Chapter 4

ASSESSMENT OF NATIONAL ACCOUNTS

4. ASSESSMENT OF NATIONAL ACCOUNTS

4.1. Introduction

4.1. This chapter describes methods appropriate for the second line of action in the NOE measurement strategy, namely the assessment of the national accounts from the perspective of non-observed activities. The ultimate goal of the assessment is to have a systematic and comprehensive list of all non-observed activities, broken down by type according to an appropriate analytical framework, and to have recorded their likely magnitudes and the ways in which they are currently being included (if at all) in the GDP estimates. Achievement of this objective would be exemplified by completion of all the tables in the Eurostat Tabular Framework.

4.2. Section 4.2 deals with analysis of the incoming basic data by comparison with data from other sources, through data confrontation and discrepancy analysis. Section 4.3 outlines the approaches that can be used to estimate upper bounds to the extents of non-observed and non-measured activities. Section 4.4 describes the use of supplementary and special surveys for assessment. The last two sections present more general national accounts assessment tools, namely the Data Quality Assessment Framework and the National Accounts Process Table, being developed by the IMF and Eurostat, respectively. In each case exhaustiveness is seen as just one of several quality aspects that collectively define the overall quality of the national accounts.

4.3. As exemplified in the following sections, many of these assessment methods can also be used to provide adjustments for the NOE using indirect compilation methods as described in Chapter 5. Equally, the adjustment procedures can be used as a basis for assessment. Thus Chapter 5 is an additional rich source of assessment methods.

4.2. Data Confrontation and Discrepancy Analysis

4.4. Confrontation of data from different sources is an integral part of the national accounts compilation. It can also be used to identify the remaining errors and gaps in and between these data. Ideally, data confrontation should take place prior to national accounts compilation, with the aim of checking the statistics and increasing their quality. Examples of possible data confrontation are:

- Enterprise survey data *versus* taxation data; wages paid *versus* taxes raised; sales of goods and services due to VAT *versus* VAT raised; and production *versus* production related taxes.
- Enterprise survey data about the production of commodities *versus* enterprise survey data about purchases of commodities; supply of goods and services *versus* the use of them.
- Expenditure survey data *versus* retail trade survey data; household expenditures *versus* retail trade.
- Expenditure survey data *versus* income or taxation data; household expenditures *versus* available income.
- Enterprise survey data *versus* labour force survey data; use of labour *versus* supply of labour; turnover, value added, intermediate consumption, etc. *versus* the use of labour.

The following paragraphs illustrate the general approach through some specific examples.

Actual value added tax versus *theoretical value added tax*

4.5. This analysis is useful if a significant amount of value added tax (VAT) is collected. The actual amount of VAT collected by the government is compared with the theoretical amount of VAT that should have been raised, calculated from the supply and use tables. In principle, the latter equals the VAT revenues that a government should raise if no VAT was evaded. In practice, there are some statistical and legal reasons for differences, such as bankruptcies, discharges and individual arrangements. Therefore, the theoretical VAT should always exceed the actual VAT by perhaps as much as 5%, depending on the particular situation in the country. If the difference is less, or even worse, if the theoretical VAT is less than the actual VAT, it may safely be assumed that there are NOE activities not included in the national accounts.

Actual tax returns versus *theoretical income tax*

4.6. The income available to households computed from income tax returns is compared with that calculated from the national accounts. Apart from some differences for statistical and legal causes such as bankruptcies, discharges and individual arrangements the differences between the figures can be assumed to be a consequence of tax evasion on the one hand and NOE on the other. Such an analysis is especially useful if the national accounts do not use the income tax statistics as a data source, as in the Netherlands, for example. Van de Laan and De Waard (1985) describe how Statistics Netherlands used this type of analysis to arrive at an estimate of tax fraud. They calculated that about three quarters of the difference between the two income estimates was due to the differences in their definitions. The remaining difference averaged around 6% over the period 1977 to 1985, as summarised in Table 4.1.

Table 4.1. **Primary income of households according to the national accounts and income tax statistics for the Netherlands 1977-1985**
(in 1 000 million guilders)

	1977	1979	1981	1983	1985
a) Income tax statistics	192.2	223.9	246.0	264.6	283.8
b) National accounts	203.0	235.3	258.2	282.9	301.6
Difference *b*) – *a*) as % of *b*)	+5.3	+4.8	+4.7	+6.5	+5.9

Source: Van de Laan and De Waard, 1985.

4.7. In principle, the difference between the national accounts and the income tax statistics should be an indication of the size of economic activity that is concealed for income tax reasons and that has been included in the national accounts. However, at a more detailed level in this analysis, some of the differences proved to be negative, *i.e.*, had the wrong sign for such a conclusion. This shows that a detailed analysis is preferable to an aggregate one, as aggregation may hide significant details and lead to the wrong conclusions.

4.8. A few years later, the research was repeated. The legal situation had changed. From 1987, interest payments made by financial institutions had to be reported to the tax authorities. As a consequence, the size of the interest income concealed by households from the tax records decreased dramatically from almost 30% of the total in the period 1983-1987 to about 15% in 1988-1990. This indicates the importance of checking and updating the assumptions built into adjustment models.

I*ncome based* versus *expenditure based estimates of national income*

4.9. In many countries, the national accounts lend themselves very well to discrepancy analysis as they provide both income based and expenditure based estimates of national income (MacAfee, 1980; O'Higgins, 1989). The first estimate does not include income that has been concealed for tax reasons,

whereas the second estimate does. Thus, the difference between both estimates can be partly attributable to tax evasion. Other causes for discrepancy are errors in timing and statistical errors. However, according to MacAfee (1980), the long run average of these latter causes is zero. Of course, there are likely to be omissions from the expenditure based estimates, for example expenditure on narcotics, gambling and alcohol.

