 |
| Totalt | 324 | 737 | 30 | 103 | 14 | 1208 |

Vi får $\hat{\gamma} = 0.113$, och konfidensintervallet (95%) blir [0.04 ; 0.186]. Vid test av $H_0: \gamma = 0$ blir p-värdet 0.003. Vi förkastar alltså nollhypotesen om att $\gamma = 0$, dock kan vi inte påstå att det finns någon korrelation mellan variablerna eftersom $\hat{\gamma}$ är så pass litet och eftersom båda gränser av konfidensintervallet ligger i [-0.2 ; 0.2].

ALLMÄN UPPLEVELSE – SÄKERHET I VALET AV VARUKODER

Vi tabulerar data avseende dessa variabler i Tabell 33.

Programpaketet SAS ger $\hat{\gamma} = 0.426$, 95% konfidensintervall är [0.346 ; 0.506]. P-värdet beträffande testet av $H_0: \gamma = 0$ blir, som förväntat, mindre än 0.001. Vi ser att variabeln Säkerhet i valet av varukoder har en mycket starkare korrelation med variabeln Allmän upplevelse än med variabeln Tidsåtgång. Detta kan vara ett tecken på att svårigheter i samband med framtagning av varukoder bidrar mycket till att hela Intrastatrapporteringen ses som betungande.

Tabell 33: Allmän upplevelse – Säkerhet i valet av varukoder

| <i>Allmän upplevelse / Säkerhet i valet av varukoder</i> | <i>Mycket säker</i> | <i>Ganska säker</i> | <i>Ingen åsikt</i> | <i>Ganska osäker</i> | <i>Mycket osäker</i> | <i>Totalt</i> |
|--|-------------------------|-------------------------|------------------------|--------------------------|--------------------------|---------------|
| <i>Mycket lätt</i> | 73 | 67 | 3 | 1 | 0 | 144 |
| <i>Ganska lätt</i> | 189 | 464 | 14 | 40 | 6 | 713 |
| <i>Ingen åsikt</i> | 6 | 5 | 1 | 1 | 0 | 13 |
| <i>Ganska svårt</i> | 49 | 173 | 8 | 37 | 4 | 271 |
| <i>Mycket svårt</i> | 7 | 28 | 4 | 24 | 4 | 67 |
| <i>Totalt</i> | 324 | 30 | 30 | 103 | 14 | 1208 |

Vi avslutar detta kapitel med att sammanställa resultat av den genomförda dataanalysen.

Tabell 34: Sammanfattning av resultat från korrelationstest

| <i>Test</i> | $\hat{\gamma}$ | <i>95 % konfidens- intervall</i> | <i>Mått på korrelation</i> |
|---|----------------|--------------------------------------|--------------------------------|
| Tidsåtgång – Allmän upplevelse | 0.558 | [0.502 ; 0.615] | Starkt positiv |
| Tidsåtgång – Framtagning av uppgifter | 0.350 | [0,287 ; 0,414] | Måttligt positiv |
| Tidsåtgång – Antal medverkande | 0.312 | [0.233 ; 0.390] | Måttligt positiv |
| Allmän upplevelse – Antal medverkande | 0.347 | [0.256 ; 0.437] | Måttligt positiv |
| Tidsåtgång – Säkerhet i valet av varukoder | 0.113 | [0.040 ; 0.186] | Ingen |
| Allmän upplevelse – Säkerhet i valet av varukoder | 0.426 | [0.346 ; 0.506] | Måttligt positiv |

6 UNDERSÖKNING AV ORSAKER TILL EXCEPTIONELLT LÅNG RAPPORTERINGSTID FÖR VISSA UPPGIFTSLÄMNARE: EN STATISTISK MODELL

6.1 INLEDANDE ANMÄRKNINGAR

Tidigare har vi noterat att variation mellan rapporteringstider för olika uppgiftslämnare är mycket stor. Det kan ta från 5 minuter till ca 1 vecka att rapportera till Intrastat. Samtidigt konstaterar vi att över 87 % av uppgiftslämnarna behöver 4 timmar eller mindre för detta. De få företag och organisationer som inte hör till dessa 87 % behöver däremot mycket längre tid för att rapportera.

Att en uppgiftslämnare behöver mer än 4 timmar för att rapportera kan bero på en mängd olika faktorer. Dessa faktorer kan vara sådana som vi inte kan observera, såsom icke-perfekta rutiner för databearbetning, som inte täcktes av undersökningsfrågorna¹⁴, eller den mänskliga faktorn. Samtidigt kan långa rapporteringstider bero på faktorer som vi har inkluderat i vår undersökning genom stratifiering. Nämligen anser vi att flöde, rapporteringsmedium samt antal rapporterade varuposter per månad kan ha större effekt på uppgiftslämnarbördan. I Tabell 9 ser vi att tidsåtgången för att rapportera inte är samma för de olika strata. Om man dock statistiskt kunde visa att viss stratatillhörigheten har samband med exceptionellt långa rapporteringstider, skulle detta öka vår förståelse av vad sådana tider orsakas av.

Enligt ovanstående resonemang har vi bestämt oss för att skapa en modell för analys av sambandet mellan långa rapporteringstider och strata. Tidsåtgången, som är i det här fallet responsvariabeln, kommer att delas in i två kategorier: 0 – 240 minuter (kort och normal tid) och 241+ minuter (lång och exceptionellt lång tid). Vi kommer med hjälp av en typ av logistisk regressionsmodell undersöka om uppgiftslämnarnas risk att bli exponerad, d.v.s. att hamna i den andra kategorin påverkas av deras stratatillhörighet. Denna chans kommer att indirekt mätas med oddskvot, som är ett mått på samband mellan kategoriska variabler.

Innan vi fortsätter, bör vissa antaganden göras:

- Observationerna (uppgiftslämnare) är iid (oberoende och likafördelade). Detta anser vi stämma, med ett undantag (se nedan).
- De uppgiftslämnare som har valt att inte svara på undersökningen skiljer sig inte från dem som har valt att besvara frågan om tidsåtgång med avseende på varje stratum. Alltså antal vi att bortfallet har varit MAR (Missing at random).

När det gäller det första antagandet, noterar vi att detta kan gälla samtliga strata förutom ett. Uppgiftslämnare som tillhör stratumet "flera tilläggsnummer" är nämligen inte av

¹⁴ Exempelvis har en uppgiftslämnare (med rapporteringstid lika med 1 vecka) konstaterat att respektive Intrastatsansvarig får besöka lager varje gång det sker en varuleverans för att ta reda på vad det är för varor som har kommit in.

samma typ som de andra. Så är det inte enbart för att det inte är känt för oss hur rapportering på sådana företag fördelar sig mellan olika tilläggsnummer utan även på grund av att detta stratum enbart innehåller tilläggsnummer med störst antal varuposter för respektive organisationsnummer. På grund av detta exkluderar vi stratumet "flera tilläggsnummer" ur modelleringen.

