

# STATISTIKENS FRAMSTÄLLNING

## Industrins förbrukning av inköpta varor (INFI)

### **Ämnesområde**

*Näringsverksamhet*

### **Statistikområde**

*Näringslivets struktur*

### **Produktkod**

*NV0106*

### **Referenstid**

*2020*

### **Kontaktuppgifter**

<b>Statistikansvarig myndighet</b>	<i>Statistiska centralbyrån</i>
<b>Kontaktinformation</b>	
<b>E-post</b>	<i>gumaj.jussubova@scb.se</i>
<b>Telefon</b>	<i>010-479 65 17</i>

## Innehåll

1	Statistikens sammanhang .....	3
2	Undersökningsdesign .....	4
	Inledning Nya VE .....	<b>Fel! Bokmärket är inte definierat.</b>
	Testurval .....	<b>Fel! Bokmärket är inte definierat.</b>
	Insamling .....	<b>Fel! Bokmärket är inte definierat.</b>
	Skattningar .....	<b>Fel! Bokmärket är inte definierat.</b>
	Sammanfattning .....	<b>Fel! Bokmärket är inte definierat.</b>
2.1	Målstorheter .....	4
2.2	Ramförfarande .....	4
2.3	Förfaranden för urval och uteslutning .....	4
2.3.1	Urvalsförfarande .....	4
2.3.2	Uteslutning från insamling (cut-off) .....	7
2.4	Insamlingsförfarande .....	7
2.4.1	Datainsamlingsmetoder .....	7
2.4.2	Mätning .....	8
2.4.3	Bortfallsuppföljning .....	9
2.5	Bearbetningar .....	9
2.6	Granskning .....	10
2.6.1	Granskning under insamlingen .....	10
2.6.2	Granskning av mikrodata .....	10
2.6.3	Granskning av makrodata .....	11
2.6.4	Granskning av redovisning .....	11
2.7	Skattningsförfarande .....	11
2.7.1	Principer och antaganden .....	11
2.7.2	Skattningsförfarande för målstorheter .....	12
2.7.3	Skattningsförfarande för tillförlitlighet .....	23
2.7.4	Röjandekontroll .....	24
3	Genomförande .....	24
3.1	Kvantitativ information .....	24
3.2	Avvikelser från undersökningsdesignen .....	25

## 1 Statistikens sammanhang

Den ekonomiska företagsstatistiken står inför stora förändringar från och med referensår 2022.

Under flera decennier har stora, oftast multinationella, företag i allt större utsträckning utvecklat mer komplexa organisationer för sin verksamhet, och det har fått konsekvenser för statistiken. Genom förändringen kan vi producera mer relevant statistik av högre kvalitet. Utvecklingsarbetet inom detta område pågår.

Information finns och kommer att uppdateras på SCB:s webbplats ([www.scb.se](http://www.scb.se)), under rubriken [Det ekonomisk-statistiska systemet \(scb.se\)](#).

INFI ingår som en betydelsefull del i den ekonomiska statistiken. Det övergripande syftet med undersökningen är att ge Nationalräkenskaperna (NR) bättre underlag för beräkning av bruttonationalprodukten (BNP) i tillgångs- och användningstabeller. Med hjälp av statistiken erhålls den detaljerade information som krävs av underlaget till input- och outputtabeller för industrin. Behovet av inputdata gäller samtliga industriföretag från 0 anställda och uppåt. NR är i dagsläget den enda användaren och undersökningen har ännu aldrig publicerats.

INFI är en årlig undersökning och har som egen undersökning genomförts för åren 1999–2018 och 2020. Olika branscher undersöks olika år enligt ett rullande schema. Eftersom uppgiftslämnandet roterar mellan branscherna innebär det att ett enskilt företag som mest behöver lämna uppgifter vart tredje år. De branscher som undersöks, indelning enligt Svensk Näringsgrensindelning 2007 (SNI 2007), är 05–33. På grund av pandemi genomfördes inte INFI 2019. De branscherna undersöktes i stället årgång 2020 tillsammans med de branscher som ursprungligen var planerade. Tidigare var INFI en delundersökning till Företagens Ekonomi (FEK).

INFI är ett fristående urval, men utgår från FEK:ens population. Den stora skillnaden är att INFI numera går ut till företag som inte blankettundersöks i FEK. Företag med flera verksamhetsenheter (VE) får frågeformulär på våren, men sedan 2008 får resterande företag frågeformulär på hösten. Detta görs bland annat för att underlätta samgranskningen med FEK.

Den del av populationen som inte blankettundersöks modellberäknas. I modellberäkningarna används uppgifter från FEK och Industrins Varuproduktion (IVP).

## 2 Undersökningsdesign

### 2.1 Målstorheter

INFI avser att mäta industrins förbrukning av inköpta varor. På en detaljerad varunivå samlas företagens råvarukostnad in per kalenderår vilket slutligen redovisas på NR:s branschaggregat.

### 2.2 Ramförfarande

INFI:s målpopulation är de råvaror som industriföretagen förbrukar i sin produktion. För att komma åt råvaruförbrukningen tillfrågas de VE:er inom företag med industriell verksamhet om värdet på inköpta och förbrukade råvaror. Så även om målpopulationen är råvaror, utgör VE:er rampopulationen. INFI använder samma ram av VE:er som FEK med begränsningen att INFI endast undersöker VE:er inom tillverkningsindustrin. Rampopulationen utgörs av objekt från Företagsdatabasen (FDB) i november månad det aktuella året (November-SAMU).

För utförligare detaljer rörande vilka objekt som ingår i FDB och andra detaljer avseende ramen se dokumentationen för [FEK](#).

Observationsobjekt och uppgiftskälla är VE, där finns en kontaktperson som fungerar som uppgiftslämnare, med kontaktuppgifter så som email adress, telefonnummer samt företags adress.

För den del av målpopulationen som modellberäknas hämtas uppgifter från andra undersökningar och register på SCB. Observationsobjekt är VE men uppgiftskällorna är FEK och IVP.

### 2.3 Förfaranden för urval och uteslutning

#### 2.3.1 Urvalsförfarande

Urvalet är inget sannolikhetsurval. Urvalsförfarandet kan delas upp i två steg där det första steget kallas för täckningsgradsansatsen och det andra steget för modellfellsansatsen. Företag med fler än 100 anställda eller kostnader för råvaror och handelsvaror större än 200 miljoner kronor anses vara stort och ska alltid vara med i urvalet. Om ett så stort företag inte kommer med i urvalsdragningen med täckningsgradsansatsen eller modellfellsansatsen läggs det till i urvalet manuellt. Det finns ingen cut off-gräns utan samtliga företag i industribranscherna kan komma med i urvalet.

Steg 1 utgår från att alla branscher ska ha en täckningsgrad på minst 70 procent med avseende på variabeln Summa råvaror och handelsvaror (vgr500). Det är variabeln vgr500 i första hand och fgr500 i andra hand som används. Variabeln hämtas från SRU och finns för i princip alla VE:er i ramen (ett visst bortfall förekommer).

Om vgr500 inte finns för aktuellt år används värdet för föregående år,  $vgr500_{t-1}$ .

