statisticon
STATISTICS & RESEARCH

Företagens villkor och verklighet 2020 – Coronakrisens effekter

Beskrivning av genomförande och metoder

Teknisk rapport

Version 1

Datum: 2021-01-08

Klass: -

Statisticon AB Östra Ågatan 31, SE-753 22 Uppsala Klara Södra Kyrkogata 1, SE-111 52 Stockholm +46 (0)10 130 80 00, info@statisticon.se, www.statisticon.se Sweden

Kund	Tillväxtverket
Mottagare	Olof Dunsjö, Olle Grünewald
Projektnummer	TVV_6355
Författare	Mats Nyfjäll
Granskare	Pernilla Tollin

Innehållsförteckning

1 Bakgrund och inledning 5

2 Innehåll och utformning 6

2.1 Objekt och population 6
2.2 Ramförfarande 7
2.3 Urval 7
2.4 Variabler 7
2.5 Tidsperioder i statistiken 8
2.6 Statistikens presentation 8

3 Datainsamling 9

3.1 Mätmetod och mätinstrument 9
3.2 Svarsinflöde och svarsandelar 9
3.3 Bortfallsanalys 12

4 Databearbetningar 17

4.1 Kontroller, rättningar och imputeringar 17
4.2 Härledda variabler 20

5 Estimationsförfarande 21

5.1 Parameter/målstorheter 21
5.1.1 Domäner (redovisningsgrupper) 22
5.2 Urvalsdesign 22
5.3 Punktestimator 23
5.3.1 Bortfallskompensation – kalibrering 24
5.4 Variansestimator 29

6 Osäkerhetskällor 30

6.1 Urval 30
6.2 Mätning 31
6.3 Bortfall 31
6.4 Ramosäkerhet 32
6.5 Bearbetning 32
6.6 Modellantaganden 32

7 Databaser 34

7.1 Enkätdata med registervariabler 34

7.2 Kalibreringstotaler 36

7.3 Urvalsfil från SCB 37

7.4 Format-fil 37

Referenser 38

Bilageförteckning 39

1 BAKGRUND OCH INLEDNING

Företagens villkor och verklighet (Fvov) är en enkätundersökning som riktar sig till små och medelstora företag i Sverige. Undersökningen har genomförts sju gånger tidigare, vart tredje år sedan 2002.

Den senaste undersökningen (Fvov 2020) genomfördes av SCB på uppdrag av Tillväxtverket mellan november 2019 och januari 2020. Det vill säga strax innan coronakrisen slog till i Sverige. I och med coronakrisen finns det ett behov av att titta på hur svaren på vissa utvalda frågor, som ställts tidigare, har förändrats och med krisen. Men det finns också ett behov av att ställa helt nya frågor, som inte ställts tidigare, som varar på olika typer av frågeställningar kopplat till corononakrisen.

Tillväxtverket upphandlade utförandet av den nya *Fvov – Coronakrisens effekter* i augusti 2020. Statisticon AB vann upphandlingen och har under hösten 2020 genomfört undersökningen. Undersökningen har i mångt och mycket genomförts på samma sätt och med samma avgränsningar av målpopulationen som den ordinarie Fvov. Det som skiljer är följande:

- Urvalsstorleken i den ordinarie Fvov är omkring 30 000 företag. I *Fvov – Coronakrisens effekter* var urvalsstorleken 2 500 företag.

Datainsamlingen har genomförts från mitten av oktober till och med första veckan i december. Av 2 500 utvalda företag har 969 svarat, vilket motsvarar en svarsandel på ca 39 procent.

Detta dokument utgör en teknisk beskrivning av undersökningen. I kapitel 2 *Innehåll och utformning* beskrivs undersökningens objekt, målpopulationen och ramförfarandet. Urval och urvalsmetod beskrivs översiktligt här. I kapitel 3 *Datainsamling* beskrivs mätmetoden och mätinstrumentet. Även svarsinflöde och svarsandelar beskrivs samt en bortfallsanalys görs. I kapitel 4 *Databearbetningar* beskrivs vilka databearbetningar som genomförts, exempelvis rättningar och rättningsprinciper. I kapitel 5 beskrivs estimationsförfarandet (skattningsförfarande) i tekniska termer. Parametern formuleras och urvalsdesignen beskrivs i tekniska termer. I kapitel 6 *Osäkerhetskällor* diskuteras undersökningens osäkerhetskällor. I kapitel 7 *Databaser* beskrivs vilka databaser som Statisticon levererar till Tillväxtverket.

2 INNEHÅLL OCH UTFORMNING

2.1 Objekt och population

I kvalitetsdeklarationen i den ordinarie undersökningen Fvov förs resonemang kring intressepopulation. Vi hänvisar till kvalitetsdeklarationen för den intresserade.

Målobjektet i undersökningen *Fvov – Coronakrisens effekter* är företag (juridisk enhet). Observationsobjekt, eller uppgiftslämnare, i undersökningen är funktionen som enkäten ställs till vilket är ”VD, företagsledare eller motsvarande”.

Målpopulationen avgränsas främst i termer av bransch och antal anställda. I tabell 1 redovisas vilka branscher, på avdelningsnivå enligt SNI 2007, som ingår i målpopulationen för. I termer av antal anställda ingår samtliga företag med **0-249 anställda.**

Tabell 1. Branscher (avdelningsnivå) enligt SNI 2007 som ingår i målpopulationen för *Fvov – Coronakrisens effekter*

Avdelning	Aktivitetsart
A	Jordbruk, skogsbruk och fiske
B	Utvinning av mineral
C	Tillverkning
D	Försörjning av el, gas, värme och kyla
E	Vattenförsörjning; avloppsrening, avfallshantering och sanering
F	Byggverksamhet
G	Handel; reparation av motorfordon och motorcyklar
H	Transport och magasinering
I	Hotell- och restaurangverksamhet
J	Informations- och kommunikationsverksamhet
M	Verksamhet inom juridik, ekonomi, vetenskap och teknik
N	Uthyrning, fastighetsservice, resetjänster och andra stödtjänster
P	Utbildning
Q	Vård och omsorg; sociala tjänster
R	Kultur, nöje och fritid
S	Annan serviceverksamhet

Förutom avgränsningen i termer av bransch och antal anställda finns det ytterligare avgränsningsvillkor rörande målpopulationen. Nedan sammanfattas samtliga avgränsningsvillkor:

- Företaget ska vara en aktiv juridisk enhet enligt FDB i augusti 2020 (augusti-SAMU)
- Följande juridiska former (JurForm) ingår
    - Personliga företag (JurForm = 10)
    - Handels- och kommanditbolag (JurForm = 31)
    - Aktiebolag utom banker och försäkringsbolag (JurForm = 49)

  - Ekonomiska föreningar (JurForm = 51)
- Företaget ska vara svenskkontrollerad (variabeln AgKat={41, 42} i företagsdatabasen) med begränsning till endast ett företag per koncern
- Företaget ska vara verksamt inom någon av branscherna A–J, M (exklusive M7010), N, P–S enligt SNI 2007, se tabell 1.
- Företag ska ha 0 till högst 249 anställda (Anst)
- Företaget ska ha en omsättning på minst 400 000 kr

2.2 Ramförfarande

Som urvalsram för målpopulationen användes företagsdatabasen (FDB) på SCB (version 2020 augusti-SAMU). SCB upprättade urvalsramen baserat på en specifikation från Tillväxtverket/Statisticon. Urvalsramen bestod av 347 443 företag. Som en jämförelse kan nämnas att när den ordinarie Fvov 2020 genomfördes bestod ramen av 356 201 företag.

Urvalsramen stratifierades på antal anställda (4 grupper) och bransch (13 grupper) till 52 strata. Se vidare avsnitt 5.2 där urvalsdesignen beskrivs i mer detaljer.

2.3 Urval

Urvalsstorleken var 2 500 företag och drogs via ett obundet slumpmässigt urval (OSU) inom respektive stratum. SCB genomförde urvalsdragningen. Fördelningen av de 2 500 företagen över strata var likformig allokering, dvs. lika många utvalda inom varje stratum. Detta innebar 48 utvalda företag inom varje stratum. Ett fåtal undantag kring denna princip fanns, se avsnitt 5.2 för mer detaljer.

Som en jämförelse kan nämnas att i den ordinarie Fvov 2020 var urvalsstorleken 28 540 företag.

2.4 Variabler

I och med coronakrisen finns det ett behov av att titta på hur svaren på vissa utvalda frågor, som ställts tidigare, har förändrats i och med krisen. Men det finns också ett behov av att ställa helt nya frågor, som inte ställts tidigare, som svarar på olika typer av frågeställningar kopplat till corononakrisen. Frågorna i undersökningen bestämdes helt och hållet av Tillväxtverket.

I bilaga 1 (separat fil) redovisas enkäten i sin helhet. Där framgår vilka variabler som efterfrågades i undersökningen.

2.5 Tidsperioder i statistiken

Tidsperioden varierar mellan olika frågor i undersökningen. Vissa frågor är tillbakablickande och rör tiden under coronapakrisen, dvs. i stora drag våren och hösten 2020. Vissa frågor rör jämförelser mellan 2019 och 2020. Vissa frågor är framåtblickande och rör framtiden i både kort och medellångt perspektiv (upp till tre år).

I bilaga 1 redovisas enkäten i sin helhet och där framgår tidsperioden för respektive fråga.

2.6 Statistikens presentation

Statisticon har inte producera några skattningar i undersökningen. Alla resultat kommer Tillväxtverket att själva ta fram. Statisticon levererar en databas med samtliga enkätfrågor, design- och registervariabler samt uppräkningsvikt till Tillväxtverket. Baserat på denna databas kan Tillväxtverket själv producera de skattningar som önskas. I kapitel 7 beskrivs databasen mer i detalj.

