



Konjunkturbarometern Fördjupning

Urvalsminskning i Konjunkturbarometern företag

FÖRDJUPNING

Urvalsminskning i Konjunkturbarometern företag

I en urvalsundersökning är urvalsstorleken en viktig faktor att beakta. Inför urvalsbytet i Konjunkturbarometern företag i maj 2019 genomförde Konjunkturinstitutet (KI) en analys av möjligheten att minska urvalet i undersökningen. I denna fördjupning beskrivs denna analys kortfattat.

Ett stort urval minskar risken för slumpmässiga fel men medför samtidigt ökade kostnader för genomförandet av undersökningen. I en tid där antalet undersökningar i samhället ökar är det också önskvärt att begränsa företagens uppgiftslämnarbörda. Det är viktigt att potentiella effekter av minskade urvalsstorlekar studeras innan en sådan åtgärd genomförs eftersom urvalsundersökningar inte kan göras om.

Den befintliga urvalsmetodiken studerades i syfte att identifiera stratum (se marginalruta) där urvalsstorlekarna förväntades kunna minskas utan att resultatet från undersökningen påverkas nämnvärt. Minskningar i dessa stratum (se marginalruta) användes sedan som ingångsvärden i simuleringar på data i Konjunkturbarometerns faktiska undersökning för perioden maj 2018 till och med mars 2019. För var och en av dessa 11 månader genomfördes 10 000 simuleringar per bransch, där olika, men samma antal, företag slumpmässigt uteslöts vid varje simulering. Resultaten av simuleringarna jämfördes sedan med publicerade resultat baserade på det faktiska urvalet.

Skillnaderna mellan de simulerade resultaten med begränsat urval och de publicerade resultaten var i allmänhet små. Detta ledde till beslutet att minska storleken på urvalet. Effekten blev att antalet företag i urvalet minskade med cirka 20 procent, från ca 6 800 till ca 5 500 företag i samband med urvalsbytet i maj 2019.

Skäl till att minska urvalsstorleken

Det huvudsakliga argumentet för att minska storleken på urvalet i KI:s företagsundersökning är att minska företagets uppgiftslämnarbörda. I en undersökning bör inte fler uppgifter än nödvändigt samlas in och det finns i Sverige ett uttalat mål att myndigheter ska sträva efter att minska uppgiftslämnarbördan för

Begrepp

Cut-off gräns

Det minsta antalet anställda ett företag måste ha för att ingå i urvalsramen.

Målpopulation

Företag med verksamhet i Sverige i de branscher som Konjunkturbarometern företag undersöker.

Stratum

Icke-överlappande delmängder som tillsammans täcker hela urvalsramen. Stratum definieras med syfte att företagen inom dessa ska ha liknande egenskaper. Stratifierat urval används för att urvalet ska vara representativt för målpopulationen. I Konjunkturbarometern företag stratifieras på antal anställda och branschtillhörighet.

Täckningsgrad

Avser hur väl ramen täcker målpopulationen.

Urvalsbransch

Bransch definieras enligt Svensk Näringsgrensin- delning (SNI) på 2 till 4 siffrors nivå, som urvalet stratifieras på. Urvalsbranscherna är till viss del styrda av vilka branscher Konjunkturinstitutet själva vill publicera, samt vilka som ska redovisas till EU.

Urvalsram

Företag som ingår i målpopulationen och som har fler anställda än cut-off gränsen i dess urvals- bransch.

företag.¹ Det finns också andra fördelar med ett minskat urval. För KI medför ett minskat urval lägre kostnader, exempelvis genom mindre arbete med Konjunkturbarometerns företagsregister, färre frågor från uppgiftslämnare och mindre arbete vid urvalsbyte. Dessutom minskar kostnaden för pappersutskick och färre utskick innebär även minskad miljöpåverkan.