Utgå från de branschvisa totalsummorna av  $x_i = vgr500$  för bransch  $g$

$$t_{x,g} = \sum_{i \in g} x_i$$

Varje VE:s andel av den industribransch som den ingår i beräknas som kvoten  $x_i/t_{x,g}$  och den önskade täckningsgraden (TG) för den givna branschen ges av att summera upp andelar

$$TG_g = \sum_{i=1}^{\min(n)} (x_i/t_{x,g})$$

Minsta antalet VE:er,  $\min(n)$ , väljs så att  $TG_g \geq 0,7$ . Notera att summeringen alltid börjar med branschens största VE (med högst andel av branschen) och sedan i fallande ordning. Antalet VE:er som behöver summeras för att uppnå 70 procent av branschtotalen varierar mycket mellan olika branscher, i vissa fall behövs bara något eller ett par företag och i vissa fall krävs flera hundra. Om ett företag själv står för mer än 70 procent av sin bransch väljs alltså bara ett enda företag.

Att ha 70 procent täckningsgrad garanterar dock inte att kvaliteten i skattningarna blir densamma för alla branscher. Därför sker i Steg 2 en omfördelning av urvalsenheter mellan homogena och heterogena branscher med hjälp av modellfelsansatsen. I en homogen bransch behövs färre utvalda företag för att beskriva hela branschen på ett bra sätt än i en heterogen bransch. Således räcker det med att undersöka färre VE:er i en homogen bransch till förmån för fler VE:er i heterogena branscher. Principen med 70 procents täckningsgrad i alla branscher frångås därmed, men endast då en god kvalitet kan garanteras. På detta sätt är förhoppningen att få en smartare fördelning av urvalsobjekt.

Med homogena och heterogena branscher avses branschens fördelning av insatsvaror över storleksklasser. För att skatta denna fördelning används produktionsvaror från undersökningen Industrins Varuproduktion (IVP). I IVP undersöks alla industriella VE:er med över 20 anställda. Anledningen till att IVP:s produktionsvaror används i stället för exempelvis en tidigare produktionsomgång av INFI är att IVP omfattar en mycket större svarmängd och kan således ge mer pålitliga skattningar. En annan fördel är att uppgifterna i IVP är mer aktuella och således mer relevanta. Sedan antas även att värdet av VE:ens insatsvaror och produktionsvaror är högt korrelerade. Detta testas i en

strukturjämförelse mellan IVP och tidigare produktionsomgångar av INFI.

För att påvisa att en bransch är homogen eller heterogen skattas ett fel från produktionen i IVP. Utgå ifrån

$$e_{ij} = Y_{ij} - \hat{\beta}_j * X_i$$

där  $Y_{ij}$  är VE i:s faktiska produktionsvärde av vara j och  $X_i$  är VE i:s totala produktionsvärde. Vidare gäller att

$$\hat{\beta}_j = \frac{\sum_{j=1}^n y_j}{\sum_{j=1}^N y_j} \quad \text{och} \quad \sum_g \hat{\beta}_j = 1$$

Täljaren summerar värdet av vara j i en given bransch g och nämnaren summerar värdet av samtliga varor i bransch g. I modellansatsen skattas alltså en modell där  $\hat{\beta}_j$  är vara j:s andel av hela branschens värde.  $\hat{\beta}$  fungerar således som en fördelningsnyckel av varor för VE:er i en specifik bransch medan  $e$  är ett mått på hur väl modellen har lyckats anpassa (skatta  $\hat{\beta}$ ). Summeras samtliga  $\hat{\beta}_j$  för en bransch täcks 100 procent av branschens varor in.

För att summan av felen inte ska bli större bara för att en bransch är stor beräknas ett relativt fel. Summera upp felen för bransch g och dividera med g:s totala produktionsvärde enligt

$$FKV_{IVP,g} = \frac{\sum_g \sum_g \sqrt{e_{ij}^2/2}}{\sum_g \sum_g y_{ij}}$$

Felkvoten ger ett relativt mått på variationen i branschen. En hög felkvot betyder att felet är stort i relation till branschens storlek. Tillsammans med täckningsgraden beräknas det relativa felet för bransch g som

$$RF_{INFI,g} = (1 - TG_g) * FKV_{IVP,g}$$

Målet är att få ett minsta gemensamma relativa fel för samtliga branscher och detta fås genom att justera upp täckningsgraden i branscher där det relativa felet är högt. På motsvarande sätt justeras täckningsgraden ner om det relativa felet redan från början är lågt.

Ansatsen med att justera täckningsgraden efter det relativa felet används endast då variationen av varuproduktion i IVP överensstämmer med variationen av insatsvaror i INFI. En variansanalys utförs för att strukturjämföra IVP och tidigare årsversioner av INFI. Varierar fördelningen av varor inom företag ungefär lika mycket i IVP som i INFI så används modellansatsen med dess justering. Om ingen bra likhet kan påvisas används istället täckningsgradsansatsen.

Strukturjämförelsen är alltså en variansanalys och kan beskrivas som en jämförelse av den relativa medelkvadratsumman av felet (RMSE) mellan IVP och INFI. För både IVP och INFI räknas RMSE för varje bransch som

$$RMSE = \frac{SSE/(a * n - n)}{SST/(a * n - 1)}$$

där de n stycken varorna ses som behandling och a stycken VE:er ses som replikat. Detta innebär något förenklat att RMSE ger ett mått på hur stor del av den totala variationen ( $y_{ij} - \bar{y}_{..}$ ) som är variation mellan VE:er för varje vara ( $y_{ij} - \bar{y}_{i.}$ ). En mer utförlig och teknisk framställning av variansanalys finns i kapitel 3 i *Design and analysis of experiments* (Douglas C. Montgomery, 1976 3rd edition). Då RMSE beräknats för varje bransch i både IVP och INFI jämförs deras värden. Ett gränsvärde ( $\delta$ ) kan anges av användaren för att bestämma hur olika variationerna i IVP och INFI får vara. Sätt

$$|RMSE_{IVP} - RMSE_{INFI}| \leq \delta$$

Ett högt  $\delta$  tillåter större skillnader i variation och kommer därmed att justera fler urvalsstorlekar med hjälp av modellansatsen. Risken med detta är att den modell som används och grundar sig på IVP inte är särskilt bra på att förklara responsvariabeln i INFI som är det ursprungliga målet. Då RMSE ofta är ett tal mellan 0 och 1 kan  $\delta$  initialt väljas godtyckligt till 0,3 och sedan justeras om inte önskat resultat uppnås.

### 2.3.2 Uteslutning från insamling (cut-off)

Det görs ingen uteslutning från insamling. Den andel av Summa råvaror och handelsvaror som inte täcks av urvalet modellberäknas.

## 2.4 Insamlingsförfarande

### 2.4.1 Datainsamlingsmetoder

I INFI görs både direktinsamling och en insamling av uppgifter från andra undersökningar och register. Direktinsamling genomförs på det urval som dragits medan resterande del modellberäknas med hjälp av uppgifter från undersökningarna FEK och IVP för det aktuella referensåret. Motiveringen till kombinationen är att med direktinsamlade uppgifter som förväntas omfatta 70 procent av målstorheten i varje bransch, säkerställa en hög kvalitet på skattningarna. Den återstående delen som modellberäknas kompletterar direktinsamlat material.

Uppgifterna från FEK och IVP hämtas från respektive undersöknings produktionsdatabas under våren år T+2. För referensår 2020 hämtades uppgifterna i februari 2022.

I maj år T+1 skickas introduktionsbrev ut till företag med flera VE:er, företag med sent bokslut får introduktionsbrev i augusti år T+1 och till resterande företag skickades introduktionsbrev ut i oktober år T+1. Introduktionsbrev innehåller webbadress, användarnamn samt lösenord till det elektroniska frågeformuläret (se Bilaga 2). Eftersom detaljerade uppgifter om råvaruförbrukningen många gånger saknas i företagens bokföringssystem har uppgiftslämnandet ofta inneburit stora svårigheter för många företag. Tidigare ställdes frågeformulären vanligtvis till ekonomiansvarig, men numera är det önskvärt att inköpsansvarig lämnar uppgifter. Ofta tvingas uppgiftslämnaren att göra uppskattningar.