3 DATAINSAMLING

3.1 Mätmetod och mätinstrument

Datainsamlingen genomfördes genom att en inbjudan till undersökningen skickades ut postalt. En pappersenkät följde med i huvudutskicket. I bilaga 1 framgår enkäten med de frågor som ställdes. I inbjudan framgick att enkäten också kunde besvaras via en webbenkät. En unik inloggningsskod till webbenkäten framgick av missivet.

I bilaga 2-4 framgår missiven för huvudutskick och påminnelseutskick.

I tabell 2 framgår viktiga datum under datainsamlingen. Det kan nämnas att Statisticon genomförde ett ytterligare begränsat påminnelsearbete via telefon efter påminnelse 2 (som var postal). Denna insats riktades mot strata med minst antal svar[1].

Tabell 2. Viktiga datum under datainsamlingen 2020

Datum	Moment
15-okt	Enkäten skickas ut
04-nov	Påminnelse 1
27-nov	Påminnelse 2
13-dec	Datainsamlingen avbröts

Statisticon anlitade företaget Origo Group för utsändning, insamling och skanning av pappersenkäterna. Statisticon ansvarade för programmeringen av webbenkäten. Inför påminnelserna skickade Statisticon en fil till Origo Group med de företag som hade svarat på webben för att dessa företag inte skulle påminnas. Efter avslutad datainsamling skickade Origo Group en datafil till Statisticon med inskannade enkätsvar. Dessa svar, tillsammans med svaren på webbenkäten, bearbetade Statisticon till en final databas, se vidare kapitel 7.

3.2 Svarsinflöde och svarsandelar

Sammanlagt var det 969 företag som svarade på enkäten. I tabell 3 redovisas fördelningen över de avprickningskoder som användes för att kategorisera inflödet. En komplett förteckning över avprickningskoder finns i bilaga 5.

[1] Telefonpåminnelser var inte med som påminnelsemetod i upphandlingen men genomfördes ändå av Statisticon (i begränsad omfattning).

Tabell 3. Antal företag per avprickningskod

Kategori	Avprickningskod	Förklaring	Antal	Andel (%)
Svar	10	Svar	969	38,8
Okänd status	31	Ej anträffad (information saknas)	1 489	59,6
Övertäckning	43	Företaget har upphört/konkurs	1	0,0
Bortfall	213	Frivilligheten/vill ej svara	1	0,0
Bortfall	215	Tycker att man inte passar in i målgruppen	1	0,0
Bortfall	230	Postretur (tillhör målpopulationen)	35	1,4
Bortfall	235	Språksvårigheter	1	0,0
Bortfall	236	Oanvändbara svar	3	0,1
		Summa	2 500	100,0

Förutom 969 svar var den vanligaste avprickningskoden okänd status. Ingen information hördes från 1 489 företag. Detta är ett bortfall av svar, men eftersom ingen information kring företagen finns klassas deras status som okänd. I ett fall identifierades ett företag som hade gått i konkurs och alltså utgjorde övertäckning. Vidare var det ett 40-tal företag som klassades som bortfall.

När svarsandelen i undersökningen ska beräknas behöver man normalt sett ta hänsyn till en potentiell mängd övertäckning bland företag med okänd status. Ett vanligt antagande är att anta att övertäckningen bland de företag med okänd status är samma som den observerade övertäckningen. Men eftersom den observerade övertäckningen var mycket begränsad (endast ett företag) så påverkas svarsandelen inte nämnvärt av den observerade övertäckningen. Den ovägda svarsandelen kan därför beräknas till 38,8 procent, vilket framgår av tabell 3.

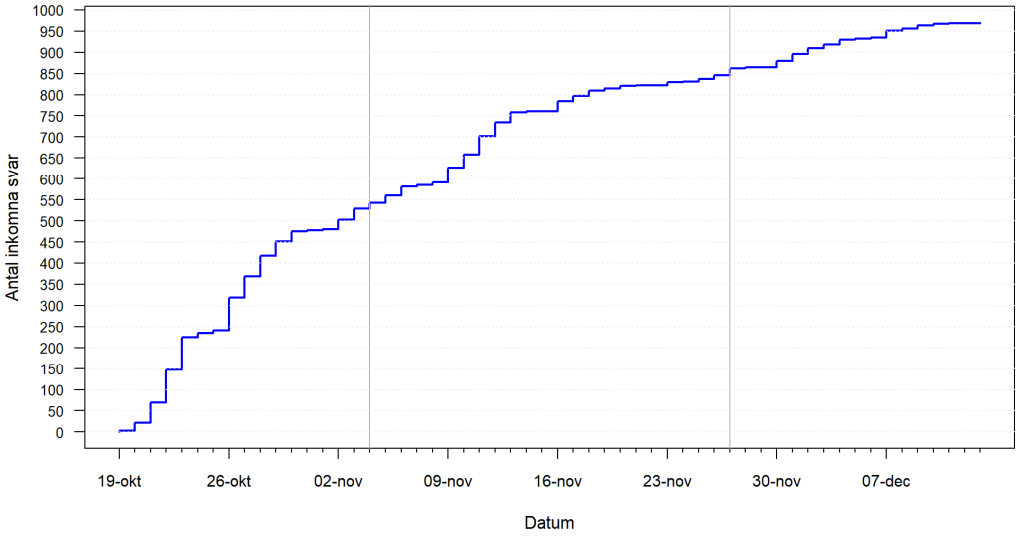
Den vägda svarsandelen, när hänsyn tas till olika urvalssannolikhet, är 34,6 procent. Detta är en skattning av vilken svarsandel som hade erhållits om en totalundersökning hade genomförts, och om benägenheten att svara hade varit samma bland samtliga företag som bland de 2 500 utvalda. Anledningen till att detta mått är lägre än den ovägda svarsandelen är för att svarsandelen är lägre bland små företag och om en totalundersökning hade genomförts hade små företag blivit representerade i högre grad än i urvalet.

I valet mellan den vägda och ovägda svarsandelen som mått är den ovägda ett bättre mått på hur datainsamlingen faktiskt har fungerat. I kvalitetsdeklarationen för den ordinarie Fvov redovisas den ovägda svarsandelen. I Fvov 2020 var svarsandelen 33,5 procent vilket kan jämföras med 38,8 procent i *Fvov – Coronakrisens effekter*.

Svarsandelar per olika kategorier av registervariabler och stratumvariabeln redovisas i avsnitt 3.3 Bortfallsanalys.

I figur 1 redovisas inflödet av svar kumulativt över den tidsperiod datainsamlingen pågick. Det framgår att det var ett ganska jämnt inflöde av svar efter huvudutskicket. Därefter avtog svarsbenägenheten. De två grå lodräta linjerna markerar datum när påminnelserna genomfördes. Det syns inget tydligt mönster att antalet svar ökade i någon nämnvärd utsträckning i samband med påminnelserna. Det sist inkomna svaret var den 13 december. Efter detta datum beaktades inga eventuellt ytterligare inkomna svar.

Figur 1. Kumulativt antal inkomna svar per dag. De två grå lodräta linjerna markerar datum för påminnelser.



Undersökningen kunde besvaras antingen via en pappersenkät eller via webbenkät. I tabell 4 redovisas antal och andel svar per svarssätt. Det svar som inkom via telefon var under telefonpåminnelserna och detta företag visade sig ha gått i konkurs (övertäckning). Det framgår att det var ungefär lika stor andel som svarade via pappersenkät som via webbenkät.

Tabell 4. Svarssätt i undersökningen

Svarssätt	Antal	Andel (%)
Postalt svar	494	50,9
Telefon	1	0,1
Webbsvar	475	49,0
Summa	970	100,0

3.3 Bortfallsanalys

Nedan presenteras ett antal tabeller med svarsandelar per olika kategorier för olika variabler. Via dessa underlag framgår det vilka grupper som har hög respektive låg benägenhet att svara. Den bortfallskompensation som genomförs via uppräkningsvikten har sin utgångspunkt i dessa analyser. För att enklare se skillnaderna mellan olika svarsandelar har den kolumnen färgkodats.

Förutom objektsbortfall förekommer även ett visst partiellt bortfall. Detta redovisas i avsnitt 4.1 i samband med imputeringar.

Några kommentarer tabellerna:

- Tabell 5 med svarsandel per stratum innehåller även variablerna antal företag totalt, antal utvalda och antal svarande. Det framgår att det är stor skillnad i benägenhet att svara mellan olika strata. Olikheter i svarsbenägenhet att svara mellan strata är dock något som (den kalibrerade) uppräkningsvikten justerar för, se vidare i kapitel 5 Estimationsförfarande.
- Tabell 8 visar stora skillnader i svarsandel per län. En förhoppning var att kunna kompensera för dessa olikheter via uppräkningsförfarandet. Det var dock för få svarande i vissa län, vilket gjorde en sådan bortfallskompensation olämplig. Istället användes Tillväxtverkets geografiska indelning, se vidare i kapitel 5 Estimationsförfarande.
- I tabell 9 och 10 redovisas svarsandel enligt Tillväxtverkets indelning i 3 regioner (tabell 9) och 6 regioner (tabell 10). Det framgår att skillnaderna i svarsandel mellan de tre regionerna inte är så stor. När man däremot delar in Sverige i 6 olika regioner blir det större skillnader i svarsandelar. Det vore därför önskvärt att genomföra en bortfallskompensation, via uppräkningsvikten, med avseende på denna geografiska indelning. Det visade sig dock inte vara lämpligt då vissa uppräkningsvikter blev negativa. Detta orsakades av att det endast var 13 företag i mycket avlägsna landsbygdskommuner som hade svarat, vilket var ett för litet underlag. Det var 65 företag i avlägset belägna landsbygdskommuner som hade svarat. En möjlighet var därför att slå ihop dessa två grupper, se nästa punkt.
- I tabell 11 redovisas svarsandelarna geografiskt efter att grupperna med mycket avlägset belägna och avlägset belägna landsbygdskommuner slagits samman. Detta leder till 5 grupper där svarsandelarna skiljer sig åt. Denna geografiska indelning har använts för att kompensera för bortfallet vid skapandet av uppräkningsvikten, se vidare i kapitel 5 Estimationsförfarande.
- I tabell 12 redovisas svarsandelarna per juridisk form. Det framgår att företag i den juridiska formen ”fysiska personer” har lägst svarsandel och ekonomiska föreningar har högst svarsandel. Eftersom det är en skevhet över svarsandelar med avseende på juridisk form är denna variabel en bra kandidat att använda för

bortfallskompensation i uppräkningsvikten. Dock var det så få svarande (tillika utvalda) handelsbolag och kommanditbolag att detta inte var möjligt. Däremot skulle en gruppering kunna göras, vilket framgår i nästa punkt.

