

## Chapter 5 -Calculation of Indices: the Elementary Level

As seen above, the IPI is based on a survey conducted at a detailed level with the aim of obtaining relatively homogeneous elementary sets. The (unpublished) elementary level is the building block of the IPI. This chapter presents the computation method of the elementary indices that will then be used to calculate the upper levels (see Chapter 6).

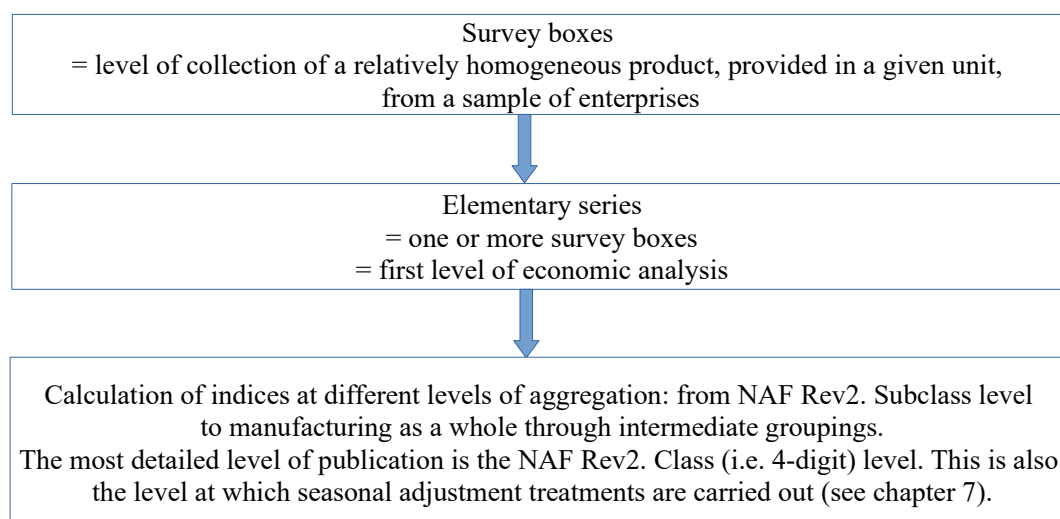
### 1- Collection of Individual Data: The Principle of the Survey Box

As set out in previous chapters, individual data on enterprises are collected using well-defined boundaries termed “cases-enquêtes” in French (survey box hereafter), corresponding exactly to the intersection of a product (ProdEMB), an observation variable (quantities, invoicing, hours worked) and a unit (tonnes, kilos, euros, etc.). The term survey box refers directly to the box of the questionnaire where, for a given product, the enterprise (in the sense of legal unit) fills in its production for the month in question. The IPI collection process is based on a large number of survey boxes. For example, within NAF 2932Z (“Manufacture of other parts and accessories for motor vehicles”), there is a survey box relating to the “manufacture of motor vehicle seats” with invoicing used as an observation variable, but also a survey box on steering and braking devices, one on parts and accessories made of thermoplastic materials, and so on.

This collection nomenclature (all survey boxes) may be revised in line with changes made to the European “ProdCom” classification of products. The nomenclature is managed in a reference system grouping all the products surveyed as part of the EAP<sup>25</sup> (and serves as a reference for the EMBs). As a reminder, the products surveyed in EMBs (referred to as ProdEMB) are generally groupings of several products surveyed under the EAP (or ProdEAP).

The construction of the sample of enterprises surveyed (approximately 4,500 and 5,000 enterprises in total) is conducted on a case-by-case basis. Data are collected through INSEE’s COLTRANE portal used for responses to business surveys. They are then being checked and integrated into the calculation of the indices.

*Figure 1: General Outline of the Index Computation Process*



<sup>25</sup>See Chapter 2 for a presentation of this source.

## 2- Method of Calculation of Elementary Indices

### 2.1- Different types of elementary series

#### *Different Monitoring Methods*

The construction of the IPI begins with the calculation of elementary series. Depending on the importance and complexity of the sectors in question, elementary series may reflect different levels of detail and use different measurement variables. They are composed of one or more homogeneous survey boxes (see above) (invoicing and hours worked are not combined, for example). The indices of primary series are set to an annual average of 100 in the reference year.

Since the IPI is an indicator of production volume, series monitored in invoicing terms must be deflated (by a price index) while series monitored in terms of hours worked need to be adjusted by a technical coefficient taking into account the change in productivity in the activity studied.

#### *Internal and External Series*

The bulk of the monitoring and data collection work involved in compiling the industrial production index is conducted by INSEE (internal series). However, in some cases, the calculation is conducted directly by an external partner (external series): these include the food and agriculture activity series and the energy series.

### 2.2- Definition of an Elementary Index

An elementary index corresponds to a development indicator (or indicator of change) calculated as the ratio between a numerator and a denominator. In the case of internal series, the denominator is equal to the average of the base year (the reference). Therefore, the difficulty is to estimate a numerator representing the change in activity relative to this reference point. In other words, an elementary index is written:

$$Ind_i^{m,A} = \frac{Num_i^{m,A}}{\left( \sum_{k \in Aref} \frac{Num_i^{k,Aref}}{12} \right)} * 100$$

where:

- *i* denotes the branch monitored
- *m* and *A* denote the month and year of calculation of the index, respectively
- *Aref* is the reference year (or base year), which is currently 2015

As noted above, in the case of series monitored in terms of invoicing or hours worked, this index must also be adjusted by a technical coefficient taking into account changes in prices or productivity.