En skriftlig påminnelse skickas ut till de uppgiftslämnare som inte inkommit med uppgifter när sista svarsdatum har passerat. Ytterligare 2-3 påminnelser görs om uppgifter inte inkommer. De som då fortfarande inte har svarat i undersökningen kontaktas av medarbetare på SCB:s insamlingsavdelning per telefon. I många fall krävs flera telefonkontakter.

De företag som i slutet av produktionsperioden inte har svarat på undersökningen imputeras. Detta görs antingen med fördelning från samma objekt från en tidigare produktionsomgång eller med fördelning från ett eller flera objekt i årets omgång med likartad produktion. Går ingen av dessa metoder att genomföra görs imputering utifrån medelfördelningen inom aktuell bransch.

#### **2.4.2 Mätning**

Från årgång 2007 samlas uppgifterna i första hand in via SCB:s webbaserade statistikinsamlingsverktyg (SIV). I SIV finns, förutom fasta variabler, förslag på varugrupper som är vanliga inom respektive branschgrupp. Efter SNI-omläggningen i årgång 2009 blev de tidigare 109 branschspecifika frågeformulären 106 stycken, alla med olika förtryckta varugrupper (se Bilaga 1 för frågeformulär som skickas till VE:er inom bransch 05-09 - Utvinning av malm och mineral). Det finns även möjlighet för uppgiftslämnaren att lägga till egna varor. För samtliga frågeformulär finns automatiska kontroller som signalerar om uppgiftslämnaren fyller i inkonsekventa uppgifter i formuläret.

Det detaljerade innehållet i statistiken beskrivs i SCB:s mikrodataregister (MetaPlus).

Uppgifterna från FEK samlas in via sju olika typer av frågeformulär som skickas ut vid olika tillfällen under produktionsprocessen. Uppgifterna från IVP samlas in via SIV. Frågeformuläret består av fyra avdelningar. Mer information finns i respektive undersöknings dokumentation.



### 2.4.3 Bortfallsuppföljning

Det är enligt lag obligatoriskt att delta i undersökningen. Uppgiftsskyldighet gäller enligt lagen (2001:99) om den officiella statistiken. Statistiken regleras även av förordningen (2001:100) om den officiella statistiken och SCB:s föreskrifter (SCB-FS 2017:5 och SCB-FS 2018:11). För INFI 2020 var det ovägda bortfallet ca 21 procent. Frågeformulär som inte inkommit efter upprepade påminnelser imputeras med hjälp av INFI:s produktionssystem Årsam.

## 2.5 Bearbetningar

Det finns tre typer av imputeringar i INFI: manuell imputering, tvillingimputering och maskinell imputering. Vid all imputering har totalen av råvaror och förnödenheter, insamlad eller imputerad, varit känd antingen från tabellen GemVE i FEK-databasen eller från företagets årsredovisning. Imputeringen i INFI har endast syftet att fördela ut denna total.

Vid manuell imputering lägger granskaren själv in värden i INFI:s produktionssystem Årsam. Om företaget har varit med i en tidigare omgång kan denna information användas vid imputering efter att avstämning har gjorts mot IVP eftersom produktionen kan förändras mellan åren. Annat som kan vara till hjälp är att diskutera fram en rimlig fördelning med uppgiftslämnaren, kunskap om branschen eller information från företagets webbplats.

Tvillingmetoden jämför produktionen, hämtad från IVP, hos objektet som ska imputeras med produktionen hos andra objekt som har inkommit med uppgifter till INFI. Antingen används ett objekt ensamt eller så viktas flera objekt som har samma struktur på produktionen ihop. Objektet eller objektens insatsvarufördelning används sedan i imputeringsarbetet. Sedan anges produktionsvikter och ett underlag skapas som har liknande produktionsstruktur som objektet som ska imputeras.

Maskinell imputering används för värden som inte har imputerats med hjälp av de andra metoderna. Medelvärdesimputering inom branschgruppen används eftersom imputering med fördelning från tidigare år, om sådan finns, har använts vid den manuella imputeringen. Vid medelvärdesimputering inom branschgruppen undersöks hur många objekt det finns med fördelning i årets produktionsomgång av INFI, detta görs på 5-siffer SNI. Finns minst tre objekt summeras värdena per variabel eller varugrupp inom branschgruppen och fördelningen som blir resultatet av detta appliceras på det objekt och det avsnitt som ska imputeras. Finns det färre än tre objekt används i första hand 4-siffer SNI och därefter 3-

siffer SNI upp till branschgruppsnivå tills en nivå med minst tre objekt hittas.

## **2.6 Granskning**

### **2.6.1 Granskning under insamlingen**

Värden från de inkomna elektroniska frågeformulären registreras automatiskt och granskas därefter i INFI:s produktionssystem Årsam. Det finns inbyggda kontroller i frågeformulären som bl.a. summeringskontroller och vissa rimlighetskontroller. I de fall uppgiftslämnaren har fyllt i stora värden ber vi dem kommentera dessa uppgifter. Tanken med kontrollerna är att uppgiftslämnaren redan vid inlämningstillfället ska göras uppmärksam på om de lämnade värdena avviker och därmed kunna rätta eventuella fel innan frågeformulären skickas in. Helt ofyllda blanketter kan inte skickas in. Ofullständiga blanketter hamnar på en lista och återkontakt tas med företagen. Företag som skickar in uppgifter via pappersblankett registreras manuellt samt kontrolleras i samband med brevöppningen. Bristfälligt ifyllda blanketter skickas tillbaka till företagen för komplettering.

### **2.6.2 Granskning av mikrodata**

I produktionssystemet Årsam finns även en flik där granskaren kan kontrollera vad som finns registrerat i IVP, IVP-variabler. Här redovisas intäkter och varor för aktuell VE och år samt tidigare år.

Om en VE, definierad med hjälp av organisationsnummer och löpnummer, var med i INFI då branschen undersöktes senast och dessutom lämnade värden så visas dessa i en kolumn bredvid årets värden.

Efter att inkomna värden har registrerats och granskats körs ett kontrollprogram. I detta program finns 5 typer av kontroller:

1. Uppgifter saknas som ska eller bör finnas med.
2. Summeringskontroll.
3. Rimlighetskontroll.
4. Samgranskning mot FEK och IVP.
5. Samgranskning mot tidigare lämnade uppgifter.

De kontroller som faller ut kan åtgärdas genom rättning eller genom att granskaren accepterar kontrollen. Vissa kontroller som till exempel summeringskontroller kan dock inte accepteras utan måste rättas upp.

2022-03-07

### **2.6.3 Granskning av makrodata**

Makrodata granskas branschvis inför leveransen till NR. Skillnader och förändringar i förbrukningen av råvaror och strukturen sedan förra gången branscherna undersöktes (2016–2017) förklaras på VE- och produktnivå. Fokus ligger främst på VE:n med förändring större än 4–5 procent. I vissa branscher kan man behöva titta på ännu lägre procentuella förändringar. Till hjälp finns produktionsvärden i IVP, varor som importerats i UH samt värden i FEK. Om förändringen stämmer är det viktigt att dokumentera vad förändringen beror på.