En ytterligare anmärkning bör göras beträffande ett annat unikt stratum, övriga media. Eftersom uppgiftslämnare i detta stratum inte delades in med avseende på flöde eller antal varuposter, kan vi inte inkludera detta stratum i modellen heller. I stället kommer vi att tabulera data avseende detta stratum mot samtliga ej-unika strata i en kontingenstabell.

6.2 BESKRIVNING AV MODELLEN

Det faktum att både responsvariabeln och de förklarande variablerna är kategoriska betyder att det är lämpligast att tillämpa en logistisk regressionsmodell. Emellertid kan vi inte tillämpa en vanlig logistisk modell på grund av att vissa observationer i vårt dataset "väger mer" än andra. Detta är på grund av att vissa strata är antingen under- eller överrepresenterade, vilket är en följd av urvalet och bortfallet. På grund av detta bör man använda en modell som kommer att ta hänsyn till de olika vikterna som varje stratum har. Detta implementeras genom att vikta olika komponenter av (pseudo) log-likelihood funktion, som används för att göra skattningar av oddskvoter samt deras (asymptotiska) varianser. Modellen tar även hänsyn till att populationen (antal uppgiftslämnare) är ändlig. En procedur i SAS som vi kommer att använda för modelleringen heter PROC SURVEYLOGISTIC (Se (17)¹⁵ för detaljer). Nedan följer en detaljerad beskrivning av modellen. Beskrivningen förutsätter vissa kunskaper i matematisk statistik, och kan skippas av läsare som inte är intresserade av modellens detaljer.

Vid beskrivning av modellen kommer följande beteckningar att användas:

Låt Y vara en binär responsvariabeln, dvs. $Y=0$ om en uppgiftslämnare behöver 4 timmar eller mindre för att rapportera, eller $Y=1$ om en uppgiftslämnare behöver exceptionellt lång tid för att rapportera (mer än 4 timmar). Låt vidare \mathbf{x} beteckna en vektor av förklarande variabler, som består av p dimensioner (alltså har vi p förklarande variabler). Notera att modellen innehåller intercept, därav $\mathbf{x}_{(1)} = 1$. Förklarande variabler kommer att kodas på följande sätt:

Flöde: 1 dummyvariabel; 0 – införsel (referens); 1 – utförsel.

Antal varuposter: 1 dummyvariabel; 0 – litet antal varuposter (referens); 1 – medel antal varuposter; 2 – stort antal varuposter.

Rapporteringsmedium: 2 dummyvariabler: 00 för IDEP (referens); 10 för webb; samt 01 för papper. Det används 2 dummy eftersom rapporteringsmedium inte är någon ordinal variabel.

¹⁵ <http://www2.sas.com/proceedings/sugi27/p258-27.pdf> [2011-09-20]

$i = 1, 2, \dots, I$ innebär stratum, och I innebär det totala antalet strata. De tre stratifieringsfaktorerna (flöde, antal varuposter och medium) slås ihop här, alltså innebär varje i ett enskilt stratum. Alltså kan exempelvis $i = 7$ innebära utförsel; IDEP samt litet antal varuposter. Det totala antalet strata, I , är i vårt fall 18 ($2 \times 3 \times 3$).

$j = 1, 2, \dots, J_i$ innebär löpnummer av en observation inom stratum i , med totala antalet enheter lika med J_i i stratum i .

w_{ij} betecknar stickprovsvikt för enhet j i stratum i .

y_{ij} är en indikatorvariabel för responsen Y . Alltså är $y_{ij} = 1$ om tidsåtgång för enhet j i stratum i är längre än 4 timmar, och $y_{ij} = 0$ annars.

\mathbf{x}_{ij} innebär p -dimensionell radvektor av förklarande variabler för element j i stratum i .

$N = \sum_{i=1}^I J_i$ är den totala urvalsstorleken; och f_i är samplingsfrekvensen för stratum i .

Vidare definierar vi π_{ij} som ett väntevärde av responsvariabeln givet de förklarande variablerna:

$$\pi_{ij} = E[y_{ij} | \mathbf{x}_{ij}] \quad [6]$$

I fortsättningen kommer vi att operera med begrepp i samband med generaliserade linjära modeller (GLM). För utförlig beskrivning av denna typ av statistiska modeller hänvisar vi till (15), (18) eller (19).

Låt $g(\cdot)$ vara en länkfunktion sådan att

$$\pi_{ij} = g^{-1}(\mathbf{x}, \boldsymbol{\theta}) = \mathbf{f}(\mathbf{x}, \boldsymbol{\theta}) \quad [7]$$

Här innebär $\boldsymbol{\theta}$ en p -dimensionell kolumnvektor av regressionskoefficienter, $\boldsymbol{\theta} = (\beta_1, \beta_2, \dots, \beta_p)^T$. Vi väljer att använda logit-länken, som är den kanoniska länken för binomialfördelade data. Inversen till länkfunktionen blir därmed följande:

$$\pi_{ij} = \frac{\exp\{\mathbf{x}_{ij}\boldsymbol{\theta}\}}{1 + \exp\{\mathbf{x}_{ij}\boldsymbol{\theta}\}} \quad [8]$$

Vi konstaterar vidare att Pseudo log-likelihood funktionen blir

$$L(\boldsymbol{\theta}) = \sum_{i=1}^I \sum_{j=1}^{J_i} w_{ij} [y_{ij} \log \pi_{ij} + (1 - y_{ij}) \log(1 - \pi_{ij})] \quad [9]$$

Kom ihåg att log-likelihood för en Bernoulli-fördelad stokastisk variabel, t.ex. $Z \sim Be(\mu)$, är

$$L(\mu) = z \log \mu + (1 - z) \log(1 - \mu)$$

Det underlättar förståelsen av varför log-likelihooden ser ut som den gör. $L(\boldsymbol{\theta})$ är alltså en viktad log-likelihoodfunktion för samtliga N observationer. Vi fortsätter med att derivera $L(\boldsymbol{\theta})$ med avseende på $\boldsymbol{\theta}$ för att få fram Scorefunktionen :

$$\mathbf{U}(\boldsymbol{\theta}) = \frac{dL(\boldsymbol{\theta})}{d\boldsymbol{\theta}} = \sum_{i=1}^I \sum_{j=1}^{J_i} w_{ij}(\mathbf{d}_{ij}) \left(\frac{1}{\pi_{ij}(1-\pi_{ij})} \right) (y_{ij} - \pi_{ij}) \quad [10]$$

Här betecknar \mathbf{d}_{ij} en vektor av partiella derivat av länkfunktionen \mathbf{f} med avseende på $\boldsymbol{\theta}$. Denna vektor visar sig vara lika med följande uttryck för multipel logistisk regressionsmodell:

$$\mathbf{d}_{ij} = \pi_{ij}(1 - \pi_{ij})\mathbf{x}_{ij} \quad [11]$$

Skattning av parametervektorn $\boldsymbol{\theta}$ sker på följande sätt. Man väljer något startvärde för $\boldsymbol{\theta}$, som vi kallar för $\boldsymbol{\theta}^{(0)}$. Låt vidare skattning av $\boldsymbol{\theta}$ under steg l betecknas som $\boldsymbol{\theta}^{(l)}$. Skattning under steg $l+1$ blir då

$$\boldsymbol{\theta}^{(l+1)} = \boldsymbol{\theta}^{(l)} + [\mathbf{I}^{(l)}]^{-1}\mathbf{U}^{(l)} \quad [12]$$