### 2.3- Calculation of the Numerators of the Internal Elementary Series

**Chaining of elementary series by month:** based on a reference month, the numerators are constructed from one month or period to the next (chaining) by applying the monthly changes calculated from the data provided by enterprises present in two consecutive months<sup>26</sup> (referred to below as “present-present”)<sup>27</sup>.

Let:

$Num_S^m$  = the numerator of elementary series ‘S’ for month m of year A;

---

<sup>26</sup>Here, “consecutive” does not always mean two consecutive months in the calendar. For certain special cases or if a change between two consecutive months cannot be calculated, it is chain-linked to the first appropriate month.

<sup>27</sup>The methodology presented here is the methodology implemented since 2015. The methodology used previously was slightly different and was based on the calculation of a denominator common to every month of a year, so the numerators were not directly chain-linked from one month or period to the next.

$EvolPP_S^{m/m-1}$  = the change calculated on all products included in series “S” for enterprises present-present in m-1 and m (enterprises in the sample for dates m and m-1 and for which either a response or an imputation is available for both dates);

$M_{j \times box}^m$  = amount of legal unit “j” for a product corresponding to survey box “box” of series “S” (being accepted that a given enterprise may be surveyed for different products within series S) and  $p_{j \times box}^m$  its sampling weight;

$(j, box) \in [S \times [m; m-1]]$  all present-present *entreprise x survey-box* pairs surveyed in both m and m-1 for the survey boxes of series S.

Then, by calculating the weighted change between m and m-1

$$EvolPP_S^{m/m-1} = \frac{\sum_{(j, case) \in [S \times [m; m-1]]} p_{j \times case}^m M_{j \times case}^m}{\sum_{(j, case) \in [S \times [m; m-1]]} p_{j \times case}^m M_{j \times case}^{m-1}},$$

We can calculate the numerators by chain-linking from one period to the next:

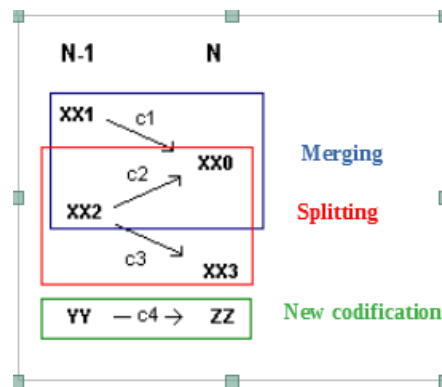
$$Num_S^m = Num_S^{m-1} \times EvolPP_S^{m/m-1}$$

Beyond this general formula, a number of special cases need to be considered.

### Special case 1: calculation of the change between December and January and management of the survey boxes

Each year, the survey boxes can be modified (groupings, recoding, splitting).

*Figure 2: Possible Changes to Survey Boxes*



The chaining of the amounts reported by enterprises from one month to the next can cause difficulties when changing years since declarations are not comparable on different survey boxes. In such cases, correspondence coefficients must be estimated to move (in cases where there has been a change) from survey boxes in one year to those in the following year.

### Special case 2: Chaining Relative to the Atypical Months of July and August

July and August being atypical months, it is not desirable to chain-link changes by taking them as a reference. Therefore, for the methods used for the IPI, when calculating the August and September months, the comparison is generally made with the month of June (in the above calculation formulas, it is then sufficient to replace “m-1” by June).

### Special Cases 3: Management of “0” Cases

When calculating  $EvolPP_S^{m/m-1}$ , it is possible that the denominator equals zero. In this case, chaining will be relative to the first non-zero month observed (m-2, m-3, etc.). The same applies if the numerator in relation to which the chaining is applied is zero.

## 2.4- Imputation of Non-Responses

The method of imputation of an individual datum  $j$  (undertaken in box survey  $C$ ) is based on the combination of the following quantities:

- ratio m/m-1 of the total of the survey box (all responding enterprises for this particular survey box in m and present in m-1 (respondent or imputed)):  $EvolR_C^{m/m-1}$  ;
- $EvolSerie_S^{m-12/m-13}$  the m-12/m-13 ratio of the elementary series  $S$  in which the survey box is located: ;
- weighted by a response rate in the survey box that allows the “quality” of the change to be taken into account:  $TxR_C^m$  .

$$M_j^m = M_j^{m-1} \times \left[ (EvolR_C^{m/m-1})^{TxR_C^m} \cdot (EvolSerie_S^{m-12/m-13})^{(1-TxR_C^m)} \right]$$

### Special Case 1: Management of “0” cases

If in the above formula one of the two ratios is zero (or if no enterprise responds): the formula becomes:

$M_{i,ce(an_m)}^m = M_{i,ce(an_{m-1})}^{m-1} \times Evol$  with  $Evol$  being the non-zero or calculable ratio (in the case of a zero response rate). If both changes equal zero, the amount imputed is 0.