- I tabell 13 redovisas svarsandelar per juridisk form efter att tre av grupperna slagits samman till en grupp. Det framgår att det även med två grupper är stora skillnader i svarsandel. Denna variabel, med denna indelning, användes därför vid framtagandet av uppräkningsvikt för att kompensera för skevheter i svarsbenägenhet.

Tabell 5. Svarsandel per stratum

Stratum	Förklaring	Antal företag			Svars-andel (%)
		Totalt	Urval	Svarande	
10	Jordbruk, skogsbruk och fiske. 0 anställda	25 769	49	16	32,7
11	Jordbruk, skogsbruk och fiske. 1-9 anställda	8 931	48	20	41,7
12	Jordbruk, skogsbruk och fiske. 10-49 anställda	498	65	35	53,8
13	Jordbruk, skogsbruk och fiske. 50-249 anställda	31	31	14	45,2
20	Utvinning av mineral, Tillverkning. 0 anställda	5 665	48	12	25,0
21	Utvinning av mineral, Tillverkning. 1-9 anställda	11 867	48	16	33,3
22	Utvinning av mineral, Tillverkning. 10-49 anställda	3 353	48	26	54,2
23	Utvinning av mineral, Tillverkning. 50-249 anställda	601	48	28	58,3
30	Energi, vatten, el. 0 anställda	668	48	26	54,2
31	Energi, vatten, el. 1-9 anställda	493	48	12	25,0
32	Energi, vatten, el. 10-49 anställda	131	77	35	45,5
33	Energi, vatten, el. 50-249 anställda	19	19	10	52,6
40	Byggverksamhet. 0 anställda	18 482	48	13	27,1
41	Byggverksamhet. 1-9 anställda	35 964	49	15	30,6
42	Byggverksamhet. 10-49 anställda	5 132	48	18	37,5
43	Byggverksamhet. 50-249 anställda	430	48	19	39,6
50	Handel. 0 anställda	14 232	48	10	20,8
51	Handel. 1-9 anställda	32 547	49	17	34,7
52	Handel. 10-49 anställda	4 659	48	20	41,7
53	Handel. 50-249 anställda	593	48	21	43,8
60	Transport och magasinering. 0 anställda	5 362	48	16	33,3
61	Transport och magasinering. 1-9 anställda	10 964	48	16	33,3
62	Transport och magasinering. 10-49 anställda	1 950	48	19	39,6
63	Transport och magasinering. 50-249 anställda	264	48	22	45,8
70	Hotell- och restaurangverksamhet. 0 anställda	3 244	48	11	22,9
71	Hotell- och restaurangverksamhet. 1-9 anställda	13 003	48	19	39,6
72	Hotell- och restaurangverksamhet. 10-49 anställda	2 622	48	19	39,6
73	Hotell- och restaurangverksamhet. 50-249 anställda	153	48	24	50,0
80	Informations- och kommunikationsverksamhet. 0 anställda	5 466	48	13	27,1
81	Informations- och kommunikationsverksamhet. 1-9 anställda	15 320	48	23	47,9
82	Informations- och kommunikationsverksamhet. 10-49 anställda	1 310	48	13	27,1
83	Informations- och kommunikationsverksamhet. 50-249 anställda	217	48	10	20,8
90	Juridik, ekonomi, vetenskap, teknik. 0 anställda	17 706	48	18	37,5
91	Juridik, ekonomi, vetenskap, teknik. 1-9 anställda	41 726	49	18	36,7

Stratum	Förklaring	Antal företag Totalt	Urval	Svarande	Svars-andel (%)
92	Juridik, ekonomi, vetenskap, teknik. 10-49 anställda	2 383	48	16	33,3
93	Juridik, ekonomi, vetenskap, teknik. 50-249 anställda	241	48	24	50,0
100	Uthyrning, fastighetsservice, resetjänster och andra stödtjänster. 0 anställda	4 490	48	8	16,7
101	Uthyrning, fastighetsservice, resetjänster och andra stödtjänster. 1-9 anställda	8 951	48	16	33,3
102	Uthyrning, fastighetsservice, resetjänster och andra stödtjänster. 10-49 anställda	1 759	48	17	35,4
103	Uthyrning, fastighetsservice, resetjänster och andra stödtjänster. 50-249 anställda	343	48	25	52,1
110	Utbildning. 0 anställda	2 104	48	14	29,2
111	Utbildning. 1-9 anställda	3 005	48	15	31,3
112	Utbildning. 10-49 anställda	234	48	26	54,2
113	Utbildning. 50-249 anställda	49	48	19	39,6
120	Vård och omsorg; sociala tjänster. 0 anställda	2 742	48	17	35,4
121	Vård och omsorg; sociala tjänster. 1-9 anställda	4 614	48	27	56,3
122	Vård och omsorg; sociala tjänster. 10-49 anställda	417	48	22	45,8
123	Vård och omsorg; sociala tjänster. 50-249 anställda	113	48	18	37,5
130	Kultur, nöje och fritid och annan serviceverksamhet. 0 anställda	14 093	48	14	29,2
131	Kultur, nöje och fritid och annan serviceverksamhet. 1-9 anställda	11 777	48	22	45,8
132	Kultur, nöje och fritid och annan serviceverksamhet. 10-49 anställda	692	48	20	41,7
133	Kultur, nöje och fritid och annan serviceverksamhet. 50-249 anställda	64	48	26	54,2
	Summa	347 443	2 500	970	38,8

Tabell 6. Svarsandel per bransch

Branschkod för stratumindelning	Förklaring	Svars-andel (%)
1	A Jordbruk, skogsbruk och fiske	44,0
2	BC Utvinning av mineral samt Tillverkning	42,7
3	DE Försörjning av el, gas, värme och kyla, samt Vattenförsörjning; avloppsrening, avfallshantering och sanering	43,2
4	F Byggverksamhet	33,7
5	G Handel; reparation av motorfordon och motorcyklar	35,2
6	H Transport och magasinering	38,0
7	I Hotell- och restaurangverksamhet	38,0
8	J Informations- och kommunikationsverksamhet	30,7
9	M Verksamhet inom juridik, ekonomi, vetenskap och teknik	39,4
10	N Uthyrning, fastighetsservice, resetjänster och andra stödtjänster	34,4
11	P Utbildning	38,5
12	Q Vård och omsorg; sociala tjänster	43,8
13	RS Kultur, nöje och fritid samt Annan serviceverksamhet	42,7

Tabell 7. Svarsandel per antal anställda

Storleksklass för stratumindelning	Förklaring	Svarsandel (%)
0	0 anställda	30,1
1	1-9 anställda	37,6
2	10-49 anställda	42,7
3	50-249 anställda	45,0

Tabell 8. Svarsandel per län

Länskod	Län	Svarsandel (%)
01	Stockholms län	35,4
3	Uppsala län	31,9
04	Södermanlands län	42,9
05	Östergötlands län	39,6
06	Örebro län	39,8
07	Västmanlands län	44,4
08	Jönköpings län	53,8
09	Kronobergs län	31,6
10	Kalmar län	27,6
12	Gotlands län	37,1
13	Blekinge län	41,7
14	Skåne län	44,9
17	Hallands län	42,2
18	Västra Götalands län	36,9
19	Värmlands län	30,0
20	Dalarnas län	43,5
21	Gävleborgs län	37,7
22	Västernorrlands län	37,1
23	Jämtlands län	46,9
24	Västerbotten län	44,6
25	Norrbottens län	33,3

Tabell 9. Svarsandel per geografisk indelning enligt Tillväxtverkets kommungruppsindelning. 3 nivåer.

Tillväxtverkets kommungruppsindelning (3 nivåer)		
Kod	Förklaring	Svarsandel (%)
1	Landsbygdskommuner	40,1
2	Täta kommuner	38,9
3	Storstadskommuner	38,2

Tabell 10. Svarsandel per geografisk indelning enligt Tillväxtverkets kommungruppsindelning. 6 nivåer.

Tillväxtverkets kommungruppsindelning (6 nivåer)		
Kod	Förklaring	Svarsandel (%)
10	Landsbygdskommuner mycket avlägset belägna	36,1
11	Landsbygdskommuner avlägset belägna	43,6
12	Landsbygdskommuner nära en större stad	38,5
21	Täta kommuner avlägset belägna	46,5
22	Täta kommuner nära en större stad	37,6
33	Storstadskommuner	38,2

Tabell 11. Svarsandel per geografisk indelning enligt Tillväxtverkets kommungruppsindelning. 5 nivåer.