Här innebär $\mathbf{U}^{(l)}$ Scorefunktionen enligt [10], och $\mathbf{I}^{(l)}$ betecknar den förväntade (Fishers) informationen. Denna är en $p \times p$ matris, där varje element visar sig vara följande:

$$\mathbf{I}_{km}^{(l)} = \sum_{i=1}^I \sum_{j=1}^{J_i} w_{ij} \mathbf{d}_{ij(k)}^{(l)} \frac{1}{\pi_{ij}^{(l)}(1-\pi_{ij}^{(l)})} \mathbf{d}_{ij(m)}^{(l)}; \quad k, m \in [1 \dots p] \quad [13]$$

Notera att $\mathbf{d}_{ij}^{(l)}$ och $\pi_{ij}^{(l)}$ evalueras i $\boldsymbol{\theta}^{(l)}$. Proceduren upprepas tills konvergens är uppnådd. Denna metod är känd som Fishers-Scoring algoritim, dock är den identisk med Newton-Raphsons algoritim eftersom länkfunktionen är kanonisk (Se ovan). Detta följer av att en kanonisk länkfunktion innebär att den förväntade (Fishers) och den observerade informationerna blir lika med varandra.

Följande uttryck används för skattning av den estimerade kovariansmatrisen av $\hat{\boldsymbol{\theta}}$:

$$\widehat{Var}(\hat{\boldsymbol{\theta}}) = \widehat{\mathbf{I}}^{-1} \hat{\boldsymbol{\theta}} \widehat{\mathbf{I}}^{-1} + \kappa \lambda \widehat{\mathbf{I}}^{-1} \quad [14]$$

Här innebär

$$\hat{\boldsymbol{\theta}} = \frac{N-1}{N-p} \sum_{i=1}^I \frac{J_i(1-f_i)}{J_i-1} \sum_{j=1}^{J_i} (\mathbf{e}_{ij} - \mathbf{e}_{i\cdot})^T (\mathbf{e}_{ij} - \mathbf{e}_{i\cdot})$$

$$\mathbf{e}_{ij} = w_{ij} \widehat{\mathbf{d}}_{ij} \frac{1}{\widehat{\pi}_{ij}(1-\widehat{\pi}_{ij})} (y_{ij} - \widehat{\pi}_{ij})$$

$$\mathbf{e}_{i\cdot} = \frac{1}{J_i} \sum_{j=1}^{J_i} \mathbf{e}_{ij}$$

Notera att $\widehat{\mathbf{d}}_{ij}$ och $\widehat{\pi}_{ij}$ evalueras i $\hat{\boldsymbol{\theta}}$. Vi kan se att [14] fås genom att summera två element. Det första elementet heter Sandwich estimator, p.g.a. Fishers informationen på bägge sidorna av uttrycket kan ses som "bröd", och komponentet $\hat{\boldsymbol{\theta}}$ innebär "smör". Metoden går

ut på att använda Taylorapproximeringen för att skatta stickprovsvariansen. För en utförligare beskrivning av metoden hänvisar vi till (20).

Komponenten $\kappa \lambda \hat{I}^{-1}$ heter Morel-justering. Vi går inte in i detaljer angående denna heller, men hänvisar en intresserad läsare till beskrivningen av SURVEYLOGISTIC-proceduren på SAS hemsida eller till (21). En av fördelarna med Morel-tillägget är att den reducerar risken för Typ I-fel om stickprovet är för litet.

6.3 MODELLVAL SAMT RESULTAT

Modellframställningen kommer att genomföras genom att konstruera ett antal modeller med olika förklaringsvariabler, och välja den modellen som anses passa bäst. Modellvalet kommer att grunda sig dels på de förklarande variabelernas rimlighet dels på ett kriterium som heter AIC (Akaike's Information Criterion). Detta kriterium används för att undersöka lämplighet av en modell och definieras som

$$AIC = 2k - 2 \ln L(\theta), \quad [15]$$

Här betecknar k antal parametrar i en modell och $L(\theta)$ – likelihoodfunktionen av den skattade modellen. Alltså ju lägre AIC desto bättre anpassning av modellen. Det vanliga tillvägagångssättet är att först konstruera en modell med maximalt antal förklarande variabler, och därefter exkludera variabelerna stegvis för att till slut komma till den minimala modellen. Man tittar senare på modellernas AIC och bestämmer vilken modell som ska väljas.

I vårt fall har vi 3 förklarande variabler: flöde; rapporteringssätt; samt antal varuposter. Vi inkluderar dessutom samspel mellan dessa för att inte missa eventuella samspelseffekter. Dessa är flöde*rapporteringssätt; flöde*antal varuposter; och rapporteringssätt*antal varuposter.

MODELL A: SAMTLIGA EFFEKTER OCH SAMSPEL

Programpaketet SAS ger AIC = 3588.628. Se tabell nedan för analys av effekter.

Tabell 35: Modell A – Analys av effekter

| Variabel | df | Wald χ^2 | P-värde |
|-------------------------|----|---------------|---------|
| Flöde | 1 | 0.618 | 0.43 |
| Rapporteringsmedium | 2 | 17.474 | < 0.01 |
| Antal varuposter | 1 | 15.062 | < 0.01 |
| Antal varuposter*Medium | 2 | 27.137 | < 0.01 |
| Flöde*Medium | 2 | 2.549 | 0.28 |
| Antal varuposter *Flöde | 1 | 0.715 | 0.40 |

Notera att denna tabell enbart innehåller design-baserade χ^2 –statistikor för samtliga effekter men inte skattningar på oddskvoter. Dessa illustreras inte för denna modell, dock kommer vi att illustrera skattningar på oddskvoter för de andra modellerna. Vi ser att

effekterna flöde, flöde*medium samt storlek*flöde inte är statistiskt signifikanta. Vi exkluderar de två samspelseffekterna men inte flöde.

MODELL B: FLÖDE; MEDIUM; ANTAL VARUPOSTER; ANTAL VARUPOSTER*MEDIUM

AIC blir 3599.089, vilket är betydligt högre än för modell A. Analysen av effekter avseende denna modell illustreras i Tabell 36.

Som förväntat, är samtliga effekter förutom flöde statistiskt signifikanta. Vi fortsätter med att illustrera maximum-likelihood skattningar för logoddskvoter. I Tabell 37 nedan ser man att samspelseffekten antal varuposter*medium är signifikant med avseende på webb, men inte med avseende på papper. Dessutom är webb signifikant men ej papper. Notera att införsel och IDEP används som referensvärden.

