

KVAL

Validitet av VFU 2009 för validering av Samgods

2017-11-06



vti

SWECO



TRAFIKVERKET





Innehåll

1	INLEDNING	5
2	EXTERN VALIDITET FÖR VFU 2009.....	6
2.1	Jämförelser med VFU 2005	7
2.2	Jämförelser med utrikeshandelsstatistiken	8
2.3	Jämförelse med PWC 2009	9
2.4	Bedömning av extern validitet för VFU 2009	9
3	INTERN VALIDITET FÖR VFU 2009.....	12
3.1	Metod för validering och krav på VFU 2009	12
3.2	Dataprotein i VFU 2009	15
3.3	Bedömning av intern validitet för VFU 2009 vid tillämpning av bootstrap-metoden för validering	22
4	SLUTSATSER	24

1 Inledning

Innevarande dokument är en delrapport från projektet KVAL. Övergripande syfte med projektet KVAL är att förbättra den svenska nationella godstrafikmodellen Samgods för att tillhandahålla pålitliga prognoser av godsflöden på väg, järnväg och sjö. Detta görs genom att sammanställa data från olika undersökningar, validera modellen mot både slutresultat och resultat från de olika delmodellerna, identifiera brister och föreslå förbättringar som kan både innefatta kalibrering och vidareutveckling.

En del av KVAL-projektet består av att använda Trafikanalys varuflödesundersökningar (VFU) för att validera Samgodsmodellen. Valideringar av Samgodsmodellen har utförts tidigare men VFU har inte använts för detta syfte.

Syftet med innevarande PM är att beskriva lämpligheten i att använda varuflödesundersökningen från 2009, VFU 2009, (Trafikanalys, 2011). för att validera Samgodsprognoser framtagna för 2009. I detta fall innebär det att använda VFU 2009 för att validera, i första hand, så kallade PWC-matriser framtagna för 2009 (Berglund, M. och D. Pettersson, 2016), och om det bedöms lämpligt, även resultat avseende transportkedjor från en Samgodsprognos för 2009.

En validering med hjälp av VFU 2009 är bara lämpligt att utföra om undersökningen tillförlitligt mäter verkliga godsvolymer härledda från varusändningar och om data från undersökningen är konsistent med den metod som används för valideringen. Den första förutsättningen, att undersökningen tillförlitligt mäter verkliga godsvolymer benämns som *extern validitet* för undersökningen. Den andra förutsättningen, att data från undersökningen är konsistent med metoden för validering, benämns som *intern validitet*.¹

I kapitel 2 görs en bedömning av den externa validiteten för VFU 2009 och kapitel 3 tar upp den interna validiteten för undersökningen. Rapporten avslutas med ett kapitel om slutsatser.

Dokumentet har sammanställts av Christer Persson, WSP och granskats av Leonid Engelson, KTH. Utöver detta har Trafikanalys givits möjlighet att komma med kommentarer på dokumentet. Deras kommentarer på slutsatserna avseende extern validitet har lagts in i texten i kapitel 4 Slutsatser och kommentarerna om intern validitet har lagts in i fotnot 7 i samma kapitel.

¹ Begrepp och teorier om validitet finns inom ett antal ämnesområden, exempelvis sociologi. När begreppen extern och intern validitet används i detta dokument så avses dock enbart det som förklaras här.

2 Extern validitet för VFU 2009

För att bedöma den externa validiteten för VFU 2009 så jämförs godsvolymer från undersökningen med, i första hand, två andra statistikkällor: (1) VFU 2005² se SIKA (2006b), och (2) SCB:s utrikeshandelsstatistik (UHS) för åren 2005 och 2009 se SCB (2017a). Utöver detta så jämförs även volymer från VFU 2009 med godsvolymer från PWC-matriser framtagna för 2009 (PWC 2009) se Berglund och Pettersson (2016). PWC 2009 är inte statistikkälla i sig utan en prognos baserad på ett antal statistikkällor av godsflöden korsklassificerade efter produktion (P), förbrukning (C) och partihandel (W). Matriserna är framtagna på uppdrag av Trafikverket i syfte att användas som input till Samgods. En av de statistikkällor som används för att ta fram PWC 2009 var UHS.

VFU 2005 och UHS är officiell statistik och är därmed relevant att jämföra VFU 2009 med. PWC 2009 är inte officiell statistik, däremot är både metod för framtagande av matriser och själva framtagandet av PWC 2009 väl dokumenterade (se avs. 1. för ref.) och datakällan är tillgänglig. Därmed är den också relevant för jämförelse mot VFU 2009. Datakällorna är relevanta i den meningen att alla syftar till att mäta populationen av godsflöden på en nationell nivå, även om UHS definitionsmässigt är avgränsad till godssändningar in och ut ur landet.

Att datakällorna är relevanta innebär inte att de av nödvändighet är jämförbara. Även om syftet är att mäta samma population i olika undersökningar så behöver i praktiken inte urvalsramen stämma överens. Trafikanalys metodrapport (Trafikanalys, 2011) för VFU 2009 diskuterar möjligheterna för jämförelse mellan VFU 2009 och VFU 2005 (och även tidigare undersökningar) samt UHS. För VFU 2005 konstaterades följande (Trafikanalys, 2011, sid. 33):

Resultat, definitioner och variabelinnehåll är inte fullt jämförbara mellan varuflödesundersökningarna avseende år 2001, 2004/2005 och 2009. Stora betydelsefulla förändringar har skett i varukoderna och i branschindelningarna. Det handlar om omläggning till olika Standard för näringsgrensindelning och Samgods 12 indelningen efter då gällande varunomenklatur. Dessutom har variabler lagts till och tagits bort.

En aggregerad indelning av godsvolymer i tre kategorier: Export, Import och Inrikes (både avsändare och mottager inom landet), påverkas inte av förändringar i varukoder eller branschindelningar, ej heller av att variabler har lagts till och tagits bort ur undersökningarna. Därmed bör jämförbarheten vara tämligen god mellan VFU 2009 och VFU 2005 för godsvolymer enligt de tre aggregerade kategorierna³.

² Egentligen VFU 2004/2005 men vi behåller den korta benämningen VFU 2005.

³ En ytterligare aspekt som påverkar jämförbarheten mellan olika VFU är de så kallade cut-off gränserna för hur små (i antal anställda) arbetsställen som ingår i undersökningens population (intressepopulation). Cut-off gränserna, och hur de har beräknats, skiljer sig åt för de olika år som

Metodrapporten (sid. 33-34) tar upp jämförelser mellan VFU 2009 och UHS. Mellan dessa två datakällor bör överensstämmelsen vara bättre än för VFU mellan olika år. Till skillnad från jämförelsen mellan VFU 2009 och 2005 där ingen konkret jämförelse görs i metodrapporten så redovisar den en jämförelse mellan VFU 2009 och UHS (sid. 34). Den jämförelsen refereras till i avsnitt 2.2 nedan.

I alla tabeller i kapitlet så innebär nivån Totalt godsvolymer för både avgående och ankommande godssändningar i landet, det vill säga totalen för inrikes, export och import. Uppgifterna för VFU 2009 är hämtade från Trafikanalys (2010) och från SIKAs rapportering för VFU 2005 (SIKA, 2006a), uppgifterna för UHS är hämtade från SCB:s hemsida se SCB (2017a) för metoddokumentation och SCB (2017b) för data. Uppgifterna för PWC 2009 är material från Trafikverket, se Berglund och Pettersson (2016).