Den största nackdelen med ett minskat urval är en ökad risk för slumpmässiga fel. Om urvalet minskas genom att cut-off gränsen för antalet anställda höjs, minskar även urvalsramen och täckningsgraden blir lägre. Låg täckningsgrad kan i sin tur öka risken för att resultatet inte speglar målpopulationen på ett korrekt sätt. En minskning i urvalet riskerar därmed att få oönskade konsekvenser, till exempel i form av tidsseriebrott.

Befintlig urvalsmetodik

Konjunkturbarometern är en panelundersökning uppdelad i fyra olika sektorer: tillverkningsindustri, tjänstesektor, handel och byggindustri. Urvalet är ett totalurval för företag med fler än 100 anställda. Av de företag som har färre än 100 anställda dras en delmängd till undersökningen. Urvalsbyte görs en gång per år.² Alla företag som ingår i urvalet tillfrågas varje månad till nästa urvalsbyte. Denna analys gjordes på urvalet som började användas i maj 2018.

Vid framtagande av urvalsramen används Statistiska centralbyråns (SCB) system för samordning av rampopulationer och urval (SAMU) för mars månad.³ SAMU är en ögonblicksbild av SCB:s företagsregister (FDB), som är heltäckande. Användningen av SAMU innebär att urvalsramen är alla företag som har fler anställda än cut-off gränsen.

Urvalet stratifieras på bransch och storleksklass. Urvalsbranscherna är definierade utifrån Svensk Näringsgrensindelning (SNI). Storleksklasserna är definierade utifrån antalet anställda i företagen. Cut-off gränsen varierar mellan 5 och 50 anställda beroende på sektor och bransch. Totalt dras urval i 96

¹ Förordningen om den officiella statistiken (SFS 2001:100).

² I samband med urvalsbyten sker naturligt förändringar i urvalsbranscher, på grund av att sammansättningen av företag av olika storlekar ändras, samt att företag omklassificeras till en annan bransch.

³ SAMU uppdateras fyra gånger per år. Uppdateringen i mars används för att många organisatoriska omstruktureringar görs vid årsskiftet. Dessa har hunnit bli bokförda i FDB i början av februari.

urvalsbranscher och 6 storleksklasser.⁴ Inom varje sektor dras urval i mellan 3 och 40 urvalsbranscher.

SCB framställer, utifrån KI:s önskemål, en urvalsram och tar fram underlag, som KI använder för att bestämma urvalsstorlekar i olika branscher och storleksklasser. KI väljer därefter själva hur många företag som ska dras till undersökningen i varje stratum. Till sist drar SCB företagen ur SAMU och levererar urvalet till undersökningen.

Tabell 1 Urvalet maj 2018, antal företag

Sektor	Urvalsram	Urvalsstorlek	Andel av urvalsram
Tillverkningsindustri	3 511	1 704	49%
Byggindustri	5 306	478	9%
Handel	8 486	1 477	17%
Tjänstesektorn	7 860	3 129	38%
Totalt	24 901	6 526	26%

Tabell 2 Urvalet maj 2018, vikt i antal anställda

Sektor	Urvalsram	Urvalsstorlek	Andel av urvalsram
Tillverkningsindustri	437 252	367 423	84%
Byggindustri	200 954	91 009	45%
Handel	387 316	282 404	73%
Tjänstesektorn	828 612	681 136	82%
Totalt	1 854 914	1 421 972	76%

Strategi för minskning av urval

Urvalsmetodiken som redovisas ovan analyserades i syfte att identifiera stratum där urvalsstorlekarna förväntades kunna minskas utan att resultatet från undersökningen påverkas nämnvärt på aggregerad nivå.

Den strategi KI har valt för hur urvalsstorleken kan minskas baseras på fyra olika angreppssätt. I denna del redogörs först för dessa fyra olika angreppssätt, därefter beskrivs den totala

⁴ Storleksklasserna som används är 5–9, 10–19, 20–49, 50–99, 100–199 och 200 eller fler anställda.

effekten av dessa tillsammans. För varje angreppssätt redovisas det förslag till förändringar som lades fram.

