

Slutrapport Bättre primärstatistik BAST

Innehåll

1	Sammanfattning	3
2	Projektet	4
2.1	Projektets bakgrund och uppgift.....	4
2.2	Deltagare.....	4
3	Kunder och kundönskemål	5
4	Projektarbete och erfarenheter	5
4.1	Vad som var positivt.....	6
4.2	Vad som kan förbättras.....	6
5	Resultat	6
5.1	Handledning och blankett.....	7
5.1.1	Blankett - variabelöversyn.....	7
5.1.2	Handledning	8
5.1.3	Kontakt med användare	8
5.2	Metodöversyn	9
5.2.1	Disposition av metodöversyn	9
5.2.3	Förutsättningar för BAST.....	10
5.2.4	Genomgång av metodområden.....	11
6	Fortsättning	20

1 Sammanfattning

Projektet inleddes med en genomgång av vilka variabler i undersökningen Balansstatistiken (BAST) som används av Finansräkenskaperna (FiR).

I förslaget till ny blankett togs de sex variabler som fanns på tillgångssidan under nyteckning, inlösen och aktieägartillskott bort. Dessutom togs motsvarande tre variabler samt personalkonvertiblerna bort på skuldsidan. Skuldsidan utökades med ett avsnitt om eget kapital och obeskattade reserver med sex nya variabler. Finansräkenskapernas behov av en bättre uppdelning av aktier och andelar samt koncern och intresseföretag gav sex nya variabler i aktiedelen på tillgångssidan i BAST. Samtidigt togs fyra variabler bort från aktiedelen på tillgångssidan. Nettoförändringen av antalet variabler i blanketten blev en minskning med två variabler. Ingen anpassning av blanketten till företagens redovisning bedömdes vara nödvändig.

Den befintliga handledningen har genomgått en förbättring och anpassning till det nya blankettförslaget. Den handledning AP-fonderna använder har också varit uppe till diskussion. Frågan är vilken av de två som har bäst upplägg och är mest användarvänlig för uppgiftslämnarna. Detta avgörande skjuts upp till etapp 2.

Projektet har genomfört en metodöversyn där vissa områden med utredningsbehov pekas ut inför nästa projektetapp. Allmänt kan man säga att det kritiska karaktärsdraget för undersökningen är den skeva populationen med avseende på vissa variabler, där variablerna dessutom är lågt inbördes korrelerade med varandra. Det innebär att det är svårt att välja en urvalsdesign som är effektiv för samtliga variabler i undersökningen. De områden som bör utredas vidare är:

- Utredda storleken på cutoff-andelar i undersökningen
- Kartlägga bristerna i rampopulationen
- Studera om stratifieringen och allokeringen blir mer effektiv om undersökningen delas upp på två eller flera undersökningar
- Undersöka hur hanteringen av få förekomster kan förbättras för att åstadkomma mer robusta skattningar över tiden
- Klargöra om systemet i praktiken utför den imputeringsmetod som beskrivs i dokumentationen. Eventuellt kan metoden utvärderas för att studera hur den fungerar i de olika bortfallssituationerna
- Ta fram kriterier för att bestämma en outlier
- Utredda om det går att hitta en estimator som ger stabilare skattningar över tiden

- Studera om skattningarna blir mer stabila över tiden genom att slå samman stratum genom att testa på historiska data

2 Projektet

2.1 Projektets bakgrund och uppgift

Det finns två huvudsyften med Balansstatistiken. Det ena är att belysa de icke-finansiella företagens betydelse på olika finansiella marknader genom att visa storleken på och förändringarna av företagens finansiella tillgångar och skulder. Det andra syftet är att förse finansräkenskaperna med underlag för kvartalsvisa beräkningar av företagssektorns finansiella sparande.

Syftet med projektet som helhet (etapp 1-3) var att förbättra undersökningen Balansstatistiken, i fortsättningen kallad BAST, när det gäller metoden för ramkonstruktion, urvalsdesign och estimation för att få bättre kvalitet i data som idag levereras till finansräkenskaperna. Målet är att förändringarna ska vara färdiga så att man kan använda sig av den förbättrade metoden inför urvalet år 2010.

Projektet skulle även se över variabelinnehållet i BAST-blanketten för att se vilka variabler som är viktigast för finansräkenskaperna och vilka variabler som eventuellt kan tas bort. Projektet skulle även försöka anpassa blanketten till företagens redovisning för att minska uppgiftslämnarbördan. Till följd av eventuella förändringar av variabelinnehållet skulle handledningen till BAST-undersökningen ses över.

På ett tidigt stadium delades projektet in i tre etapper. Etapperna är:

1. Se över variabelinnehållet, handledningen, blanketten och metodöversyn
2. Genomföra metodutredning och testa blanketten i samarbete med PCA/MÄT
3. Fastställa förändringar inom metod, urval och ram inför dragningen av urval till år 2010.

Föreliggande slutrapport gäller etapp 1.

2.2 Deltagare

Projektleddare: Anna Hallqvist U/PRO-S

Projektdeltagare:
Barbro von Hofsten ES/NS

Nicolai Nystrand NR/FSR
Lisa Allemo ES/NS
Robert Jäverlind NR/LEDN
Pär Lindholm PCA/MFFM-Ö

3 Kunder och kundönskemål

Projektet är beställt av Mattias Björling, enhetschef på NR/FSR. Slut användare är Finansräkenskaperna (NR/FSR) och Betalningsbalansen (ES/BFM). Betalningsbalansen har informerats om projektet och har lämnat synpunkter och önskemål.