2.1 Jämförelser med VFU 2005

Tabell 1. Godsvolymer mätt i 1000-tals ton enligt VFU 2009 och 2005.

	VFU 2005	VFU 2009		Diff. Volym	Diff.(%) Volym
	(1000*ton)	(1000*ton)	Konf.int.(95%)		
Totalt	349 673	229 699		-119 973	-34%
<i>Därav*</i>					
Export	75 135	54 666	± 3 167	-20 469	-27%
Import	67 282	39 511	± 1 746	-27 771	-41%
Inrikes	206 675	134 896	±7 611	-71 779	-35%

* På grund av partiellt bortfall som berör inrikes/export så summerar inte delgrupperna Export, Import och Inrikes till Totalt

Tabell 1 ovan redovisar godsvolymer för de tre kategorierna Export, Import och Inrikes samt deras total. Volymerna för VFU 2009 i alla tre kategorierna ligger betydligt under, mellan 41 procent och 27 procent under, motsvarande volymer för VFU 2005. Totalt ligger godsvolymer för VFU 2009 34 procent under totalvolymen för VFU 2005. Tabell 1 ger också konfidensintervall på 95-procentsnivå till export-, import- och inrikesvolymerna för VFU 2009, dessa uppgifter kommer att användas i avsnitt 2.4 nedan.

VFU har genomförts. Inom projektet kommer vi att återvända till frågan om cut-off gränser men vår preliminära bedömning är att de inte kvalitativt påverkar slutsatser som förs fram i denna text.

2.2 Jämförelser med utrikeshandelsstatistiken

Tabell 2. Exportvolymen i miljontals ton för varuflödesundersökningarna (VFU) och utrikeshandelsstatistiken (UHS).

År	VFU (milj. ton)	UHS (milj. ton)	Diff. (%)
2005	75	84	-11%
2009	55	79	-30%
Diff. (%)	-27%	-6%	

Varuflödesundersökningen som här benämns VFU 2005 genomfördes egentligen sista halvåret 2004 och första halvåret 2005. För att öka jämförbarheten så består uppgifterna från UHS för år 2005 i tabell 2 och 3 av genomsnittet för år 2004 och 2005.

Tabell 2 ovan visar en jämförelse avseende exportvolymen mellan VFU och UHS för båda åren 2005 och 2009. För år 2005 ligger exportvolymerna i VFU 11 procent under UHS, för 2009 ökade skillnaden till 30 procent. Detta motsvaras av att utrikeshandelsstatistiken redovisar en minskning av exportvolymerna mellan 2005 och 2009 på 6 procent, medan varuflödesundersökningarna visar en minskning på 27 procent.

Tabell 3. Importvolymen i miljontals ton för varuflödesundersökningarna (VFU) och utrikeshandelsstatistiken (UHS).

År	VFU (milj. ton)	UHS (milj. ton)	Diff. (%)
2005	67	77	-13%
2009	40	70	-43%
Diff. (%)	-41%	-9%	

Tabell 3 nedan ger samma jämförelse som Tabell 2 men för importvolymen. Den tabellen följer också mönstret från Tabell 2. Nämligen, att VFU visar en betydligt större minskning av importvolymerna över de två åren än vad UHS visar och skillnaden består i att importvolymerna i VFU 2009 ligger mer under UHS 2009 än VFU 2005 ligger under UHS 2005.

Trafikanalys (2011, sid. 34) innehåller en jämförelse mellan VFU 2009 och UHS 2009. Den är gjord efter det monetära värdet för de totala export- respektive importsändningarna. För exporten redovisade metodrapporten ett värde på 718 miljarder kronor enligt VFU och 998,9 miljarder kronor enligt UHS, det vill säga VFU ligger 28 procent under UHS för år 2009. För importen redovisade metodrapporten ett värde på 425 miljarder kronor enligt VFU och 912 miljarder kronor enligt UHS, det vill säga VFU ligger 53 procent under UHS för år 2009. Den procentuella skillnaden för de två datakällorna mätt i exportsändningarnas värde är därmed ungefär lika stor som den procentuella skillnaden mätt i exportsändningarnas volym (se Tabell 2). För importsändningarnas värde ligger VFU 2009 än mer under UHS 2009 än vad som är fallet för importsändningarnas volym.

2.3 Jämförelse med PWC 2009

Tabell 4. Godsvolymer mätt i 1000-tals ton enligt PWC 2009 och VFU 2009.

	PWC 2009 (1000*ton)	VFU 2009 (1000*ton)	Diff.	Diff.(%)
Totalt	315 978	229 699	-86 279	-27%
<i>Därav</i>				
Export	68 395	54 666	-13 729	-20%
Import	65 445	39 510	-25 935	-40%
Inrikes	182 137	135 522	-46 615	-26%

Tabell 4 ovan visar godsvolymer enligt VFU 2009 och PWC 2009. Totalvolymen ligger i VFU 2009 27 procent under PWC 2009. Volymerna i VFU 2009 för de tre delgrupperna ligger under motsvarande volymer i PWC 2009 i ett intervall mellan 20 och 40 procent under.

Det är också illustrativt att jämföra PWC 2009 med VFU 2005, detta görs i Tabell 5 nedan. I detta sammanhang är det noterbart att export- och importvolymerna i PWC 2009 är framtagna genom att använda uppgifter från nationalräkenskaperna för att prognostisera UHS (med vissa justeringar och utgående från monetärt värde i.s.f. godsvolym) för 2010 bakåt ett år i tiden till 2009. Definitionsmässigt så kan därför export- och importvolymerna i PWC 2009 förväntas ligga tämligen nära volymerna enligt UHS 2009.

Tabell 5. Godsvolymer mätt i 1000-tals ton enligt VFU 2009 och PWC 2009.

	PWC 2009 (1000*ton)	VFU 2005 (1000*ton)	Diff.	Diff.(%)
Totalt	315 978	349 673	-33 695	-10%
<i>Därav</i>				
Export	68 395	75 135	-6 740	-9%
Import	65 445	67 282	-1 837	-3%
Inrikes	182 137	206 675	-24 538	-12%

En jämförelse mellan Tabell 1 och Tabell 5 visar att PWC 2009 ligger betydligt närmare VFU 2005 än VFU 2009 gör. För totalvolymerna ligger PWC 2009 10 procent under VFU 2005 medan VFU 2009 ligger 34 procent under. Vidare så ger jämförelser med Tabell 2 och Tabell 3 att minskningarna i export- och importvolymerna från VFU 2005 till PWC 2009 ligger mer i nivå med UHS 2005 och 2009 än vad minskningarna från VFU 2005 till VFU 2009 gör.

2.4 Bedömning av extern validitet för VFU 2009

Enligt Tabell 1 från avsnitt 2.2 så ligger godsvolymer 2009 från de totala godssändningarna enligt VFU 34 procent under de totala volymerna för 2005. Detta är en betydande nedgång i transporterat gods mellan de två åren. Som jämförelse ligger BNP 2009

mätt i fasta priser 2 procent *över* BNP för 2005. BNP-förändringar är ett mått som inte ger någon säker indikationen över hur godsvolymer bör förändras men måttet har använts i ett stort antal studier för att förklara förändringar i godsflöden på nationell och regional nivå. Därmed är nedgången i totala volymer mellan de två åren enligt VFU så stor att tveksamhet väcks om den avspeglar en verklig nedgång.