### **2.6.4 Granskning av redovisning**

Slutliga resultatet sammanställs och summeringskontrolleras innan leverans till NR.

## **2.7 Skattningsförfarande**

Skattningsförfarandet är oförändrat men skattningarna levererades till NR i ett nytt format. Formatet är utformat för att passa det nya systemet NR-fiber. Nya koder för näringsgrenar och produktgrupper används. Skattningarna laddas upp i tabellen INFI\_NRprodgrp på databasen NRFiber\_NS.

### **2.7.1 Principer och antaganden**

För att få fram totalen för råvaror och förnödenheter har, förutom de insamlade och imputerade fördelningarna i blankettpopulationen, även modellberäknade fördelningar lämnats för objekt som lämnar SRU. Modellberäkningarna utgår ifrån antaganden att det finns ett generellt samband mellan vilka varor som produceras och vilka varor som förbrukas. Den metod som har utvecklats försöker därför att ta hänsyn till skillnader i produktionsstrukturen mellan verksamheter i blankett- och SRU-populationen.

För alla objekt som ska modellberäknas har det totala värdet för råvaror och förnödenheter samt legoarbete och underentreprenad hämtats från FEK (variablerna v3003, v3004, v3005 och v3021). För att få fram värden på dessa variabler för icke blankettundersökta företag har variabeln vgr503 från SRU använts. Därefter har kostnader för energi med mera som inte ska ingå räknats bort under antagandet att SRU-företagen har samma kostnadsfördelning som blankettföretagen. Uppgifter som har samlats in från de blankettundersökta objekten i INFI har naturligtvis varit en viktig informationskälla. Uppgifter om producerade varor har hämtats från IVP för företag som ingår i denna undersökning.

Produktionen summeras för de objekt som ska modellberäknas per 5-siffer SNI. Med hjälp av en algoritm skapas sedan ett underlag bestående av objekt som har inkommit med uppgifter om

råvaruförbrukningen i INFI och vars produktionsfördelning så väl som möjligt överensstämmer med den produktionsfördelning som finns bland de objekt som ska modellberäknas.

Den algoritm som används för att söka efter lämpliga objekt som ska ingå i underlaget börjar med att söka efter objekt som har en produktion som överensstämmer med de modellberäknade på finaste varunummernivå. Saknas överensstämmelse på denna nivå för någon vara söker algoritmen igen på en något grövre varunummernivå. Denna procedur upprepas på allt grövre varunummernivåer till dess att ett komplett underlag fås. För att begränsa risken för att stora mätfel ska slå igenom i underlaget är det inte bara det objekt som matchar bäst som får ingå i underlaget utan även andra objekt som matchar nästan lika bra. På alla varunummernivåer söks alltid objekt bland dem som tillhör samma NR-bransch som aktuell 5-siffer SNI. Detta för att förhindra att det dyker upp för många udda insatsvaror i redovisningen av de modellberäknade värdena.

Andra metoder än den ovannämnda algoritmen för att ta fram underlag har också tillämpats. Dessa utgår från de uppgifter som blankettundersökta objekt lämnat i INFI. Ena metoden tar fram ett underlag baserat på totalfördelningen i närmaste bransch (enligt SNI 2007) för de objekt som modellberäknas. Den andra metoden får fram ett underlag genom att beräkna medelfördelningen där stora och små objekt har samma vikt och applicerar denna.

Ett mått på hur väl de olika underlagens produktion matchar de modellberäknade objektens produktion på olika varunummernivåer tas sedan fram för de olika metoderna. Den metod som ger det underlag som i genomsnitt matchar bäst på 8-, 6-, 4- och 2-siffer KN-nummernivå väljs ut för varje bransch (5-siffer SNI).

### **2.7.2 Skattningsförfarande för målstorheter**

De formler som beskrivs nedan är tillämpliga för tvillingimputering med INFI:s produktionssystem Årsam och modellberäkningen av insatsvarufördelningen i INFI.

I flertalet av de metoder som praktiseras för modellberäkning och imputering försöker ett objekt, Imputeringskandidaten (IK), tillskrivas värden genom att titta på värden hos andra objekt, Underlagskandidater (UK), där fullständig information finns. En erkänd metod för detta är *Nearest Neighbour imputation* där imputering för IK:er görs med värden från den UK som ligger närmast (mest liknar) IK:en. Nearest Neighbour-metoden har inte använts för modellberäkningar i INFI. Istället har en av metoderna, Bästa matchning-metoden, försökt att vikta ihop värden från fler Nearest Neighbour-UK:er.

2022-03-07

Genom att vikta ihop värdet från flera UK:er till ett underlag kan följande fördelar uppnås:

1. Beroendet av att en enskild UK:s värden är korrekta blir inte lika stort.
2. Två eller flera UK:er i kombination kan ha egenskaper som bättre överensstämmer med en IK än vad varje UK gör för sig.

I INFI:s fall kombineras  $n$  UK:er ihop som producerar  $m$  olika varor så att ett underlag med en produktionsmix fås som så bra som möjligt överensstämmer med den som finns hos IK:en.

**Formel 1. Linjär kombination av flera UK : er till ett underlag, produktionssidan**

$$\begin{pmatrix} \text{PrViktUK}_1 \\ \text{PrViktUK}_2 \\ \vdots \\ \text{PrViktUK}_n \end{pmatrix}^T \begin{pmatrix} \text{PrAndV}_1\text{UK}_1 & \text{PrAndV}_2\text{UK}_1 & \cdots & \text{PrAndV}_m\text{UK}_1 \\ \text{PrAndV}_1\text{UK}_2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \text{PrAndV}_1\text{UK}_n & \cdots & \cdots & \text{PrAndV}_m\text{UK}_n \end{pmatrix} = \begin{pmatrix} \text{PrAndV}_1\text{U} & \text{PrAndV}_2\text{U} & \cdots & \text{PrAndV}_m\text{U} \end{pmatrix}$$

där :

$\text{PrViktUK}_i$  = Produktionsvikt för  $\text{UK}_i$

$\text{PrAndV}_j\text{UK}_i$  = Vara  $j$ :s andel av den totala produktionen för underlagskandidat  $i$

$\text{PrAndV}_j\text{U}$  = Vara  $j$ :s andel av underlagets totala produktion

$$0 \leq \text{PrViktUK}_i \leq 1 \text{ för } i = 1, 2, \dots, n \text{ och } \sum_{i=1}^n \text{PrViktUK}_i = 1$$

$$0 \leq \text{PrAndV}_j\text{UK}_i \leq 1 \text{ för } i = 1, 2, \dots, n \text{ } j = 1, 2, \dots, m \text{ och } \sum_{j=1}^m \text{PrAndV}_j\text{UK}_i = 1 \text{ för } i = 1, 2, \dots, n$$

Genom att tillämpa denna formel kommer ett underlag (U) fås vars fiktiva produktionsandelar av olika varor är en viktad linjär kombination av alla ingående UK:ers produktionsandelar. Villkoren som ställs för vikterna ( $\text{PrViktUK}$ ) är att de ska summera till 1 över alla UK:er och måste ligga mellan 0 och 1 för varje enskild UK samt att en enskild UK:s produktionsandelar för respektive vara summerar till 1 och ligger mellan 0 och 1 gör så att underlagets produktionsandelar ( $\text{PrAndVU}$ ) också lever upp till dessa villkor.

För att se hur bra den fiktiva produktionen hos U:et motsvarar den faktiska produktionen hos IK:en används ett mått som har döpts till gemensam produktion (GP).