Tabell 36: Modell B – Analys av effekter

| Variabel | Df | Wald χ^2 | P-värde |
|-------------------------|----|---------------|---------|
| Flöde | 1 | 1.546 | 0.21 |
| Rapporteringsmedium | 2 | 18.35 | < 0.01 |
| Antal varuposter | 1 | 24.75 | < 0.01 |
| Antal varuposter*Medium | 2 | 27.84 | < 0.01 |

Tabell 37: Modell B - Parameterskattningar

| Parameter | Skattning av logodds | Standardavvikelse | Wald χ^2 | P-värde |
|-------------------------------|----------------------|-------------------|---------------|---------|
| Intercept | - 3.519 | 0.41 | 75.01 | < 0.01 |
| Flöde: Utförsel | - 0.258 | 0.21 | 1.55 | 0.21 |
| Medium: Papper | 0.61 | 0.85 | 0.51 | 0.47 |
| Medium: Webb | - 2.767 | 0.71 | 15.15 | < 0.01 |
| Antal varuposter | 0.786 | 0.16 | 24.75 | < 0.01 |
| Antal varuposter*Medium Webb | - 0.463 | 0.35 | 1.77 | 0.18 |
| Antal varuposter *Medium Webb | 1.228 | 0.27 | 20.85 | < 0.01 |

Exponentiering av skattningen av logodds för flöde ger Odds = 0.773; 95% konfidensintervall $\in [0.515 ; 1.160]$. Alltså kan vi inte påstå att rapportering av utförsel innebär lägre risk att ha lång rapporteringstid jämfört med de som rapporterar införsel. Detta trots att oddset är betydligt lägre än 1.

Vi exkluderar variabeln flöde och fortsätter vidare.

MODELL C: MEDIUM; ANTAL VARUPOSTER; ANTAL VARUPOSTER*MEDIUM

SAS ger AIC = 3604.79, vilket är det högsta värdet hittills. Samtliga effekter blev signifikanta. Vi illustrerar skattningar av logodds i tabellen nedan.

Tabell 38: Modell C – Parameterskattningar

| <i>Parameter</i> | <i>Skattning av logodds</i> | <i>Standardavvikelse</i> | <i>Wald χ^2</i> | <i>P-värde</i> |
|-----------------------|-----------------------------|--------------------------|---------------------------------|----------------|
| Intercept | -3.609 | 0.40 | 80.9 | < 0.01 |
| Medium: Papper | 0.658 | 0.85 | 0.59 | 0.44 |
| Medium: Webb | -2.733 | 0.71 | 14.72 | < 0.01 |
| Storlek | 0.784 | 0.16 | 24.35 | < 0.01 |
| Storlek*Medium Papper | -0.483 | 0.35 | 1.92 | 0.17 |
| Storlek*Medium Webb | 1.210 | 0.27 | 20.22 | < 0.01 |

Borttag av variabeln "flöde" har inte haft stort inflytande på skattningarna av logodds. Igen ser vi att någon statistisk skillnad saknas mellan strata papper och IDEP (som är ett referensstratum). Däremot noterar vi att rapportering via webb innebär i regel kortare rapporteringstider (den negativa skattningen av webb är signifikant), dock ej om antal varuposter är stort (den positiva skattningen av papper är också signifikant).

Vi kan se att samtliga 3 modellerna har sina för- och nackdelar. Den första modellen har lägst AIC, dock innehåller den variabler som inte är statistiskt signifikanta. Den andra modellen innehåller en icke-signifikant term också, dock anser vi att flöde inte kan anses vara helt obetydlig. När det gäller modell C, noterar vi att den har högst AIC av alla. Därför faller vårt val på modell B.

Det återstår att undersöka om tillhörighet till stratum "övriga media" innebär högre risk för exponering. Ovanstående modellen kan vi inte tillämpa på grund av att stratomet varken är indelad m.a.p. flöde eller antal varuposter. Å andra sidan har vi inte heller möjlighet att statistiskt undersöka sambandet mellan exponering och tillhörighet till detta stratum på grund av stratas olika vikter. Därför får vi nöja oss med en deskriptiv undersökning av data. Se Tabell 39.

Den procentuella andelen uppgiftslämnare som behöver mer än 4 timmar för att rapportera är 25.8% för övriga media och 11,6% för övriga strata. Notera att data ur stratomet "flera tilläggsnummer" ingår i tabellen. Detta är ett tydligt tecken på att tillhörigheten till stratomet "övriga media" ökar risk för exponering rejält.

Tabell 39: Övriga rapporteringsmedia och tidsåtgång

| <i>Strata</i> | <i>Tidsåtgång</i> | | <i>Totalt</i> |
|--|---------------------|------------------|------------------|
| | <i>0 – 240 min.</i> | <i>241+ min.</i> | |
| Samtliga ej-unika, samt "flera tilläggsnummer" | 1028 | 135 | 1163 |
| Övriga media | 46 | 16 | 62 ¹⁶ |
| Totalt | 1074 | 151 | 1225 |

¹⁶ 2 uppgiftslämnare som tillhör stratomet Övriga media och som besvarade undersökningen valde att inte besvara fråga 3. Därför är det totala antalet 62 i stället för 64.

Vi avslutar detta kapitel med att sammanfatta den information som vi fick utifrån den framställda modellen (B).

- Antal varuposter har en tydlig effekt på risk för exponering, som också är statistiskt signifikant: Ju fler varuposter desto större risk att det tar lång tid för att rapportera
- Rapportering via webb medför lägre risk för exponering. Detta gäller dock enbart de företag som inte har många varuposter att rapportera. Om en uppgiftslämnare rapporterar många varuposter och gör det via webb, ökar dess risk för exponering.
- Tillhörighet till stratumet "övriga media" kan kopplas till långa rapporteringstider.

7 DISKUSSION OCH SLUTSATSER

7.1 SLUTSATSER

Enligt Regleringsbrevet för år 2007, ska företagens kostnader i samband med uppgiftsredovisningen minska med minst 25 % fram till år 2010. Definitionen "kostnader" tolkar vi här som finansiella kostnader, först och främst löner, dock även IT-stöd, förbrukning av kontorsmaterial osv. De finansiella kostnaderna har i sin tur en uppenbar koppling till tidsåtgång för Intrastatrapportering. När det gäller den genomsnittliga tidsåtgången, konstaterar vi att denna studie har visat en minskning av tidsåtgången med nästan 30 % i jämförelse med den studie som genomfördes år 2006. Om man enbart tar hänsyn till de finansiella kostnaderna samt antar att genomsnittligt antal personer som är involverade i rapporteringsprocessen inte har förändrats under dessa år, och att personalkostnaderna innebär huvuddelen av utgifterna i samband med rapporteringen till Intrastat, kan det påstås att kravet i regleringsbrevet för år 2007 är uppfyllt med avseende på Intrastatundersökningen.

De finansiella utgifterna i samband med Intrastatrapporteringen är dock inte den enda parametern som reflekterar den totala uppgiftslämnarbördan. När det gäller upplevelsen i samband med rapportering till Intrastat, som är en annan viktig parameter i vår studie, konstaterar vi att andel uppgiftslämnare som tycker att det är lätt att rapportera minskat från 81,4% till 70,8% under de senaste fem åren. Upplevelsen kan inte kopplas till kostnad direkt, dock är den en viktig komponent av den subjektiva delen av uppgiftslämnarbördan. Därför är denna parameter starkt korrelerad med tidsåtgången, vilket bevisas av vår studie. Någon direkt förklaring till varför upplevelsen har försämrats kan vi inte hitta. Det borde dock poängteras att det fortfarande är mer än 2/3 av uppgiftslämnare som anser att det är lätt att rapportera till Intrastat.