Finansinspektionen är SAM-myndighet för Finansräkenskaperna och är indirekt uppdragsgivare och intressent för detta projekt.

4 Projektarbete och erfarenheter

Idéer om att genomföra ett projekt med syfte att förbättra kvaliteten i undersökningen BAST har funnits i många år. När Emma-projekten startades under 2006 kom översynen av BAST med som en liten del i det större projektet "Bättre primärstatistik". På vägen föll detta projekt mellan stolarna och huvudprojektet fokuserade på andra ämnesområden. Våren 2008 skickades två olika projektansökningar avseende BAST in till PLG. Den ena avsåg metodöversyn och den andra avsåg översyn av handledning och blankett. Beslut togs då om att driva dessa delar tillsammans i ett gemensamt projekt.

Efter en inledande diskussion på ett BAST-möte i Stockholm den 28 april, startades projektet formellt den 19 augusti 2008. Då samlades hela projektgruppen plus Marianne Biljer, Mårten Pappila och Christina Rääf för ett uppstartsmöte i Örebro. Mötet ägnades åt att brainstorma angående idéer och förväntningar på vad projektet skulle åstadkomma. Dessutom hade Mårten och Nicolai en genomgång av hur BAST-uppgifterna läggs in i Finansräkenskaperna och vilka andra källor som används. En tidplan upprättades med aktiviteter och datum för projektets olika delar.

Under hösten har projektet haft regelbundna möten och avstämningar, både heldagsmöten i antingen Örebro eller Stockholm och avstämningsmöten på video eller telefon. Oftast har hela projektgruppen deltagit vid mötena. Mellan mötena har arbetet delats upp i två områden. Arbetet med fokus på Finansräkenskapernas behov och handledningen har NR/FSR ansvarat för, medan arbetet med fokus på metod och blankett har genomförts huvudsakligen av projektdeltagare från ES/NS och PCA/MFFM. Metodutredningen blev något försenad pga. andra prioriterade arbetsuppgifter

utanför projektet. Den första etappen har därför ägnats åt kartläggning och identifiering av problem på metodområdet. I etapp två kommer fler studier av materialet att kunna genomföras och förslag på lösningar att kunna lämnas.

Problem och frågor som dykt upp under resans gång har antingen tagits upp vid nästa möte eller genom kontakt via e-post och telefon mellan mötena. Projektgruppen har löpande hållit varandra informerade när andra kontakter, som haft anknytning till projektet, ägt rum. Exempel på detta är kontakter med ES/BFM och PCA/MÄT.

4.1 Vad som var positivt

- Bra fart i arbetet tack vare täta möten och avstämningar
- Fördjupad kunskap inom andra områden än det egna
- Det har varit en trevlig prestigelös projektgrupp som har arbetat bra tillsammans
- Det har varit högt i tak
- Bra att dela upp projektet i flera etapper
- Utökat kontaktnät på SCB
- Tydlig avgränsning av projektet efter hand

4.2 Vad som kan förbättras

- Det behövs tydligare mål redan från början. Det måste finnas en projektbeställning när projektet startar. En konsekvens av avsaknad beställning har i detta projekt blivit att flera idéer som kom upp i början inte kunnat tas om hand inom projektet
- Beställare bör vara lite mer aktiv
- Strukturen i ViP bör förenklas. Det är svårt att hitta och det finns inga logiska sökvägar. Dokument har sparats på lite olika ställen som var och en tyckt har passat
- Det har varit lite svårt att hålla ihop projektet när projektgruppen är spridd över två orter
- Det behövs tydligare prioritering från cheferna av projektmedlemmarnas arbetsuppgifter. Det har inte varit möjligt att avsätta så mycket tid som egentligen skulle ha behövts för projektet pga. "ordinarie" arbetsuppgifter

5 Resultat

Projektets delmål att se över variabelinnehållet utifrån Finansräkenskapernas behov har uppnåtts i stort sett enligt plan. Handledningen och blanketten har

setts över enligt projektets uppsatta mål. Kartläggning och identifiering av problem på metodområdet har genomförts.

En försvårande omständighet var Finansdepartementets krav att SCB ska minska uppgiftslämnarbördan med 25 procent. Projektet har försökt att ta hänsyn till detta, men beslut om användarnytta kontra uppgiftslämnarbörda måste tas högre upp i SCB:s organisation.

Rapportens resultatdel är indelad i två huvudavsnitt, ett där arbetet med handledning och blankett beskrivs (5.1) och ett där metodöversynen beskrivs (5.2). Arbetet har genomförts som två separata utredningar och har dokumenterats på skilda sätt. Resultatdelen är i första hand skriven för personer som är insatta i undersökningen och som en input till etapp två.

5.1 Handledning och blankett

I denna första etapp av projektet har förbättringar av variabelinnehållet, blanketten och handledningen föreslagits. Föreslagna förbättringar kommer att implementeras först när hela projektet (etapp 1-3) är slutfört.