**Formel 2. Beräkning av gemensam produktion**

$$GP = \min(\text{PrAndV}_1 U, \text{PrAndV}_1 \text{IK}) + \min(\text{PrAndV}_2 U, \text{PrAndV}_2 \text{IK}) + \dots + \min(\text{PrAndV}_m U, \text{PrAndV}_m \text{IK})$$

där :

$\min(a, b)$  = Minsta värdet av a och b

$\text{PrAndV}_j U$  = Underlagets produktionsandel av vara j

$\text{PrAndV}_j \text{IK}$  = Imputeringskandidatens produktionsandel av vara j

$$0 \leq \text{PrAndV}_j U \leq 1 \text{ för } j = 1, 2, \dots, m \text{ och } \sum_{j=1}^m \text{PrAndV}_j U = 1$$

$$0 \leq \text{PrAndV}_j \text{IK} \leq 1 \text{ för } j = 1, 2, \dots, m \text{ och } \sum_{j=1}^m \text{PrAndV}_j \text{IK} = 1$$

GP kommer att ligga någonstans mellan 0, för U som inte har någon vara gemensam med IK:en, och 1 för U vars fiktiva produktion andelsmässigt exakt motsvarar den produktion som finns hos IK:en. GP kan användas för att jämföra hur olika underlag passar en viss IK eller som ett slags "kvalitetsmått" på hur bra ett specifikt underlag passar i modellberäkningen.

IK:s insatsvarufördelning modellberäknas med hjälp av den fiktiva insatsvarufördelningen som behövs för att skapa U:ets fiktiva produktion. Ett sätt att göra detta på är att i U:et vikta UK:ernas insatsvarufördelning på samma sätt som vid viktningen av U:ets produktionsfördelning. Detta förfarande tar dock inte hänsyn till att produktion av olika varor är olika insatsvaruintensiva. För att korrigera för detta räknas varje produktionsvikt som väljs vid skapandet av U:et om till en förbrukningsvikt (FbViktUK).

**Formel 3. Omräkning av produktionsvikter till förbrukningsvikter**

$$\text{FbViktUK}_i = \frac{\text{PrViktUK}_i \times \frac{\text{Fb}_{\text{tot}} \text{UK}_i}{\text{Pr}_{\text{tot}} \text{UK}_i}}{\sum_{i=1}^n \left( \text{PrViktUK}_i \times \frac{\text{Fb}_{\text{tot}} \text{UK}_i}{\text{Pr}_{\text{tot}} \text{UK}_i} \right)}$$

där :

$\text{FbViktUK}_i$  = Förbrukningsvikt för underlagskandidat i

$\text{PrViktUK}_i$  = Produktionsvikt för underlagskandidat i, väljs då underlag skapas

$\text{Fb}_{\text{tot}} \text{UK}_i$  = Underlagskandidat i : s totala förbrukningsvärde

$\text{Pr}_{\text{tot}} \text{UK}_i$  = Underlagskandidat i : s totala produktionsvärde

$$0 \leq \text{PrViktUK}_i \leq 1 \text{ för } i = 1, 2, \dots, n \text{ och } \sum_{i=1}^n \text{PrViktUK}_i = 1$$

2022-03-07

Efter att förbrukningsvikter har skapats för alla UK:er i U:et kan U:ets fiktiva insatsvarufördelning beräknas.

**Formel 4. Linjär kombination av flera UK : er till ett underlag, förbrukningssidan**

$$\begin{pmatrix} \text{FbViktUK}_1 \\ \text{FbViktUK}_2 \\ \vdots \\ \text{FbViktUK}_n \end{pmatrix}^T \begin{pmatrix} \text{FbAndV}_1\text{UK}_1 & \text{FbAndV}_2\text{UK}_1 & \cdots & \text{FbAndV}_m\text{UK}_1 \\ \text{FbAndV}_1\text{UK}_2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \text{FbAndV}_1\text{UK}_n & \cdots & \cdots & \text{FbAndV}_m\text{UK}_n \end{pmatrix} = \begin{pmatrix} \text{FbAndV}_1\text{U} & \text{FbAndV}_2\text{U} & \cdots & \text{FbAndV}_m\text{U} \end{pmatrix}$$

där :

$\text{FbViktUK}_i$  = Förbrukningsvikt för  $\text{UK}_i$

$\text{FbAndV}_j\text{UK}_i$  = Vara  $j$ : s andel av den totala förbrukningen för underlagskandidat  $i$

$\text{FbAndV}_j\text{U}$  = Vara  $j$ : s andel av underlagets totala förbrukning

$0 \leq \text{FbViktUK}_i \leq 1$  för  $i = 1, 2, \dots, n$  och  $\sum_{i=1}^n \text{FbViktUK}_i = 1$

$0 \leq \text{FbAndV}_j\text{UK}_i \leq 1$  för  $i = 1, 2, \dots, n$   $j = 1, 2, \dots, m$  och  $\sum_{j=1}^m \text{FbAndV}_j\text{UK}_i = 1$  för  $i = 1, 2, \dots, n$

Efter att Formel 4 har tillämpats fås en insatsvarufördelning fram som lätt går att applicera på IK:s kända totala förbrukning av råvaror och förnödenheter från FEK.

Det totala produktionsvärdet och totala förbrukningsvärdet har generellt inte någon betydelse vid skapandet av underlag. Ibland är det dock av intresse att se förhållandet mellan någon eller båda av dessa storheter hos underlaget och imputeringskandidaten. Begreppet uppräkningsstal definieras med hur många gånger en storhet i underlaget måste räknas upp (vilken faktor den ska multipliceras med) för att motsvara samma storhet hos imputeringskandidaten.

Uppräkningstalet för produktionssidan beräknas med följande formel:

**Formel 5. Uppräkningstal**

$$\text{UpprPr}_{\text{tot}} \text{U} = \sum_{i=1}^n \left( \text{PrViktUK}_i \frac{\text{Pr}_{\text{tot}} \text{IK}}{\text{Pr}_{\text{tot}} \text{UK}_i} \right)$$

där :

$\text{UpprPr}_{\text{tot}} \text{U}$  = Uppräkningstal för totala produktionsidan hos underlaget

$\text{PrViktUK}_i$  = Produktionsvikten för underlagskandidat  $i$

$\text{Pr}_{\text{tot}} \text{IK}$  = Imputeringskandidatens totala produktion

$\text{Pr}_{\text{tot}} \text{UK}_i$  = Underlagskandidat  $i$  : s totala produktion

För förbrukningssidan beräknas uppräkningsstalet analogt. Uppräkningsstal syns i fliken *IVP/VE* i INFI:s produktionssystem Årsam där tvillingimputeringar görs. Skiljer sig uppräkningsstalen mellan produktionssidan och förbrukningssidan åt väldigt mycket kan det antas att underlaget och imputeringskandidaten har stora skillnader i produktionsprocessen, till exempel stora skillnader i bearbetningsgrad på insatsvarorna.

Individuella IK:er har inte modellberäknats var för sig. Istället slås produktionen samman för alla IK:er som har samma SNI-kod. Denna grupp av IK:er har sedan behandlats på samma sätt som en individuell IK behandlas i formlerna ovan, den totala förbrukningen av råvaror och förnödenheter har slagits samman. I varje grupp av IK:er finns också ett antal VE:er vars produktion är okänd på grund av att IVP har en cut-off. Dessa VE:er lämnar inget bidrag till insatsvarufördelningen, men är ändå med och bidrar till den totala förbrukningen av råvaror och förnödenheter. Följden blir att de teoretiskt modellberäknas med samma insatsvarufördelning som gäller för objekten ovanför IVP:s cut-off.