U*se of labour* versus *supply of labour*

4.10. Wages and employment measured from the use (employer) side and from the supply (employee) side can be compared. With appropriate allowances for conceptual differences, the measures should be the same. Examination of the discrepancies can provide an indication of the size and distribution of activities that are missing from the enterprise data. In fact this approach is so effective that analysis of labour inputs is one of the mechanisms specified by the European Commission (1994) for assessing the exhaustiveness of GDP estimates. European Union Member States are required to make a systematic comparison of the estimates of employment which underlie their estimates of GDP with the alternative estimates of employment obtained from household based sources, as described by Hayes and Lozano (1998). The elements of the method are as follows.

- E*stimate the labour input underlying* GDP *estimates*. This means deriving the labour input that is present (explicitly or implicitly) in the data sources used to derive GDP estimates. If the data are derived from enterprise surveys, then the labour input to the production covered by these surveys must be estimated.
- E*stimate the labour input based on household survey data*. Typically these data are obtained from a labour force survey, supplemented by census data if available.
- S*tandardise the labour input estimates*. Convert the enterprise based (use) and household based (supply) estimates of labour input to the same units of labour input, such as hours worked or full-time equivalent employment units, so that they can be meaningfully compared.
- C*ompare the two sets of estimates*. Analyse any discrepancies taking into account the reliability of the different sources. A surplus of labour input derived from the household source over that from the enterprise source is an indication of non-observed production. It provides a lower bound as some labour input could be missing from both sources. No difference or a surplus of the enterprise-based labour input over the household-based input suggests that the household data are not providing extra coverage.

4.11. The final step in the EU exhaustiveness procedure is to compute a multiplication factor by which to adjust the output and value added estimates to account for non-observed production. This is further elaborated in Sections 5.2 and 5.4, which provide, respectively, a more comprehensive description of the labour input method as a compilation and adjustment tool, and an example of its application by Istat.

Micro discrepancy analysis

4.12. Discrepancy analysis can be carried out at micro as well as macro level. For example, data for individual persons or enterprises, retrieved from tax files or from other administrations, can be compared with data from surveys.

4.13. An example of micro discrepancy analysis is described by the US Internal Revenue Service (1979) and Parker (1984). A sample of 50 000 income tax returns was very closely examined and compared with data from *information returns*. Information returns are reports that must be completed by payers of certain types of income, for example wages and salaries paid by employers and interest paid by banks. It was found that approximately 25% of the incomes that were covered by the information returns were not reported on the tax returns.

4.14. A similar research was conducted in the Netherlands, in 1985 in a research study on concealed interest income on giro, bank and savings accounts (Kazemier, 1991). The Dutch Central Bank collected the 1981 data on interest received by 1 369 households as far as it could be traced in the accounts of the giro and the banks. The Department of Finance collected the data on interest as reported in the tax

returns by the same households. Both data files were put in the same order, made anonymous and sent to Statistics Netherlands for further analysis. The results of this research were almost similar to those of the macro-discrepancy analysis between the national accounts and the income tax statistics, described above.

4.15. Another example of a micro discrepancy analysis, from the United Kingdom, was reported by Dilnot and Morris (1981). Of a sample of households, they compared the expenditures and income. The expenditures were measured in the 1977 Family Expenditure Survey. The income data were retrieved from tax forms. If expenditures inexplicably exceeded income, it was assumed that the difference between both figures equalled concealed income.

Partial integration

4.16. The extension of simple confrontation of two data sources to produce new data on a particular topic in more detail than is normal during compilation of the national accounts is commonly termed *partial integration*. For example, information on construction and the production of construction materials may be integrated, as these are two closely linked branches of industry.

4.17. Another example of partial integration is the preparation of labour accounts in which labour data from the supply and demand sides are integrated. Leunis and Verhagen (1999) provide an example. Other examples are energy accounts, which combine the results of surveys on the production, transformation and use of energy, data on the imports and exports of energy and fuels et cetera, resulting in energy balances per branch of industry and per type of fuel. More information on energy accounts can be found in Eurostat (2000).

4.3. Upper Bound Estimation

4.3.1. *Introduction*

4.18. The aim of upper bound estimation (sometimes termed "sensitivity analysis") is to derive an upper limit to the extent of NOE activities. The essence of the procedure is to consider systematically for each component of GDP the maximum possible amount of NOE activities and to total the results to obtain an upper bound. The procedure can be applied to GDP compiled by any or all of the expenditure, income and production approaches. It can be applied for all NOE problem areas, or just for specific types, for example, underground production or the informal sector. Analyses of this sort have been conducted in many countries. The following sections illustrate upper bound estimation for underground production using the expenditure and the production approaches, respectively, by examples from Canada and the Netherlands. Willard (1989) describes other applications in France and Italy.

4.3.2. *Upper Bound – Expenditure Approach*

4.19. In Statistics Canada (1994) Gervais describes the procedures used by Statistics Canada in estimating an upper bound for the underground economy in Canada in 1992. In the second part of the paper, Gervais works systematically through all the components of GDP compiled by the expenditure approach. For each component in turn an upper bound on underground (legal) production is computed and compared with the allowance for such production (if any) already made in compiling the official estimates. The differences are summed to provide an upper bound on the extent to which the published estimates underestimate the GDP due to underground production.

4.20. The following paragraphs give a flavour of the sort of reasoning that is used in this analysis. Whilst the particular reasons for underground activity and sources of data are likely to be different in other countries, the same general approach and type of reasoning will still be applicable.