UTESLUTA STRATUM MED LITEN VIKT

Historiskt har KI försökt att i så stor utsträckning som möjligt använda generella regler för cut-off gränser för branscher inom respektive sektor.⁵ På så vis blir det lättare att förstå hur undersökningen är uppbyggd och beställningen av urvalet förenklas. Nackdelen med generella cut-off gränser är att urvalet därmed inkluderar företag som har mycket liten betydelse för resultatet inom enskilda branscher. Dessutom medför mer generella cut-off gränser, på grund av branschstrukturen, att vissa stratum blir så små att rotation i urvalet försvåras. När ett stratum utesluts påverkas urvalsramen och därmed minskar branschens vikt (mätt med antal anställda).

Tabell 3 visar hur de föreslagna förändringarna i cut-off gränser påverkar urvalsstorleken. Exempelvis minskas antalet dragna företag i urvalet för handeln med nästan 14,8 procent.

Tabell 3 Föreslagna minskade urvalsstorlekar i stratum med liten vikt

Sektor	Antal företag i urval	Antal uteslutna företag	Andel uteslutna företag
Tillverkningsindustri	1 704	126	7,4%
Byggindustri	478	-	-
Handel	1 477	218	14,8%
Tjänstesektorn	3 129	55	1,9%
Totalt	6 526	399	6,1%

Tabell 4 visar hur stor andel av urvalsramen som utesluts till följd av höjda cut-off gränser. I handeln utesluts 4,9 procent av sektorns vikt till följd av detta.

⁵ Exempelvis har alla företag inom sektorn tillverkningsindustri haft 20 anställda som cut-off gräns. För handeln var gränsen olika för tre delar, 50 anställda inom partihandel, 20 inom handel med motorfordon och 5 i den övriga detaljhandeln.

Tabell 4 Urvalsramens förändring till följd av höjda cut-off gränser

Sektor	Antal anställda i urvalsram	Utesluten vikt urvalsram	Andel utesluten vikt
Tillverkningsindustri	437 252	9 735	2,2%
Byggindustri	200 954	-	-
Handel	387 316	18 981	4,9%
Tjänstesektorn	828 612	12 798	1,6%
Totalt	1 854 914	41 514	2,3%

MINSKA URVALSSTORLEKAR I BRANSCHER MED LITEN VIKT

En del branscher, där resultaten särredovisas, har jämförelsevis liten betydelse i ett eller flera branschaggregat. De mindre storleksklasserna kan i detta fall ändå ha en jämförelsevis stor påverkan på den enskilda branschens resultat. Därför vill vi inte höja cut-off gränsen för branschen, utan istället minska storleken på urvalet i de mindre storleksklasserna. I detta, och följande, angreppssätt, påverkas inte urvalsramen.⁶ Tabell 5 visar hur många företag som utesluts till följd av detta angreppssätt.

Tabell 5 Föreslagna minskade urvalsstorlekar i branscher med liten vikt

Sektor	Antal företag i urval	Antal uteslutna företag	Andel uteslutna företag
Tillverkningsindustri	1 704	60	3,5%
Byggindustri	478	-	-
Handel	1 477	12	0,8%
Tjänstesektorn	3 129	249	8,7%
Totalt	6 526	321	4,9%

MINSKA URVALET I STRATUM DÄR EXTRA FÖRETAG DRAS

I Konjunkturbarometern företag är det stor skillnad i antalet företag i olika stratum. I de stratum som har flest företag har vi tidigare haft större urvalsstorlekar än vad som motiverats av att kunna uttala oss om den enskilda branschen. Detta för att få bättre skattningar i branschaggregat. Det är dock möjligt att det går att minska urvalet i dessa stratum utan att det får några större

⁶ Detta eftersom kvarvarande företag i storleksklassen räknas upp med motsvarande vikt som de uteslutna företagen hade. Därmed bibehålls branschens relativa vikt i de branschaggregat där branschen ingår.

negativa konsekvenser (den föreslagna neddragningen redovisas i tabell 6).