5.1.1 Blankett - variabelöversyn

Projektet inleddes med en genomgång av vilka variabler i BAST-undersökningen som används av Finansräkenskaperna (FiR). En excelfil med BAST-värden och värden från andra källor sammanställdes för åren 2005-2007. Källorna var: Monetära Finansinstitut (MFI), Värdepapperscentralen (VPC) och undersökningen Företagens Ekonomi (FEK). Detta underlag användes för att jämföra uppgifter från de olika källorna och för att kartlägga vilka värden som idag används för de olika posterna i FiR.

Uti från excelunderlaget gjordes en variabelöversyn för att se vilka variabler som skulle kunna tas bort och vilka som eventuellt skulle behöva läggas till. Förslagen på förändringar lämnades sedan till ES/NS som förde in förändringarna i ett utkast till ny blankett. Nuvarande blankett bifogas som Bilaga 1, utkast till ny blankett bifogas som Bilaga 2.

Ingen anpassning av blanketten till företagens redovisning bedömdes vara nödvändig. Variablerna i BAST är mycket mer detaljerade än uppställningen av företagens balansräkning. En förändring av ordningen i blanketten skulle därför inte innebära någon förenkling för uppgiftslämnarna.

I förslaget till ny blankett togs de sex variabler som fanns på tillgångssidan under nyteckning, inlösen och aktieägartillskott (rad 40-43) bort, eftersom

den detaljnivån på uppgifter inte används i Finansräkenskaperna. Av samma anledning togs tre variabler (rad 60, 302 och 303) på skuldsidan också bort.

Även personalkonvertiblerna (rad 24) har tagits bort. Variabelns betydelse har minskat under de senaste åren eftersom företagen idag använder andra typer av belöningsystem för sina anställda.

Finansräkenskapernas behov av en bättre uppdelning av aktier och andelar samt koncern och intresseföretag resulterade i sex nya variabler i aktiedelen på tillgångssidan i BAST. Samtidigt togs fyra variabler bort vilket resulterade i att aktiedelen nu består av åtta variabler istället för sex. Bokförda värden efterfrågas nu för alla variabler i den föreslagna aktiedelen, utom för tillgångarna i noterade aktiebolag där istället marknadsvärden efterfrågas. Anledningen till detta är att FiR tror att företagen gör marknadsvärderingen på olika sätt. Projektet föreslår därför att samla in bokförda värden för att sedan marknadsvärdera dem på ett enhetligt sätt.

Skuldsidan utökades med ett avsnitt om eget kapital och obeskattade reserver. Eget kapital i onoterade aktiebolag är mycket viktigt för Finansräkenskaperna, eftersom det med denna information i kombination med uppgifter från andra källor kan bli möjligt att residualt beräkna hushållens tillgångar i onoterade aktiebolag. Totalt består den nya delen på skuldsidan av sex nya variabler.

Kolumnen "Omvärdering" i blanketten har döpts om till "Orealiserade värdeförändringar". Nettoförändringen av antalet variabler i blanketten efter översynen blev en minskning med två variabler.

5.1.2 Handledning

Den befintliga handledningen är utformad som en uppslagsbok där varje variabel beskrivs detaljerat. För varje variabel anges hur transaktionen eller omvärderingen (den orealiserade värdeförändringen) skall beräknas. De föreslagna variablerna har lagts till i handledningen och förklarats. Dessutom har en allmän översyn med mindre förändringar skett, där vissa förklaringar har förtydligats. Projektet har tagit hjälp av Jonas Ehrnst på ES/BFM för att beskriva den föreslagna delen om eget kapital.

En annan idé har varit att använda samma upplägg som i AP-fondernas handledning. Den har en lång beskrivande del i inledningen, där även exempel på transaktioner och orealiserade värdeförändringar (omvärderingar) finns. Variablerna beskrivs i denna handledning mer översiktligt än i den befintliga handledningen.

Ett förslag som ska diskuteras är att båda dessa handledningar ska testas av Enheten för Mätteknik tillsammans med den nya blanketten för att se vilket upplägg företagen föredrar. Ett annat tillvägagångssätt kan vara att projektgruppen kontakter några uppgiftslämnare och ber dem läsa båda handledningarna och lämna synpunkter.

5.1.3 Kontakt med användare

Arbetsgruppen som arbetar med den totala beräkningen av Betalningsbalansen har informerats om projektet och vilka effekter projektets arbete kan få för de utlandsposter i BAST som ingår i statistikundersökningen Betalningsbalansen (BB BAST). BB BAST ingår som en del i den totala beräkningen av Betalningsbalansen. Arbetsgruppen fick komma med förslag och ombads att se över handledningen som används i BB BAST. Inga förändringar av denna handledning behövde göras, däremot vill de att vi ändrar namnet på BB BAST till "Balansstatistik mot utlandet".

5.2 Metodöversyn

I en undersökning kan det förekomma flera typer av fel av varierande storlek som härrör från olika delområden i undersökningen. Vilka områden som genererar de största felen i BAST undersökningen vet vi i dagsläget inte. Pga. undersökningens karaktär och erfarenheter från statistikproduktionen och användningen av statistiken kan vi dock misstänka vilka områden som är problematiska.