Tillväxtverkets kommungruppsindelning (5 nivåer)		
Kod	Förklaring	Svarsandel (%)
1	Landsbygdskommuner avlägset + mycket avlägset belägna (grupperad)	42,2
2	Landsbygdskommuner nära en större stad	38,5
3	Täta kommuner avlägset belägna	46,5
4	Täta kommuner nära en större stad	37,6
5	Storstadskommuner	38,2

Tabell 12. Svarsandel per juridisk form

Juridisk form		
Kod	Förklaring	Svarsandel (%)
10	Fysiska personer	29,7
31	Handelsbolag, kommanditbolag	38,6
49	Övriga aktiebolag	40,5
51	Ekonomiska föreningar	64,3

Tabell 13. Svarsandel per juridisk form. Grupperad till 2 grupper.

Juridisk form kodad till 2 grupper		
Kod	Förklaring	Svarsandel (%)
1	10 Fysiska personer	29,7
2	31 Handelsbolag, kommanditbolag: 49 Övriga aktiebolag: 51 Ekonomiska föreningar	40,9

4 DATABEARBETNINGAR

Uppgiftslämnarna kunde svara antingen på pappersenkäten eller på webbenkäten. Respektive insamlingssätt gav upphov till två olika primärdatabaser. Dessa slogs samman och bearbetades till en final databas. Nedan beskrivs kortfattat olika moment som genomförts.

4.1 Kontroller, rättningar och imputeringar

Värdemängderna för samtliga variabler kontrollerades. I den postala insamlingen användes en kod 999 för felaktigt svar[2]. Dessa kodades om till saknat svar.

Vissa frågor har en logik att man exempelvis ska hoppa över efterföljande fråga om man svarat Ja på huvudfrågan. Webbenkäten programmerades på ett sådant sätt att inga inkonsistenta svarsmönster uppstod. I pappersenkäten kan dock uppgiftslämnare svara på ett inkonsistent sätt. Tillväxtverket och Statisticon beslutade i samråd om hur förekommande inkonsistenser skulle hanteras. Nedan redogörs för principerna.

Fråga A5

På fråga A5, se nedan, är det ologiskt att svara ”Nej” (sista svarsalternativet) tillsammans med något av de övriga svarsalternativen. Detta inträffade vid 5 tillfällen och dessa svar betraktades som ogiltiga, dvs. både Ja- och Nej-svaret blankades ut.

A5. Har företaget tagit del av någon form av stöd eller åtgärd i samband med coronapandemin? *Markera med kryss de stöd/åtgärder företaget tagit del av.*

- ☐ Korttidspermittering/korttidsarbete (via Tillväxtverket)
- ☐ Hyresstöd via din hyresvärd / Tillfällig rabatt för fasta hyreskostnader i utsatta branscher
- ☐ Omställningsstöd för företag med kraftigt minskad omsättning (via Skatteverket)
- ☐ Tillfällig nedsättning av arbetsgivaravgifter och egenavgifter (via Skatteverket)
- ☐ Skjutit upp betalning av skatt och/eller avgifter (via Skatteverket)
- ☐ Utökad avsättning till periodiseringsfond (via Skatteverket)
- ☐ Lånegaranti för små och medelstora företag (via bankerna/Riksgälden)
- ☐ Brygglån för små och medelstora företag (via Almi)
- ☐ Annat, vad? _____
- ☐ Nej, företaget har inte tagit del av stöd eller åtgärder

[2] Exempelvis två kryss angivna på en envalsfråga

Fråga D1

Om man svarar Nej på fråga D1 ska man hoppa över fråga D2. Efter samråd beslutades att låta fråga D1 styra. Alltså, om man svarat Nej på D1 samt något av svarsalternativet på D2 så blankades svaren på D2 ut. Motivet till detta förfarande bygger på tolkningen att företaget kanske har haft export till dessa länder (därav svaret på D2), men inte under 2020 (vilket var svaret på D1). 11 sådana rättningar genomfördes.

D1. Har företaget haft någon export under 2020?
Uppskatta exportens andel av företagets totala omsättning under 2020

- ☐ Ja, mindre än 5 %
- ☐ Ja, 5–25%
- ☐ Ja, 26–50%
- ☐ Ja, 51–100 %
- ☐ Nej --> **Gå direkt till fråga D3**

D2. Till vilken/vilka delar av världen exporterar företaget sina varor/tjänster?
Markera ett eller flera alternativ.

- ☐ Norden
- ☐ EU utom Norden
- ☐ Övriga Europa (inkl. Ryssland)
- ☐ Övriga världen

D3. Har företaget haft någon egen import under 2020?
Med egen import menas import som företaget genomfört direkt med företag i annat land, dvs. inte med återförsäljare i Sverige.

- ☐ Ja
- ☐ Nej

Fråga E1

Fråga E1 är en flervalsfråga, men det är ologiskt att svara både Ja och Nej. För de hade svarat på Ja och Nej blankades hela svaret på fråga E1 ut. 3 sådana rättningar gjordes.

E1. Har företaget utvecklat och sålt nya eller väsentligt förbättrade varor/tjänster under de senaste tre åren?

☐ Ja, varor ☐ Ja, tjänster ☐ Nej, varken varor eller tjänster

Flervalsfrågor

Flervalsfrågor är principiellt kodades så att om ett svarsalternativ är angivet antar variabeln koden 1 och 0 om svaret inte är angivet. Om en uppgiftslämnare inte har svarat på något svarsalternativ överhuvudtaget på flervalsfrågan betraktas detta som ett partiellt bortfall och en sådan uppgiftslämnare har missing (istället för 0) på alla svarsalternativ på flervalsfrågan. På exempelvis fråga A5 var det 13 personer som hoppade över hela frågan. Dessa är alltså kodade som missing på denna fråga.

Svarat både postalt och via webb

Det var sju företag som svarade *både* postalt och via webben, alltså två svar från samma företag. Endast ett svar per företag ska beaktas. Principen som användes var att ta det svar där flest frågor var besvarade. Om lika antal frågor var besvarade valdes det senast inkomna svaret. De som svarade på båda sätten hade svarat på nästan samtliga frågor. Principen med senast inkomna användes i ett av fallen. I de övriga fallen skiljde sig antalet svar.

Imputeringar – partiellt bortfall

Det förekommer ett visst partiellt bortfall. Om vi bortser från vissa enskilda svarsalternativ på flervalsfrågor (t.ex. alternativet ”Annat” på fråga A5) är det frågorna D4 a och b som har störst partiellt bortfall. Det är drygt 50 företag som hoppat över dessa två frågor, vilket motsvarar drygt 5 procent av alla svar.

D4. Hur har följande förändrats gentemot 2019?

	Minskat	Oförändrad	Ökat	Har ej haft någon
Företagets export?	☐	☐	☐	☐
Företagets import?	☐	☐	☐	☐

Därefter är det fråga B1b med ett partiellt bortfall på 28 företag, motsvarande knappt 3 procent, som har högst partiellt bortfall.

B1. Hur har följande påverkats för företaget under 2020, jämfört med 2019?
Markera ett kryss på varje rad

	Minskat	Oförändrat	Ökat
a) Omsättning?	☐	☐	☐
b) Antal anställda?	☐	☐	☐
c) Lönsamhet?	☐	☐	☐

Överlag är omfattningen på det partiella bortfallet relativt begränsat i undersökningen.

Det partiella bortfallet har inte hanterats via någon form av imputering. Det partiella bortfallet framgår i databasen som missingvärden för respektive variabel.

4.2 Härledda variabler

Den finala databasen består av 154 variabler. Det är både enkätvariabler, registervariabler och designvariabler (stratum, urvalsstorlekar m.m.). I kapitel 7 kommenteras databasen mer i detalj. I detta avsnitt nämns kortfattat några av de härledda variablerna.

Kalvikt

Variabeln Kalvikt är den uppräkningsvikt som ska användas vid framtagande av skattningar av deskriptiva målstorheter/parametrar. Hur uppräkningsvikten är framtagen beskrivs detaljerat i kapitel 5 Estimationsförfarande. I kapitel 7 ges ytterligare beskrivningar.

Svar

Variabeln svar är en indikator som antar värdet 1 för de som svarat och 0 för de som inte svarat. Hela databasen består av samtliga 2500 utvalda så variabeln svar identifierar de svarande. Notera att det företag som utgör övertäckning (ett sådant förekom) betraktas som svarande och har alltså svar = 1[3].

Source

Variabeln source är en textvariabel som indikerar om svaret är ett postalt svar eller svar via webb.

Datapoints

Variabeln datapoints är en numerisk variabel som anger hur många frågor en uppgiftslämnare har svarat på. Öppna frågor ingår inte i beräkningen. På flervalsfrågor ingår samtliga svarsalternativ i beräkningen.

[3] Det är företag med username = 3324 som utgör övertäckning.

5 ESTIMATIONSFÖRFARANDE

Detta kapitel är mer tekniskt till sin karaktär än föregående kapitel. Statisticon har inte producerat några skattningar i undersökningen. Istället har en uppräkningsvikt skapats som möjliggör skattningar för målstorheter i målpopulationen. I korta ordalag kan uppräkningsvikten beskrivas enligt:

- Uppräkningsvikten utgår från urvalssannolikheten inom varje stratum. Därefter har den kalibrerats för att justera för potentiell skevhet orsakad av bortfallet. Som kalibreringsvariabler har stratifieringsvariablerna använts (korsning av bransch och antal anställda, 52 strata), och en geografisk indelning samt juridisk form. För geografisk indelning och juridisk form sker kalibrering mot marginalfördelningen. Den geografiska indelningen baseras på Tillväxtverkets 6 regioner där de två grupperna avlägset samt mycket avlägset belägna landsbygdskommuner slagits samman till en grupp. Juridisk form har grupperats till två grupper: (i) Personliga företag (JurForm = 10) samt (ii) övriga företagsformer (Jurform = {31, 49, 51})

Nedan beskrivs detaljer kring hur uppräkningsvikten har skapats samt teoretiskt ramverk för skattningsförfarandet.