Tabell 6 Föreslagna minskade urvalsstorlekar i stratum där extra företag dras

Sektor	Antal företag i urval	Antal uteslutna företag	Andel uteslutna företag
Tillverkningsindustri	1 704	99	5,8%
Byggindustri	4 786	6	1,3%
Handel	1 477	156	10,6%
Tjänstesektorn	3 129	169	5,9%
Totalt	6 526	430	6,6%

MINSKAT URVAL I BRANSCHEN FÖRVALTARE AV EGNA FASTIGHETER

I branschen förvaltare av egna fastigheter används företagets omsättning som en proxy för antalet anställda.⁷ I denna bransch har KI tidigare dragit oproportionerligt många företag relativt branschens storlek.⁸

Tabell 7 Föreslagna minskade urvalsstorlekar i förvaltare av egna fastigheter

Bransch	Antal företag i urval	Antal uteslutna företag	Andel Uteslutna företag
Förvaltare av egna fastigheter	262	177	67,5 %

SAMMANTAGEN EFFEKT PÅ URVALET

När de fyra ovanstående angreppssätten genomförs samtidigt leder det till att antalet företag i urvalet minskar med cirka 20 procent, se tabell 8. I termer av antalet anställda stannar dock neddragningen på 2,5 procent, se tabell 9.

⁷ Många företag i denna bransch har få anställda, trots hög omsättning. Omsättningen antas därför bättre spegla företagets storlek än antal anställda.

⁸ Den tidigare metoden innebar att alla företag över cut-off gränsen ingick i urvalet. Minskat urval i denna bransch föreslogs att göras genom att dra ett slumpmässigt urval bland de företag som hade mellan 20 och 99 "anställda" enligt proxyvariabeln.

Tabell 8 Föreslagna minskningar, antal företag

Sektor	Faktisk 2018		Simulerad 2018		Andel uteslutna företag
	Urvalsram	Urvalsstorlek	Urvalsram	Urvalsstorlek	
Tillverkningsindustri	3 511	1 704	3 297	1 419	16,7%
Byggindustri	5 306	478	5 306	472	1,3%
Handel	8 486	1 477	6 254	1 091	26,1%
Tjänstesektorn	7 860	3 129	7 561	2 479	20,8%
Totalt	24 901	6 526	22 156	5 376	19,5%

Tabell 9 Föreslagna minskningar, antal anställda

Sektor	Faktisk 2018		Simulerad 2018		Andel uteslutna anställda
	Urvalsram	Urvalsstorlek	Urvalsram	Urvalsstorlek	
Tillverkningsindustri	437 252	367 423	427 517	353 195	3,3%
Byggindustri	200 954	91 009	200 954	90 704	0,2%
Handel	387 316	282 404	368 335	273 561	2,3%
Tjänstesektorn	828 612	681 136	815 814	637 799	2,9%
Totalt	1 854 914	1 421 972	1 812 620	1 355 259	2,5%

AVGRÄNSNINGAR

En annan möjlighet för minskning av urvalsstorleken är att frångå principen om totalurval i storleksskassen som utgörs av företag med 100–199 anställda. För vissa branscher är vikten, i antal anställda, i den storleksskassen relativt liten jämfört med de övriga storleksskasserna. KI har valt att avvakta med att analysera en sådan förändring, dels för att avgränsa projektets storlek, dels för att begränsa antalet förändringar i urvalsmetodiken vid ett och samma tillfälle.

Beräkning av netttotal

Majoriteten av frågorna i Konjunkturbarometern företag redovisas i form av så kallade netttotal.⁹ Ett netttotal är skillnaden mellan andelen respondenter som exempelvis svarat ”ökat” respektive ”minskat” på en fråga. I denna fördjupning används

⁹ Vissa frågor har fler svarsalternativ eller besvaras numeriskt.

jämförelser av netttotal för att redovisa effekten av den föreslagna urvalsminskningen.