De synpunkter som har framförts, både från de personer som producerar statistiken och primära användare av statistiken, NR/FSR, är att nivåskattningarna generellt är för låga och att de är skakiga över tiden samt att det blir hopp i skattningarna vid urvalsbyten. Vissa skattningar anses vara så osäkra att de inte kan användas av NR/FSR.

Det kritiska karaktärsdraget för undersökningen är den skeva populationen med avseende på vissa variabler, där variablerna dessutom är lågt inbördes korrelerade med varandra. Det innebär att det är svårt att välja en urvalsdesign som är effektiv för samtliga variabler i undersökningen. Det blir även skakigt vid uppräknings till populationsnivå då urvalet innehåller få förekomster med nollskilda värden för vissa variabler.

5.2.1 Disposition av metodöversyn

I projektets etapp 1 har för undersökningen relevanta metodområden beskrivits och diskuterats. Projektet pekar på problematiska områden och

vidare utredningsbehov till etapp 2. Alla områden har inte beskrivits lika utförligt, vilket indikerar att utrednings-/förbättringsbehoven i de mindre utförliga beskrivningarna inte är lika akuta. Före metodgenomgången ges i avsnitt 5.2.3 förutsättningarna för undersökningen.

Följande metodområden har gåtts igenom:

- Ramkonstruktion (avsnitt 5.2.4.1)
 - Cutoff-ansats och registerpopulation
 - Rampopulation
- Urvalsdesign (avsnitt 5.2.4.2)
- Mätning (avsnitt 5.2.4.3)
- Granskning (avsnitt 5.2.4.4)
- Estimation (avsnitt 5.2.4.5)
 - Bortfallskompensation (imputering)
 - Outlierhantering
 - Val av estimator
- Röjandekontroll (avsnitt 5.2.4.6)
- Kvalitetsvärdering (avsnitt 5.2.4.7)

Under vissa punkter har projektgruppens kommentarer och frågeställningar angivits kursivt. I de fall det föreligger utredningsbehov har dessa beskrivits under en särskild rubrik i anslutning till området.

5.2.3 Förutsättningar för BAST

5.2.3.1 Syfte

Statistiken skall dels belysa de icke-finansiella företagens betydelse på olika finansiella marknader genom att visa storleken på och förändringarna av företagens finansiella tillgångar och skulder, dels förse finansiella räkningskaperna med underlag för kvartalsvisa beräkningar av företagssektorns finansiella sparande.

5.2.3.2 Målobjekt

Målobjekten för undersökningen är företag.

5.2.3.3 Målpopulation

Målpopulationen utgörs av alla företag inom näringslivet, exklusive de finansiella företagen och stiftelser, som någon gång under referensåret är verksamma och har finansiella tillgångar och/eller skulder.

Målpopulationen borde vara densamma som i undersökningen Företagens Ekonomi m a p finansiella tillgångar och skulder med undantag för stiftelser.

5.2.3.4 Målvariabler

Undersökningens målvariabler utgörs av finansiella tillgångar (t.ex. banktillgodohavande, derivat, certifikat) och finansiella skulder (t.ex. derivat, certifikat, upplåning). Se blanketten i bilaga 1 eller Beskrivning av statistiken (BAS) för en fullständig variabelförteckning.

För en lekman utgörs undersökningen av ekonomiska variabler som är komplicerade. Som vi tidigare har betonat så är populationen väldigt skev med avseende på vissa variabler och variablerna är väldigt olika sinsemellan i sin karaktär. T.ex. har de flesta företag i populationen banktillgodohavanden medan en liten andel av företagen har värdepapper i form av certifikat mot en viss sektor.

5.2.3.5 Statistiska mått

Totaler av variabler för finansiella tillgångar/skulder i form av ställningsvärde samt transaktioner och omvärderingar.

5.2.3.6 Redovisningsgrupper

Resultat redovisas för den totala sektorn icke-finansiella företag, dvs. ingen redovisning på finare grupper.

5.2.3.7 Periodicitet

Statistiken framställs och redovisas kvartalsvis. Statistiken levereras till NR/FSR ungefär 55 dagar efter referenskvartalets utgång.

5.2.3.8 Primärt användningsområde

Finansräkenskaperna indaterar via ett inläsningsunderlag alla BAST-uppgifter till FiR-systemet. Sedan jämförs dessa siffror med motsvarande siffror från andra källor och FiR bestämmer vilken källa som ska användas. De variabler från BAST som är viktigast för FiR är de där det saknas information från andra källor. Dessa är framförallt certifikat, obligationer, onoterade aktier och andra andelar på tillgångssidan. Eget kapital, andra ägarandelar, certifikat och obligationer är de viktigaste variablerna på skuldsidan.

5.2.4 Genomgång av metodområden

5.2.4.1 Ramkonstruktion

Cutoff-ansats och registerpopulation

BAST undersökningen genomförs som en cutoff-undersökning. Förteckningen över de företag (juridiska enheter) som kan bli aktuella för

ramppopulationen baseras på FDB:s företagsregister (*marsSAMU*) och tillhör sektorerna icke-finansiella företag och statliga affärsverk. Även inaktiva företag ingår i förteckningen. Till denna förteckning läggs företag till som är "stora" enligt annan källa. Även manuella tillägg av företag förekommer. Denna förteckning av företag kallar vi fortsättningsvis lite slarvigt för registerpopulation.

