PART TWO:
SURVEY EXECUTION

Survey execution comprises several phases involving the team responsible for the survey at the national level (located at the INSEE Department of Short-Term Economic Analysis) and a team of computer specialists (based at several INSEE National Computing Centres).

The stages of survey execution in the narrow sense (excluding dissemination, interpretation, and application of results) are the following:

- sample formation and replacement;

- data collection and keying;

- preparation and editing of individual data;

- survey processing;

- seasonal adjustment;

- storage of individual data and aggregate results.

2.1. - Sample formation and replacement

The initial sample of wholesalers to be surveyed is chosen by means of stratified sampling without replacement from the Annual Enterprise Survey (Enquête Annuelle d'Entreprise: EAE). Most of the sample—which is an enterprise panel—is maintained from one Wholesale-Trade Survey to the next. However, owing to economic developments (death or restructuring of enterprises) and the need to preserve a sample of adequate size, the sample is updated periodically.

The Wholesale-Trade Survey sample is thus partially replaced once a year, using identical procedures from year to year. The sampling frame is always the latest available Annual Enterprise Survey. For the replacement, we calculate a theoretical sample of 4,000 enterprises, using a stratified sampling technique that combines two criteria: (1) "enterprise size" measured by workforce size and (2) "sector".

Secondly, we determine the number of enterprises to be picked to fill out the previous year's sample after "cleaning", i.e., removal of enterprises that have gone out of business, been restructured, and so on. The new enterprises are chosen by simple random sampling without replacement in each stratum[1].

Between two replacement operations, INSEE performs ongoing sample maintenance by continuously updating the addresses of surveyed enterprises and the correspondents in each enterprise, tracking the restructuring and sale of enterprises, and other procedures.

More details on samples are provided in the summary technical description of the Survey in Appendix 1.

---

[1] Reminder: this sample-construction method is used only for enterprises in size brackets 0-4. In bracket 5 (200 or more employees, or sales above a threshold determined by the economic-activity classification), all units are sampled.

2.2. - Data collection and keying

2.2.1. - Data collection

The questionnaires for the month-*m* Survey are sent to the enterprises in the sample around the 25th of month *m*-1. The enterprises return the forms to INSEE, where the questionnaires are keyed. The responses must reach the Institute by around the 7th of the Survey month. In each Survey, follow-up questionnaires are sent to enterprises that have not responded by the return date printed on the initial questionnaire.

In January 2004, a new procedure was introduced as an alternative to the standard collection procedure by postal mail. It consists of an electronic questionnaire available online at http://conjoncture.entreprises.insee.fr. For each survey, respondents can choose to answer on paper forms or on the Internet.

2.2.2. - Keying and editing individual data

The data are captured by INSEE's Business Surveys Division as the questionnaires arrive.

The Division performs several data edits:

- consistency edits on structural data: these focus on consistency of total sales and product sales figures, as well as on their year-to-year changes;

- edits on the codings allowed for keying responses to the "short-term" questions: specific values are allowed depending on enterprise responses[2];

- classification edits: two keying classification tables have been introduced in the keying application, one for activities, the other for products.

2.2.3. - Online collection

Since January 2005, business owners have the option of responding to the Wholesale-Trade Survey online. This simply requires logging on to the website and entering the access code and password supplied by INSEE. They can then fill in the questionnaire exactly as they would with the paper form.

The website offers explanations and services to facilitate response, as well as enhanced information feedback (see §3.2.5). This second collection procedure will be gradually extended to all business surveys[3].

2.3. - Organisation of individual data

The individual data are stored in a database common to all business surveys. It contains:

- information describing the enterprises, such as SIREN numbers, names, and addresses;

- structural data, such as sales, number of employees, the enterprise's NAF code, and the CPF codes for its products;

- short-term data: responses to qualitative questions.

---

[2] See Appendix 2 for the detailed tables of the main questions, along with the authorised responses and codes.

[3] See Dutrech et al. (2002).

2.4. - Processing

As noted in §1.5, most short-term questions are asked every two months. Others are asked only every six months.

Short-term questions are processed in two stages. First, we aggregate responses in each primary stratum (see §2.4.1). Second, we aggregate the results thus obtained, in keeping with the structure of the Survey field.