Exempel på uträkning av netttotal: 40 procent av företagen uppger att produktionsvolymen har ökat och 10 procent att den har minskat (50 procent uppger att den är oförändrad). Nettotalet blir i det här fallet 30 ($40 - 10 = 30$). Således antar netttotal alltid värden mellan -100 (alla respondenter har angett ett negativt svarsalternativ) och +100 (alla respondenter har angett ett positivt svarsalternativ).¹⁰

Simulering

För att undersöka hur stor påverkan ett minskat urval skulle ha haft på redan publicerade resultat genomförde vi simuleringar utifrån den presenterade strategin. Vi använde oss av en tidsperiod med samma urval i undersökningen, maj 2018 till mars 2019. För var och en av dessa 11 månader genomförde vi 10 000 simuleringar per urvalsbransch. För varje simulering uteslöt vi slumpmässigt företag så att den slutliga urvalsstorleken var enligt tabell 8. Därefter räknade vi samman resultaten för branscher där nytt resultat hade simulerats med branscher där inga förändringar i urvalsstorlekar föreslagits, för att kunna se påverkan även på sektornivå.

RESULTAT

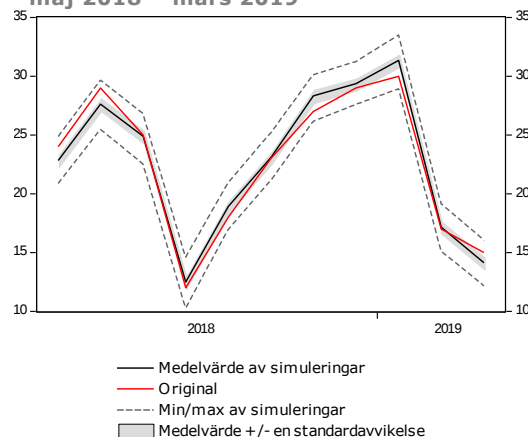
I detta avsnitt redovisas resultaten av simuleringarna med minskade urvalsstorlekar på icke säsongsrensade netttotal för en fråga per sektor. För övriga frågor som har simulerats, men som inte redovisas här, blev slutsatserna liknande. Effekterna av minskningarna i urvalet i byggindustrin var så små att dessa resultat ej redovisas här. Från resultatet av simuleringarna redovisas medelvärde, standardavvikelse, max- och minvärde tillsammans med publicerade resultat.

I diagram 1 jämförs den publicerade originalserien för frågan om produktionsvolym inom tillverkningsindustrin¹¹ med en serie som visar medelvärdet av simuleringarna. Resultatet av simuleringarna visar på små skillnader mot originalserien. Standardavvikelsen för simuleringarna av nettotalet för en period är ungefär

Diagramförklaring

I diagrammen jämförs den publicerade icke säsongsrensade originalserien med ett medelvärde av de 10000 simuleringarna. Max och min är det högsta respektive lägsta värdet av alla simuleringar. Standardavvikelse för de simulerade värdena används för att visa på variationen i simuleringarna. Det grå bandet är ett intervall, medelvärde +/- en standardavvikelse.

Diagram 1 Tillverkningsindustri, Netttotal, maj 2018 – mars 2019



¹⁰ För en mer utförlig beskrivning av hur netttotal beräknas, se appendix.

¹¹ Frågeformulering med svarsalternativ: Produktionsvolymen har de senaste 3 månaderna (ökat, varit oförändrad, minskat)

0,5 enheter och är i stort sett den samma för alla månader. En halv enhet är jämförelsevis lite. Det kan jämföras med att nettotalet för frågan i genomsnitt har rört sig med 5,3 enheter från en månad till en annan under den period vi simulerade. Max- och minvärden från simuleringarna visar att även för de mest extrema utfallen blir resultaten nära originalserien.

I diagram 2 redovisas frågan om utvecklingen av företagets verksamhet inom tjänstesektorn.¹² Även för tjänstesektorn visar resultaten på små skillnader gentemot originalserien. Standardavvikelsen är också för denna sektor i genomsnitt 0,5 enheter.

Slutligen redovisas resultaten för frågan om försäljningsutveckling i handeln¹³ i diagram 3. För handeln blev standardavvikelsen i de simulerade resultaten större än för tillverkningsindustrin och tjänstesektorn, i genomsnitt 1,5 enheter över hela perioden. Även denna variation är dock låg, jämfört med hur tidsserien normalt rör sig.