4.21. Viewing GDP through the expenditure approach, Gervais argues that the main areas for underground activities are residential construction and final consumption expenditure of households. The effects of underground activities on imports and exports are small and on the other expenditure

components of GDP are negligible. The reasoning goes as follows, starting with the least significant components.

Government current expenditure on goods and services and government investment

4.22. The upper bound is zero as there is no scope for underground activities.

Business investment in inventories

4.23. The upper bound is zero. Businesses may have a tax incentive to exaggerate holding losses and to understate holding gains. However, holding gains or losses are removed in the calculation of inventory changes, as they do not relate to current production. Thus if businesses exaggerate their holding losses or understate holding gains this should not affect the national accounts measure of inventory changes.

Business fixed capital investment

4.24. The upper bound is zero. Capital investment is measured on the demand side. Businesses have no incentive to hide expenditures.

Imports

4.25. In Canada the only legal goods smuggled into the country in sufficient volume to be worth considering are tobacco, alcoholic beverages and precious jewellery. Commodity related data sources provide estimates for the quantities. For example, estimates of the quantities of smuggled alcohol are obtained using data from the Liquor Board of Ontario. The black-market price in Canada is assumed to be 60% of the regular price and the import price is assumed to be between 40-60% of the black-market price.

Exports

4.26. In Canada, businesses have little or no incentive to hide exported goods. There are no export duties. Ninety per cent of Canadian exports involve vehicles and parts, crude petroleum, natural gas, lumber, pulp and paper, ores, metals and alloys. None of these are likely to be exported "under the table".

Residential construction

4.27. The ratio of domestic sales of lumber and building materials to reported value of residential construction put in place provides, over time, a broad indication of the growth of unrecorded residential construction. However, increases in this ratio may also be due to reduced profit margins and/or reduced wages, as well as to increase in underground transactions. An upper bound on underground activities is determined in terms of three components:

- New construction. Estimates for new construction are based on housing starts, average value of building permits and work put in place. It is more or less impossible to construct a house without a permit. Builders have a vested interest in understating the values, but local authorities base tax liabilities on values and will thus not issue a permit if they believe the value is too low. Hence, under-estimation is probably not more than 15% on average, plus another 5-10% for additional work done after the permit has been issued. To get upper bounds for seasonal dwellings (*i.e.*, those used only during the summer months) the number of permits issued is multiplied by five and 25% is added to stated permit values. The number of permits issued for conversion from seasonal to all-season dwelling is multiplied by two. The larger conversions have to be reported, as it is difficult to avoid detection.
- Alterations and improvements. The data source for alterations and improvements is the (Canadian) *homeowner repair and renovation survey*. Homeowners have no particular incentive to hide repairs and renovations, as they are not responsible for payment of taxes on the income

earned. Thus, a maximum of 20% is assumed in the under-reporting of the total payments for materials and labour. This upper bound can be checked for plausibility by first estimating the total amount of home improvements put in place. These estimates are based on householders' purchases of materials (assuming all purchases are made legally) and the estimated ratios of material content to the total value of work. Next the estimated value of the homeowners' own work (based on time use survey data) is deducted to arrive at an estimate of the value of underground contract work.

- Transfer costs. Real estate commissions are computed by applying average commission rates to reported sales. Average commission rates are obtained from the Canadian Real Estate Association. The value of sales is not thought to be underestimated, as real estate agents prefer to advertise rather than hide their success in making sales.

Household expenditure on goods and services

4.28. There are three aspects to underground activities in this area:

- *under the table* purchases, known by other authors as *cash settlements*;
- *skimming*, whereby businesses fail to declare part of their legitimate business income to the tax authorities and (presumably) to the statistical agency as well; and
- *tips*.

4.29. Expenditures are divided into 140 different categories, which are grouped according to type of goods and services and the opportunities for underground transactions. For some groups of goods and services, underground transactions are virtually impossible. An important example is purchases of new motor vehicles. The expenditure groups that are singled out for special treatment are tobacco, alcoholic beverages, rent, room and board, professional services, food, childcare, and domestic and household services. Within each of these groups the maximum conceivable values of under the table purchases, skimming and tips are considered and compared with the figures already included in the official GDP estimates.

4.30. Understatement of business receipts (skimming) does not necessarily translate into underestimation of GDP. In the case of businesses selling to other businesses, the businesses doing the purchasing will find a way of passing purchase costs onto their customers whether or not the purchases have been under the table. Thus the market prices of good and services sold to final consumers automatically embody any skimming that may have gone on at an intermediate level. This reasoning suggests that only skimming by businesses selling to households needs to be considered.

4.31. Self-employed individuals and small businesses have more opportunity to engage in skimming than large businesses. It is very unlikely that large retail organisations with hundreds or thousands of employees indulge in skimming at all as the risk of damage to their reputation (should the skimming be discovered) would far outweigh the benefits of it. Large businesses are thus likely to focus on tax avoidance rather than tax evasion and it can be assumed that skimming is restricted to small businesses.

4.32. Input-output tables provide a general picture because the effects of skimming can show up in balancing the tables as implausibly low values of retailers' profit.

4.33. Skimming levels are estimated by category: 25% of gross receipts for services; 15% for taxicabs and most retail trade; 25% for vending machine operators, direct sellers, and repair shops.

4.34. On top of skimming there are direct sales of goods and services by individuals that are not registered as businesses. These are also considered on a category-by-category basis. In the case of food, for example, direct sales by farmers reported by the provincial departments of agriculture are inflated by 20%.

4.35. Tables 4.2 and 4.3, respectively, indicate the upper limits thus obtained of personal expenditures missing from the published estimates and of underground transactions missed from expenditure based GDP.