Jämförelse mellan VFU och UHS för de två åren 2005 och 2009 visar att VFU redovisar en betydligt större nedgång för både exporten och importen, nedgång med 27 respektive 41 procent, än vad UHS gör, som ger nedgång med 7 respektive 9 procent för exporten och importen. Samtidigt består större delen av avvikelserna mellan VFU och UHS i att export- och importvolymerna för VFU 2009 ligger betydligt mer under motsvarande volymer enligt UHS 2009, än vad volymerna i VFU 2005 ligger under UHS 2005. Jämfört med UHS så är det alltså framför allt VFU 2009 som avviker med låga export- och importvolymer. Det stärker en hypotes att det är tveksamt om godsvolymer i VFU 2009 avspeglar verkliga godsvolymer tillräckligt väl.

Faktum är om VFU 2009 läggs åt sidan så ger föregående avsnitt 2.3 att, vid en jämförelse med VFU 2005 och utrikeshandsstatistiken, ligger godsvolymererna för PWC 2009 på en till synes rimlig nivå. Men de avvikelserna som enligt Tabell 4 finns mellan PWC 2009 och VFU 2009 ger också följdslutsatsen att: godsvolymererna i VFU 2009 inte ligger på en godtagbart rimlig nivå. Den slutsatsen stärks av att godsvolymererna för de tre delgrupperna (export, import och inrikes) av PWC 2009 alla enligt Tabell 4 och Tabell 1 ligger utanför konfidensintervallen på 95-procentsnivå kring motsvarande delgrupper i VFU 2009. De tre delgrupperna av godsvolymer i PWC 2009 är alltså alla statistiskt signifikant skilda från motsvarande godsvolymer i VFU 2009. Konfidensintervallen för VFU 2009 är också snäva kring godsvolymererna i VFU 2009.

På grund av att vi inte vet de verkliga godsvolymererna så kan statistiska signifikanser endast i särskilda fall avgöra frågan om extern validitet. I frågan som studeras här skulle ett sådant fall vara att vi var säkra på att PWC 2009 underskattar de verkliga godsvolymererna. Eftersom vi vet att VFU 2009 än mer underskattar de verkliga volymerna, så ger faktumet att volymerna i VFU 2009 är signifikant skilda från PWC 2009 att vi i detta fall med automatik bör förkasta den externa validiteten för VFU 2009 (och detta oavsett om vi anser att PWC 2009 uppfyller extern validitet eller inte). Men om det i stället är så att PWC 2009 överskattar de verkliga godsvolymererna så kan det mycket väl vara så att godsvolymererna i VFU 2009 inte är signifikant skilda från de verkliga godsvolymererna, och i så fall i stället dra slutsatsen att VFU 2009 har godtagbar extern validitet. Nu vet vi inte om PWC 2009 underskattar eller överskattar de verkliga godsvolymererna. För att kunna fälla ett avgörande om den externa validiteten för VFU 2009 måste vi då godta en subjektiv bedömning. För vår del är den bedömningen att med stöd från datakällorna VFU 2005 och UHS för 2005 och 2009 så ligger godsvolymererna för PWC 2009 närmare de verkliga godsvolymererna än godsvolymererna gör i VFU 2009.

Uppgifterna om att godsvolymererna i PWC 2009 är signifikant skilda från motsvarande volymer i VFU 2009 *och* att konfidensintervallen är snäva kring godsvolymererna för VFU 2009 understödjer den subjektiva bedömningen att PWC ligger närmare de verkliga

godsvolymerna än VFU 2009. För att se detta, notera att med ledning av Tabell 1 och 4 går det att bestämma att den övre gränsen för konfidensintervallet kring exportvolymerna i PWC ligger på en punkt som är 23 procent av sträckan från exportvolymen för VFU 2009 till exportvolymen för PWC 2009. För import- och inrikesvolymerna ligger de motsvarande gränserna för konfidensintervallen på 7 respektive 16 procent. För att VFU 2009 ska betraktas som externt valid bör de verkliga godsvolymerna inte vara signifikant skilda från volymerna i VFU 2009. Det vill säga, de verkliga godsvolymerna måste ligga närmare volymerna i VFU 2009 än de ovan bestämda procentuella gränserna. Om så vore fallet är vår bedömning att detta skulle ge utslag i volymerna för PWC 2009 så att de också skulle ligga betydligt närmare volymerna i VFU 2009. Nu är detta dock inte fallet, därmed drar vi slutsatsen att godsvolymerna avseende export-, import- och inrikessändningar för VFU 2009 inte uppfyller en godtagbar extern validitet i syfte att validera Samgodsprognoser för 2009.

3 Intern validitet för VFU 2009

Utgångspunkten för intern validitet i detta sammanhang är att datakällan (VFU 2009) är konsistent med metoden som används för att validera en Samgodsprognos för 2009. Metoden som valts ställer därmed ett antal krav på datakällan som är nödvändiga för att kunna tillämpa valideringsmetoden på ett konsistent sätt. Avsnitt 3.1 redogör för metoden och vilka krav den ställer på datakällan, avsnitt 3.2 går igenom ett antal dataproblem som har identifierats för VFU 2009 som har betydelse för kraven som ställs på datakällan, slutligen så görs i avsnitt 3.3 en bedömning av den interna validiteten för VFU 2009 relativt valideringsmetoden. Det bör understrykas att problem som diskuteras i detta kapitel gäller för den valda valideringsmetoden. I princip sägs inget om hur problemen påverkar resultat från VFU 2009 framtagna med andra metoder.

3.1 Metod för validering och krav på VFU 2009

3.1.1 Metod för validering

Utgångspunkten för valideringen med avseende på ett givet valideringsmått som går att skapa från VFU och Samgodsprognosen är att Samgods anses vara valid om måttet från VFU inte är statistiskt signifikant skild från motsvarande mått från Samgodsprognosen. Därmed blir det avgörande att kunna bestämma signifikanser för mått skapade från VFU. För rena totaler eller medelvärden över redovisningsgrupper av sändningar i VFU finns det slutna formler som går att använda för att beräkna signifikans. Men en del av valideringen består snarare av regressionsbaserade analyser för vilka det är svårt att finna enkla formler för att beräkna trovärdiga signifikanser för en undersökning med urvalsutformning enligt VFU. Därför bör så kallad bootstrap-metodik (ibland kallad återsampling på svenska) att användas.

Givet att det går att replikera urvalsutformningen med hjälp av bootstrap går det i princip att beräkna signifikanser för alla typer av mått som går att skapa ur undersökningen (VFU i detta fall), även regressionskoefficienter och till exempel lack-of-fit-statistikor. Utöver detta förbättrar bootstrap-metoden även skattningen av signifikanser för totaler och medelvärden jämfört med de traditionellt använda slutna formlerna för det syftet. Detta på grund av att formlerna för totaler och medelvärden beror på så kallade uppräkningsvikter. Om det förekommer bortfall i undersökningen (t.ex. som i VFU) så är även uppräkningsvikterna slumpmässiga storheter som påverkar signifikansen för totaler och medelvärden, vilket inte tas hänsyn till i de traditionella slutna formler som används för signifikansberäkning till totaler och medelvärden för redovisningsgrupper. Men med bootstrap metoden går det att ta hänsyn till denna källa till slumpmässighet.