Both "enterprise" questions (e.g., cash flow) and "product" questions (e.g., past sales) are processed using a "sector" approach: responses are consolidated in accordance with the enterprise's NAF code.

The first processing stage (primary aggregation) solely concerns Survey data, i.e., the responses to the short-term questions and the structural data (which serve as weighting coefficients). The second processing stage draws on external data (from the Annual Enterprise Survey), which serve as imputation coefficients.

All Survey questions are processed twice: once for the publication of the Survey for the current month m, and again for the publication of the following Survey in $m+2$. There are two basic reasons for this procedure:
- the application of the constant-sample method, described in §2.4.2;
- the incorporation of late responses.

2.4.1. - Stratification

A primary stratum is defined by an activity sector (industry) at the NAF700 level combined with a size bracket[4]. There are 42 different sectors and 5 size brackets: the number of primary strata in the "sector" approach is therefore $42 \times 5 = 210$.

2.4.2. - The constant sample

The interpretation of the results of a Survey is largely based on their comparison with those of the previous Survey. This is because—as explained in greater detail in §3.1.1—the only meaningful aspect of the results of such Surveys is their variations, not their absolute levels. Given the qualitative nature of most of the questions and the small number of enterprises surveyed, it is preferable to perform such a comparison using an identical sample for two successive Surveys, i.e., on a "constant-sample" basis. This ensures that a change of results between two Surveys will exclusively reflect a change in the enterprises' responses and not a change in the sample structure.

However, for a given question, if we were to take into account only the opinions of enterprises having responded to two successive Surveys, we would overlook the potentially instructive information provided by the enterprises having responded to only one of the two Surveys. In any event, because of the smallness of the samples, we prefer to include as many responses as possible. This implies several methodological choices. First, we construct a constant sample for each question rather than for an entire Survey: this offers maximum coverage of the responses to each question. Second, to handle partial responses, we define completion procedures. These are very simple, and consist in imputing an enterprise's missing response to a particular question from its most recent responses to the question.

The "constant sample" method encompasses all these operations, i.e., determination of the constant sample and processing of non-responses. It is described in the following sub-sections.

---

[4] Recall that size brackets are defined by combining two criteria: workforce size and sales volume.

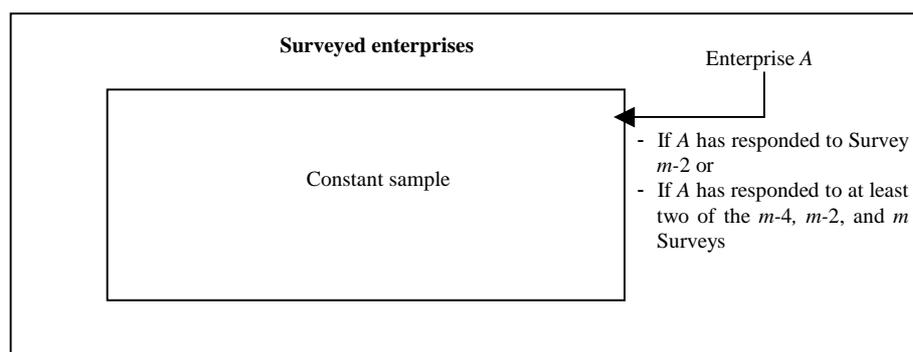## 2.4.2.1. - General principles for determining constant sample

The constant-sample method is a fairly simple way to process non-response, based on a set of rules for "carrying forward" the earlier responses of non-respondent enterprises. Changes in results can thus be explained solely as the consequence of a shift in enterprise opinion. For a given question, the responses taken into account in the results are those of the enterprises in the constant sample.

The general principles of the constant-sample method are as follows:

- The constant sample is formed for individual questions.

- The constant-sample method is applied only to the two-monthly qualitative questions with three multiple choices. The questions in this category are those on sales, ordering intentions, merchandise received, export orders, past and expected workforce size, and cash flow. No six-monthly question is processed with the method[5].

- For each question processed with the constant-sample method, an enterprise is included in the constant sample only if it has responded to that particular question in at least two of the past three Surveys or if it has responded to the intermediate Survey. For the two-monthly Wholesale-Trade Survey, the Surveys comprise those of the current month m and of the months m-2 and m-4. The procedures are summarised in the following diagram.