Tre metoder har använts för att skapa underlag. Dessa beskrivs nedan var för sig:

#### 1. Medelfördelningsmetoden

U skapas av alla UK:er som har samma SNI-kod som modellberäkningsgruppen enligt Formel 4. Saknas helt UK:er som har samma SNI-kod på 5-siffernivå som modellberäkningsgruppen tas istället alla UK:er med samma SNI-kod på 4-siffernivå med. Saknas UK:er på 4-siffernivå tas alla UK:er på 3-siffernivå. Saknas dessa tas slutligen alla UK:er på 2-siffernivå. Samtliga förbrukningsvikter för de UK:er som ingår i U:et görs lika stora.

#### **Formel 6. Förbrukningsvikter för medelfördelningsmetoden**

$$FbViktUK_i = \frac{1}{n}$$

där :

$FbViktUK_i$  = Förbrukningsvikt för  $UK_i$

n = Antalet ingående UK i U

För att kunna beräkna GP måste en omräkning göras från förbrukningsvikter till produktionsvikter. Detta görs analogt som i Formel 3 med ombytta platser för produktionsvikter och förbrukningsvikter, produktionsvärden och förbrukningsvärden.



**Formel 7. Omräkning av förbrukningsvikter till produktionsvikter**

$$PrViktUK_i = \frac{FbViktUK_i \times \frac{Pr_{tot} UK_i}{Fb_{tot} UK_i}}{\sum_{i=1}^n \left( FbViktUK_i \times \frac{Pr_{tot} UK_i}{Fb_{tot} UK_i} \right)}$$

där :

$FbViktUK_i$  = Förbrukningsvikt för underlagskandidat i

$PrViktUK_i$  = Produktionsvikt för underlagskandidat i

$Fb_{tot} UK_i$  = Underlagskandidat i : s totala förbrukningsvärde

$Pr_{tot} UK_i$  = Underlagskandidat i : s totala produktionsvärde

$0 \leq FbViktUK_i \leq 1$  för  $i = 1, 2, \dots, n$  och  $\sum_{i=1}^n FbViktUK_i = 1$

**2. Medelvärdesmetoden**

Denna metod fungerar på samma sätt som medelfördelningsmetoden förutom att förbrukningsvikternas storlek är proportionell mot UK:ens förbrukningsvärdes relativa storlek.

**Formel 8. Förbrukningsvikter för medelvärdesmetoden**

$$FbViktUK_i = \frac{Fb_{tot} UK_i}{\sum_{i=1}^n Fb_{tot} UK_i}$$

där :

$FbViktUK_i$  = Förbrukningsvikt för  $UK_i$

$n$  = Antalet ingående UK i U

**3. Bästa matchning-metoden**

Detta är den enda metod som aktivt försöker att skapa ett underlag som matchar produktionen i modellberäkningsgruppen. Metoden bygger på en algoritm som stegvis bygger upp ett underlag genom att ta med nya UK:er eller öka på ingående UK:ers vikter genom en jämförelse med modellberäkningsgruppens produktionsfördelning på allt grövre varunivåer. Metoden ska försöka ge ett högt värde på GP samtidigt som det får en bra spridning på UK:erna i underlaget (med tanke på eventuella mätfel är det inte bra om en UK blir alltför dominerande i underlaget).

Efter varje steg läggs den del av underlaget som skapats i detta steg till den del av underlaget som skapats i tidigare steg. Det beräknas hur stor del av underlaget som återstår att skapa. Detta värde kallas för restvikt. Det beräknas även hur stora andelar av modellgruppens

producerade varor som täcks av underlaget och dessa andelar dras av från modellgruppens andelar inför nästa steg.

Som UK väljs alla VE:er med inkomna varuuppgifter i INFI som har samma SNI-kod. Var och en av dessa UK:ers varufördelning jämförs med modellberäkningsgruppens resterande fördelning vara för vara. Jämförelsen går till så att det räknas fram hur stor vikt som behövs för att få max värde på GP för en specifik vara.

**Formel 9. Beräkning av den produktionsvikt som maximerar GP - värdet för en vara**

$$\text{PrViktV}_j\text{UK}_i = \min(\text{PrAndV}_j\text{UK}_i / \text{PrAndV}_j\text{IK}, 1)$$

där :

$\text{PrViktV}_j\text{UK}_i$  = Produktionsvikt som maximerar GP - värdet för vara j och UK i

$\min(a, b)$  = Minsta värdet av a och b

$\text{PrAndV}_j\text{UK}_i$  = UK i : s produktionsandel av vara j

$\text{PrAndV}_j\text{IK}$  = Gruppen av modellberäkningskandidaters resterande produktionsandel av vara j

När denna vikt har räknats ut för alla varor för en enskild UK väljs den som är minst bland alla k vikter  $> 0$ .

**Formel 10. Val av produktionsvikt för en UK**

$$\text{PrViktUK}_i = \min(\text{PrViktV}_1\text{UK}_i, \text{PrViktV}_2\text{UK}_i, \dots, \text{PrViktV}_k\text{UK}_i)$$

där :

$\text{PrViktUK}_i$  = Vald produktionsvikt för UK<sub>i</sub>

$\text{PrViktV}_j\text{UK}_i$  = Produktionsvikt som maximerar GP - värdet för vara j och UK i

$\text{PrViktV}_j\text{UK}_i > 0$

Med denna vikt beräknas sedan den enskilda UK:ens bidrag till GP.

**Formel 11. Beräkning av en enskild UK : s bidrag till GP**

$$\text{GPUK}_i = \min(\text{PrAndV}_1\text{UK}_i, \text{PrAndV}_1\text{IK}) + \min(\text{PrAndV}_2\text{UK}_i, \text{PrAndV}_2\text{IK}) + \dots + \min(\text{PrAndV}_m\text{UK}_i, \text{PrAndV}_m\text{IK})$$

där :

$\text{PrAndV}_j\text{UK}_i$  =  $\text{PrViktUK}_i \times \text{PrAndV}_j\text{UK}_i$

$\text{PrViktUK}_i$  = Vald produktionsvikt för underlagskandidat i

$\text{PrAndV}_j\text{UK}_i$  = Underlagkandidats i produktionsandel av vara j

$\text{PrAndV}_j\text{IK}$  = Gruppen av modellberäkningskandidatens resterande produktionsandel av vara j

$0 \leq \text{PrAndV}_j\text{UK}_i \leq 1$  för  $j = 1, 2, \dots, m$  och  $\sum_{j=1}^m \text{PrAndV}_j\text{U} \leq 1$

$0 \leq \text{PrAndV}_j\text{IK} \leq 1$  för  $j = 1, 2, \dots, m$  och  $\sum_{j=1}^m \text{PrAndV}_j\text{IK} = 1$

För att se vilka UK:er som ger mest till GP med minst vikt beräknas något av dessa två jämförelsetal: en effektivitetskvot eller en relativ effektivitetskvot. Dessa beräkningar görs på följande sätt:

**Formel 12. Beräkning av jämförelsetal**

$$jmfrlstal_i^{eff} = GPUK_i / PrViktUK_i \quad (\text{Alt. I, effektivitetskvot})$$

$$jmfrlstal_i^{rel} = GPUK_i / GPUK_{max} \quad (\text{Alt. II, relativ effektivitetskvot})$$

där :

$GPUK_i$  = Underlagskandidats i bidrag till gemensam produktion

$PrViktUK_i$  = Produktionsvikt för underlagskandidat i

$GPUK_{max}$  = Maximala bidraget till gemensam produktion bland alla UK

Alla UK:er vars jämförelsetal överstiger ett visst värde kommer att tas med i underlaget. Vilket jämförelsetal som används och hur stort gränsvärdet är ska betraktas som parametrar som kan ställas in för varje steg. Innan de UK:er som har kvalificerat sig kommer med i underlaget görs ytterligare ett par justeringar.