Rent konkret simulerar bootstrap-metoden den empiriska fördelningsfunktionen för ett givet mått genom att från det givna urvalet dra obundna slumpmässiga urval med samma urvalsstorlek som det givna urvalet men med återläggningen av urvalsenheterna. Bootstrap-metoden återskapar alltså variationen i population genom att en given urvalsenhet slumpmässigt förekommer, antingen inte alls, eller ett upprepat antal gånger i ett av de

simulerade bootstrap-urvalen. I varje simulerat bootstrap-urval beräknas sedan det givna måttet. Måttets spridning i de simulerade urvalen ger den empiriska fördelningsfunktionen för måttet, som därefter kan användas för att beräkna signifikanser. I en stratifierad undersökning, som VFU, dras separata bootstrap-urval i varje stratum. I en flerstegsundersökning, som VFU, dras successiva bootstrap-urval i varje steg.⁴

3.1.2 Viktning av observationer i VFU 2009

Uppräkningsvikterna för att räkna upp observationer i urvalet till populationsnivå VFU 2009 beräknas som det sedvanliga ”ett genom urvalssannolikheten”, se Trafikanalys (VFU 2011, sid. 71). Urvalet i VFU 2009 sker i tre steg (1) stratifiering av arbetsställen, (2) urval av mätperiod (veckor) från ett givet kvartal, och (3) urval sändningar från arbetsställets totala antal sändningar under perioden. Urvalssannolikheten är sannolikheten att en observation i populationen, i detta fall en sändning från/till ett arbetsställe, kommer att ingå i urvalet. Urvalet av mätveckor i steg 2 slumpades ut likformigt och oberoende fördelat över det givna kvartalet. I steg 3 ombads arbetsstället att dra ett urval av sändningar under mätperioden där urvalssannolikheten berodde på antalet totala sändningar under mätperioden för arbetsstället. Det går därmed att betrakta urvalen i de tre stegen som oberoende, och urvalssannolikheten p_{urval} kan därmed beräknas som en produkt av sannolikheterna för var och en av urvalsstegen.

$$p_{urval} = p_{stratifiering} p_{mätperiod} p_{sändningar}$$

Information finns i urvalsdata för att beräkna sannolikheterna för de tre urvalsstegen. Vikten som används för att räkna upp enskilda observationer i data till populationsnivå blir därmed

$$(1) \quad vikt = \frac{1}{p_{urval}} = \frac{1}{p_{stratifiering}} \cdot \frac{1}{p_{mätperiod}} \cdot \frac{1}{p_{sändningar}}$$

I undersökningsdata finns ett antal hjälpvariabler som används för att beräkna vikterna, dessa visas i Tabell 6 nedan.

Tabell 6. Hjälpvariabler för beräkning av uppräkningsvikter.

Variabel	Beskrivning
StoraN	Antal arbetsställen i observationens stratum.
LillaN	Antal arbetsställen i bruttourvalet i observationens stratum
NSVAR	Antal svarande arbetsställen (nettourvalet) i observationens stratum.
Matveckor	Antalet mätveckor för arbetsstället.

⁴ När bruttourvalet i ett stratum är en totalundersökning, alltså det är identiskt med population (eller snarare den så kallade urvalsramen), så måste bootstrap-metoden modifieras så att den tar hänsyn till att bortfallet utgör den enda källan till slumpmässighet.

AvgAnt	Totalt antal avgående sändningar under mätperioden för arbetsstället som observationen tillhör.
AVGPOSTER	Antal poster (sändningar) som arbetsstället, som observationen tillhör, har lämnat.

Variabelnamnen är visade för datamängden med avgående sändningar; för ankommande sändningar byt Avg/AVG till Ank/ANK i namnen för de två sista variablerna.

Genom att använda dessa variabler i ekvation (1) ovan så beräknas vikterna som

$$(2) \quad vikt = \frac{StoraN}{NSVAR} \cdot \frac{13}{Matveckor} \cdot \frac{AvgAnt}{AVGPOSTER}$$

Den första faktorn i viktberäkningen är *stratifieringsvikten*. Den är helt enkelt, för varje observation i ett stratum, antalet arbetsställen som ingår i populationen för stratomet delat med antalet svarande arbetsställen i som ingår i stratums urval (nettourvalet). Talet 13 i den andra faktorn står för antalet veckor i ett kvartal. Ekvation (2) stämmer överens med motsvarande ekvationer i Trafikanalys (2011, sid 71). Vi har tillämpat ekvation (2) på datamängden för VFU 2009 och kommit fram till resultat som stämmer överens med resultat publicerade i Trafikanalys (2010).

3.1.3 Krav på datakällan vid användning av bootstrap-metoden

För att genomföra en bootstrap-baserad analys av VFU 2009, som tar hänsyn till bortfall och därmed att uppräkningsvikter i sig är slumpmässiga storheter behöver innehållet i hjälpvariabeln *NSVAR* förtydligas. Detta är inte tydligt beskrivet i metodrapporten för undersökningen, men kontakter med Trafikanalys har gett att *NSVAR* innefattar följande:

1. arbetsställen som har besvarat undersökningen och har utfört minst en sändning under mätperioden,
2. arbetsställen som har besvarat undersökningen men inte utfört någon sändning under mätperioden, och
3. arbetsställen som har besvarat undersökningen men tillhör den så kallade överteckning

Endast den första gruppen av arbetsställen finns med som observationer i datamängden för undersökningen. De två sista grupperna representeras enbart av hjälpvariabeln *NSVAR*. Överteckning är arbetsställen som inte tillhör populationen men som var med i den så kallade urvalsramen (populationsregistret) och därmed också kom med i urvalet.

För att genomföra en bootstrap-baserad analys där uppräkningsvikterna räknas om för att ta hänsyn till bortfall behövs åtminstone för varje stratum tre kategorier av arbetsställen i urvalet kunna identifieras, nämligen:

- A. arbetsställen i urvalet som har besvarat undersökningen och ingår i datamängden,
- B. arbetsställen i urvalet som har besvarat undersökningen men inte ingår i datamängden, samt
- C. arbetsställen i urvalet som är bortfall.

För arbetsställen av kategori B och C, som inte ingår i datamängden räcker det att känna till antalet arbetsställen i respektive kategori. Med ledning av punkt 1-3 ovan går det att konstatera att hjälpvariabeln $NSVAR$ ger antalet arbetsställen av kategori A och B. Om antalet arbetsställen i kategori A betecknas med n_{resp} , så ges antalet arbetsställen i kategori B av skillnaden $NSVAR - n_{resp}$ och skillnaden $LillaN - NSVAR$ ger antalet arbetsställen i kategori C.

Variabeln n_{resp} blir alltså viktig för en bootstrap-baserad analys. Konkret så beräknas den med hjälp av följande två steg.

1. identifiera stratumet i urvalet, och
2. för varje stratum räkna antalet arbetsställen i urvalets datamängd som ingår i stratumet.

Totalt sett så innebär detta att, för att på ett konsistent och tillförlitligt sätt kunna tillämpa bootstrap-metoden är det viktigt att variablerna $LillaN$, $NSVAR$ och n_{resp} är väldefinierade samt korrekta i data från VFU 2009. Fokus i den följande texten kommer dock att ligga på variabeln $LillaN$. Utöver den variabeln kommer också $StoraN$ (se Tabell 6) att studeras. Den variabeln har inte betydelse i en bedömning av den interna validiteten. Däremot används $StoraN$ i beräkningen av uppräkningsvikter (se ekv. 2), och kommer att påverka resultatet av de flesta (om inte alla) bootstrap-baserade analyser.