5.1 Parameter/målstorheter

De deskriptiva parametrar (eller målstorheter) som undersökningen främst ska skatta är av typen totaler eller kvoten mellan två totaler, t.ex. en procentandel.

Låt mängden av alla företag i målpopulationen betecknas $U = \{1,2, \dots, k, \dots, N\}$ där N är populationsstorleken och k är löpindex. Låt y representera en undersökningsvariabel (till fråga B4 *Vilka åtgärder har företaget vidtagit under coronakrisen?*) och y_k värdet för objekt (företag) k.

Vi formulerar de parametrar som är av intresserade av. En total i populationen ges av

$$t = \sum_{k \in U} y_k \tag{1}$$

De flesta variabler i enkäten är av karaktären skalfrågor, antingen mätta på nominalskale- eller ordinalskalenivå. Exempelvis skalan ”Inte alls”, ”I viss grad” och ”I stor grad” på fråga B4 *Vilka åtgärder har företaget vidtagit under coronakrisen?* a) *Etablerat nya försäljningskanaler…* För denna typ av variabler kan exempelvis svarsalternativet ”I stor grad” kodas som 1 och övriga alternativ kodas som 0. Med en sådan kodning blir totalen t enligt (1): *totalt antal företag* i populationen som svarar ”I stor grad”.

En annan viktig parameter är

$$\bar{y}_U = \frac{t}{N} = \frac{\sum_{k \in U} y_k}{N} \tag{2}$$

dvs. populationsmedelvärdet. Notera att om en skalfråga kodas med 1 för ett visst skal-alternativ och 0 i övrigt blir populationsmedelvärdet enligt (2) populations*andelen* som svarar detta svarsalternativ, t.ex. andelen ”I stor grad” på fråga B4 a). Eftersom andelar kan formuleras genom att koda skal-alternativen 1 och 0 på lämpligt sätt behövs ingen separat notation för procentandelar.

5.1.1 Domäner (redovisningsgrupper)

Tillväxtverket kommer att ta fram skattningar för olika redovisningsgrupper. Formellt sett hanteras sådana skattningar inom ramverket för domänskattningar.

Låt U_d beteckna mängden av alla företag i populationen i redovisningsgrupp (domän) d, t.ex. alla företag i Skåne med fler än 100 anställda. Låt N_d beteckna antalet företag i domänen. Parametern för redovisningsgrupper bildas genom att forma variabeln

$$y_{dk} = \begin{cases} y_k & \text{om } k \in U_d \\ 0 & \text{i övrigt} \end{cases} \tag{3}$$

och använda variabeln y_{dk} i uttryck (1) och (2) ovan. Vi har alltså

$$\bar{y}_{U_d} = \frac{t_d}{N_d} = \frac{\sum_{k \in U} y_{dk}}{N_d} \tag{4}$$

5.2 Urvalsdesign

Urvalsramen stratifierades på $H = 52$ strata, $h = 1,2, \dots, H$. Låt $U_h = \{1,2, \dots, k, \dots, N_h\}$ beteckna mängden av företag inom stratum h där N_h betecknar totalt antal företag. Alla stratummängder sammantaget bilder populationsmängden, dvs. $U = \bigcup_{h=1}^{H} U_h$.

Inom varje stratum drogs ett obundet slumpmässigt urval (OSU) $s_h = \{1,2, \dots, n_h\}$ där n_h betecknar urvalsstorleken. Den totala urvalsstorleken $n = \sum_{h=1}^{H} n_h$ var 2 500 företag. I tabell 5 redovisas stratumindelningen och antal företag per stratum, antal utvalda samt antal svarande.

Allokeringen, som framgår av tabell 5, gjordes likformigt över samtliga 52 strata. Detta innebär $\frac{2500}{52} = 48$ företag per stratum initialt. I två strata var dock N_h mindre än 48

(stratum 13 och 33) så där genomfördes en totalundersökning. Eftersom färre företag då valdes ut kunde urvalsstorleken ökas något i andra strata. Principen som användes var att öka antalet utvalda inom samma bransch. I stratum 12 valdes 65 företag ut, vilket medförde att bransch 1 (jordbruk, skogsbruk och fiske) fick lika många utvalda företag som övriga branscher. Motsvarande gäller för bransch 3 (energi, vatten och el) och stratum 32 där 77 företag valdes ut.

För att den totala urvalsstorleken skulle bli 2 500 och 2 500 är inte jämnt delbart med 48 ökades urvalsstorleken till 49 i ett fåtal strata. Detta gjordes i strata 10, 41, 51 och 91. I tabell 5 framgår samtliga urvalsstorlekar.

5.3 Punktestimator

Urvalssannolikheten (inklusionssannolikheten) inom varje stratum ges vid OSU av $\pi_k = n_h/N_h$. Inversen av urvalssannolikheten ger designvikten enligt $d_k = N_h/n_h$. Denna designvikt (uppräkningsvikt) kan användas för att skapa punktestimatorer av typen Horvitz-Thompson-estimatorer eller kvoter mellan två sådana estimatorer, se exempelvis Särndal m.fl. (1992). Vi presenterar först hur estimatorn ser ut om samtliga svarar, därefter presenteras estimatorn när hänsyn tas till bortfall. Estimator för totalen t ges av

$$\hat{t}_d = \sum\nolimits_{k \in s} d_k y_{dk} \tag{5}$$

Notera att vi utgår från variabeln y_{dk} för att möjliggöra skattningar inom redovisningsgrupper. Vi kan betrakta skattningen för hela populationen som ett specialfall av en redovisningsgrupp, dvs. när $y_{dk} = y_k$ för alla företag i urvalet s.

Estimatorn kan skrivas mer explicit, genom att utveckla designvikten, på följande sätt

$$\hat{t}_d = \sum_{h=1}^{H} \frac{N_h}{n_h} \sum\nolimits_{k \in s_h} y_{dk} \tag{6}$$

På motsvarande sätt kan punktestimatorn för populationsmedelvärdet (och populationsandelen) formuleras via

$$\hat{\bar{y}}_{U_d} = \frac{\hat{t}_d}{\hat{N}_d} = \frac{\sum_{k \in s} d_k y_{dk}}{\sum_{k \in s} d_k I_{dk}} \tag{7}$$

där I_{dk} är en 0-1 indikator för om företag k tillhör redovisningsgrupp d enligt

$$I_{dk} = \begin{cases} 1 & \text{om } k \in U_d \\ 0 & \text{i övrigt} \end{cases} \tag{8}$$

Notera att för skattning av hela populationen har vi följande likhet

$$\widehat{N} = \sum_{k \in s} d_k I_{dk} = \sum_{k \in s} d_k \times 1 = \sum_{k \in s} \frac{N_h}{n_h} \times 1 = \sum_{h=1}^{H} \frac{N_h}{n_h} \sum_{k \in s_h} 1 = N \tag{9}$$

För skattningar i redovisningsgrupper är det vanligare att $\widehat{N}_d \neq N_d$. I en sådan situation är det bättre (ur varianssyvinkel) att använda skattningen $\widehat{N}_d$ i nämnaren.

Anmärkning: om man använder en statistisk mjukvara, t.ex. SAS, som kan hantera uppräkningsvikter på ett korrekt sätt och beräknar medelvärdet är det uttryck (7) som används[4].

5.3.1 Bortfallskompensation – kalibrering

Bortfall har inträffat i undersökningen. Låt $r_h = \{1,2, \dots, m_h\}$ beteckna svarandemängden (respons set) i stratum h där m_h betecknar antalet svarande. Hela svarsmängden ges av $r = \cup_{h=1}^{H} r_h$ och antalet svarande totalt av $m = \sum_{h=1}^{H} m_h$. En tänkbar estimator för att ta hänsyn till bortfallet skulle vara att genomföra s.k. rak uppräkning inom strata. Vid denna justering skapas en bortfallskompenserande vikt $b_k = n_h / m_h$. Denna vikt tillsammans med designvikten ger uppräkningsvikten

$$w'_k = d_k \times b_k = \frac{N_h}{n_h} \times \frac{n_h}{m_h} = \frac{N_h}{m_h} \tag{10}$$

Vikten w'_k kallas vanligtvis för rak uppräkning inom strata, vilket beror på att de svarande i ett stratum betraktas som om de vore de utvalda. I Tillväxtverkets upphandlingsunderlag specificerades inte närmare hur uppräkningsvikten skulle konstrueras, endast att en uppräkningsvikt skulle beräknas.

I Statisticons offert beskrevs att ett möjligt sätt att kompensera för bortfallet var via rak uppräkning inom strata. Denna kompensation skulle då kompensera för eventuella skevheter bland de svarande företagen över bransch och antal anställda, eftersom dessa variabler användes som stratumvariabler. Det finns, som visats i avsnitt 3.3

[4] I exempelvis proc surveymeans eller proc surveyfreq i SAS används denna estimator.

Bortfallsanalys, skevheter över andra variabler. Det är därför önskvärt att också kompensera för dessa skevheter för att på så sätt minska risken för skevheter i skattningarna. Att kompensera för många variabler samtidigt genom att korsklassificera dem leder snabbt till problem med celler med för få svarande. En metod, vetenskapligt beprövad och mycket använd av bland andra SCB, är tekniken med kalibrering av vikter. Via den metoden kan man kompensera för exempelvis korsningen av bransch och antal anställda (stratum) och fördelningen geografiskt. Denna kompensation kan göras på flera sätt. Om man korsklassificera den geografiska variabeln inom varje stratum får man det problem som nämnts med få svarande i enskilda celler. Tekniken möjliggör dock att kompensation kan ske genom korsklassificeringen av bransch och antal anställda *samtidigt* som kompensation sker för *marginalfördelningen* över geografi.