### Special Case 2: Incoming and Non-Responding Enterprises:

Enterprises that are new to the sample and do not respond cannot be imputed. Therefore, they are not used for the calculation in the first month.

## 2.5- Management of Infra-Annual Movements in the Sample (Chaining Corrections)

Drawn once a year, the sample of enterprises may undergo changes every month, in particular because of restructuring (takeover of one company by another, separation, etc.). Thus, new companies may be created that incorporate one or more enterprises surveyed in previous months. Enterprises may also cease their activity or operate in a completely different field of activity.

Changes in samples (taken into account through the application used to manage the monthly branch surveys) affect the calculation of the elementary series. Since the calculation of the indices is chained between months m-1 and m, several boundary changes can affect it:

- for restructured companies retaining the same identifier but which see changes in scope or response behaviour, comparable amounts between m and m-1 are desirable;
- for new restructured companies incorporating one or more enterprises in the sample, comparable amounts between m and m-1 are also desirable, even if not all companies within the new boundaries of the restructured company were surveyed in m-1;

- removals, on the other hand, are not involved in the chaining and, therefore, in the calculation of the index<sup>28</sup>.

In all these cases, it is necessary to implement chain corrections allowing for relevant changes to be maintained. Where possible, corrections are implemented in tandem with companies.

Let:


$Y_{m-1}$ : chaining correction amount for the intersection (enterprise (SIREN) x survey box) for month m-1 and comparable in field terms to a declaration in m, written  $X_m$ ;

$X_m$ : amount reported for an intersection (enterprise (SIREN) x survey box) in m (enterprise response or imputation).

### Figure 1: Enterprise Birth

In the case of an enterprise birth, an adjustment is made in m-1 to calculate the change based on a comparable field: the chaining adjustment amount is used to calculate the numerator of month m (chaining between m and m-1). However, it should not be used in the calculation of m-1 (there is no response in m-2).

m-2	m-1	m
	$Y_{m-1}$	$X_m$


 chaining month to month

If a new company is created but no chaining correction amount,  $Y_{m-1}$  has been inputted (for example because no reliable estimate of the amount is available), the company is not taken into account in calculation for month m. It will then be incorporated for the first time in the calculation carried out the following month.

### Case 2: Restructuring or Change in Response Behaviour

In the case of a change in a company (restructuring or change in response behaviour), the chaining correction amount is used to calculate the numerator of month m. By contrast, in the calculation of m-1, the amount used is always that of the previous month, entered by the company (or modified by the Insee clerk, or imputed).

m-2	m-1	m
$X_{m-2}$	$X_{m-1}$	$X_m$
	$Y_{m-1}$	

 chaining month to month

If a company is restructured or changes its response behaviour but no chaining correction amount has been entered, the amount used is  $X_{m-1}$ .

<sup>28</sup>Generally, if this is not already the case, the product concerned is set to 0 before the actual removal of the calculation sample to take the removal into account.

## 2.6- Prioritising the Analysis of Individual Responses Within an Elementary Series

The elementary series are assessed by clerks at the Insee Normandy regional directorate with a view to detecting possible erroneous responses or atypical changes. Because of the very tight production deadlines, it is not possible to check all data. Priorities for the analysis of series are therefore determined based on two main criteria:

- individual contributions to the year-over-year change (between m/m-12) of the elementary series;
- individual contributions to the revision of changes between m-1 and m of the elementary series.

Re-using the notations already used and if  $\{j \in [S \times (m - 12)]\}$ , the contribution of enterprise  $i$  (present in m and m-12) to the “m/m-12” year-on-year change in an elementary series is written:

$$\text{Contribution}_{i \rightarrow S}^{m/m-12} = \frac{p_i^m * M_i^m - p_i^m * M_i^{m-12}}{\sum_{j \in [S \times (m-12)]} p_j^m * M_j^{m-12}}$$

In addition, let:

- “dca” be the current campaign date and “dca-1” the previous campaign date, i.e. for the indices calculated the previous month;
- $M_i^{m;dca}$  be the amount declared by enterprise  $i$  for month  $m$  during calculation campaign “dca”;
- $\{j \in [S \times (m; m-1); dca]\}$  all (enterprises x survey boxes) present-present in  $m$  and  $m-1$  for the  $S$  series during the “dca” monthly campaign;
- $\{j \in [S \times (m; m-1); dca-1]\}$  all (enterprises x survey boxes) present-present in  $m$  and  $m-1$  for the  $S$  series during the previous monthly campaign.

The contribution of enterprise  $i$  to the revision of the monthly change “m/m-1” of an elementary series is written:

$$\text{Contribution}_{i \rightarrow S}^{\text{Révision } m/m-1, \text{ campagne } dca} = \frac{p_i^m * M_i^{m;dca}}{\sum_{j \in [S \times [m,m-1]; dca]} p_j^m * M_j^{m-1;dca}} - \frac{p_i^m * M_i^{m;dca}}{\sum_{j \in [S \times [m,m-1]; dca-1]} p_j^m * M_j^{m-1;dca}}$$