Table 4.2. **Upper limit of underground transactions missing from personal expenditure**
Canada 1992

	Underground $ million	Published estimates $ million	Proportion %
Skimming by businesses	10 836		
Contraband tobacco	1 057		
Contraband spirits	768		
Illegally manufactured wine	515		
Mark-ups on illegal alcohol	565		
Rent, room and board	269		
Tips	312		
Professional services	208		
Food	50		
Domestic and household services	250		
Subtotal	14 830	393 053	3.8
Goods and services tax and provincial sales tax	0	26 483	
Total	14 830	419 536	3.5

Source: (Gervais) Statistics Canada (1994).

Table 4.3. **Upper limit of underground transactions missing from expenditure based GDP**
Canada 1992

	Underground $ million	Published $ million	Proportion %
Personal expenditure on goods and services	14 830	419 536	3.5
Government current expenditure on g&s	0	148 377	
Government investment	0	16 508	
Business investment in fixed capital	3 578	113 440	3.2
Residential construction	3 578	43 992	8.1
New residential construction	1.883	20 934	9.0
Alterations and improvements	1 695	12 153	13.9
Transfer costs	0	10 905	
Non-residential construction	0	30 189	
Machinery and equipment	0	39 259	
Business investment in inventories	–15	–2 258	–0.6
Exports	1 100	181 948	0.6
Merchandise	800	156 567	0.5
Non-merchandise	300	25 381	1.2
Less Imports	1 003	185 751	0.5
Merchandise	1 003	147 588	0.7
Non-merchandise	0	38 163	0
GDP at market prices	**18 408**	**688 541**	**2.7**

Source: (Gervais) Statistics Canada (1994).

4.3.3. *Upper Bound – Production Approach*

4.36. Broesterhuizen (1985) describes upper bound estimation (referred to in the paper as "sensitivity analysis") for underground production in the Netherlands using the production approach. The following paragraphs outline the procedure used.

4.37. GDP is broken down into six categories, according to the estimation method and/or the sector of origin. The categories differ in the extent to which they are susceptible to underground production.

Category 1: industries where estimation is by indirect methods

4.38. Usually a significant part of GDP is measured indirectly, *i.e.*, not based on reports from producers. In the Dutch case this applies to agriculture, crude petroleum and natural gas production and exploration, the petroleum refining and the operation of real estate.

- The production of agriculture is calculated by multiplying production volumes (measured by independent observers) by known market prices.
- The gross value added of the petroleum industry in the Netherlands is almost completely dependent on imports and exports of crude petroleum and coal products. Data on these imports and exports are retrieved from the customs declaration forms. (The case described is the situation as of 1979. Since then, things have slightly changed because intra-EC trade customs forms are no longer available.)
- The domestic crude petroleum and natural gas production and exploration are so strictly controlled by the government that concealed activities are not very likely.
- The value added of real estate is almost entirely measured by indirect methods.

The researchers concluded that whatever size underground activities in these sectors might be they had no significant effect on GDP.

Category 2: estimates for government and government controlled units

4.39. Category 2 consists of the gross value added of all sectors that are subject to strict government supervision. These include general government, public utilities, railways, tramway and regular bus services, subsidised motor coach services, banking and insurance, hospitals, mental homes and nursing houses, subsidised welfare services, social and cultural institutions, co-operative business organisations, and communications. Underground activities are assumed to be zero. (Note that, since this study, the railways, many tramways and bus services have been privatised and government control over the communications industry has become less strict, so these services should now be analysed within the following categories, 3-5.)

Categories 3, 4 and 5: estimates for large, small and very small enterprises

4.40. Categories 3 to 5 refer to large, small and very small businesses respectively. Estimates of the proportions of underground activities are made separately within each category. These categories contain the gross value added of enterprises that are not included in categories 1 and 2 and for which the estimates of gross value added are not based on fiscal data. Typically the data are obtained by annual surveys. The sampling frame for these surveys is a business register based on the administrative register maintained by the Chambers of Commerce. Each enterprise has to register in order to obtain licences and a value added tax-number.

4.41. For some industries there is a cut off in the sample frame and only enterprises with over 10 employees receive a questionnaire. Other industries are fully included. For large (category 3) enterprises, *i.e.*, those with over 100 employees, the sampling rate is 100%. For the smaller (category 4) enterprises, *i.e.*, the remainder of those included within the sampling frame, the sampling rates are lower. In general, the questionnaires used in this category are shorter and ask for less detail than the questionnaires sent to the larger enterprises. More detailed information than is provided in the questionnaires is borrowed from the category 3 enterprises in the form of ratios and percentage shares. The estimates for the very small enterprises (category 5) are all obtained as extrapolations of the category 4 estimates.

Category 6: estimates based on fiscal data

4.42. For some own account workers, the national account estimates are based on information derived from tax files. This includes, for example, the value added of own account workers in hotels, restaurants and cafés, repair of consumer goods, business services, renting of machinery and other movables,

health and veterinary services, social-cultural institutions and private households with domestic servants. (Since the study, other data sources have been introduced in place of tax estimates for several of these activities.)

Summary of results

4.43. In arriving at estimates of undercoverage of underground activities in GDP, several alternative sets of assumptions were made. The scheme giving the largest values is presented in Table 4.4 and indicates that the highest possible (but very unlikely) estimate of GDP understatement in 1979 is 5.7%.

Table 4.4. **NOE upper bound for the Netherlands, 1979**

		Categories						
		1	2	3	4	5	6	Total
GDP	1 000 million guilder	65.2	104.6	61.1	54.0	11.3	19.7	315.9
Share in GDP	%	20.6	33.1	19.3	17.1	3.6	6.2	100.0
NOE	% of GDP	0	0	2.0	10.0	30.0	40.0	5.7
NOE	1 000 million guilder	0	0	1.2	4.4	3.4	7.9	19.9

Source: Broesterhuizen (1985).