Företagen i ramppopulationen för år t bestäms genom att välja ut samtliga företag från registerpopulationen som har ett värde på storleksmättet större än tröskelvärdet. Tröskelvärdet är 30 mnkr.

Storleksmättet för år t bestäms enligt följande hierarkiska ordning för respektive företag:

- Källa 1. Ställningsvärde 31:a dec år t-1 i undersökningen BAST.
- Källa 2. Ställningsvärde i form av värdepapper 31:a dec år t-1 från Värdepapperscentralen (VPC), om värdet från källa 2 är större än värdet från källa 3.
- Källa 3. Ställningsvärde 31:a dec år t-2 från Standardiserade RäkenskapsUtdraget (SRU).

Ställningsvärdet är summa finansiella tillgångar och skulder. I källa 3 (SRU) kan ställningsvärdet komma från ett annat datum om företaget har ett annat bokslutsdatum. Majoriteten av företagen får ett storleksmått baserat på källa 3.

Genom en cutoff-ansats undantas en stor mängd företag från ramppopulationen. Motivet för att exkludera företagen är att de antas ha en försumbar inverkan på totalskattningen. I BAST står dessa företag för uppskattningsvis 7-8 procent av det totala storleksmättet, vilket får betraktas som en ganska stor andel. Dessutom kan den andelen vara ännu större för enskilda variabler. T.ex. är motsvarande andel för variabeln banktillgodohavanden cirka 35 procent.

I dagsläget görs ingen kompensation för cutoff-andelen vid totalskattningen av en parameter. Det innebär en systematisk underskattning.

För en fördjupad beskrivning av cutoff- och registerpopulation så hänvisas till bilaga 3.

Utredningsbehov

Det är viktigt att i nästa etapp försöka utreda storleken på dessa cutoff-andelar. För aktievariablerna har jämförelser gjorts med uppgifter från VPC. Övriga variabler kan inte stämmas av mot VPC. En framkomlig väg för

dessa variabler är att utgå från SRU-registret och jämföra registret mot gamla rampopulationer i BAST (t.ex. årgångar 2006 och 2007). Det är önskvärt att mäta dessa andelar på variabelnivå. Dock finns inte uppgifter i SRU-registret på BAST-undersökningens variabelnivå med undantag för banktillgodohavanden. Jämförelser får därför göras på grupperingar av variabler för att uppskatta cutoff-andelens storlek på en finare nivå än summa finansiella tillgångar och skulder.

Även om utredningen inte leder till en justering av tröskelvärdet eller möjlighet att finna kompensationsfaktorer så är denna kartläggning viktig för att ge en bättre förståelse för hur statistiken bör användas på variabelnivå och som en input till kvalitetsbeskrivningen.

Rampopulation

Som vi tidigare har beskrivit så innehåller rampopulationen de företag som ingår i registerpopulationen och har ett värde på storleksmättet större än 30 mnkr. Samma problem som för registerpopulationen är att informationen som anger storleksmättet vanligtvis är två år gammal. Det innebär att undersökningen riskerar att missa ”nya” företag som har betydande tillgångar och skulder. Extra påtagligt blir det för en undersökning som BAST då de finansiella tillgångarna/skulderna är lätttrörliga.

I vissa fall har undersökningen genom omvärldsbevakning eller andra källor (ES/BFM, Storföretagsgruppen, MM) identifierat betydande företag som initialt inte ingick i rampopulationen.

Utredningsbehov

På motsvarande sätt som att utreda cutoff-andelen i registerpopulationen så bör en utredning göras som kartlägger bristerna i rampopulationen i nästa etapp. Genom att jämföra SRU-registret och BAST undersökningens rampopulation för äldre årgångar (2006, 2007) kan vi försöka kvantifiera under- och övertäckningen i rampopulationen.

5.4.2.2 Urvalsdesign

Urvalsdesignen är stratifierat OSU och urvalet dras i SAMU (marsSAMU). Rampopulationen stratifieras efter storlek, där storleksmättet är samma mått som tidigare använts vid cutoff-förfarandet.

Stratifiering och allokering

Stratifiering och allokering görs utanför SAMU. Stratifieringen är utformad med hjälp av en modellbaserad metod som beskrivs t.ex. i avsnitt 12.4 i boken Model Assisted Survey Sampling av Carl-Erik Särndal, Bengt

Swensson och Jan Wretman (se även avsnitt 12.5). Allokeringen är gjord så att urvalsstorleken är ungefär densamma i varje urvalsstratum.

Totalt utgörs ramen av 10 storleksstratum varav fyra (post)stratum är totalundersökta. För undersökningsåret 2008 omfattar urvalsramen 11 282 företag och i urvalet ingår 800 företag, varav 320 ingår i ett totalundersökt stratum. Samma urval används för samtliga kvartal under kalenderåret.

Storleksmättet som används för stratifiering korrelerar dåligt med vissa variabler i undersökningen. T.ex. så har flera företag stora tillgångar men små skulder. Det medför att stratifieringen inte blir effektiv och även att vissa företag måste outlierhanteras (se avsnitt 5.4.2.5). Ett förslag som har diskuterats är att undersökningen ska delas upp i två undersökningar, en som mäter tillgångar och en som mäter skulder. Dock har ingen utredning gjorts för att värdera effekten av en sådan uppdelning.