Eftersom flera UK:er kan kvalificera sig på grund av att de har samma varor kan en eller flera produktionsandelar bli mer än täckta. Detta skulle i sådana fall innebära att produktionsandelen för någon annan vara i modellberäkningsgruppen redan i detta steg blockeras från att helt täckas av underlaget. För att undvika detta görs en proportionell justering av produktionsvikterna så att alla UK:er som kvalificerat sig till underlaget tillsammans inte täcker mer än nödvändigt av alla varor.

**Formel 13. Justering av vikter för att undvika överskott av någon vara i underlaget**

$$\text{Overskottsjust}_{j,i} = \frac{PrAndV_{j,IK}}{\sum_{i=1}^n (PrViktUK_i \times PrAndV_{j,UK_i})}$$

$$PrViktUK_i^{\text{öjust}} = \min(\text{Overskottsjust}_{1,i}, \text{Overskottsjust}_{2,i}, \dots, \text{Overskottsjust}_{m,i}) \times PrViktUK_i$$

där :

$\text{Overskottsjust}_{j,i}$  = Justeringsfaktor m.a.p. vara j och underlagskandidat i

$PrViktUK_i^{\text{öjust}}$  = Den överskottsjusterade produktionsvikten för underlagskandidat i

$\text{Overskottsjust}_{j,i} > 0$  för alla j och i

2022-03-07

För att få spridning i underlaget finns också en parameter som begränsar hur stort bidrag till underlaget som ett enskilt steg får bidra med. Om denna gräns överskrids görs en proportionell justering av vikterna ner till denna gräns.

**Formel 14. Justering av vikter för att se till att gränsen för stegets bidrag inte överskrids**

$$\text{MaxBidrJust} = \frac{\text{MaxBidr}}{\sum_{i=1}^n \text{PrViktUK}_i^{\text{öjust}}}$$

$$\text{PrViktUK}_i^{\text{BidrJust}} = \text{MaxBidrJust} \times \text{PrViktUK}_i^{\text{öjust}}$$

där :

MaxBidrJust = Justeringsfaktor m.a.p. maxbidragsgränsen

MaxBidr = Hur stor andel av underlaget som tillåts komma från aktuellt steg

Slutligen görs en justering så att stegets bidrag till underlaget inte överstiger underlagets restvikt.

**Formel 15. Justering av vikter för att se till att restvikten inte överskrids**

$$\text{RestViktJust} = \frac{\text{RestVikt}}{\sum_{i=1}^n \text{PrViktUK}_i^{\text{BidrJust}}}$$

$$\text{PrViktUK}_i^{\text{RestJust}} = \text{RestViktJust} \times \text{PrViktUK}_i^{\text{RestJust}}$$

där :

RestViktJust = Justeringsfaktor m.a.p. underlagets restvikt

RestVikt = Hur stor andel av underlaget som som ej blivit täckt i tidigare steg

De restjusterade produktionsvikterna för varje UK läggs till de i tidigare steg eventuellt skapade produktionsvikterna. En ny restvikt beräknas och underlagets produktionsandelar dras ifrån produktionsandelarna från modellberäkningsgruppen inför nästa steg.

Stegen som körs är följande modellberäkningsnivåer:

Steg 1

- Modellberäkningsnivå 9.
- Sökning på finaste varunummernivå, KNnr + eventuell tilläggskod från IVP.
- Jämförelsetalet måste vara större än 0,85 och beräknas som alternativ I, effektivitetskvot.
- Maxbidrag till underlaget från detta steg är 0,5.

#### Steg 2

- Modellberäkningsnivå 8.
- Sökning på finaste (8-siffrig) KNnr-nivå.
- Jämförelsetalet måste vara större än 0,8 och beräknas som alternativ I, effektivitetskvot.
- Maxbidrag till underlaget från detta steg är 0,75.

#### Steg 3

- Modellberäkningsnivå 8\_2.
- Ny sökning på finaste (8-siffrig) KNnr-nivå.
- Jämförelsetalet måste vara större än 0,9 och beräknas som alternativ II, relativ effektivitetskvot.
- Maxbidrag till underlaget från detta steg är 0,8.

#### Steg 4

- Modellberäkningsnivå 6.
- Sökning på 6-siffrig KNnr-nivå.
- Jämförelsetalet måste vara större än 0,75 och beräknas som alternativ I, effektivitetskvot.
- Maxbidrag till underlaget från detta steg är 1.

#### Steg 5

- Modellberäkningsnivå 6\_2.
- Ny sökning på 6-siffrig KNnr-nivå.
- Jämförelsetalet måste vara större än 0,9 och beräknas som alt. II, relativ effektivitetskvot.
- Maxbidrag till underlaget från detta steg är 1.

#### Steg 6

- Modellberäkningsnivå 4.
- Sökning på 4-siffrig KNnr-nivå.
- Jämförelsetalet måste vara större än 0,8 och beräknas som alternativ I, effektivitetskvot.
- Maxbidrag till underlaget från detta steg är 1.

#### Steg 7

- Modellberäkningsnivå 4\_2.
- Ny sökning på 4-siffrig KNnr-nivå.
- Jämförelsetalet måste vara större än 0,9 och beräknas som alternativ II, relativ effektivitetskvot.
- Maxbidrag till underlaget från detta steg är 1.

#### Steg 8

- Modellberäkningsnivå 2.

- Sökning på 2-siffrig KNnr-nivå.
- Jämförelsetalet måste vara större än 0,9 och beräknas som alternativ I, effektivitetskvot.
- Maxbidrag till underlaget från detta steg är 1.

#### Steg 9

- Modellberäkningsnivå 2\_2.
- Ny sökning på 2-siffrig KNnr-nivå.
- Jämförelsetalet måste vara större än 0,9 och beräknas som alternativ II, relativ effektivitetskvot.
- Maxbidrag till underlaget från detta steg är 1.

Är inte hela underlaget skapat (restvikten > 0) efter steg 9 används medelfördelningsmetoden för den återstående delen.

Alla tre metoder tillämpas på alla modellberäkningsgrupper. En jämförelse görs sedan av metoderna genom att GP med modellberäkningsgruppen för de olika underlagen beräknas på 2-, 4-, 6- och 8-siffrig KN-nummernivå. För varje modellberäkningsgrupp på 5-siffer SNI används den metod som ger högst värde på summa GP i modellberäkningen.

### 2.7.3 Skattningsförfarande för tillförlitlighet

Skattningsgarnas tillförlitlighet beskrivs med andelen modellberäknad population i kombination med matchningen mot produktionen.