KÄNSLIGHETSANALYS

För att undersöka vad som händer om urvalsstorlekarna istället minskas relativt mycket i en enskild bransch, istället för att minska i olika typer av branscher, gör vi ett test med branschen ”Detaljhandel med brett sortiment”. Branschen har jämförelsevis stor vikt i storleksklasser med mindre företag.

Diagram 2 Tjänstesektorn, Nettotal, maj 2018 – mars 2019

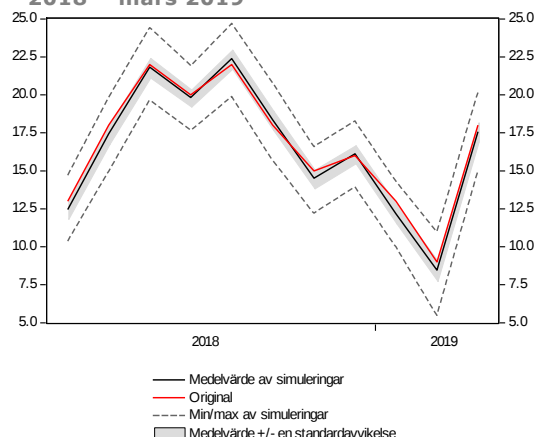
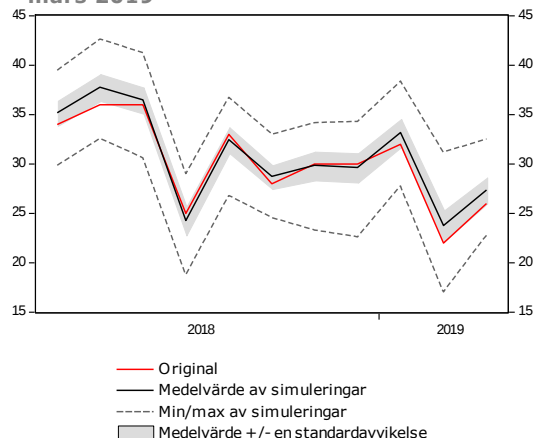


Diagram 3 Handeln, Nettotal, maj 2018 – mars 2019



¹² Hur har företagets verksamhet utvecklats de senaste 3 månaderna (förbättrats, varit oförändrad, försämrats).

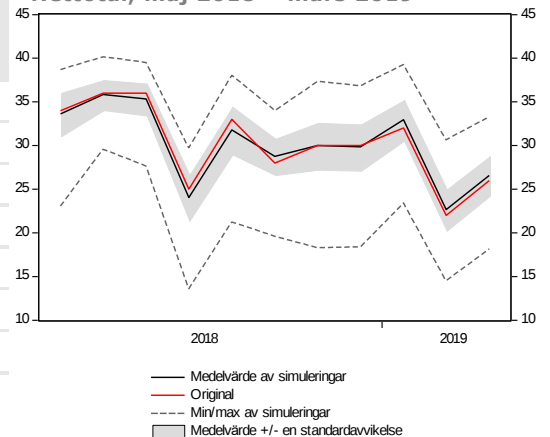
¹³ Försäljningsvolymen har de senaste 3 månaderna (ökat, varit oförändrad, minskat)

Tabell 10 Testsimulering, detaljhandel med brett sortiment

Antal anställda	Antal företag i urvalsram	Antal företag i urval 2018	Antal företag i simulering	Minskat urval	Antal anställda
5-9	592	30	10	20	3 991
10-19	634	30	10	20	8 804
20-49	465	39	10	29	14 052
50-99	160	30	10	20	11 062
100-199	64	64	64	0	8 482
>=200	22	22	22	0	34 167
Totalt	1 937	215	126	89	80 558

I testet minskas urvalet med 89 företag (se tabell 10), vilket kan jämföras med den tidigare simuleringen, där det totalt uteslöts 386 företag i olika branscher i handeln.