4.44. This analysis also shows that, although category 6 estimates are quite small as a proportion of GDP, this category contributes the largest part of the upper bound estimate. This suggests that category 6 might be the place to invest more effort in improving estimates, thus illustrating how such analyses can help in setting priorities.

4.4. Analytical Material from Supplementary, Special Purpose Surveys

4.45. Supplementary, special purpose surveys are another tool for assessing the exhaustiveness of the national accounts. These are surveys that are not part of the basic data collection programme. They can take a variety of forms, being special surveys of expenditure, income, labour, time use, and opinion surveys. They can be designed to target any or all of the NOE problem areas – underground, illegal, informal sector and household production for own use. They can be conducted by the national statistical office or by other agencies including, for example, sample audits by the tax authorities. However, the results must be interpreted very carefully, especially for those surveys focussed on sensitive subjects. For example, in surveys relating to tax evasion, it is very likely that the non-response is selective because people who are involved in tax evasion are more likely to refuse to co-operate than people who are not. Such non-response is difficult to reduce or to adjust for.

4.46. Supplementary surveys in the context of assessment of GDP estimates are summarised in this section. They are further elaborated in later chapters in association with the other purposes for which they are used. In particular, they are described in considerably more detail in Chapters 8-11 in the context of standalone measurements of the individual NOE problem areas. Also, in so far as the results of such surveys provide parameters or adjustment factors for compilation of the national accounts, they are further described in Chapter 5. To the extent that they eventually become integrated within the basic statistical data collection programme, they are further discussed in Chapter 6. For example a survey of retail sales in city markets may begin as an *ad hoc* survey intended to check on the relative size of such sales. From this survey, adjustment factors may subsequently be derived for use in the national accounts expenditure estimates, and finally the survey may become a regular part of the basic programme of retail trade statistics.

Surveys of expenditure on goods and services from underground production

4.47. Expenditure on underground production is a less sensitive topic than underground income. In the latter case the respondent is being asked to report on fraud whereas expenditure on underground production is usually not forbidden nor prosecutable. Indeed, the purchaser may well not know whether a seller is operating underground or not. For example, if one orders a beer in a café, one does not know whether the owner of the café pays all value added tax or not.

4.48. Kazemier and Van Eck (1992) provide an example of a survey on expenditure on underground production. The survey was on the subject of home maintenance and home repair and included questions on both underground labour and underground expenditure, *i.e.*, expenditure on building materials without paying value added tax.

Surveys of labour input to underground production

4.49. In surveying the input of labour to underground production, there are two alternative approaches, namely surveying the *demand* and surveying the *supply*. Questions on demand are less sensitive than those on supply and can thus be expected to yield larger numbers. Zienkowski (1996) provides an example.

4.50. Experiences suggest that, in an anonymous interview, many people are willing to admit part or all of their underground (but not illegal) production activities. However, surveying supply is prone to item non-response and incorrect response and so needs careful design. Kazemier and Van Eck (1992) show how a sequence of questions, that gradually lead to the key questions on underground activities, gives the best results. The methods they describe are further elaborated in Chapter 8 on underground production.

Surveys of time use

4.51. Time use surveys are generally used to measure the time spent on activities like household work, do-it-yourself, neighbour help and voluntary work. For examples, see Goldschmidt-Clermont (1987), Goldschmidt-Clermont and Pagnossin-Aligisakis (1995) and Organisation for Economic co-operation and Development (Editor Ann Chadeau, 1995). However, they can also be used to measure parts of the NOE that should be included in the national accounts. In particular they can provide insight into the size and structure of household production for own use, as further elaborated in Chapter 11.

4.52. If used in the context of the national accounts, special care must be taken in classification of the relevant activities. For a number of activities, additional questions are necessary. For example, it must be possible to distinguish between time spent on working as an employee in an enterprise and time spent as an own account worker. If respondents acknowledge that time is spent working on own account, they should be asked some additional questions to enable a classification of their work by branch and an estimate of the turnover and the amount of income earned. If the latter information cannot be obtained, questions on the living conditions and the wealth of the responding household, which can be part of the block containing general questions to classify the household, may provide an indication of the profits earned.

4.53. Similar information can be collected in a regular labour force survey. However, the advantage of the time use approach is that the questionnaires require respondents to account fully for their time. This improves the reliability of the results if everything else (geographic coverage, sample size, response rates, etc.) is equal.

Surveys of the informal sector and household production for own use

4.54. Surveys specifically designed to measure the informal sector or household production for own use can shed light on the extent of activities in these NOE problem areas. As with surveys of time use, they have to be carefully designed so that the results they provide regarding the extent of the NOE can

be blended with the results of other surveys in the sense that there is no overlap between them, or, if there is overlap, it is known. These surveys are discussed in detail in Chapters 10 and 11.

Qualitative surveys

4.55. Qualitative surveys of enterprises and of households can also provide information about non-observed activities. They have a number of advantages relative to quantitative surveys. They are quicker and easy to answer. They can be easily changed or supplemented to deal with new circumstances. They can also address questions regarding the causes as well as prevalence of non-observed activities. They can be addressed to very senior staff in surveys of large enterprises but are equally effective with very small enterprises. They can be designed to be less threatening than quantitative surveys by using questions referring to an industry or population group rather than the particular respondent. For example in place of asking a business respondent whether they accept "under the table" payments, the question can be phrased along the lines "What proportion of payments do you think are made in cash in your industry?"