Upplevelsen i samband med rapportering till Intrastat är också korrelerad med antal personer som medverkar i Intrastatrapporteringen: fler medverkande personer kan kopplas till sämre upplevelse. En anmärkning som bör göras är att majoriteten av uppgiftslämnarna (72%) enbart har en anställd som medverkar i Intrastatrapporteringen. Mer än 90% av uppgiftslämnare har inte fler än 2 personer som medverkar i rapporteringsprocessen. Vi konstaterar även att det finns samband mellan antal medverkande och tidsåtgång – dessa är positivt korrelerade. Det är inte uppenbart om det är hög belastning i samband med rapporteringen som gör att fler personer anställs eller om det är antal medverkande som påverkar tidsåtgången och rapporteringsupplevelsen.

En annan variabel som är korrelerad med upplevelsen i samband med rapporteringen är säkerheten i valet av varukoder. Korrelationen mellan dessa variabler är positiv och statistiskt signifikant (5% signifikansnivå), alltså ju mer osäker man är i valet av varukoder desto svårare upplevs rapportering till Intrastat. Generellt har situationen blivit mycket bättre för denna fråga: denna studie har visat att 87,5 % av rapportörer är säkra i valet av varukoder, jämfört med 75,6 % år 2006. Alltså har den procentuella andelen uppgiftslämnare som inte är säkra i valet av varukoder minskad med ca 50 % under de senaste 5 åren.

När det gäller variabeln tidsåtgång, har denna studie påvisat att denna också är korrelerad med tillgänglighet av de uppgifterna som bör rapporteras. Att ha uppgifterna direkt tillgängliga i ett företags datorsystem innebär i regel mycket kortare rapporteringstider i jämförelse med om uppgifterna behöver bearbetas, tas fram på ett annat sätt eller uppskattas utifrån annan kunskap.

De flesta uppgiftslämnare behöver inte mer än några timmar per månad för att rapportera till Intrastat. Dock finns det ett antal företag som spenderar flera dagar på detta. En analys av vad som orsakar så långa rapporteringstider har visat att de främsta orsakerna är typ av rapporteringsmedium och antal rapporterade varuposter. Beroendet mellan antal varuposter och rapporteringstid är uppenbart: ju fler varuposter en uppgiftslämnare har desto längre tid behövs för att rapportera. När det gäller beroendet mellan rapporteringsmedium och tidsåtgång, har det visat sig att förekomst av uppgiftslämnare med mycket långa rapporteringstider är betydligt högre bland de uppgiftslämnare som använder övriga rapporteringsmedia än samtliga andra. De uppgiftslämnare som rapporterar via webblankett och som rapporterar få varuposter har lägst snittlig rapporteringstid av alla. Å andra sidan, tenderar de företag som rapporterar via webb och som har många varuposter (mer än 25) att ha mycket långa rapporteringstider. Vi drar därför slutsatsen att rapportering via webb passar bäst för små företag som rapporterar få varuposter. De företag som har många varuposter att rapportera bör helst använda ett annat rapporteringssätt.

När det gäller sätt att ta fram varukoder, konstaterar vi att det vanligaste sättet är egen kunskap eller bedömning (31,4 % av uppgiftslämnare tar fram varukoder på så sätt antingen alltid eller ofta). Sedan kommer framtagning av varukoder via fakturan (27,6%), SCB (20,2%) och Tullverket (20,2%). 19,9 % av uppgiftslämnare brukar ta fram varukoder via sina leverantörer/partnerföretag, och 7,2 % använder övriga framtagningssätt.

Förslag om användning av en förenklad varukod om fakturavärde är lågt är nu mindre populärt jämfört med år 2006. Denna studie har visat att 50,3 % av uppgiftslämnare är positiva till förslaget, jämfört med 57,8% år 2006, vilket ändå är ca hälften av samtliga uppgiftslämnare.

När det gäller införande av möjligheten att rapportera i Euro, konstaterar vi att 27,3 % av uppgiftslämnare är positiva till förslaget. Detta värde inkluderar inte IDEP användare, som redan har möjlighet att rapportera i Euro. År 2006 var enbart 16,7 % av uppgiftslämnarna positiva till förslaget.

Cirka en tredjedel (34%) av uppgiftslämnarna är intresserade av att ta del av statistiken som framställs med hjälp av Intrastatundersökningen. Detta intresse är högre bland uppgiftslämnare som rapporterar via övriga media (48,3 %). En förklaring till detta kan vara att de företag som rapporterar via övriga media är i regel de allra största företagen, som kan vara intresserade i att ta del av Intrastatstatistiken för att t.ex. genomföra undersökningar av marknadskonjunktur, etc.

Uppgiftslämnarna hade även möjlighet att skriva kommentarer och förslag angående rapportering till Intrastat. Dessa sammanfattas i avsnitt 4.6.

7.2 DISKUSSION

Huvudresultatet av denna studie är att vi har fått en mer eller mindre fullständig nutidsbild av uppgiftslämnarbördan (år 2011). Vi har fått en bild över hur uppgiftslämnarbördan har förändrats under de senaste åren, vad den består av, hur den påverkas av varierande faktorer såsom rapporteringsmedium eller flöde samt hur uppgiftslämnarna är inställda till olika förbättringsförslag. Vi konstaterar vidare att svarsfrekvensen blev 63%. Detta är mindre än svarsfrekvensen för undersökningen år 2006 (85%).

Det fanns dock en del begränsande moment i undersökningen som behöver belysas. Det moment som kan anses vara allra viktigast är omöjligheten att inkludera de uppgiftslämnare som rapporterar via ombud i studien. Antal sådana företag är stort (688 st.), och det är nästan säkert att sådana uppgiftslämnare behöver mycket mindre tid för att rapportera än de företag och organisationer som inte anlitar ombud. Dock har sådana företag en extrakostnad i och med att de behöver betala ut ersättningar till ombud för arbetet i samband med rapporteringen. Dessa ersättningar bör på ett eller annat sätt räknas med eftersom man annars kommer att göra ett felaktigt antagande att det är gratis att anlita ombud. Den största svårigheten här är att inte många företag kommer att vara villiga att dela med sig informationen om ersättningar till ombud. Dessutom vore det kanske inte korrekt att ställa en sådan fråga till uppgiftslämnarna i en frivillig statistisk undersökning. Ett alternativ som möjligtvis skulle fungera vid eventuell studie av uppgiftslämnarbördan i framtiden är att försöka göra en skattning av ombudens pris för rapportering av en varupost. I och med att antal rapporterade varuposter är information som är tillgänglig för SCB, skulle man kunna göra ungefärliga skattningar på uppgiftslämnarnas utgifter i samband med att anlita ombud, konvertera dessa till tidsåtgång och summera den imaginära tidsåtgången med den faktiska.