Question Q

Survey m

Surveyed enterprises

Enterprise A

Constant sample

- If A has responded to Survey m-2 or
- If A has responded to at least two of the m-4, m-2, and m Surveys

---

[5] Accordingly, the constant-sample method is not applied, for example, to the questions on the general business outlook and prices. This is because the carry-forward of earlier responses to the current Survey seems rather inappropriate for questions on prices (given their volatility). General sentiment is published in unweighted form, which substantially reduces the need for the constant-sample approach.

2.4.2.2. - Details on method for completion of partial responses

The two tables below describe the specific procedure applied under the constant-sample method for the Wholesale-Trade Survey.

Table 1: Decision to include an enterprise in the constant sample and procedure for completing partial responses to a given question Q

Enterprise A has responded (R) or not (NR) to question Q in past 3 Surveys			Formation of constant sample for question Q in Survey m: Procedures for enterprise A	
Survey m-4	Survey m-2	Survey m	Enterprise A included constant sample?	Processing of enterprise A's responses and procedure for completing its non-responses
R^6	R	R	**YES**	Surveys m-2 and m: enterprise A's responses are included in results.
R	R	NR	**YES**	Survey m-2: enterprise A's response is included in results. Completion of response to question in Survey m: enterprise A's response to Survey m-2 is copied to Survey m.
R	NR	R	**YES**	Survey m-2: specific procedure applied is described in Table 2 below. Survey m: enterprise A's response is included in results.
R	NR	NR	NO	Surveys m-2 and m: enterprise A is excluded from constant sample.
NR	R	R	**YES**	Surveys m-2 and m: enterprise A's responses to both Surveys are included in results.
NR	R	NR	**YES**	Survey m-2: enterprise A's response is included in results. Completion of response to question in Survey m: enterprise A's response to Survey m-2 is copied to Survey m.
NR	NR	R	NO	Surveys m-2 and m: enterprise A is excluded from constant sample.
NR	NR	NR	NO	Surveys m-2 and m: enterprise A is excluded from constant sample.

An enterprise may therefore be excluded from the constant sample for a question to which it has responded in the current Survey m (this is the next-to-last configuration in Table 1). In this case, its response to the question is not included in the processing of preliminary results, i.e., for the publication of the current Survey m. However, under the rules set out in Table 1, this response is added back in the processing of final results, i.e., for the publication of the following Survey, m+2.

---

[6] The symbol R denotes that the enterprise has responded to this particular question in the Survey indicated. The symbol NR indicates that it did not respond to the question at the time the Survey was processed (questionnaire not received, or response missing from questionnaire).

Table 2: Procedure for completion of partial responses by enterprise *A* in (*R*, *NR*, *R*) configuration

In Survey *m*-4, enterprise *A* responded:	In Survey *m*-2, enterprise *A* did not respond:	In Survey *m*, enterprise *A* responded:	In Survey *m*-2, the following response ("interpolated" between responses to Surveys *m*-4 and *m*) is recorded for enterprise *A*:
(+)	*NR*	(+) or (=)	(+)
(+)	*NR*	(−)	(=)
(=)	*NR*	(+), (=) or (−)	(=)
(−)	*NR*	(+)	(=)
(−)	*NR*	(=) or (−)	(−)

Key:
(+) = "increasing", "above normal" response
(=) = "unchanged", "normal" response
(−) = "decreasing", "below normal" response.

An enterprise that has not responded to a question, but whose response is completed using one of the procedures described in the preceding tables, is said to be "interpolated" for this question.

2.4.3. - Question processing

2.4.3.1. - Processing of "enterprise" questions

In this entire section, "respondent enterprises" include interpolated enterprises (if any[7]), i.e., after application of the constant sample for the question concerned. By extension, "respondent enterprises" will be used to designate the aggregation of enterprises responding to at least one question, as defined above.

2.4.3.1.1. - Primary aggregation of results

Primary aggregation consists in calculating the weighted averages of individual responses within each primary stratum (a sector × size pair). "Individual" here denotes a respondent enterprise. The weighting coefficient is a datum supplied by the enterprise. The coefficient is based on workforce size for questions on past and expected workforce size, and on sales for questions on the general business outlook, cash-flow position, settlement times, and late payments.