4.56. Whilst qualitative surveys do not often lead directly to quantitative values that can be directly incorporated into national accounts estimates, they can provide impressions of the size of specific types of NOE activities. For example, approximate ratios of observed to concealed production can be obtained for each of the various branches of industry. They can also provide guidance in assigning priorities for subsequent more precise quantitative assessment.

4.57. The Russian Federation Centre for Economic Analysis (2000) describes opinion surveys of businesses in retail, construction and manufacturing conducted on an experimental basis in the Russian Federation. In a study of the underground economy as viewed by households, the Hungarian Central Statistical Office (1998) included some qualitative questions aimed at collecting citizens' views. Further details are provided in Chapter 8.

Analyses of tax audit data

4.58. Quantitative surveys of tax evasion are unlikely to yield reliable results because of the delicate nature of the subject, even if anonymity is guaranteed. Tax audits by their very nature are more compelling than surveys. The "respondents" are obliged to provide their complete accounts, not simply information derived from them. However, because they are designed for tax auditing not statistical purposes, tax audit samples have limitations for estimating undercoverage of the GDP, typically including the following:

- the definitions they use may not be consistent with 1993 SNA;
- they do not include all undeclared income, only what the auditors can find based on their examination of the accounts;
- they refer only to small enterprises;
- they are usually clustered in certain activity sectors and/or geographic areas;
- they are rarely selected on a probability basis.

4.59. Nevertheless, in the absence of any better source, tax audit samples can provide useful information on some types of non-observed activities, in particular those associated with underreporting. This is illustrated in Section 5.2.5 in the description of the procedures used by INSEE to derive adjustments for underreporting.

4.5. IMF Data Quality Assessment Framework

4.60. Exhaustiveness must also be seen within a broader context, namely that of the overall quality of the national accounts. A tool that provides a structure and a common language for assessing data quality in general, including that of the national accounts, is being developed by the International Monetary Fund (IMF). Referred to as the Data Quality Assessment Framework (DQAF), it complements

the quality dimension of the IMF's Special Data Dissemination Standard (SDDS)[1] and General Data Dissemination System (GDDS). It also aims to assess even-handedly the quality of the data covered by the IMF Reports on the Observance of Standards and Codes.[2] Its construction and general structure are summarised in the following paragraphs.

4.61. As noted by Carson (2001), a data quality assessment framework should be:

- comprehensive as regards the dimensions of quality and the elements (indicators) that might represent quality;
- balanced between the rigor desired by an expert and the bird's-eye view desired by a general data user;
- structured but flexible enough to be applicable across a broad range of stages of statistical development;
- structured but flexible enough to be applicable (at least) to the major macroeconomic datasets;
- designed to lead to transparent results; and
- arrived at by drawing on best practices of national statisticians.

4.62. Consideration of these criteria leads to a five level hierarchical structure, in two parts, going from the general to the specific, as shown in Figure 4.1. First, recognising that data quality, in the sense of fitness for use and meeting users' needs, is multidimensional, there is a general purpose set of *dimensions of quality* which are broken down into *elements* and *indicators* that constitute *pointers to quality*. These three levels are collectively termed the *generic framework*. A full copy of the generic framework is reproduced in Annex 5. The generic framework embodies internationally accepted core principles for official statistics. It provides the basis for the fourth and fifth levels of the DQAF, which contain more detailed and concrete sets of pointers termed *focal issues* and *key points*. These are collectively referred to as the *data specific framework* and vary according to the type of data being assessed. There are data specific frameworks for national accounts, balance of payments, monetary, government finance, price, and employment statistics, etc. For the purposes of this Handbook, the framework of most interest is that for the national accounts.

4.63. The first level of the DQAF defines five *dimensions* of quality: integrity, methodological soundness, accuracy and reliability, serviceability, and accessibility, as well as a set of prerequisites, or institutional preconditions,[3] for quality. They are described in more detail in the paragraphs below. For each dimension, element, and indicator, the generic framework presents a brief statement of good practice. The data-specific frameworks provide more detail in the form of focal issues for each indicator that are tailored to the dataset in question. Further, the bullet points below for each focal issue are the key points that describe quality features that may be considered in connection with the focal issues. Although they are considerably more specific than the generic framework, the data-specific frameworks cannot, and indeed are not meant to cover all quality issues exclusively.