Ett annat kriterium som gjorde att många uppgiftslämnare exkluderades ur rampopulationen var kravet på att enbart erfarna uppgiftslämnare ska vara med. I vår studie bestämde vi oss att enbart inkludera uppgiftslämnare som började rapportera under oktober 2009 eller senare, vilket ledde till att mer än 1000 uppgiftslämnare uteslöts. Detta är nästan 13 % av samtliga rapporteringsskyldiga företag. Fördelning av svar på frågan om erfarenhet av rapportering till Intrastat visar att uppgiftslämnarna är ganska snabba när det gäller inlärning av rapporteringsprocedurerna. Därför konstaterar vi att det skulle kunna väljas ett senare gränsdatum i detta fall.

En del uppgiftslämnare hör till en grupp som behövde behandlas på ett särskilt sätt. Dessa är företag som använder flera tilläggsnummer för Intrastatrapportering. Det har nämnts tidigare att det inte går att tillämpa hierarkiskt urval på sådana uppgiftslämnare eftersom tilläggsnummer inom ett företag inte är oberoende. Det går inte heller att behandla tilläggsnummer som separata företag. Fördelen med den strategin för behandling av sådana uppgiftslämnare som vi har valt (att välja tilläggsnummer med maximalt antal varuposter) är att man inte kan underskatta tidsåtgången för sådana företag. Å andra sidan leder denna strategi till en uppenbar överskattning av den genomsnittliga rapporteringstiden, vilket reflekteras av undersökningsresultatet.

När det gäller undersökningsfrågorna, borde det noteras att svarsalternativ avseende frågan angående erfarenhet av rapportering skulle kunna formuleras annorlunda.

Nämligen anser vi att tidsintervall angivna i svarsalternativ 1 och 2 (jan 2011 eller senare, respektive sep – dec 2010) borde vara smalare, t.ex. apr 2011 eller senare (1), och jan – apr 2011 (2). Detta är på grund av att det har visat sig att personal på företag och organisationer lär sig att rapportera väldigt snabbt, vilket framgår av uppgiftslämnarnas svar på respektive fråga. Dock hade vi inte vetat det då frågorna formulerades.

Ett fenomen som vi inte hittar någon förklaring till är den betydliga skillnaden i tidsåtgång mellan stratum "övriga media" i denna studie och stratum "alternativa media" i studien år 2006. Snitttid för rapportering är 302 minuter för "övriga media" vs. 150 minuter för "alternativa media". Enligt beskrivning borde strata innehålla samma typ av företag, och det är väldigt osannolikt att den genomsnittliga rapporteringstiden verkligen har fördubblats under dessa 5 år.

Slutligen, har en uppgiftslämnare konstaterat att hon missförstod frågan om tidsåtgång i att hon trodde att tiden borde anges i timmar i stället för i minuter. Svaret för just denna uppgiftslämnare korrigerades omedelbart, dock finns det en risk att fler uppgiftslämnare har missförstått denna fråga.

APPENDIX

A. EXAKT FORMULERING AV FRÅGOR (SVENSKA)

| Frågenr | Frågeformulering | Svarsalternativ |
|---------|---|---|
| 1 | Hur upplever ni att det är att ta fram och rapportera uppgifterna till Intrastat? | a) Mycket lätt b) Ganska lätt c) Ganska svårt d) Mycket svårt e) Ingen åsikt |
| 2 | Hur många personer medverkar normalt i ert arbete med att ta fram och rapportera uppgifterna till Intrastat? | a) 1 person b) 2 personer c) 3 personer eller fler d) Kan inte uppskatta antal medverkande |
| 3 | Uppskatta hur mycket tid ert företag använder i genomsnitt per månad för att ta fram och rapportera uppgifter till Intrastat? | XX minuter |
| 4 | Sedan hur länge medverkar du personligen i att rapportera uppgifterna till Intrastat på företaget där du arbetar nu? | Jag började rapportera ... a) Under januari 2011 eller senare b) Under september – december 2010 c) Under augusti 2010 eller tidigare d) Kan ej ta ställning |
| 5 | Hur hittar ni de varukoder ni använder? | Se nästa tabell |
| 6 | Hur säkra känner ni er på att ni väljer rätt varukod? | a) Mycket säker b) Ganska säker c) Ganska osäker d) Mycket osäker e) Ingen åsikt |
| 7 | Hur tas uppgifterna till Intrastat fram? [Flera alternativ tillåts] | a) Direkt ur företagets datorsystem/affärssystem b) Direkt ur företagets datorsystem/affärssystem, men behövde bearbetas c) Hämtades från andra källor d) Uppskattades utifrån annan kunskap |
| 8 | Hur skulle ni ställa er till att SCB inför en förenklad varukod för låga fakturavärden (mindre än 2000 kronor)? | a) Positiv b) Varken positiv eller negativ c) Negativ |
| 9 | Hur skulle ni ställa er till att själva kunna välja om ni vill rapportera det fakturerade värdet i SEK eller i Euro? | a) Positiv b) Varken positiv eller negativ c) Negativ d) Möjligheten finns redan eftersom vi använder IDEP |
| 10 | Hur intresserade skulle ni på ert företag vara av att ta del av resultatet från Intrastatundersökningen? <i>D.v.s. aggregerad statistik för de svenska företagen som redovisar</i> | a) Mycket intresserade b) Ganska intresserade c) Inte särskilt intresserade d) Helt ointresserade |

| | | |
|----|--|------|
| | <i>Intrastat på olika varunivåer</i> | |
| 11 | Kommentarer och förslag på förbättringar | Text |

Fråga ang. sätt att hitta varukoder: detaljbeskrivning

| <i>Metod för framtagning</i> | <i>Svarsalternativ</i> | | | | | |
|--------------------------------------|------------------------|---------|-----------|-----------|-----------|-----------|
| På fakturan | a) Alltid | b) Ofta | c) Ibland | d) Sällan | e) Aldrig | f) Vet ej |
| Via SCB | a) Alltid | b) Ofta | c) Ibland | d) Sällan | e) Aldrig | f) Vet ej |
| Via Tullverket | a) Alltid | b) Ofta | c) Ibland | d) Sällan | e) Aldrig | f) Vet ej |
| Egen kunskap/bedömning | a) Alltid | b) Ofta | c) Ibland | d) Sällan | e) Aldrig | f) Vet ej |
| Från leverantören (partnerföretaget) | a) Alltid | b) Ofta | c) Ibland | d) Sällan | e) Aldrig | f) Vet ej |
| Övrigt | a) Alltid | b) Ofta | c) Ibland | d) Sällan | e) Aldrig | f) Vet ej |