We then calculate the weighted percentages of "increasing", "unchanged", and "decreasing" responses for each question and each primary stratum[8]:

$$(1) \qquad \% \, INCREASE = 100 \times \frac{\text{"increasing" } SW}{SW}$$

---

[7] That is, if the constant-sample method is used to complete the responses to the question concerned.

[8] Throughout §2.4.3, to simplify notations, "increasing" denotes the positive response (increasing, above normal, etc.), "unchanged" the "intermediate" response (unchanged, normal, etc.), and "decreasing" the negative response (decreasing, below normal, etc.).

$$(2) \qquad \% \, UNCHANGED = 100 \times \frac{\text{"unchanged" } SW}{SW}$$

$$(3) \qquad \% \, DECREASE = 100 \times \frac{\text{"decreasing" } SW}{SW}$$

With:

"increasing" SW = Sum of weights of "increasing" responses

"unchanged" SW = Sum of weights of "unchanged" responses

"decreasing" SW = Sum of weights of "decreasing" responses

SW = "increasing" SW + "unchanged" SW + "decreasing" SW.

For the general business outlook, we also determine the unweighted percentages of "increasing", "unchanged", and "decreasing" responses for each primary stratum:

$$(4) \qquad \% \, INCREASE_{unweighted} = 100 \times \frac{\text{"increasing" } N}{N}$$

$$(5) \qquad \% \, UNCHANGED_{unweighted} = 100 \times \frac{\text{"unchanged" } N}{N}$$

$$(6) \qquad \% \, DECREASE_{unweighted} = 100 \times \frac{\text{"decreasing" } N}{N}$$

With:

"increasing" N = Number of "increasing" responses

"unchanged" N = Number of "unchanged" responses

"decreasing" N = Number of "decreasing" responses

N = "increasing" N + "unchanged" N + "decreasing" N.

We then determine the weighted average called *balance of opinion*[9]. This is the summarised opinion of enterprises in the stratum, defined as the difference between the proportion of enterprises voicing a positive opinion of the situation and the proportion of those with a negative opinion:

$$(7) \qquad BALANCE = \% \, INCREASE - \% \, DECREASE$$

For the general business outlook, we also compute an unweighted balance of opinion:

$$(8) \qquad BALANCE_{unweighted} = \% \, INCREASE_{unweighted} - \% \, DECREASE_{unweighted}$$

Lastly, we calculate the weighted percentage of non-responses for each question and for each stratum (relative to total Survey respondents in stratum):

$$(9) \qquad \% \, NR = 100 \times (1 - \frac{SW}{TSW})$$

With:

SW = Sum of weights of stratum respondents to particular question (defined above)

[9] Balances of opinion are interpreted with greater precision in §3.1.1.

TSW = Total sum of weights of Survey respondents in stratum:

TSW = *SW* + *SWN*

where *SWN* = Sum of weights of non-responses to question in stratum (among Survey respondents in stratum).

For the general business outlook, we also calculate an unweighted percentage of non-responses for each stratum (relative to total Survey respondents in stratum):

$$(10) \qquad \% NR_{unweighted} = 100 \times (1 - \frac{N}{TN})$$

With:

N = Number of stratum respondents to particular question (defined above)
TN = Total number of Survey respondents in stratum:

TN = *N* + *NN*

where *NN* = Number of non-respondents to particular question (among Survey respondents in stratum).

Lastly, for all the questions to which the constant-sample method is applied, we calculate the weighted percentage of interpolated responses for each stratum (relative to total stratum respondents to question after constant-sample procedure):

$$(11) \qquad \% INT = 100 \times \frac{SWI}{SW}$$

With:

SWI = Sum of weights of interpolated enterprises in stratum for question examined
SW = Sum of weights of respondents to question after application of constant sample in stratum (defined above):

SW = *SWR* + *SWI*

where *SWR* = Sum of weights of actual responses to question (excluding interpolations).

2.4.3.1.2. - *Secondary aggregation of results*

For a given question, the balance of opinion for a primary-stratum grouping is the weighted average of the primary balances of opinion. Weighting enables us to make allowance for the stratum's size relative to the total Survey field.