Om man gör en sådan bortfallskompensation kommer de svarande att erhålla ”rätt” uppräkningsvikter i förhållande till två aspekter: (i) stratumindelningen och (ii) den geografiska indelningen. Med rätt avses här att om man för de svarande tar fram en frekvensfördelning över stratum med uppräkningsvikten får man det sanna populationsantalet inom varje stratum. På samma sätt, om man tar en *vägd* frekvensfördelning över den geografiska indelningen erhålls det sanna populationsantalet inom varje geografisk grupp.

Det finns en omfattande teoribyggnad för kalibreringsestimatorn. En bra referens är Lundström och Särndal (2001) som är utgiven som en CBM (Current Best Methods) på SCBs förlag.

De registervariabler som finns till förfogande är hjälpvariabler vilka kan användas i ett kalibreringsförfarande. I Lundström och Särndal (2001) anges tre olika villkor för att kalibreringsansatsen ska vara framgångsrik för att reducera bortfallsskevheten

(i) Hjälpvariabeln eller hjälpvariablerna ska samvariera med svarsbenägenheten (svarssannolikheten). Detta är det viktigaste kriteriet eftersom det leder till en minskning av bortfallsskevheten för samtliga skattningar.
(ii) Hjälpvariabeln eller hjälpvariablerna ska samvariera med viktiga målvariabler (frågor i enkäten). Om detta kriterium uppfylls minskar dels bortfallsskevheten dels osäkerheten (variansen) för dessa målvariabler.
(iii) Hjälpvariabeln eller hjälpvariablerna ska identifiera de viktigaste redovisningsgrupperna. Det leder framför allt till minskad osäkerhet (varians) i skattningarna för dessa redovisningsgrupper.

Via bortfallsanalysen i avsnitt 3.3 framgår att de svarande är skevt fördelade över variablerna:

- Stratum (bransch och antal anställda)

- Geografi: främst över län, men även över Tillväxtverkets indelning i 6 regioner
- Juridisk form

Statisticon genomförde ett flertal olika försök att skapa kalibreringsvikter som tog hänsyn till skevheter över dessa variabler. Variabeln stratum (som avser bransch och antal anställda) var med i samtliga varianter. Nedan redogörs för hur variablerna geografi och juridisk form har hanterats.

Geografi
Rörande geografi genomfördes försök att kalibrera gentemot län men det visade sig att de var för få svarande i vissa län, vilket medförde att uppräkningsvikten i vissa fall blev negativ. Detta är inte acceptabelt.

Därför fick ambitionsnivån för geografisk bortfallskompensation sänkas något från län till Tillväxtverkets geografiska indelning. Det gick utmärkt att använda indelningen i 3 grupper i kalibreringen. Däremot, när 6-gruppsindelningen provades var det för få svarande i framförallt gruppen *mycket* avlägset belägna kommuner (13 svarande). Detta ledde till att vissa företag fick negativa uppräkningsvikter. Att genomföra bortfallskompensation med avseende på indelningen i 3 grupper är dock inte helt optimalt eftersom det inte är så stor skillnad i svarsandelar mellan de tre grupperna. En kompromiss prövades därför genom att slå samman grupperna

- Mycket avlägset belägna kommuner, och
- Avlägset belägna kommuner

till en och samma grupp. Då erhålls 5 geografiska grupper. Svarsandelarna för dessa 5 grupper redovisas i tabell 11. Där framgår att det är skillnader i svarsandelar mellan dessa 5 grupper. Vid kalibrering mot denna geografiska indelning med 5 grupper fungerade kalibreringen tillfyllest.

Juridisk form
I tabell 12 framgår att det är stora skillnader i svarsandelar mellan de fyra olika juridiska former som ingår i undersökningen. Tyvärr var det för få svarande ibland handelsbolag och kommanditbolag (27 svar) och ekonomiska föreningar (också 27 svar) för att en denna variabel skulle kunna användas för bortfallskompensation. För att ändå kunna använda juridisk form, där stora skillnader i svarsandelar finns, slogs följande tre grupper samman

- Handelsbolag, kommanditbolag
- Övriga aktiebolag
- Ekonomiska föreningar

till en grupp. I tabell 13 redovisas svarsandelarna för denna grupp samt för den resterande gruppen Fysiska personer. Det framgår att det är stor skillnad i svarsbenägenhet och att denna variabel därför är en bra kandidat att använda i kalibreringen. Denna variabel, med 2 grupper, ledde inte till några problem i kalibreringen.

Sammanfattningsvis kan de variabler som används för att skapa uppräkningsvikt som kompenserar för bortfallet beskrivas enligt

- **Bransch × antal anställda (52 grupper) + geografisk indelning (5 grupper) + juridisk form (2 grupper)**

Bransch och antal anställda utgör stratumindelningen och där görs en bortfallskompensation via korsningen av variabler (multiplikationstecken mellan dem). Plustecknet mot geografisk indelning och juridisk form indikerar att kalibrering sker på *marginal*fördelningen av dessa variabler.

Den uppräkningsvikt som erhålls vid denna kalibrering är <u>inte</u> samma som vikten vid rak uppräkning inom strata, enligt uttryck (10). Nedan beskrivs de tekniska aspekterna kring hur kalibreringen genomförts och hur den slutliga uppräkningsvikten Kalvikt i databasen är framtagen.

Låt $\mathbf{x}_k$ representera hjälpvektorn som anger till vilken kombination av kategorierna

- bransch × antal anställda + geografi + juridisk form

företag k tillhör. Om t.ex. δ_{ik} representerar vilken av $i = 1,2,3,4,5$ geografiska grupper företag k tillhör och λ_{jk} representerar vilken av $j = 1,2$ juridiska former företag k tillhör kommer den delen av vektorn att skrivas[5] $\mathbf{x}_k = (\dots, \delta_{1k}, \delta_{2k}, \delta_{3k}, \delta_{4k}, \delta_{5k}, \lambda_{1k}, \lambda_{2k})'$. För exempelvis ett företag i den geografiska gruppen 33 Storstadskommuner ($i = 5$) och juridisk form övrigt aktiebolag ($j = 2$) blir hjälpvektorn $\mathbf{x}_k = (\dots,0,0,0,0,1,0,1)'$. Uppräkningsvikten (den kalibrerade) skapas därefter på följande sätt[6]

$$w_k = d_k \times v_k \tag{11}$$

där justeringsvikten v_k ges av

[5] För att undvika singularitet vid matrisinvertering i ett senare steg måste en kategori för respektive hjälpvariabel (utom för en av dem) droppas från $\mathbf{x}_k$
[6] Detta sätt att skapa den kalibrerade vikten benämns ibland Calibration for Nonrespons.

$$v_k = 1 + \left(\sum_{k\in U} \mathbf{x}_k - \sum_{k\in r} d_k \mathbf{x}_k\right)' \left(\sum_{k\in r} \mathbf{x}_k \mathbf{x}_k'\right)^{-1} \mathbf{x}_k \tag{12}$$

Den kalibrerade vikten har egenskapen

$$\sum_{k\in r} w_k \mathbf{x}_k = \sum_{k\in r} d_k v_k \mathbf{x}_k = \sum_{k\in U} \mathbf{x}_k \tag{13}$$

Detta betyder att om man använder den kalibrerade vikten och tar fram frekvenstabeller för de variabler som kalibrering sker gentemot erhålls de sanna populationsantalen. Uttryck (13) är det s.k. kalibreringsvillkoret.

Nu när uppräkningsvikten är etablerad presenteras punktestimatorerna för totalen t och medelvärdet y_{U_d} igen, men med rätt uppräkningsvikt. Om den kalibrerade vikten används för att skatta totalen t enligt (1) har vi kalibreringsestimatorn enligt

$$\hat{t}_{dw} = \sum_{k\in r} w_k y_{dk} \tag{14}$$

Notera att summeringen görs i svarandemängden r, inte hela urvalet s. Notera även subindex w i $\hat{t}_{dw}$ som betecknar att den kalibrerade vikten w_k används. Punktestimatorn för populationsmedelvärdet (och populationsandelen) $\bar{y}_{U_d}$ enligt (2), ges av kalibreringsestimatorn enligt

$$\hat{\bar{y}}_{wU_d} = \frac{\hat{t}_{dw}}{\hat{N}_{dw}} = \frac{\sum_{k\in r} w_k y_{dk}}{\sum_{k\in r} w_k I_{dk}} \tag{15}$$

Vikten w_k heter Kalvikt i databasen och ska användas för skattningar i populationen (eller redovisningsgrupper).

Anmärkning: Om Tillväxtverket använder exempelvis SAS och tar fram skattningar med vikten Kalvikt påkopplad är det estimatorn (14) som SAS använder om det är en total som beräknas och estimatorn (15) som SAS använder om det är ett medelvärde som beräknas.

Anmärkning: I den ordinarie Fvov 2020 användes inte ett kalibreringsförfarande utan metoden med rak uppräkning inom strata för att kompensera för bortfallet. I den ordinarie Fvov används dock betydligt fler strata än i *Fvov – Coronakrisens effekter*, närmare bestämt 385 strata i Fvov 2020. När rak uppräkning används inom dessa strata sker en bortfallskompensering med hänsyn till den använda stratifieringen. Exempelvis finns en geografisk stratifieringsvariabel, men även andra variabler. Detta betyder att det

sker en kompensation för eventuella skevheter geografiskt i den ordinarie Fvov 2020. I den bemärkelsen har den använda kalibreringen, där t.ex. geografi ingår, mer likheter med den bortfallskompensation som används i den ordinarie Fvov 2020.