B. EXAKT FORMULERING AV FRÅGOR (ENGELSKA)

| <i>Frågenr</i> | <i>Frågeformulering</i> | <i>Svarsalternativ</i> |
|----------------|---|---|
| 1 | Do you think it is easy or difficult to produce and report data to Intrastat? | a) Very easy b) Quite easy c) Quite difficult d) Very difficult e) No opinion |
| 2 | How many people are normally involved in your effort to produce and report information to Intrastat | a) 1 person b) 2 persons c) 3 persons or more d) Cannot estimate the number of participants |
| 3 | Estimate how much time your company uses on average per month to produce and transmit information to Intrastat | XX minutes |
| 4 | Since how long are you personally involved in the reporting of trade statistics to Intrastat in the company where you work now? | I started to report ... a) During January 2011 or later b) During September – December 2010 c) During august 2010 or earlier d) Cannot answer |
| 5 | How do you find the commodity codes you use? | Se nästa tabell |
| 6 | Are you certain that you indicate the correct commodity codes when you fill in the form? | a) Very certain b) Quite certain c) Quite uncertain d) Very uncertain e) Cannot answer |
| 7 | How is the data collected? (Flera alternativ tillåts) | a) Directly from the company's computer system/ERP system b) Directly from the company's computer system/ERP system, but needed to be processed c) From other sources |

| | | |
|----|--|---|
| | | d) Estimated by using other knowledge |
| 8 | Do you think it would be good or not good if SCB would introduce a simplified commodity code for the low invoice values (2000 SEK or less)? | a) Good b) Somewhat good c) Not good |
| 9 | Do you think it would be good or not good if you could choose a currency to report the invoice value in (SEK or Euro)? | a) Good b) Somewhat good c) Not good d) The possibility exists already while we use IDEP |
| 10 | How interesting would it be for your company to see the results from the Intrastat survey? <i>I.e. aggregated data for the Swedish companies that have submitted Intrastat on different production levels</i> | a) Very interesting b) Quite interesting c) Not very interesting d) Completely uninteresting |
| 11 | Comments and suggestions | Text |

Fråga ang. sätt att hitta varukoder: detaljbeskrivning

| <i>Metod för framtagning</i> | <i>Svarsalternativ</i> | | | | | |
|-------------------------------------|------------------------|----------|-----------------|--------------|-------------|---------------|
| On the invoice | a) Always | b) Often | c) Sometimes | d) Seldom | e) Never | f) Don't know |
| Via Statistics Sweden (SCB) | a) Always | b) Often | c) Sometimes | d) Seldom | e) Never | f) Don't know |
| Via Swedish Customs (Tullverket) | a) Always | b) Often | c) Sometimes | d) Seldom | e) Never | f) Don't know |
| Own knowledge/assessment | a) Always | b) Often | c) Sometimes | d) Seldom | e) Never | f) Don't know |
| From the supplier (partner company) | a) Always | b) Often | c) Sometimes | d) Seldom | e) Never | f) Don't know |
| Other | a) Always | b) Often | c) Sometimes | d) Seldom | e) Never | f) Don't know |

C. LISTA ÖVER TABELLER

| | |
|---|----|
| Tabell 1: Effekt av avgränsningar på rampopulationen..... | 9 |
| Tabell 2: Fördelning av uppgiftslämnare med avseende på flöde (ej flera tilläggsnummer, övriga media)..... | 9 |
| Tabell 3: Procentuellt antal varuposter som rapporteras via tilläggsnummer med störst antal varuposter; avser uppgiftslämnare som använde flera tilläggsnummer för att rapportera under dec 2010..... | 10 |
| Tabell 4: Fördelning av rampopulationen med avseende på antal varuposter och rapporteringsmedium (ej unika strata)..... | 12 |
| Tabell 5: Fördelning av urvalet..... | 14 |
| Tabell 6: Fördelning av urvalet med avseende på strata (ej flöde)..... | 14 |
| Tabell 7: Bortfall m.a.p. ursprung..... | 16 |
| Tabell 8: Variabelbortfall avseende frågor 3 och 5..... | 18 |
| Tabell 9: Tidsåtgång med avseende på strata..... | 20 |
| Tabell 10: Antal medverkande personer..... | 21 |
| Tabell 11: Erfarenhet av rapportering..... | 22 |
| Tabell 12: Sätt att hitta varukoder..... | 22 |
| Tabell 13: Sätt att hitta varukoder och tidsåtgång..... | 23 |
| Tabell 14: Tillgänglighet av uppgifterna..... | 24 |
| Tabell 15: Allmän upplevelse i samband med uppgiftsrapporteringen..... | 25 |
| Tabell 16: Allmän upplevelse vs. Tidsåtgång..... | 24 |
| Tabell 17: Säkerhet i valet av varukoder..... | 26 |
| Tabell 18: Möjlighet att rapportera i Euro, samtliga rapporteringsmedia..... | 26 |
| Tabell 19: Möjlighet att rapportera i Euro, svar från IDEP-användare exkluderade..... | 26 |
| Tabell 20: Införande av förenklad varukod för låga fakturavärden..... | 27 |
| Tabell 21: Användbarhet av statistiken..... | 27 |
| Tabell 22: Jämförelse av svarsskalor avseende denna undersökning och undersökning år 2006: Frågor 8 och 9..... | 29 |
| Tabell 23: Förändring av genomsnittliga rapporteringstiden m.a.p. flöde, unika strata ej inkluderade..... | 29 |
| Tabell 24: Förändring av den genomsnittliga rapporteringstiden: ordinära strata, unika strata, och totalt..... | 30 |
| Tabell 25: Jämförelse mellan undersökningarna: Allmän upplevelse..... | 30 |
| Tabell 26: Jämförelse mellan undersökningarna: Säkerhet i valet av varukoder..... | 30 |
| Tabell 27: Jämförelse mellan undersökningarna: Införande av förenklad varukod för låga fakturavärden..... | 31 |
| Tabell 28: Jämförelse mellan undersökningarna: Möjlighet att rapportera det fakturerade värdet i Euro..... | 31 |
| Tabell 29: Gruppering av variabeln Tidsåtgång..... | 33 |
| Tabell 30: Tolkning av γ | 34 |
| Tabell 31: Allmän upplevelse – Antal medverkande..... | 37 |
| Tabell 32: Tidsåtgång – Säkerhet i valet av varukoder..... | 37 |
| Tabell 33: Allmän upplevelse – Säkerhet i valet av varukoder..... | 38 |
| Tabell 34: Sammanfattning av resultat från korrelationstest..... | 38 |
| Tabell 35: Modell A – Analys av effekter..... | 43 |
| Tabell 36: Modell B – Analys av effekter..... | 44 |
| Tabell 37: Modell B - Parameterskattningar..... | 44 |

| | |
|---|----|
| Tabell 38: Modell C – Parameterskattningar | 45 |
| Tabell 39: Övriga rapporteringsmedia och tidsåtgång | 45 |

D. LISTA ÖVER FIGURER

| | |
|---|----|
| Figur 1: Bortfall per stratum, % | 17 |
| Figur 2: Bortfall per stratum, absoluta tal..... | 17 |
| Figur 3: Histogram för det ursprungliga och det logaritmerade data avseende variabeln | |
| Tidsåtgång..... | 32 |



Statistiska centralbyrån
Statistics Sweden

Datum
2011-04-26

Beteckning
2011/583

«Namn»
«Kontaktperson»
«gata»
«PostNr» «postadr»
«land»

Organisationsnummer: «OrgNr1»

Enkät om uppgiftslämnarbördan i Intrastat 2011

Bästa uppgiftslämnare!