The imputation coefficients are taken from the latest available Annual Enterprise Survey. The weighting system varies according to the questions: workforce size is used for past and expected workforce size; sales are used for other questions.

The same procedure is applied to the percentages of "increasing", "unchanged", and "decreasing" responses as well as to the percentage of interpolated responses.

2.4.3.2. - Processing of "product" questions

2.4.3.2.1. - *Primary aggregation of results*

Primary aggregation consists, as before, in computing weighted averages of individual responses in each primary stratum (a sector × size pair). Here, "individual" denotes the product of an enterprise in the primary stratum. For all the questions processed, the weighting coefficient is the sales by that particular enterprise of the product concerned (the product sales figure is supplied by the enterprise).

As previously, the weighted percentages of "increasing", "unchanged", and "decreasing" responses are calculated for each question and for each primary stratum, using formulas (1), (2), and (3). With:

"increasing" SW = Sum of weights of "increasing" responses

"unchanged" SW = Sum of weights of "unchanged" responses

"decreasing" SW = Sum of weights of "decreasing" responses

SW = Sum of weights of "increasing", "unchanged", and "decreasing" responses to particular "product" question in the primary stratum (after application of constant sample, if appropriate).

The balance of opinion is still written as:

BALANCE = % INCREASE - % DECREASE

We also calculate the weighted percentage of non-responses for each question and each stratum (relative to total Survey respondents in stratum), using equation (9), with:

TSW = Sum of weights of Survey respondents in stratum (after application of constant sample, if appropriate).

For questions on sales, export orders, ordering intentions, stocks, and merchandise received, we determine the weighted percentage of interpolated responses for each stratum (relative to total stratum respondents to the question after constant-sample procedure), using equation (11), with:

SWI = Sum of sales of products for which the response has been interpolated, for enterprises in that stratum.

2.4.3.2.2. - Secondary aggregation of results

The procedure is exactly the same as the one applied to "enterprise" questions. For a given question, the balance of opinion in a primary-strata grouping is the weighted average of the primary balances. For all "product" questions, the weighting system is based on sales taken from the latest available Annual Enterprise Survey.

The same procedure is applied to the percentages of "increasing", "unchanged", and "decreasing" responses, as well as to the percentage of interpolated responses.

2.5. - Seasonal adjustment (SA)

2.5.1. - General method

Selected balances of opinion are seasonally adjusted once a year[10]. The operation is performed with *TRAMO-SEATS* software for the two-monthly and six-monthly series.

To produce a seasonally adjusted series from a raw series, we subtract the seasonal-coefficient series from the raw series:

$$VAR_{SA}^{q} = VAR_{RAW}^{q} - VAR_{SC}^{q}$$

where *VAR* is the series processed, *q* the question and *SC* the seasonal coefficient.

The seasonal-coefficient series are calculated by the Survey manager. They are stored in a table updated each year to reflect the influence of the latest data points on seasonal-coefficient values. The updates require an annual recalculation of the seasonally adjusted series.

---

[10] The percentages of "increasing", "unchanged", and "decreasing" responses are not seasonally adjusted; nor are the percentages of non-responses and interpolated responses.

All qualitative questions (including the six-monthly short-term questions on settlement times) are seasonally adjusted.

We seasonally adjust the raw series processed with a constant sample. The series derived from the "general business outlook" question are seasonally adjusted in their weighted and unweighted forms.

2.5.2. - Method for applying seasonal coefficients to recent observations

Seasonal coefficients are carried forward as follows: if the coefficients are not updated, we use the latest computed coefficients. In other words, we recapture the latest immediately existing value of the coefficient for month *M*:

$$VAR_{SA}^q(M,Y) = VAR_{RAW}^q(M,Y) - VAR_{SC}^q(M,Y-i)$$

where *M* is a month of the current year *Y* and *Y-i* the year of the latest coefficient calculated for month *M*. In practice, *i* takes the value 1, or even the value 2 if two successive *SA* procedures are separated by slightly over a year.

2.6. - Storage of aggregated data

Raw series, seasonally adjusted series, results from the constant sample, and results from the current sample are stored. The series on the general business outlook are stored in their weighted and unweighted forms.