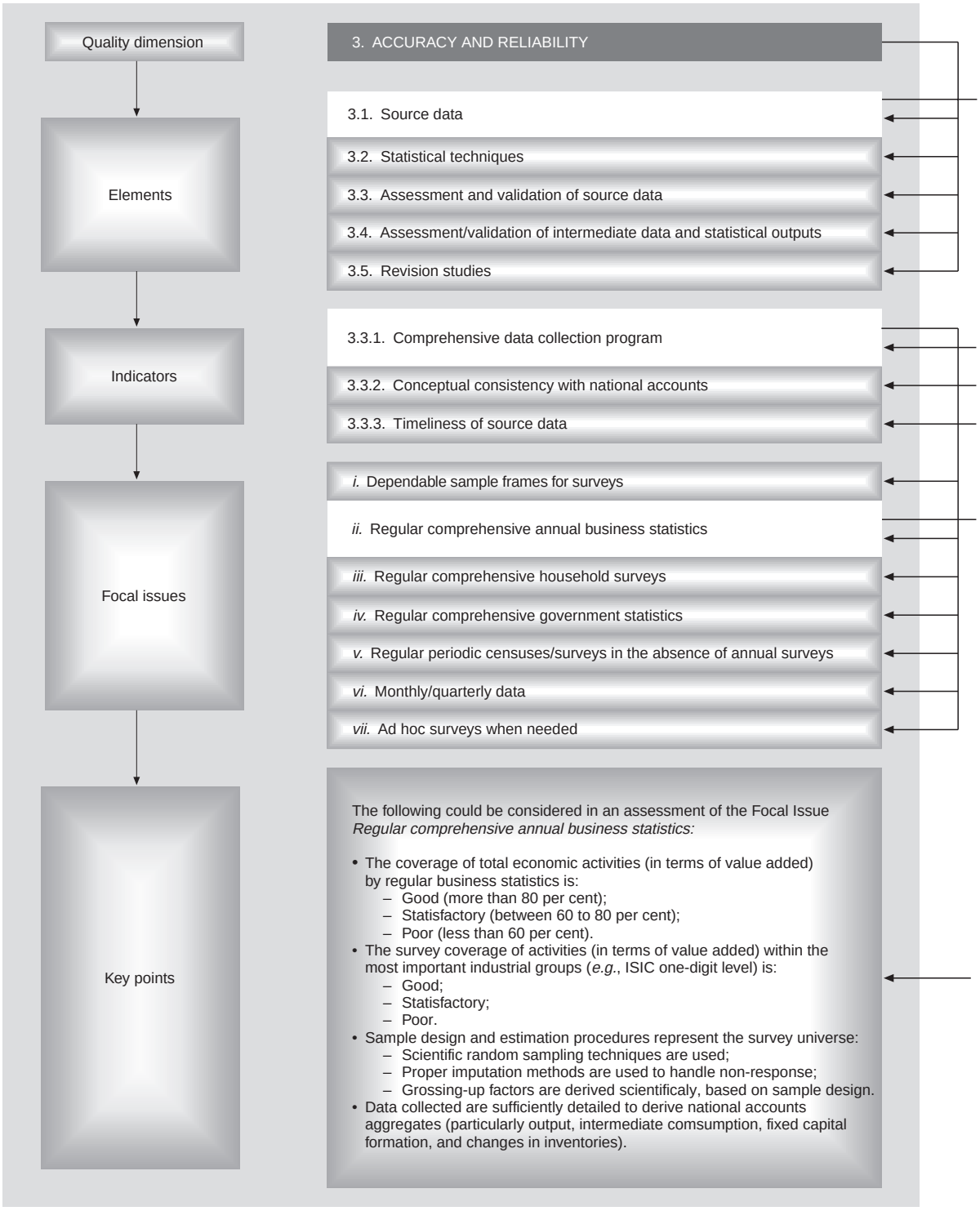
4.64. Using *accuracy and reliability* as the example of a quality dimension, Figure 4.1 shows how the framework identifies five elements that point towards quality. Within the *source data* element, the framework identifies three indicators. Within the *comprehensive data collection program* indicator there are seven focal issues, which are dataset specific. For the *regular comprehensive annual business statistics* focal issue, quality is assessed by considering four key points.

Prerequisites of quality

4.65. The quality of an individual dataset is intrinsically bound together with that of the institution producing it. The legal framework should be supportive of statistics; resources should be commensurate with the needs of statistical programmes; and quality should be recognised as a cornerstone of statistical

1. For a description of the SDDS, see *http://dsbb.imf.org/sddsindex.htm*.
2. Materials on the reports can be found on *www.imf.org/external/np/rosc/index.htm*.
3. The DQAF recognises that the quality of a dataset is intrinsically linked with that of the institution producing it. Thus, it also includes some elements and indicators that, although not constituting a quality dimension in themselves, have an important role as pointers to quality. They generally refer to desirable attributes of the agencies of the statistical system.

Figure 4.1. **Example of structure of data quality assessment framework for national accounts estimates**



work. For example, one element of institutional preconditions is whether the responsibility for compiling national accounts is clearly delineated, while another element is whether there exist mechanisms to co-ordinate the exchange of information between data producing agencies.

Integrity

4.66. This dimension is intended to capture the notion that statistical systems should be based on firm adherence to the principle of objectivity in the collection, compilation, and dissemination of statistics. The dimension encompasses the institutional foundations that are in place to ensure professionalism in statistical policies and practices, transparency, and ethical standards.

Methodological soundness

4.67. This dimension of quality covers the idea that the methodological basis for the production of statistics should be sound and that this can be attained by following international standards, guidelines, and generally accepted practices. In application, this dimension will necessarily be data-specific, reflecting differing methodologies for different datasets. For example, the 1993 SNA or ESA 1995 provide the yardstick for the overall methodological soundness of the national accounts. Specifically, the principles of these manuals are used to assess the coverage of the constituent units of an economy, and to determine the production and asset boundaries.

Accuracy and reliability

4.68. For most users, accuracy and reliability are among the most sought-after attributes of data. We are all concerned that the data we use sufficiently portray reality at all stages of dissemination, from first (or "flash") estimates to "final" estimates. Thus, this dimension relates to the notion that source data and compilation techniques must be sound if data are to meet users' needs. For national accounts, the coverage and comprehensiveness of data sources used in their compilation are evaluated. Elements on the statistical techniques cover the level of industrial detail at which the output and intermediate consumption estimates are compiled, the size of the economy that is covered by regular national accounts compilation, degree of reliance on fixed ratios derived from outdated benchmarks, the techniques used to address specific issues of GDP compilation, and the procedures for compiling volume measures of GDP according to the production approach. The measures of the accuracy of source data are tracked on the basis of different types of errors and the extent to which estimates are underpinned by observed data either from surveys/censuses, or from appropriate administrative records. One way of testing the reliability of the national accounts estimates is by analysing the magnitude and causes of revisions. The framework contains indicators on whether such studies are made routinely.