För många undersökningsvariabler är det en liten andel av företagen i ramen som har nollskilda värden. Det skapar problem vid stratifiering och allokering. I förväntan får vi för få svarande företag med nollskilda värden i urvalsundersökta stratum. T.ex. så är det ungefär 40 företag av drygt 700 svarande företag i urvalet som har tillgångar i form av certifikat, cirka 1/3 av företagen tillhör urvalsundersökt stratum. Uppdelat på sektorer är det mellan 10-15 företag som har ett nollskilt värde.

Det medför att skattningarna blir ”skakiga” beroende på vilket urval som dras. Särskilt påtagligt blir det vid urvalsbyten. Eventuellt kan de minsta storlekssklasserna slås samman för att säkerställa att vi i förväntan får X antal företag med egenskapen som undersöks. Dock kan det i sin tur leda till att stratifieringen inte blir effektiv för andra variabler. Ett sätt att lösa problemet är att kraftigt öka urvalsstorlekarna i små stratum, till priset av en hög kostnad i form av högre produktionskostnader och ökad uppgiftslämnarbörd. Projektet tror inte att den lösningen är en framkomlig väg. Om det inte går att förbättra designen så kan problemet eventuellt hanteras i estimationsfasen (se avsnitt 5.4.2.5).

Eftersom urvalet (och rampopulationen) utgår från marsSAMU så är rampopulationen idealisk för första kvartalet i kalenderåret. Sedan försämrar rampopulationen successivt under året. Det görs ingen justering för detta i estimationen. Undersökningen bör även överväga om betydande företag som upptäcks under året ska läggas till i urvalet (i ett totalundersökt stratum) och i rampopulationen.

Utredningsbehov

Det är intressant att studera om stratifieringen och allokeringen blir mer effektiv om undersökningen delades upp på två eller flera undersökningar.

Dock kvarstår problemet med skeva variabler och låg inbördes korrelation vid en grov uppdelning av undersökningen. En fin uppdelning av undersökningen låter sig inte göras då hjälpinformation saknas på variabelnivå.

Ett tydligare utredningsbehov är att undersöka hur hanteringen av få förekomster kan förbättras för att åstadkomma mer robusta skattningar över tiden. Hur säkerställer vi att vi får ett minsta antal förekomster med nollskilda värden för alla variabler i de urvalsundersökta stratumerna? (Problemet kanske hanteras mer effektivt i estimationsfasen). Tonvikten i utredningen bör ligga på de variabler som Finansräkenskaperna betraktar som viktigast.

5.4.2.3 Mätning

Uppgifterna samlas in via en pappersenkät som går ut till utvalda företag eller genom att företagen laddar ner en excelblankett från BAST-undersökningens webbplats. Företagen ska fylla i ställningsvärde för efterfrågat kvartal i marknadsvärde. Om ställningsvärde som lämnats föregående kvartal är reviderat så ska företagen ändra på det förtryckta värdet. Företagen ska också ange omvärdering och/eller transaktion som skett under kvartalet. Mätsvårigheter finns för ställningsvärde, men framförallt för omvärdering och transaktioner. Vissa företag har inte möjlighet att lämna uppgifter om dessa variabler. Ett exempel på detta är när företagen inte kan ange omvärderingar pga. att de har utländska tillgångar och skulder i många olika valutor.

Parallellt med metodöversynen görs en utredning där blankett och handledning ses över. Dessa kommer sedan att genomgå test. För mer ingående information om mätning hänvisas till den utredningen (Se avsnitt 5.1).

5.4.2.4 Granskning

Inkomna uppgifter granskas och återkontakt tas vid behov med uppgiftslämnaren. Uppgifterna dataregistreras av personal inom DFO/FU. Det finns ett antal maskinella granskningskontroller för felsignalering på variabelnivå som sedan kan leda till rättning.

För att se ett företags inverkan på totalskattningen granskas uppräknade värden (insamlade och imputerade), på objektsnivå men även på stratumnivå. Makrogranskning görs genom att jämföra referenskvartalets utfall med föregående kvartals utfall.

Granskningsområdet beskrivs inte mer ingående i denna metodöversyn. Eftersom granskning är ett omfattande område och i dagsläget inte anses vara ett kritiskt moment för undersökningen, så görs inget försök att i detta projekt förbättra metoden eller arbetsrutinerna för granskning. Ett stort granskningsprojekt (SELECT/EDIT) pågår för närvarande på SCB och det är lämpligt för undersökningen att avvakta det projektets utgång.

5.4.2.5 *Estimation*

Estimationsmetoden är en HT-estimator inom strata. Imputering görs för både partiellt- och objektsbortfall. Det innebär att vikterna som används vid uppräknings är fixa från designen.

I estimationsfasen finns ett antal kritiska moment. Bortfallskompensation är ett betydande moment för undersökningen, ett annat är outlierhantering. Dels hur en outlier identifieras, dels hur vi i estimationen väljer att hantera dessa outliers. Ytterligare ett kritiskt moment är valet av estimator. Den nuvarande estimatoren är sårbar för ett problem som vi tidigare har diskuterat, dvs. skattningar av variabler där få företag har den sökta egenskapen.

Produktionssystemet tar fram punktskattningar, dock görs inga variansberäkningar vilket är en stor svaghet för undersökningen.