3.2 Dataproblem i VFU 2009

3.2.1 Inkonsistenser mellan datamaterialet och metodrapporten

Ett antal inkonsistenser har upptäckts mellan metodrapporten (Trafikanalys, 2011, tabell 9.5 sid. 65) och variablerna $StoraN$ och $LillaN$ i datamängden för VFU 2009. Kontakter med Trafikanalys har gett att detta troligen beror på att en äldre version av datamängden har redovisats i metodrapporten, samt att det förekommer redigeringsfel⁵. Inkonsistenser påverkar inte direkt den interna validiteten, däremot försvårar de en värdering av

⁵ Enligt Trafikanalys så ”kom Statisticon (Trafikanalys producent av VFU 2016) fram till samma resultat som de officiella när de använde VFU 2009 för utvärdering och för allokering i samband med utveckling och produktion av nya VFU 2016”.

aspekter som har betydelse för den interna validiteten. Därför redovisas inkonsistenser i texten nedan.

Stratum identifieras av en kod i variabeln `Stratum` som delar in arbetsställena i 73 kategorier beroende på arbetsställets bransch, storlek och regional tillhörighet. Utöver detta så dras ett nytt urval för var och ett av de fyra kvartal som ingår i undersökningen, därför består stratifiering av alla kombinationer av variablerna `Stratum` och `Kvartal`. Detta ger $73 \cdot 4 = 292$ strata som arbetsställena är indelade i. I Metodrapporten redovisas antal arbetsställen i populationen och i bruttourvalet. Denna tabell definierar variablerna `StoraN` och `LillaN` i datamängden. Motsvarande information konstruerad från data ges av Tabell 77 och 8 nedan, tabellerna innehåller också uppgifterna från metodrapporten.

Tabell 7. Antal arbetsställen i populationen (eg. urvalsramen) hämtad från uppgifter i urvalets datamängd, variabeln `StoraN` och motsvarande uppgifter från Trafikanalys metodrapport (tabell 9.5).

Stratum	Kvartal 1		Kvartal 2		Kvartal 3		Kvartal 4	
	Data	Metodr.	Data	Metodr.	Data	Metodr.	Data	Metodr.
3199		178		172		166		165
3299		228		230		231		228
4199	9	101	9	99	9	98	9	97
4299		339		334		350		350
5199	101		99		98		97	
5299	339		334		350		348	
6199	176	8	172	8	166	7	164	7
6299	228	21	230	21	231	23	228	23
7199	48	9	45	9	43	9	44	9
7299	33	0	32	0	30	1	28	0
8199	8	36	8	37	7	37	7	36
8299	21	61	21	60	23	56	23	56
9199	8	48	8	45	8	44	8	44
9299	21	33	21	32	20	30	19	29
10199	54	54	52	52	50	50	49	49
10299	249	249	256	256	260	261	263	263
11199	36	8	37	8	37	8	36	8
11299	61	21	60	21	56	20	56	19
12199	278	279	274	274	274	274	274	274
12299	325	327	324	324	307	307	306	307
37301	510	512	506	506	512	513	509	510
37302	839	840	830	831	817	820	810	811
37303	944	944	932	932	942	942	930	932
37304	713	714	697	699	715	717	708	709
37305	1103	1103	1088	1088	1112	1113	1106	1106
37306	485	485	482	482	492	492	487	487
37307	176	176	179	179	179	180	179	179
37308	262	262	260	260	264	264	263	263
37401	536	536	531	531	516	516	518	519
37402	573	573	572	572	586	586	584	584
37403	508	508	502	504	484	485	487	488
37404	550	552	545	545	556	556	557	558
37405	829	829	832	832	777	777	779	779
37406	359	359	360	360	345	345	344	344
37407	189	189	187	187	183	183	183	183
37408	222	222	218	218	223	223	220	220
46301	1574	1576	1559	1562	1625	1629	1613	1619
46302	700	700	702	703	713	714	711	713
46303	461	461	469	471	473	476	473	473
46304	898	898	896	896	908	909	897	900
46305	1401	1403	1393	1396	1404	1410	1397	1402

46306	370	370	370	371	384	384	378	380
46307	163	163	160	160	156	157	158	158
46308	201	201	196	198	205	205	200	201
46401	1708	1711	1706	1708	1638	1638	1628	1628
46402	780	780	792	792	767	767	770	771
46403	515	515	525	525	531	531	534	534
46404	1041	1041	1045	1045	1021	1022	1026	1026
46405	1436	1437	1448	1449	1401	1401	1402	1403
46406	406	406	410	410	413	413	414	414
46407	194	194	195	195	200	200	193	193
46408	201	201	206	206	196	196	200	200
451199	22	22	22	22	23	23	23	23
453199	10	10	10	10	10	10	10	10
453299	107	107	109	109	116	116	116	117
454199	10	10	10	10	9	9	9	9
454299		26	24	24		24	24	24
461199	9	9	10	10	8	8	8	8
461299	159	159	148	148	154	154	154	154
462199	46	46	45	46	50	50	50	50
462299	53	53	47	47	44	44	45	45
463199	18	18	16	16	16	16	15	15
463299	63	63	62	62	64	64	64	66
464199	20	20	19	19	15	15	15	15
464299	131	132	137	138	143	143	144	145
465199	8	8	8	8	8	8	8	8
465299	58	58	58	58	55	55	54	55
466199	7	7	7	7	7	7	7	7
466299	73	73	68	69	68	68	66	67
467199	25	25	25	25	25	25	25	25
467299	73	73	89	89	90	90	88	88
469199	11	11	12	12	11	11	13	13
469299	29	29	30	30	31	31	29	29
479199	6	6	6	6	6	6	6	6
479299	16	17	16	16	16	16	16	16
990001	1		1		1		1	
990002	7		7		7		7	
990003	9		9		9		9	
990004	19		22		28		34	
990005	2		1		1		6	

Tabell 8. Antal arbetsställen i bruttourvalet hämtad från uppgifter i urvalets datamängd, variabeln LillaN och motsvarande uppgifter från Trafikanalys metodrapport (tabell 9.5).

Stratum	Kvartal 1		Kvartal 2		Kvartal 3		Kvartal 4	
	Data	Metodr.	Data	Metodr.	Data	Metodr.	Data	Metodr.
3199		178		172		163		162
3299		57		58		58		75
4199	9	101	9	98	9	97	9	96
4299		85		84		88		99
5199	101		98		97		96	
5299	85		84		88		97	
6199	176	8	172	8	163	7	161	7
6299	57	5	58	5	58	6	75	6
7199	48	9	45	9	43	9	44	9
7299	8	0	8	0	8	1	5	0
8199	8	36	8	37	7	37	7	36
8299	5	15	5	15	6	14	6	17
9199	8	48	8	45	8	44	8	44
9299	5	8	5	8	5	8	4	6
10199	54	54	51	51	50	50	49	49
10299	62	62	64	64	64	65	72	72
11199	36	8	37	8	37	8	36	8
11299	15	5	15	5	14	5	17	4
12199	278	279	269	266	268	268	268	268
12299	80	82	81	81	77	77	90	91
37301	41	43	36	36	33	34	47	48
37302	55	56	87	88	73	76	68	69
37303	51	51	81	81	79	79	72	74
37304	49	50	76	78	74	76	70	71
37305	64	64	102	102	102	103	96	96
37306	22	22	35	35	34	34	32	32
37307	10	10	17	17	16	17	16	16
37308	13	13	19	19	23	23	21	21
37401	10	10	11	11	21	21	19	20
37402	10	10	10	10	10	10	10	10
37403	10	10	8	10	9	10	9	10
37404	93	95	10	10	10	10	9	10
37405	10	10	12	12	12	12	11	11
37406	10	10	10	10	10	10	10	10
37407	10	10	10	10	10	10	10	10
37408	10	10	10	10	10	10	10	10
46301	227	229	238	241	244	248	234	240
46302	45	45	70	71	79	80	72	74
46303	50	50	62	64	75	78	81	81
46304	129	129	118	118	110	111	100	103
46305	206	208	181	184	178	184	177	182
46306	29	29	45	46	44	44	40	42
46307	10	10	11	11	10	11	10	10
46308	50	50	27	29	31	31	27	28
46401	219	222	121	123	89	89	84	84
46402	22	22	35	35	31	31	28	29
46403	12	12	19	19	15	15	14	14
46404	23	23	36	36	54	55	52	52
46405	34	35	56	57	89	89	81	82
46406	10	10	15	15	15	15	14	14
46407	14	14	10	10	14	14	13	13
46408	10	10	10	10	10	10	10	10
451199	22	22	22	22	20	23	18	23
453199	10	10	10	10	10	10	9	9
453299	27	27	27	27	29	29	34	35
454199	10	10	10	10	7	7	7	7
454299		7	6	6		6	8	8
461199	9	9	7	7	4	4	4	4
461299	40	40	37	37	39	39	50	50
462199	46	46	45	46	44	44	43	43
462299	13	13	12	12	11	11	14	14