**Tabell 1. Gemensam produktion enligt KN-nummer för modellberäknad population och underlag med de imputeringsmetoder som valts.**

Bransch	Modellberäknad		
	Andel	Värde	snitt
B05_06 Kol- o petroleumutv	-	-	-
B07 Utv av metallmalmer	1	132	0,91
B08_09 Ö utvinning service	40	1 499	0,83
C18 Grafisk produktion	32	1 761	0,88
C19 Stenkol raffinaderie	0	134	0,59
C20 Kemiska produkter	19	6 257	0,48
C22 Gummi- o plastvaror	32	6 723	0,81
C23 Glas porslin cement	17	3 800	0,62
C241 Järn o stål	3	919	0,43
C24A Primärbearb av stål	15	1 226	0,64
C24B Ö metaller+gjutning	2	1 110	0,42
C25A Investeringsmetallv	30	5 066	0,88
C25B Insatsmetallvaror	32	12 870	0,94
C26B Instrument ur optik	23	1 863	0,68
C275 Hushållsmaskiner	24	203	0,46
C27A Elmotorer mm	19	5 454	0,65
C28 Tillv av ö maskiner	24	20 016	0,70
C325 Medicinsk utrustn	18	1 143	0,76
C32A Ö tillv ej medicinsk	24	687	0,92

I tabellen ovan framgår hur mycket som modellberäknats per bransch, både i kronor (Mod. värde) och andelen (Mod. andel) av branschens total. Vad man bör vara uppmärksam på är om både andelen och det faktiska värdet är stort, samtidigt som matchningen mot produktionen varit låg (snitt, 2-siffer etc).

Urvalsförfarandet sätter först en gräns på andelen modellberäknat på 30 procent men man tillåter högre andel om man tror att matchningen mot produktionen är bra. I branscherna B08\_09, C18,

C22 och C25B ser vi att andelen modellberäknad är över 30 procent. För dessa branscher har matchningen fungerat ganska bra, 83, 88, 81 och 94 procent i snitt. Det verkar alltså som att antagandet från då vi drog urvalet stämde.

I en bransch som C28 Tillverkning av övriga maskiner modellberäknas ett stort värde. Där matchar produktionen och varuförbrukningen till 70 procent i snitt varför det kan vara bra att vara lite försiktig i tolkning av modellbaserade värden på varunivå i en sådan bransch.

#### 2.7.4 Röjandekontroll

Eftersom INFI inte publiceras genomförs ingen röjandekontroll. Resultatet lämnas endast vidare till NR. En bedömning av risken för röjande gjordes i samband med att en eventuell publicering av INFI utreddes 2014. Bedömningen ledde fram till att alldeles för få statistikvärden skulle vara möjliga att publicera och rekommendationen blev att inte publicera INFI.

### 3 Genomförande

#### 3.1 Kvantitativ information

För referensåret 2020 ingick branscherna B05\_06, B07, B08\_09, C18, C19, C20, C22, C23, C241, C24A, C24B, C25A, C25B, C26B, C275, C27A, C28, C32A och C325. Rampopulationen omfattade 25 913 VE:er och 990 av dessa ingick i urvalet. Det saknades företag inom B05\_06 i rampopulationen.

**Tabell 2. Täckningsgrad i urvalet avseende variabeln *Summa råvaror och handelsvaror per bransch.***

Bransch	Täckningsgrad
B05_06 Kol- o petroleumutv	-
B07 Utv av metallmalmer	99
B08_09 Ö utvinning service	70
C18 Grafisk produktion	70
C19 Stenkol raffinaderie	100
C20 Kemiska produkter	83
C22 Gummi- o plastvaror	71
C23 Glas porslin cement	80
C241 Järn o stål	98
C24A Primärbearb av stål	70



Bransch	Täckningsgrad
C24B Ö metaller+gjutning	98
C25A Investeringsmetallv	70
C25B Insatsmetallvaror	70
C26B Instrument ur optik	71
C275 Hushållsmaskiner	81
C27A Elmotorer mm	84
C28 Tillv av ö maskiner	85
C325 Medicinsk utrustn	84
C32A Ö tillv ej medicinsk	70

Urvalet för 2020 har bestått av 990 VE:er. Utskick sker i tre omgångar; april, augusti och september. Aprilutsändningen bestod av 94 st. verksamhetsenheter, augustiutsändningen bestod 5 st. verksamhetsenheter och september bestod av 894 st. företag. Dessa skickades ut via brev. Inflödet kom upp i 77,6 % vilket var lite lägre än förra gången 2018 (80,4%) och även lägre än jämförande period för tre år sedan (82,8%). Att inflödet blev lägre beror till stor del på uppsnabbningen på 6 veckor vilket gjorde att uppgiftslämnarna hade mycket kortare tid att svara. Dessutom skickades ingen fjärde påminnelse vilket även det påverkar jämförelsen med för 3 år sen. Andra saker som påverkar är Coronapandemin men kanske allra mest att vi inte "hotar" om vite längre. Det gör att vi inte längre får in alla stora företag.

**Tabell 3. Utfallet av insamlingen med avseende på svarsfrekvens.**

INFI branscher som skulle undersökas årgång 2019

05-06 *Kolutovinning och utvinning av råpetroleum och naturgas*

Inget företag i denna bransch.

07 *Utvinning av metallmalmer*

5 av 6 företag inkom (83%). 1 företag lämnades till maskinell imputering.

08 *Annan utvinning av mineral*

14 av 19 företag inkom (74%). 2 företag imputerades och 3 företag lämnades för maskinell imputering.

09 *Service till utvinning*

Inga företag

18 *Grafisk produktion och reproduktion av inspelningar*

47 av 61 företag inkom (77%). 8 företag imputerades manuellt av FU, 6 företag lämnades för maskinell imputering.

19 *Tillverkning av stenkolsprodukter och raffinerade petroleumprodukter*

7 av 8 företag inkom (88%). St1 har efter mycket om och men befriats då de inte äger råvaran utan det gör St1 Supply.

20 *Tillverkning av kemikalier och kemiska produkter*

37 av 66 företag inkom (56 %). 20 företag imputerades manuellt, 6 företag lämnades för maskinell imputering. 3 företag befriades.

22 *Tillverkning av gummi – och plastvaror*

69 av 76 företag inkom (91%). 6 företag imputerades manuellt och 1 företag lämnades för maskinell imputering.

23 *Tillverkning av andra icke-metalliska mineraliska produkter*

39 av 44 företag inkom (89%). 3 företag imputerades manuellt och 2 företag lämnades för maskinell imputering.

32A *Annan tillverkning, ej medicinsk och dental utrustning*

47 av 72 företag inkom (65%). 8 företag imputerades manuellt av FU, 15 lämnades för maskinell imputering. 2 företag befriades.

## INFI branscher 2020

### 24 *Stål- och metallframställning*

44 av 56 företag inkom (79%). 8 företag imputerades manuellt, 2 företag lämnades för maskinell imputering. 2 VE för Sanvik kom felaktigt med i urvalet och befriades.

### 25 *Tillverkning av metallvaror utom maskiner och apparater*

270 av 341 företag inkom (79%), 6 företag befriades, 45 imputerades manuellt och 20 lämnades för maskinell imputering

### 26.5-26.8 *Tillverkning av instrument, ur och optik*

26 av 36 företag inkom (72%), 4 imputerades manuellt och 6 lämnades för maskinell imputering.

### 27 *Tillverkning av elapparatur*

30 av 44 företag inkom (68%), 7 imputerades manuellt av FU och 4 lämnades för maskinell imputering. 3 företag befriades.

### 28 *Tillverkning av övriga maskiner*

113 av 134 företag inkom (84%), 15 imputerades manuellt och 3 lämnades för maskinell imputering. 3 företag befriades.

### 32.5 *Tillverkning av medicinsk och dental utrustning*

10 av 12 företag inkom (83%). 1 företag imputerades och ett företag lämnades för maskinell imputering.

## **3.2 Avvikelser från undersökningsdesignen**

Undersökningen har genomförts enligt plan.