Bortfallskompensation

Objektsbortfall avseende antal företag är ca 9% under kvartal 1 2008 (ca 6 - 12% under 2007). Objektsbortfallet brukar vara högst under första kvartalet. Bland totalundersökta stratum är bortfallet lågt. Mellan 6 och 13 företag av totalt 285 företag, som ingår i stratum som är totalundersökta, saknades år 2007. Trots att objektsbortfallet är lågt kan det ibland ändå bli alltför stora skevheter i skattningarna pga. brister i imputeringen.

Bortfallskompensationen görs genom att imputera saknade värden på företagsnivå. Imputering görs för såväl partiellt- som objektsbortfall. Mycket av det partiella bortfallet åtgärdas redan i mikrogranskningen genom att återkontakta företagen och begära in saknade uppgifter eller genom manuell imputering.

Metoderna för imputering beskrivs utförligt i dokumentet Imputering_BAST_2005.doc (Per Magnusson, Tiina Orusild, 2005-06-14).

Det finns misstankar om att imputeringen (medelvärdesimputeringen) kan leda till underskattningar vid urvalsbyten då många "nya" företag tillkommer. Framförallt för de variabler och stratum som har många nollvärden.

Imputeringen utförs i produktionssystemet (SQL-miljö). Svagheten med systemet är att det inte är transparent (för metodstatistiker) och att den inte lagrar historiskt imputerade värden vilket försvårar möjligheten till felsökning och utvärdering av effekterna av imputeringarna.

Utredningsbehov

Imputeringsmetoden i BAST är mycket genomarbetad och väl dokumenterad. Imputeringarna i undersökningen har tidigare varit föremål för en utvärdering men avstannade 2006 och fick lämna plats för annat utvecklingsarbete på SCB. Det behöver klargöras om systemet i praktiken utför den metod som beskrivs i dokumentet. Eventuellt kan metoden utvärderas för att studera hur den fungerar i de olika bortfallssituationerna. Dock är bortfallet relativt lågt varför utredningsbehovet inte är kritiskt.

På sikt borde imputeringssystemet migreras till SAS-miljö för att få det mer transparent och enligt vår bedömning enklare att förvalta.

Outlierhantering

När vi i designen inte kan skapa en effektiv stratifiering för samtliga variabler då de är lågt inbördes korrelerade så uppstår värden som kan betraktas som avvikande eller som outliers. Det är vanligtvis företagsvärden från små stratum som uppräknade får en markant inverkan på totalskattningen. Sådana värden kan genom enkla listor från systemet identifieras. I dagsläget finns ingen regel som identifierar om ett värde ska betraktas som en outlier utan bedömningen blir subjektiv.

En outlier identifieras på variabelnivå. Uppräkningsvikten för en outlier ändras i estimationen och sätts till 1. Tolkningen är att företaget fortfarande får bidra med värden till totalskattningen men representerar enbart sig själv.

Utredningsbehov

Några kriterier för att bestämma en outlier vore önskvärt. Hur mycket ska ett värde inverka på totalskattningen för att kodas till outlier? Identifiering av outliers blir oftast skev. Det är ”enklare” att upptäcka stora avvikande värden än små avvikande värden.

Val av estimator

Ett antal totaler skattas med hjälp av en HT-estimator inom strata. Skattningarna beräknas för hela populationen (icke-finansiella företag), redovisningsgrupper i form av delmängder av populationen förekommer inte.

Det vi vill skatta är totaler för olika variabler. Totalen, t_y , för en variabel y kan skrivas

$$t_y = \sum_{k \in U} y_k$$

där U betecknar populationen och y_k är värdet på variabeln y för företag k .

En skattning av t_y ges av

$$\hat{t}_y = \sum_{k \in s} w_k \cdot y_k$$

där s är urvalet och $w_k = 1 / (\text{sannolikheten att företag } k \text{ kommer med i urvalet})$. Vikten w_k bestäms av vilken urvalsdesign man använder. I BAST undersökningen dras urvalet som ett stratifierat urval med OSU (Obundet Slumpmässigt Urval) inom strata. Detta innebär bl.a. att vikten ges av

$$w_k = \frac{N_h}{n_h} \quad \text{om företag } k \text{ tillhör stratum } h$$

där N_h är storleken på stratum h (d.v.s. antalet företag i stratum h) och n_h är urvalsstorleken i stratum h .

Eftersom bortfallskompensationen hanteras i form av imputering upp till urvalsmängden, s , är vikterna givna från designen (med undantag för outlierföretagen).

Det ska betonas att uppräknings här sker till rampopulationen, U , inte till registerpopulationen. I estimationen görs ingen kompensation upp till registerpopulationen vilket medför en systematisk underskattning av totalerna.

Utredningsbehov

Ibland finns möjligheten att i estimationsfasen kompensera för en svag design. Vi har tidigare diskuterat svårigheterna i att skapa en effektiv design varför vi bör utreda om vi kan finna en estimator som ger stabilare skattningar över tiden. Vi eftersträvar en robust estimator som är mindre känslig för vilket urval som erhålls.