Liksom i den första simuleringen för handeln hamnar medelvärdet nära originalserien (se diagram 4). Däremot blir standardavvikelsen i testsimuleringen högre, i genomsnitt 2,5 enheter över hela perioden. I testet uteslöts alltså en fjärdedel så många företag som i den tidigare simuleringen, men standardavvikelsen fördubblades nästan. Det innebär att osäkerheten i skattningarna skulle bli större för handeln som helhet, om uteslutning av företag skulle göras enligt tabell 10, trots att färre företag uteslöts i denna simulering.

Diagram 4 Testsimulering, Handeln, Nettotal, maj 2018 – mars 2019

Slutsatser

Sammantaget var skillnaderna mellan simuleringarna och publicerade resultat oftast små. Detta ledde till beslutet att storleken på urvalet minskades enligt den strategi som föreslogs. Effekten blev att antalet företag i urvalet minskade med cirka 20 procent, från ca 6 800 till ca 5 500 företag i samband med urvalsbytet i maj 2019.

Appendix

För att beräkna nettotalet för en fråga beräknas en viktsumma, V_b^a , för varje svarsalternativ (a) och urvalsbransch (u). Beräkningen görs i två steg. Först räknas varje företags vikt upp med två faktorer, en populationsuppräknings och en svarsbortfallsuppräknings.

Eftersom företag i olika strata valts ut med olika sannolikhet måste man göra en uppräknings till den nivå som svaren skulle ha nått om alla företag i populationen i ett stratum hade tillfrågats. Små urvalsenheter får sin vikt uppräknad, eftersom de representerar övriga små enheter som inte kom med i urvalet. Här används det inverterade värdet av respektive stratums urvalsfraktion, det vill säga den inverterade sannolikheten med vilken varje urvalsobjekt har blivit utvald ($\frac{N_j}{n_j}$, där N_j är totalt antalet företag i storleksklass j i urvalsramen och n_j är antalet företag i urvalet i storleksklass j).

Summan av företagsvikterna för respektive svarsalternativ summeras sedan per stratum (storleksklass och urvalsbransch). Sedan räknas dessa summor upp till den nivå som de skulle ha haft om samtliga företag i stratomet i urvalet hade svarat på frågan ($\frac{\sum_i W_{ij}}{\sum_i W_{ij}^r}$, där $\sum_i W_{ij}$ summa vikt för alla företag i i storleksklass j och $\sum_i W_{ij}^r$ summa vikt för alla företag i i storleksklass j som har besvarat frågan). Detta innebär att om en fråga besvarats av företag vilkas sammanlagda vikt utgör 90 procent (90/100) av hela urvalets vikt (för ett stratum), så multipliceras företagens viktade svar med faktorn 100/90.

$$W_{ij}^* = W_{ij} \sum_{j=1}^S \frac{N_j}{n_j} \frac{\sum_i W_{ij}}{\sum_i W_{ij}^r} \quad (1)$$

Där:

j är storleksklass 1 till S , där S är antalet storleksklasser i branschen

W_{ij} vikt för företag i i storleksklass j

W_{ij}^* uppräknad vikt för företag i i storleksklass j

Därefter beräknas viktsumman för varje svarsalternativ inom urvalsbranschen som:

$$V_u^a = \sum_{j=1}^S \sum_{i=1}^n W_{ij}^{*a} \quad (2)$$

Där:

$\sum_i W_{ij}^{*a}$ summa vikt för alla företag i i storleksklass j för svarsalternativ a

För att räkna ihop viktsummorna till branschaggregat (b), så summeras för varje svarsalternativ (a) alla ingående urvalsbranschers (u) viktsummor.

$$V_b^a = \sum_u V_u^a \quad (3)$$

Därefter beräknas bruttotal, det vill säga andelen företag som svarat respektive alternativ, för branscher och branschaggregat.

$$B^a = \frac{V^a}{V^+ + V^= + V^-} \quad (4)$$

Slutligen beräknas netttotal för frågan som

$$N = B^+ - B^- \quad (5)$$