463199	18	18	16	16	16	16	14	14
463299	16	16	16	16	16	16	17	19
464199	20	20	19	19	10	10	10	10
464299	32	33	34	35	36	36	42	43
465199	8	8	7	7	3	3	3	3
465299	15	15	15	15	14	14	12	13
466199	7	7	6	6	6	6	5	5
466299	18	18	16	17	17	17	19	20
467199	25	25	25	25	24	24	21	21
467299	18	18	22	22	23	23	26	26
469199	11	11	11	11	9	9	7	7
469299	7	7	8	8	8	8	9	9
479199	6	6	6	6	6	6	6	6
479299	3	4	4	4	4	4	5	5
990001	1		1		1		1	
990002	7		7		7		7	
990003	9		9		9		9	
990004	19		22		28		34	
990005	2		1		1		6	

Kategorierna i variabeln *Stratum* i datamängden stämmer inte helt med uppgifterna i Metodrapporten. Strata nr 990001 - 990005 som innehåller uppgifter från den separata registerbaserade totalundersökningen i aggregerad form, finns inte redovisade i metodrapporten. Detta är inget egentligt problem, men utöver detta finns följande avvikelser i koderna för variabeln *Stratum* i datamängden och i tabell 9.5 i Metodrapporten (inledande nollor är uteslutna i koderna):

- (1) Stratumkoderna 3199, 3299 och 4199 som finns i metodrapporten finns inte i variabeln *Stratum* i datamängden.
- (2) Stratumkoderna 5199 och 5299 som finns i variabeln *Stratum* finns inte i metodrapporten

Utöver att stratumkoderna inte stämmer överens mellan variabeln *Stratum* och Metodrapporten så stämmer heller inte antalsuppgifterna överens mellan Tabell 7 och 8 ovan och tabell 9.5 i Metodrapporten. Speciellt gäller detta för stratumkoderna 4199 - 11299 som avser totalundersökta arbetsställen inom mineral- och tillverkningsindustrin. För dessa skiljer sig antalsuppgifterna i många fall betydligt mellan Tabell 7 och 8 ovan och tabell 9.5 i Metodrapporten. Dessa skillnader ser ut att bero på omkastningar mellan raderna i tabellerna. Som ett exempel på detta ger tabell 4 nedan dessa skillnader för de första sex första stratumkoderna i Tabell 7 ovan.

Tabell 9. Identifiering av rader i Tabell 7 ovan och Tabell 9.5 i Metodrapporten.

Stratumkod, enligt tabell 9.5		
Stratum	metodrapp.	Anmärkning
4199	7199	Dessa rader överensstämmer exakt.
5199	4199	Dessa rader överensstämmer exakt.
5299	4299	Utom kvartal 4 där tabell 9.5 i metodrapp. anger 2 extra arbetsställen
6199	3199	Utom kvartal 1 och 4 där tabell 9.5 i metodrapp. anger 2 respektive 1 extra arbetsställe
6299	3299	Dessa rader överensstämmer exakt.
7199	9199	Utom kvartal 3 där tabell 9.5 i metodrapp. anger 1 extra arbetsställe

Återigen kan orsaken till avvikelserna mellan stratumkodningen i variabeln `Stratum` i datamängden och Tabell 9.5 i Metodrapporten bero på att tabellen är felaktig. Men skillnaderna kan också vara en indikation på felaktigheter i hjälpvariablerna för uppräkningsvikterna.

Totalt sett finns det i alla $70 \cdot 4 = 280$ strata, där det går att jämföra Tabell 7 och 8 med Tabell 9.5 i Metodrapporten, avvikelser i antal arbetsställen i populationen (`StoraN`) i totalt 105 strata.

För strata 454299 gäller att det i datamängden saknas värden för alla hjälpvariabler (tabell 7) för kvartal 1 och 3. Trolig orsak till det är svarsbortfall, då dessa strata har liten storlek.

3.2.2 Problem i data

Med utgångspunkt i att stratumkodningen i variabeln är korrekt så finns det 77 stratumkategorier i variabeln `Stratum` totalt ger det $77 \cdot 4 - 2 = 306$ strata som förekommer i datamängden. För denna stratifiering gäller att variablerna `StoraN` och `LillaN` har två värden istället för ett entydigt värde i 69 stycken stratum. Förekomsten av två olika värden för dessa variabler beror på att de har olika värde i datamängderna för avgående (inrikes och export) och ankommande (import) sändningar. Det har inte varit möjligt att få någon klarhet i orsakerna till dessa avvikelser. I Tabell 7 och 8 ovan har det första värdet (i data ordning) valts för `StoraN` och `LillaN`, för de strata där det förekommer två värden. De skillnader som konstaterades mellan metodrapporten och data i föregående avsnitt kan, i de fall de inte beror på omkastningar i rader i metodrapporten, härröra från att värdena för dessa två variabler inte är unika. Ingen ytterligare utredning har utförts om detta.

Tabell 10 nedan visar alla stratum där `StoraN` och `LillaN` har två värden. Som synes är en stor majoritet av avvikelserna marginella, mindre än en procent, men det finns några Stratum där skillnaderna mellan de två värdena är betydande, och kan förväntas påverka beräkningen av uppräkningsvikter (se avsnitt 3.1.1). För `LillaN` så är dock de procentuella avvikelserna mer märkbara, endast i ett stratum så är avvikelserna mindre än en procent.

Tabell 10. Redovisning av stratum (StratumxKvartal) där StoraN och LillaN har två värden.