Att utnyttja hjälpinformation i estimationsfasen kan i många tillämpningar vara effektiv. I BAST är dock hjälpinformationen svag. Ett sätt att få en mer robust estimator (till priset av bias) är att för vissa variabler med få förekomster slå samman stratum i estimationsfasen. Dessa stratum bör på förhand vara givna så att vi inte betingar på det urval som dragits. Det

innebär i så fall att ansatsen om ett helt gemensamt viktsystem frångås. En ansats med individuella vikter (för huvudvariabler) utesluter dock fullständig konsistens och försvårar sambandsanalyser, dock ställs inte sådana krav på undersökningen.

Genom att testa på historiskt data kan vi studera om skattningarna blir mer stabila över tiden genom att slå samman stratum. Detta är en högst lämplig utredning i etapp 2.

En ny ansats värd att fundera på är att frånga samplingskattningar för urvalsundersökta stratum på sektoruppdelade variabler och istället modellberäkna deras totaler utifrån en lämplig modell.

5.4.2.6 *Röjandekontroll*

Röjandekontroll av tabelldata på SCB kan betraktas som en sammansatt process som kan delas upp i ett antal mindre delprocesser. En uppdelning i delprocesser kan göras enligt följande:

- Skadeprovning
 - Röjanderiskbedömning
 - Skaderiskbedömning
- Skydda tabell
- Kvalitetsbedömning

Skadeprovning består av två delar men vanligtvis på SCB är utgångspunkten att skaderisk föreligger varför skadeprovningen då reduceras till att utföra en röjanderiskbedömning. I praktiken är det nog röjanderiskbedömningen som brukar betraktas som en metodarbetsuppgift även om hela röjandekontrollprocessen kan ses som en "metodprocess".

I dagsläget utförs ingen röjandekontroll av tabelldata i BAST. Detta är en brist och borde på sikt åtgärdas. För primär röjanderiskbedömning bör den metod som rekommenderas som standard för magnitudtabeller användas. Vanligtvis används tröskelvärdesregeln i kombination med dominansregeln men det finns diskussioner om att ersätta dominansregeln med p%-regeln. När det gäller urvalsundersökningar på SCB så används lite olika beräkningsvarianter på dominansregeln. Den metod som bör användas är att utföra beräkningen på motsvarande sätt som för totalundersökningar men att använda urvalsvikterna.

Ett förslag är att röjanderiskberäkningen görs på ställningsvärde och att ställningsvärde sedan används som skuggvariabel för omvärdering och transaktion.

Vissa variabler i BAST tillåter negativa värden. Det skapar ett problem vid röjanderiskbedömning, men får försöka hanteras på något sätt.

Projektet ser inte något större behov av att utföra en fördjupad utredning (om inte en skaderiskbedömning ska göras) utan metoderna bör snarast implementeras i undersökningen. Då det för tillfället saknas standardverktyg för röjandekontroll på SCB får vi ta en fortsatt diskussion om hur implementeringen i undersökningen ska genomföras.

5.4.2.7 Kvalitetsvärdering

I dagsläget levereras inte några mått på osäkerhet eller redovisning av tillförlitligheten i skattningarna till NR/FSR. Vi har tidigare nämnt att varians och således inte heller osäkerhetsmarginaler beräknas. De primära användarna av statistiken är kvalificerade användare som själva måste värdera statistiken för att rätt försörja sitt statistiksystem. För att lättare kunna värdera kvaliteten i skattningarna så rekommenderas att ett antal mått redovisas i anslutning till statistiken. Exempel på sådana mått kan vara

- 95%-igt konfidensintervall i absoluta eller relativa termer
- Andelen av den skattade totalen som svarsbortfallet står för
- Förändringar mot föregående kvartal som totalundersökta stratum står för jämfört med den totala förändringen
- Antal svarande företag med nollskilda värden

Det kan även finnas andra mått som är intressanta för att värdera kvaliteten i skattningarna. Personer som arbetar med statistikleveransen och de primära användarna, NR/FSR, bör diskutera om och i så fall vilka mått som ska ingå i leveransen.

6 Fortsättning

Föreliggande projekt är den första av totalt tre etapper. Under etapp 2 kommer den nya BAST-blanketten att testas i samarbete med Enheten för Mätteknik. Etapp 2 startar i slutet av februari 2009 och pågår i tre månader till slutet av maj. I etapp 3, som genomförs under hösten 2009, kommer förändringarna att fastställas inom metod, urval och ram inför dragningen av urval till år 2010.

Metodutredningen kommer att intensifieras under etapp 2. Metodöversynen som har genomförts under etapp 1 pekar ut nedanstående områden för vidare utredning:

- utreda storleken på cutoff-andelar i undersökningen
- kartlägga bristerna i rampopulationen
- studera om stratifieringen och allokeringen blir mer effektiv om undersökningen delas upp på två eller flera undersökningar
- undersöka hur hanteringen av få förekomster kan förbättras för att åstadkomma mer robusta skattningar över tiden
- klargöra om systemet i praktiken utför den imputeringsmetod som beskrivs i dokumentationen. Eventuellt kan metoden utvärderas för att studera hur den fungerar i de olika bortfallssituationerna
- ta fram kriterier för att bestämma en outlier
- utreda om det går att hitta en estimator som ger stabilare skattningar över tiden
- studera om skattningarna blir mer stabila över tiden genom att slå samman stratum genom att testa på historiska data