Ert företag/Er organisation rapporterar för närvarande uppgifter om varuhandel inom EU (Intrastat) till Statistiska centralbyrån.

Ni har valt ut att ingå i en frivillig enkätundersökning om uppgiftslämnarbördan i Intrastat. Frågorna i enkäten belyser rapporteringen ur olika perspektiv samt hur man kan minska belastningen i samband med rapporteringen.

Observera att enbart ett handelsflöde («InfUtf» Sverige) ska beaktas vid ifyllning av enkäten.

Uppgifterna ska avse organisationsnummer:«OrgNr1», tilläggsnr. «tillnr».

Vi ber Er att besvara webbenkäten med följande inloggningsuppgifter:

Skicka in efterfrågade uppgifter
senast 6 maj 2011.
Logga in på
www.insamling.scb.se

| | |
|-------------|--------------|
| Användarid: | «AnvändarId» |
| Lösenord: | «Losenord» |

Uppgifterna som lämnas skyddas enligt 24 kap. 8 § offentlighets- och sekretesslagen (2009:400). Vid publicering kommer inga enskilda arbetsställen att kunna identifieras. Samråd har skett med Näringslivets Regelnämnd (NNR).

Frågor om enkäten:

Konstantin Kalinichenko, 08-506 941 67
Petter Samuelsson, 08-506 949 87
e-post: uppgiftslamnarbordan@scb.se

Tack på förhand för er medverkan!
Med vänliga hälsningar,

Annika Nyström
Enhetschef för Utrikeshandel och Industriindikatorer

Postadress
Box 24 300, 104 51 STOCKHOLM
701 89 ÖREBRO
Webbplats: www.scb.se
Organisationsnummer: 20 21 00 - 0837

Besöksadress
Karlavägen 100
Klostergatan 23
E-post: scb@scb.se
Momsregistreringsnummer: SE202100083701

Telefon
08 - 506 940 00
019 - 17 60 00

Telefax
08 - 661 52 61
019 - 17 70 80



Date
2011-04-26

Designation
2011/583

«Namn»
«Kontaktperson»
«gata»
«PostNr» «postadr»
«land»

Corporate id. no: «OrgNr1»

Questionnaire about the reporting burden in Intrastat 2011

Dear reporter!

Your company currently reports data on trade in goods within the EU to Statistics Sweden (SCB).

You have been selected to participate in a voluntary survey about the reporting burden in Intrastat. The questions in the survey highlight the reporting from different perspectives and how to reduce burden associated with the reporting.

Note that only one trade flow («InfUtf» Sweden) should be considered when completing the questionnaire.

The information to be submitted shall belong
to corporate identity number: «OrgNr1», supplementary no. «tillnr».

We ask you to answer the web questionnaire with the following information:

Your reply is requested
by 6 may 2011.
Log in to
www.insamling.scb.se

| | |
|------------------|--------------|
| Username: | «AnvandarId» |
| Password: | «Losenord» |

The information is held confidential according to the Secrecy Act.

Questions about the questionnaire:
Konstantin Kalinichenko, tel: +46 8 506 941 67,
Petter Samuelsson, tel: +46 8 506 949 87,
E-mail: uppgiftslamnarbordan@scb.se.

Thank you for your cooperation!
Sincerely,

Annika Nyström
Head of Foreign Trade and Industry Indicators Unit

Postal address
Box 24300, SE-10451 STOCKHOLM, SWEDEN
SE-70189 ÖREBRO, SWEDEN
Website: www.scb.se
Corporate registration no: 20 21 00 - 0837

Visiting address
Karlavägen 100
Klostergatan 23
E-mail: scb@scb.se
VAT no.: SE202100083701

Telephone
+46 8 506 940 00
+46 19 17 60 00

Telefax
+46 8 661 52 61
+46 19 17 70 80

LITTERATURFÖRTECKNING

1. **Finansdepartementet.** Förordning 2001:100 om den officiella statistiken.
2. **Bradburn, N.** Respondent Burden. *Health Survey Research Methods*. 1978, Vol. DHEW Publication No. 79-3207.
3. **Haraldsen, G.** Identifying and Reducing the Response Burden in Internet Business Surveys. *Journal of Official Statistics*. 2004, Vol. 20(2).
4. **Lorenc, B.** *Mätning av uppgiftslämnararbete i undersökningar med företag och organisationer*. Stockholm : SCB, 2007.
5. **Willeboordse, A.** The Mission of National Statistical Institutes. *Handbook on Design and Implementation of Business Surveys, Luxemburg, Office for Official Publications of the European Communities*. 1998.
6. **Hedlin, D.** *Developing Methods for Assessing Perceived Response Burden*. Luxembourg : Eurostat, 2005.
7. **Nutek.** *Manual för att mäta företagens administrativa kostnader enligt Standardkostnadsmodellen*. u.o. : Nutek, 2007. v.2.1.
8. **EC No 1982/2004.** Commission Regulation (EC) No 1982/2004 implementing Regulation (638/2004) on Community Statistics relating to the trading of goods between Member States and repealing Commission Regulations (EEC) No 3330/91 and (EC) No 1901/2000.
9. **EC No 638/2004.** Regulation (EC) No 638/2004 on Community Statistics relating to the trading of goods between Member States and repealing Council Regulation (EEC) No 3330/91. u.o. : European Parliament and European Council.
10. **Europeiska unionens officiella tidning, L 286.** 2007. ISSN 1725-2628.
11. **Sveriges Varuhandel med EU,Handledning.** u.o. : SCB, 2010.
12. **Sundberg, R.** *Urval ur ändliga populationer*. u.o. : Stockholms Universitet, 2010.
13. **Cochran, W.** *Sampling Techniques*. New York : Wiley, 1977.
14. **Weideskog, F.** *Intrastat system - Presumption for One Flow System*. u.o. : SCB, 2007.
15. **Agresti, A.** *Categorical Data Analysis*. New Jersey : John Wiley & Sons, 2002. ISBN 0-471-36093-7.
16. **Berry, K.** An APL function for the asymptotic test of significance for Goodman and Kruskal's gamma statistic. *Behavior Research Methods*. 4, 1989, Vol. 21.
17. **SAS Institute Inc., An, A.** *Performing Logistic Regression on Survey Data with the New SURVEYLOGISTIC procedure*. [<http://www2.sas.com/proceedings/sugi27/p258-27.pdf>] Cary, North Carolina, USA : SAS Institute Inc.,.
18. **Dobson, A.** *An introduction to generalized linear models*. u.o. : Chapman & Hall, 1990.

19. **Sundberg, R.** *Statistical Modelling by Exponential Families*. u.o. : Stockholms Universitet, 2010. ISSN 1650-0377.
20. **Geyer, C.** The Sandwich Estimator. [Online] 2003. [Citat: den 25 09 2011.] <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.142.8918&rep=rep1&type=pdf>.
21. **Morel, G.** Logistic regression under Complex Survey Designs. *Survey Methodology*. 15, 1989.