Stratum	Kvartal	StoraN				LillaN			
		Värde 1	Värde 2	Avvikelse	Avvikelse(%)	Värde 1	Värde 2	Avvikelse	Avvikelse(%)
6199	1	176	178	2	1,1%	176	178	2	1,1%
12199	1	277	278	1	0,4%	277	278	1	0,4%
12299	1	325	327	2	0,6%	80	82	2	2,5%
37301	1	510	512	2	0,4%	41	43	2	4,9%
37302	1	839	840	1	0,1%	55	56	1	1,8%
37304	1	713	714	1	0,1%	49	50	1	2,0%
37308	1	261	262	1	0,4%	12	13	1	8,3%
37404	1	550	552	2	0,4%	93	95	2	2,2%
46305	1	1401	1403	2	0,1%	206	208	2	1,0%
46401	1	1708	1711	3	0,2%	219	222	3	1,4%
46405	1	1436	1437	1	0,1%	34	35	1	2,9%
464299	1	131	132	1	0,8%	32	33	1	3,1%
479299	1	16	17	1	6,3%	3	4	1	33,3%
990004	1	4	19	15	375,0%	4	19	15	375,0%
990005	1	1	2	1	100,0%	1	2	1	100,0%
12199	2	272	274	2	0,7%	267	269	2	0,7%
37302	2	830	831	1	0,1%	87	88	1	1,1%
37304	2	697	699	2	0,3%	76	78	2	2,6%
37403	2	502	504	2	0,4%	8	10	2	25,0%
46301	2	1559	1561	2	0,1%	238	240	2	0,8%
46302	2	702	703	1	0,1%	70	71	1	1,4%
46303	2	469	471	2	0,4%	62	64	2	3,2%
46305	2	1393	1396	3	0,2%	181	184	3	1,7%
46306	2	370	371	1	0,3%	45	46	1	2,2%
46308	2	196	198	2	1,0%	27	29	2	7,4%
46401	2	1706	1708	2	0,1%	121	123	2	1,7%
46405	2	1448	1449	1	0,1%	56	57	1	1,8%
462199	2	45	46	1	2,2%	45	46	1	2,2%
464299	2	137	138	1	0,7%	34	35	1	2,9%
466299	2	68	69	1	1,5%	16	17	1	6,3%
990004	2	3	22	19	633,3%	3	22	19	633,3%
7199	3	43	44	1	2,3%	43	44	1	2,3%
10299	3	260	261	1	0,4%	64	65	1	1,6%
37301	3	512	513	1	0,2%	33	34	1	3,0%
37302	3	817	820	3	0,4%	73	76	3	4,1%
37304	3	715	716	1	0,1%	74	75	1	1,4%
37305	3	1112	1113	1	0,1%	102	103	1	1,0%
37307	3	179	180	1	0,6%	16	17	1	6,3%
46301	3	1625	1628	3	0,2%	244	247	3	1,2%
46302	3	713	714	1	0,1%	79	80	1	1,3%
46303	3	473	476	3	0,6%	75	78	3	4,0%
46304	3	908	909	1	0,1%	110	111	1	0,9%
46305	3	1404	1410	6	0,4%	178	184	6	3,4%
46308	3	204	205	1	0,5%	30	31	1	3,3%
46404	3	1021	1022	1	0,1%	54	55	1	1,9%
990004	3	2	28	26	1300,0%	2	28	26	1300,0%
5299	4	348	350	2	0,6%	97	99	2	2,1%
6199	4	164	165	1	0,6%	161	162	1	0,6%
7299	4	28	29	1	3,6%	5	6	1	20,0%
37301	4	509	510	1	0,2%	47	48	1	2,1%
37303	4	930	931	1	0,1%	72	73	1	1,4%
37304	4	708	709	1	0,1%	70	71	1	1,4%
37401	4	518	519	1	0,2%	19	20	1	5,3%

37403	4	487	488	1	0,2%	9	10	1	11,1%
37404	4	557	558	1	0,2%	9	10	1	11,1%
46301	4	1613	1619	6	0,4%	234	240	6	2,6%
46302	4	711	713	2	0,3%	72	74	2	2,8%
46304	4	897	900	3	0,3%	100	103	3	3,0%
46305	4	1397	1402	5	0,4%	177	182	5	2,8%
46306	4	378	380	2	0,5%	40	42	2	5,0%
46308	4	200	201	1	0,5%	27	28	1	3,7%
46402	4	770	771	1	0,1%	28	29	1	3,6%
46405	4	1402	1403	1	0,1%	81	82	1	1,2%
453299	4	116	117	1	0,9%	34	35	1	2,9%
463299	4	64	66	2	3,1%	17	19	2	11,8%
464299	4	144	145	1	0,7%	42	43	1	2,4%
465299	4	54	55	1	1,9%	12	13	1	8,3%
466299	4	66	67	1	1,5%	19	20	1	5,3%
990004	4	3	34	31	1033,3%	3	34	31	1033,3%

3.3 Bedömning av intern validitet för VFU 2009 vid tillämpning av bootstrap-metoden för validering

De problem som beskrevs i avsnitt 3.2.1 om inkonsistenser mellan data och metodrapporten påverkar inte tillämpningen av bootstrap-metoden, i och med att problemen troligen härrör från att en äldre version av data har använts för metodrapporten. Däremot påverkar problemen beskrivna i avsnitt 3.2.2 om att det finns ett betydande antal stratum där variablerna *StoraN* och *LillaN* har fler än ett värde (d.v.s. två värden), möjligheten att tillämpa bootstrap-metoden på ett konsistent och tillförlitligt sätt på VFU 2009.

Det är lämpligt att diskutera problemen med variablerna *StoraN* och *LillaN* var för sig eftersom de påverkar metoden på olika sätt. Variabeln *StoraN* som anger populationsstorleken för stratum (eller mer korrekt storleken räknad på urvalsramen) används inte direkt vid tillämpning av bootstrap-metoden, och tillhörde heller inte kraven redovisade i avsnitt 3.1.3. Därför påverkar den heller inte konsistensen och tillförlitligheten av att tillämpa metoden på data från VFU 2009. Däremot påverkas beräkningen av uppräkningsvikter (avsnitt 3.1.1) av *StoraN*. Detta påverkar i sin tur resultat av att använda bootstrap-metoden på data från VFU 2009. Problemen med *StoraN* behöver inte bero på att något fel förekommer. Orsaken kan vara att något efter-stratifieringsarbete har utförts i syfte att förbättra precisionen för uppräkningsvikterna. Om detta är fallet så är det bästa alternativet att tillämpa *StoraN* i beräkningen av uppräkningsvikter enligt ekvation (2), och helt enkelt negligera problemet med att variabeln har två värden i vissa stratum. Svårigheten är i detta fall att det inte finns någon dokumentation om *StoraN* som beskriver orsaken till att variabeln har dubbla värden i vissa stratum så osäkerhet finns om viktberäkning.

Som tas upp i avsnitt 3.1.3 så har variabeln *LillaN* en direkt påverkan på hur bootstrap-metoden tillämpas. Orsaken till detta är att bootstrap-metoden behöver kunna återskapa hur urvalsdragningen har gått och *LillaN* ger information om storleken på bruttourvalet i varje stratum. Den informationen är nödvändig om metoden ska kunna ta hänsyn till bortfall. Om problemet med dubbla värden för *LillaN* har uppstått under efter-

stratifieringsarbetet så saknas det i så fall information för att med säkerhet kunna åter-skapa storleken på bruttourvalen som användes i urvalsdragningen. Därmed blir slutsat-sen att problemen med `LillaN`, som beskrevs i avsnitt 3.2.2, i betydande grad förmins-
kar tillförlitligheten i att tillämpa bootsrap-metoden på data från VFU 2009. Bedömning
blir att VFU 2009 saknar intern validitet vid tillämpning av den valda metoden.