Anmärkning: Statisticon genomförde ett 10-tal olika varianter på hur hjälpvariablerna skulle användas innan det slutliga valet föll på bransch × antal anställda + geografi + juridisk form. En viktig analys vid jämförelse mellan olika kandidater till kalibrering är att studera justeringsfaktorn v_k enligt (12). För att kalibreringsvillkoret (13) ska bli uppfyllt kan v_k bli negativt, speciellt om det är för många kategorier av en variabel. Man vill inte heller att v_k ska bli ”alltför” stort så att enstaka företag får en oproportionerligt stor uppräkningsvikt. Under processen kontrollerades storleken på v_k hela tiden. För den finala uppsättningen variabler som användes i kalibreringen var det största värdet 6,2 och det minsta 0,8. Ett värde under 1 betyder att designvikten d_k justeras ned lite grann. Ett värde på 6,2 betyder att designvikten justeras upp med en faktor 6,2. Notera att eftersom svarsandelen var $\frac{970}{2500} = 0{,}388$ så bör designvikten justeras upp med en faktor $\frac{1}{0{,}388} = \frac{2500}{970} = 2{,}58$. Medelvärdet av v_k är 2,58 och en variation upp mot 6,2 och ner mot 0,8 är därför rimliga. Variabeln v_k finns inte med i databasen Tillväxtverket får, men kan enkelt skapas genom att dividera variabeln Kalvikt med variabeln dk (se sambandet i uttryck (11)).

5.4 Variansestimator

Statisticon producerar inga skattningar i undersökningen utan det gör Tillväxtverket på egen hand. Via den kalibrerade uppräkningsvikt som erhålls kan punktskattningar erhållas. Om Tillväxtverket önskar beräkna variansskattningar, för t.ex. konfidensintervallsberäkningar, bör en mjukvara som är rätt anpassad för dessa typer av skattningar användas. I exempelvis SAS och proc surveyfreq eller proc surveymeans beräknas variansskattningar med hänsyn till urvalsdesign och den uppräkningsvikt som används. I SAS-manualen framgår hur variansskattningarna genomförs. SCB har en mjukvara som heter CLAN, som är ett SAS-makro, som kan producera punkt- och variansskattningar med hänsyn till urvalsdesign och den kalibrering som använts.

Man får dock inte identiska variansskattningar med t.ex. SAS proc surveymeans och CLAN (punktskattningarna blir identiska). Vanligtvis blir variansskattningarna något större med SAS än med CLAN. Man kan möjligen säga att CLAN producerar en korrekt variansskattning med hänsyn taget till kalibrering, medan SAS proc surveymeans producerar en approximativ variansskattning.

6 OSÄKERHETSKÄLLOR

Statistiken är behäftad med osäkerhet. De största osäkerhetskällorna är bortfall och urval. Osäkerheten som beror på urval kan kvantifieras med hjälp av konfidensintervall. Bortfallet kompenseras i skattningarna via metoden med kalibrerade vikter och inkorporeras på så sätt i osäkerhetsskattningarna under förutsättning att metoden är adekvat.

Det finns troligtvis en viss okänd mätosäkerhet i resultaten. Denna osäkerhet kvantifieras inte i osäkerhetsskattningarna.

I avsnitten nedan diskuteras alla osäkerhetskällor och deras påverkan på statistiken.

6.1 Urval

Urvalsosäkerhet kan kvantifieras. Det vanligaste sättet är att beräkna konfidensintervall. Urvalsstorleken på 2 500 företag är ett stort urval. Svarandemängden är dock lägre med 970 företag. Även denna svarandebas är dock hyggligt stor på totalnivå.

För att ge en bild av osäkerheten på totalnivå används fråga B1a och andel som svarat 1=Minskat.

B1. Hur har följande påverkats för företaget under 2020, jämfört med 2019?
Markera ett kryss på varje rad

	Minskat	Oförändrat	Ökat
a) Omsättning?	☐	☐	☐

Baserat på estimatorn $\hat{\bar{y}}_{wU_d}$ enligt (15) erhålls

$$\hat{\bar{y}}_{wU_d} = \frac{\hat{t}_{dw}}{\hat{N}_{dw}} = \frac{188\,462.9}{343\,919.5} = 0{,}548 = 54{,}8\%$$

Skattningen är alltså att 54,8 procent i populationen anser att omsättningen minskat 2020 jämfört med 2019. Skattningen i täljaren avser totalt *antal* företag i populationen där omsättningen har minskat. I tabell 14 redovisas dessa punktskattningar, samt variansskattningen[7] och ett 95% konfidensintervall.

Tabell 14. Skattningar av antal och andel på fråga B1a avseende kategorin Minskat, i hela populationen

Storhet	Beteckning	Punktskattning	Variansskattning	95% konfidensintervall
Täljare	$\hat{t}_{dw}$	188 462,9	92 508 257,2	±18 851,5
Nämnare	$\hat{N}_{dw}$	343 919,5	283 749,7	±1 044,1
Kvot	$\hat{\bar{y}}_{wU_d}$	0,548	0,0007835	±0,055

[7] Variansskattningen är framtagen med CLAN.

6.2 Mätning

Osäkerhetskällan mätning har att göra med att det mätresultat som registreras kan avvika från det sanna mätvärdet. Det frågeformulär som användes är i viss utsträckning samma som användes i Fvov 2020. Några nya frågor har tillkommit. I kvalitetsdeklarationen för Fvov 2020 framgår att mätinstrumentet har genomgått en begränsad mätteknisk granskning av mätexperter på SCB. Statisticon har ingen kännedom om de nya frågorna har genomgått motsvarande granskning.

Det har inte kommit till Statisticons kännedom att det skulle vara någon fråga som har berett mer problem än andra. En indikation på problem med en fråga kan vara om det partiella bortfallet är stort. Tidigare konstaterades att det var ett relativt begränsat partiellt bortfall. Frågorna D4 a och b har störst partiellt bortfall på drygt 5 procent. Därefter har övriga frågor ett partiellt bortfall på mindre än 3 procent. Detta ger en viss indikation om att frågorna inte har berett några svårigheter att svara på.

Några speciella metodstudier för att studera eventuell mätfel har inte genomförts.

6.3 Bortfall

Bortfall påverkar statistiken på två sätt. Det ena är att svarandebasen minskar, vilket ökar osäkerheten. Detta är dock inte något stort problem, i alla fall för skattningar på hög aggregeringsnivå. När resultat bryts ned på små redovisningsgrupper kan detta dock medföra att resultat inte kan presenteras. Den andra effekten, som är allvarligare, är risken att resultat snedvrids på grund av bortfallet. Om de som inte svarar avviker systematiskt (rörande undersökningsvariablerna) från de som har svarat finns risk för skevheter i skattningarna.

Svarsandelen är ca 39 procent vilket ger en bortfallsandel på 61 procent. Detta kan sägas vara ett stort bortfall, vilket tyvärr är en trend i många underökningar idag. I bortfallsanalysen i avsnitt 3.3 framgår att det är skillnader i bortfall mellan olika grupper. Bortfallet har kompenserats för via metoden med kalibrerade vikter. I avsnitt 6.6 Modellantagande kommenteras vilka antaganden som görs.

6.4 Ramosäkerhet

Ramen baseras på Företagsdatabasen (augusti-SAMU 2020) och upprättades av SCB. Eftersom registret uppdateras kontinuerligt med viss eftersläpning förekommer såväl över- som undertäckning.

Övertäckning, dvs. att företaget ingår i ramen trots att det inte borde det, har observerats i ett fall. Det var ett företag som hade gått i konkurs. Den observerade övertäckningen är därmed lägre än i den ordinarie Fvov 2020. Statisticon har ingen förklaring till varför övertäckningen skulle vara lägre i *Fvov – Coronakrisens effekter* än i Fvov 2020.

Undertäckning, dvs. att ett företag tillhör målpopulationen men inte ingår i ramen, är potentiellt sett ett större problem än övertäckning. Undertäckningen har inte studerats.

Avslutningsvis bör nämnas att FDB betraktas som ett bra register med hög tillförlitlighet. Det finns viss osäkerhet kring ramtäckning, men det är troligtvis ett marginellt problem jämfört med den osäkerhet som uppstår via urval och bortfall.

6.5 Bearbetning

Data har bearbetats i flera steg innan en final databas har erhållits. Samtliga steg genomförs via programkod och resultaten av varje steg granskas. Det finns ingen garanti för att inte bearbetningsfel kan uppstå, men genom noggranna kontroller av varje datasteg minskas risken för fel.

6.6 Modellantaganden

Det viktigaste modellantagande i undersökningen rör hur kompensation för bortfallet görs. Kompensation för bortfall görs inom ramen för kalibrering av vikter. I avsnittet bortfallsanalys i avsnitt 3.3 visade analysen att det finns en samvariation mellan kalibreringsvariablerna och svarsbenägenheten. Detta är gynnsamt och genom kalibreringen finns en möjlighet att reducera eventuell bortfallsskevhet. Det variabler som användes i kalibreringen var bransch × antal anställda + geografi + juridisk form. Det var dessa variabler som fanns att tillgå som hjälpvariabler.

Snedvridningar kan naturligtvis uppstå rörande andra variabler. Om exempelvis företag som drabbas hårt av coronakrisen (t.ex. att omsättningen minskar dramatiskt) i större utsträckning väljer att inte svara på enkäten har vi ingen variabel som *explicit* tar hänsyn

till denna snedvridning. Bland de svarande företagen skulle då extra hårt drabbade företag vara relativt sett underrepresenterade. Om ovägda resultat skulle användas för att ta fram skattningar skulle den statistiken bli missvisande.

Det finns tyvärr ingen explicit variabel i något register som indikerar vilka företag som drabbats extra hårt eller inte. Det är troligtvis så att olika branscher drabbas olika hårt. Om bortfallet inom varje bransch skett någorlunda slumpmässigt (alltså inte systematiskt att en viss grupp företag avstår att svara) kommer den använda metoden för bortfallskompensation att verka i rätt riktning.