Serviceability

4.69. Another area of concern for users is whether the data that are produced and disseminated are actually useful. This dimension of quality relates to the need to ensure that data are produced and disseminated in a timely fashion, with an appropriate periodicity, provide relevant information on the subject field, are consistent both internally and with other related datasets, and follow a predictable revision policy. Timeliness and periodicity are addressed within the context of the requirements of the IMF dissemination standards. The consistency indicators are tracked on the basis of specific measures, such as availability of consistent time series, consistency between annual and quarterly accounts, and consistency with balance of payments statistics and other related datasets. The revision issue is addressed in terms of the transparency of the revision process, and in terms of release and timing of adequate documentation.

Accessibility

4.70. Users want understandable, clearly presented data and need to know how data are put together as well as be able to count on prompt and knowledgeable support from data producers for their questions. Thus, this quality dimension relates to the need to ensure that clear data and information on the sources and methods used to compile them are easily available, and that assistance to users of data is adequate.

Implications for NOE *measurement*

4.71. NOE measurement can thus be seen to have a place within the broader context of quality assessment over five dimensions of (plus prerequisites for) quality. From the NOE perspective, the key points, focal issues, and indicators associated with the *accuracy and reliability* dimension are of particular relevance.

4.6. Eurostat National Accounts Process Table

4.72. Another general-purpose national accounts measurement tool, termed the National Accounts Process Table (NAPT), is being developed and pilot tested by Eurostat (2001) in co-operation with the statistical offices of European Union (EU) Member States. It arises from the need for an objective assessment of the quality and comparability of the national accounts estimates of EU Members. This is especially important given that EU Member contributions depend upon gross national income (GNI), which is obtained by transition from GDP.

4.73. Member States provide Eurostat with detailed sources and methods metadata, a considerable amount of which will soon be available on the Eurostat website. These metadata follow a common structure and enable well-founded *qualitative* judgements of the reliability, comparability, and exhaustiveness of Members' national accounts. The NAPT is an attempt to provide the basis for a *quantitative* assessment. Although developed independently of the IMF's DQAF, it provides, in fact, a rather detailed elaboration of the DQAF accuracy and reliability dimension.

4.74. The NAPT embodies a condensed and structured quantitative description of the GNI compilation process. It is stylised to highlight the main successive stages by which national accountants make adjustments to basic data to arrive at final estimates. The table consists of three layers: a top layer containing numbers; a second layer with references to the appropriate sections in the sources and methods metadata; and a third layer with questions that summarise specific points relating to the compilation stages. For NOE measurement, the first layer, illustrated in Table 4.5, is of special interest.

4.75. The columns of Table 4.5 reflect the different stages in compiling the national accounts aggregates that are specified on the rows. Three blocks of columns are distinguished. A first block deals with the *primary data sources*, which can be from surveys or censuses or administrative files (*i.e.*, *basic data* in the nomenclature of this Handbook), or extrapolation models, or other sources. Here, *other sources* refers to data taken directly from other statistical accounting systems, such as the balance of payments, also to the results of special one-time investigations. The primary data sources also include *extrapolation models*, examples of which are those used for imputing rents for owner-occupied housing, for consumption of fixed capital, and for extrapolation from benchmark years.

4.76. The second block shows the adjustments made to the primary data sources. They are broken down into data validation adjustments, conceptual adjustments, adjustments for exhaustiveness and balancing adjustments. Data validation adjustments are corrections to the data sources that are uncovered because the various data sources used in the national accounts compilation process contradict each other. Conceptual adjustments are changes made to the data to bring them in line with the ESA 1995 definitions. Exhaustiveness adjustments include adjustments for non-observed activities that are not included in the primary data sources. Balancing adjustments are adjustments made for no other reason than to satisfy national accounting identities. The sum of all columns should equal the final estimates, which are in the third block.

4.77. The rows in the table show the three different approaches to the calculation of GDP and the transition to GNI. The classifications used and the degree of detail are indicated in the second column of the table.

4.78. The table does not provide quantitative estimates of the accuracy or reliability of the national accounts estimates. It does, however, show the extent to which these aggregates are based upon real data and the extent to which these data are modified and adjusted. Together with methods and sources metadata, comparisons over time and between countries can be expected to uncover the relative strengths and weaknesses of the various methods.

Table 4.5. **Eurostat national accounts process table, layer 1: Quantitative overview**

	Level of detail	Basis for NA Figures					Adjustments				Final estimate
		Surveys + censuses	Administrative data	Extrapolation + models	Other	Total	Data validation adjustments	Conceptual adjustments	Explicit exhaustiveness adjustments	Balancing adjustments	
Production approach											
Output of goods and services (basic prices)	17 (NACE)										
Intermediate consumption (purchasers' prices)	17 (NACE)										
Gross value added (basic prices)	17 (NACE)										
Taxes on products											
Value added type taxes											
Other taxes on products											
Subsidies on products											
Residual item											
Gross domestic product											
Expenditure approach											
Total final expenditure											
Household final consumption	COICOP 2-digit										
NPISH final consumption											
General government final consumption											
Gross capital formation											
Gross fixed capital formation											
Changes in inventories											
Acquisition less disposals of valuables											
Exports of goods and services											
Goods											
Services											
Imports of goods and services											
Goods											
Services											
Residual item											
Gross domestic product											
Income approach											
Compensation of employees											
Gross operating surplus and mixed income											
Taxes on production and imports											
Subsidies											
Residual item											
Gross domestic product											
Compensation of employees received from Rest of world (ROW)											
Compensation of employees paid to ROW											
Property income received from ROW											
Property income paid to ROW											
Taxes on production and imports subsidies											
Gross national income											