4 Slutsatser

En validering med hjälp av VFU 2009 är bara lämpligt att utföra om undersökningen tillförlitligt mäter verkliga godsvolymer härledda från varusändningar och om data från undersökningen är konsistent med den metod som används för valideringen. Den första förutsättningen, att undersökningen tillförlitligt mäter verkliga godsvolymer benämns som *extern validitet* för undersökningen. Den andra förutsättningen, att data från undersökningen är konsistenta med metoden för validering, benämns som *intern validitet*.

I avsnitt 2.4 drogs slutsatsen att godsvolymer avseende export-, import- och inrikes-sändningar för VFU 2009 inte uppfyller en godtagbar extern validitet i syfte att validera Samgodsprognoser för 2009. Slutsatsen byggde på jämförelser med VFU 2005 och utrikeshandelsstatistiken för 2005 och 2006. Som stöd för slutsatsen användes också av Trafikverket beställda prognoser för PWC-matriser för 2009 samt konfidensintervall för godsvolymer i VFU 2009. Det bör påpekas att de verkliga godsvolymer inte är kända, endast volymer från datakällor som mer eller mindre väl mäter de verkliga godsvolymer är kända. Därför kan frågan om extern validitet inte enbart avgöras utifrån objektiva kriterier såsom statistiskt signifikans (konfidensintervall). I slutändan bygger frågan om extern validitet på en rimlighetsbedömning som självklart kan ifrågasättas. Dessa bedömningar redogörs för i avsnitt 2.4.⁶

När det gäller den interna validiteten, att data från undersökningen är konsistent med metoden för validering, så konstaterades att det fanns dataproblem i VFU 2009 för variabler som är av betydelse för den valda valideringsmetoden. Dessutom finns inte metadata om undersökningen tillgängliga, som skulle kunna användas för att värdera och eventuellt korrigera dataproblemen. Därför har bedömningen gjorts att VFU 2009 inte uppfyller kravet för intern validitet med avseende på den valda valideringsmetoden. Bedömning om den interna validiteten redovisades i avsnitt 3.3.⁷

⁶ Trafikanalys har kommenterat en tidigare version av den innevarande rapporten. Trafikanalys kommentar om extern validitet har lagts in i texten i kapitlets tredje sista stycke nedan. Trafikanalys kommentar om den interna validiteten har lagts in som fotnot 7.

⁷ Trafikanalys kommentar om intern validitet: ”Trafikanalys håller med om att dokumentationen är bristfällig i vissa delar och tyvärr saknas det en bra dokumentation över databasen. Ändringar i stratifieringen och felskrivningar i dokumentationen kan vara orsaker till diskrepanserna. Diskrepanser i stratumkoder kan också bero på att registerbaserade strata finns med i databasen, men inte i dokumentationen. Produktionsfilerna som producenten har ansvarat för och använt sig av finns inte tillgängliga. Det är osäkert om producenten har sparat dessa. De produktionsfiler som skapats från .de tidigare undersökningsomgångarna innehåller också fullständig information om arbetsställen, men dessa är inte heller kända för oss. Finns inte produktionsfilerna sparade är det svårt att återskapa informationen. Det innebär att vi inte kan säga vad som är rätt eller fel dessvärre. Trots det kan vi konstatera inkonsistens, precis som påpekas.”

Slutsatserna om extern respektive intern validitet har olika status. Extern validitet är förvisso beroende av syftet för tillämpningen, i det här fallet validering av Samgodsprognoser. Däremot är inte extern validitet (så som den är här definierad) beroende av vilken metod som används för validering. Därför kan en bedömning av extern validitet även påverka synen på tillförlitlighet hos datakällan vid andra typer av analyser, beroende på syftet av analysen eller tillämpningen. Syftet för tillämpning i detta fall innebar att extern validitet enbart studerades avseende godsvolymer (ton) på total nivå för VFU 2009 samt för de tre delgrupperna inrikes-, export- och importsändningar. Därför kan slutsatsen om extern validitet i detta fall påverka bedömningen av tillförlitlighet i andra tillämpningar av VFU 2009 där godsvolymer på total nivå och de tre delgrupperna påverkar tillämpningens resultat. Vår bedömning är att dessa andra typer av tillämpningar kan utgöra ett betydande antal. Ett exempel på tillämpningar som nödvändigtvis inte påverkas av denna bedömning av den externa validiteten är studier av den geografiska fördelningen av varusändningar. Trafikanalys kommenterar extern validitet enligt följande:

Det finns begränsningar i varuflödesundersökningarna som gör att de inte fullt ut kan beskriva totala varuflöden på nationell nivå. För de mest betydelsefulla branscherna och de godstransportflöden (fördelat på varuslag) som dessa genererar, bedöms däremot undersökningarna ha god täckning. För importen bedöms täckningen vara sämre än för avgående sändningar. Detta eftersom urvalet till VFU grundar sig på en bedömning om arbetsstället/branschen har avgående sändningar av varor snarare än ankommande sändningar från utlandet.

Dessutom kommenterar Trafikanalys vidare enligt:

Förutom att användas för att validera varuslag och branscher tror vi slutligen att underlaget från VFU 2009 lämpligen används för jämförelser av olika fördelningar och transportmönster som exempelvis destinationer, transportkedjor och liknande. I de fall man vill skapa villkor, baserat på VFU data, för hur stora volymer som ska vara möjliga för exempelvis region till region i prognosmodellen, så är det möjligt att göra. VFU består också av registerundersökta branscher och de kan ses som totalundersökningar och borde vara utmärkta källor för validering.

Utifrån de uppgifter som har redovisats i texten om extern validitet tillsammans med Trafikanalys kommentarer avser projektet att arbeta vidare med validering av Samgods och PWC 2009, *men avgränsat till inrikes- och exportsändningar från registerundersökta branscherna*. Utgångspunkten för detta ställningstagande är att de redovisade problemen med extern validitet främst grundas på problem med undersökningens täckning.

Statusen för bedömningen av intern validitet (så som den är här definierad) är annorlunda. Intern validitet är starkt beroende av den specifika datakällan, metod samt syfte för tillämpning. Det kan i princip vara så att bedömningen att den interna validiteten ej var godtagbar inte kommer att påverka värderingen av resultat från VFU 2009 i något annat sammanhang.

Referenser

Berglund, M. och D. Pettersson (2016). *PWC-MATRISER 2009*. WSP 2016-11-14.

SCB (2017a). *Utrikeshandel med varor. Dokumentation*. SCB 2017.

<https://www.scb.se/hitta-statistik/statistik-efter-amne/handel-med-varor-och-tjanster/utrikeshandel/utrikeshandel-med-varor/#Dokumentation>

SCB (2017b). *Varuimport och varuexport efter handelspartner och varugrupp SITC rev3/rev4, ej bortfallsjusterat, sekretessrensad. År 1995 - 2016*. Statistikdatabasen, SCB 2017

http://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START_HA_HA0201_HA0201D/OImpExpSITC4Ar/?rxid=e28d504d-1e83-4200-a437-8a1441cea188

SIKA (2006a). *Varuflödesundersökningen 2004/2005*. SIKA Statistik 2006:12, SIKA.

SIKA (2006b). *Varuflödesundersökningen 2004/2005 Metodrapport*. Bilaga till SIKA Statistik 2006:12, SIKA.

Trafikanalys (2010). *Varuflödesundersökningen 2009*. Statistik 2010:16, Trafikanalys.

Trafikanalys (2011). *Trafikanalys Metodrapport varuflödesundersökningen 2009*. PM 2011:13, Trafikanalys.