7 DATABASER

Statisticon levererar flera filer med data till Tillväxtverket. I tabell 15 redovisas dessa och därefter kommenteras respektive databas.

Tabell 15. Databaser som levereras till Tillväxtverket

Nr	Beskrivning	Filnamn
1	Enkätdata med registervariabler och uppräkningsvariabel	Fvov_corona_2020_databas.sas7bdat
2	Kalibreringstotaler (både som SAS och Excel)	Fvov_corona_2020_AUXdata.sas7bdat
3	Urvalsfil från SCB	SCB_U019519.xlsx
4	Format till variablerna i enkätdata (både som katalog och dataset)	formats.sas7bdat, formats.sas7bcat

7.1 Enkätdata med registervariabler

Den viktigaste filen är filen med enkätdata, som även innehåller registervariabler och designvariabler. Filen består av

- 2 500 företag, dvs. hela urvalet. Variabeln Svar indikerar om företaget har svarat på enkäten eller inte.
    - Svar = 1, om företaget svarat och
    - Svar = 0, om företaget inte har svarat
- 154 variabler

I bilaga 6 finns en variabelförteckning med förklaringar[8]. I variabelförteckning finns en förklaring till varifrån variabeln härstammar. I tabellen nedan framgår att det finns fyra olika källor och antalet variabler av respektive sort.

Källa	Antal variabler
Designvariabel	11
Enkätvariabel	96
Insamling	6
Registervariabel	41
Summa	154

Kort kommentar kring de olika källorna:

- Designvariabler: detta är variabler som använts i urvalsdesignen, t.ex. variablerna stratum, stratumtxt, bransch, StoraN, *n* men även härledda variabler som variabeln antal svarande per stratum (m), designvikten (dk) och den kalibrerade vikten (Kalvikt)

[8] Förklaringarna är samma som de labels som finns på variablerna i databasen.

- Enkätvariblerna: detta är de variabler som frågorna i enkäten gett upphov till
- Insamling: detta är variabler som har skapats i samband med insamlingen. Exempelvis den unika identifikatorn (username), huruvida ett företag svarat via webb eller postalt (source), antal frågor uppgiftslämnare har besvarat (datapoints), svarsindikator (svar), avprickningskod (avprkod) och svarsdatum (svarsdatum).
- Registervariabler: detta är de registervariabler som erhölls via urvalsbeställningen från Företagsdatabasen.

Databasen med enkätdata kan användas för att producera skattningar av deskriptiva målstorheter/parametrar i populationen. Den uppräkningsvikt som ska användas för att producera skattningar för de svarande avseende enkätfrågorna heter

- Kalvikt

Det kan nämnas att det finns ytterligare en uppräkningsvikt i databasen, nämligen designvikten $d_k = N_h/n_h$ se avsnitt 5.3 Punktestimator. Denna variabel heter dk i databasen och kan användas för uppräkning till populationen baserat på hela urvalsmängden (2 500 utvalda). Detta är tillämpligt vid framtagande av statistik för registervariabler där samtliga utvalda har giltiga värden.

Om Kalvikt används för att exempelvis ta fram en frekvenstabell för stratum baserat på svarandemängden ($m = 970$, svar=1) erhålls <u>samma</u> resultat som om dk används som viktvariabel för att ta fram samma frekvenstabell baserat på samtliga utvalda ($n = 2500$). Detta är en effekt av att kalibrering sker gentemot stratumindelningen (bransch och anställda i klasser)

Kalibreringsvariabler

I databasen finns ett visst antal variabler med prefixet 'kal_'. Dessa kan kommenteras speciellt. I tabellen nedan (från variabelförteckningen i bilaga 6) framgår vilka de är.

nr	Variabel	Förklaring
146	kal_KomInd5	Kalibreringsvariabel. Tillväxtverkets kommungruppsindelning (kod) - 6 nivåer men grupp 1 och 2 är sammanslagna till 1 grupp
147	kal_KomInd5_namn	Kalibreringsvariabel. Förklaring till kodning i variabel kal_komInd5
148	kal_jurform	Kalibreringsvariabel. Juridisk form (sammanslagen till 2 grupper)
149	kal_jurform_text	Kalibreringsvariabel. Förklaring till Juridisk form i variabel kal_jurform
150	kal_bransch_anstkl	Kalibreringsvariabel. Korsning av bransch och antal anställda i klasser. 1-1 koppling med variabeln stratum
151	kal_bransch	Branschindelning. 1-1 koppling med variabeln bransch_str.
152	kal_anstkl	Branschindelning. 1-1 koppling med variabeln anstkl_str.

Variabel kal_bransch_anstkl (nr 150) är korsningen mellan branschindelningen (13 grupper) och antal anställda i 4 grupper. Denna variabel har ett 1-1-förhållande med variabeln stratum, men är annorlunda kodad. Variabeln kal_bransch_anstkl är kodad 1,2,3,…,52 medan variabeln stratum har andra koder. I proceduren med att ta fram den kalibrerade vikten är det praktiskt att ha en variabel med löpnummer 1,2,3,…. För att lätt kunna se kopplingen mellan variabeln kal_bransch_anstkl och stratum finns en sådan tabell i bilaga 7.

I bilaga 7 visas även variablerna kal_bransch och kal_anstkl vilket är bransch respektive antal anställda i klasser. Dessa två variabler har ett 1-1-förhållande mot registervariablerna bransch_str och anstkl_str, men är kodade annorlunda. Kal_-variablerna är kodade 1,2,3,…. Variablerna Kal_bransch och kal_anstkl användes dock inte i kalibreringen eftersom kalibrering skedde gentemot *korsningen* av de båda variablerna (kal_bransch_anstkl). Under framtagande av kalibreringsvikten övervägdes om kalibrerings kulle ske gentemot marginalen för bransch respektive antal anställda (i grupper). Dock kunde kalibrering ske gentemot korsning av dem båda, vilket är önskvärt.

Variablerna kal_KomInd5 och kal_jurform användes för att ta fram den kalibrerade uppräkningsvikten (tillsammans med variabeln kal_bransch_anstkl). Var och en av de båda variablerna har en motsvarande förklaringsvariabel med samma indelning (kal_KomInd5_namn och kal_jurform_text).

7.2 Kalibreringstotaler

I filen (SAS och Excel) *Fvov_corona_2020_AUXdata* finns de totaler som kalibrering har skett gentemot. Den består av variablerna

- kal_KomInd5
- kal_bransch_anstkl
- kal_jurform

Variablerna MAR1, MAR2, MAR3,… är de antalsuppgifter som kalibrering sker gentemot. Kopplingen mellan variabelvärdena och MAR-variablerna är att när variabeln antar värdet 1 motsvarar det antalsuppgiften i variabel MAR1, koden 2 motsvarar MAR2, koden 3 motsvarar MAR3 osc. För exempelvis variabeln kal_bransch_anstkl betyder kod 1 bransch Jordbruk, skogsbruk och fiske och 0 anställda. Det finns 25 769 sådana företag i urvalsramen. Variabeln MAR1 antar värdet 25 769 för variabeln kal_bransch_anstkl och det är gentemot detta värde som kalibrering sker.

7.3 Urvalsfil från SCB

SCB levererade en urvalsfil i Excelformat och Tillväxtverket får i leveransen från Statisticon denna fil i original, dvs. Statisticon har inte förändrat filen. Filen består av tre flikar:

- Urval: själva urvalet med registervariabler
- Stratifiering: en beskrivning av stratifieringen
- Tab_urvalsram: på varje rad i denna flik visas antal företag i urvalsramen per kombination av variabler

Fliken tab_urvalsram består av variablerna Stratum, bransch, storleksklass, Kommun och JurForm. För exempelvis rad 2 framgår att det finns 7 företag i just den kombinationen av de fem variablerna. Denna flik är det underlag som använts för att skapa totaler som kalibrering sker gentemot.

7.4 Format-fil

I filerna formats.sas7bdat och formats.sas7bcat finns de format som kan läggas på variablerna. Exempel: variabeln B1a har tre svarsalternativ enligt nedan

B2. Hur tror du företaget kommer att utvecklas på *tre* års sikt vad gäller…
Markera ett kryss på varje rad

	Minska	Oförändrad	Öka
a) … omsättning?	☐	☐	☐

De tre svarsalternativen i variabeln B1a är kodade 1,2,3. I format-filen finns kopplingen mellan koderna och förklaringstexten till svarsalternativen, dvs. 1= Minska, 2=oförändrad och 3=Öka. Statisticon levererar alltså format-filen i två olika filer; (i) en formatkatalog och (ii) ett dataset med formaten. Endast en av dessa behöver användas för att lägga på format på variablerna. Tillväxverket kan välja vilken källa som används, de är båda identiska.

REFERENSER

Lundström, S. and Särndal, C.E. (2001). *Estimation in the presence of nonresponse and frame imperfections*. Statistics Sweden.

Särndal C.E., Swensson B., Wretman J. (1992). *Model Assisted Survey Sampling*. Springer Verlag.

BILAGEFÖRTECKNING

Nr	Beskrivning	Filnamn
1	Enkät för Fvov - Coronakrisens effekter	Bilaga_1_Enkät Fvov_corona.pdf
2	Missiv huvudutskick	Bilaga_2_Missiv_huvudutskick.pdf
3	Missiv påminnelse 1	Bilaga_3_Missiv_påminnelse 1.pdf
4	Missiv påminnelse 2	Bilaga_4_Missiv_påminnelse 2.pdf
5	Avprickningskoder	Bilaga_5_Avprickningskoder.pdf
6	Variabelförteckning	Bilaga_6_Variableförteckning.xlsx
7	Koppling mellan kalibreringsvariabler bransch och antal anställda (grupperad)	Bilaga_7_Bransch_anstkl_kalibreringsvar.xlsx