

V Editing and Reconciliation

A. Introduction

5.1. Editing and reconciliation are essential stages of statistical production and are among the tasks in national accounts compilation that require the greatest skill. While other chapters deal with the sources of data and techniques, this chapter emphasizes reviewing and understanding the data. The process of reviewing and understanding data can be called “editing,” “checking,” or “data validation.” It should occur at all stages—before, during, and after—of the calculation of the estimates. “Reconciliation” or “confrontation” is a special kind of editing done after initial compilation, in which alternative data are checked in the context of national accounting relationships. Editing and reconciliation may involve fixing errors or adopting alternative sources and methods; these tasks should, however, never be an excuse for manipulating data without evidence or adjusting data to fit forecasts or for political reasons.

5.2. National accounts compilation is a complicated process, bringing together a wide range and large volume of data. The data cover different periods; come from varying sources; are of varying quality; and may have different units, concepts, and timing. Large volumes of data mean that mistakes are easy to make and hard to find. In addition, when a method or program has worked well in the past, the production process has gone smoothly, or the calculations are complicated, there is a natural tendency for busy compilers to accept the data without close scrutiny, resulting in a risk of errors.

5.3. Data suppliers are an integral part of national accounts compilation, so editing should be supplemented by continuing contact with suppliers to gain knowledge from them about problems they have identified or suspect. In addition, the national accounts compilation process itself may shed new light through volume measures, seasonally adjusted and trend-cycle

data, analysis of revision patterns, and reconciliation with related data sources. Thus, communication needs to be in both directions.

5.4. Many of the reconciliation and editing issues in quarterly national accounts (QNA) are the same as in annual national accounts (ANA). However, these issues are particularly important in the compilation of QNA. Deadlines for QNA are usually much tighter than for ANA, work is more rushed, and a higher proportion of source data may be preliminary or unpublished. As a result, errors are more likely to occur. There is typically less detailed information in QNA. The tight deadlines applying to quarterly compilation impose a severe limit on the amount of investigation done for the latest quarter. In the time available, it may be necessary to limit checks to known problem areas, the most recent periods, and some major ratios. In the time between the end of one quarterly compilation cycle and the beginning of the next, however, there may be opportunities to undertake further investigation.

5.5. The highest priority in editing is usually to *identify and remove errors* before publication; however, there are other benefits. Editing helps national accountants *understand the data and the economy better*. It also helps national accountants *anticipate queries from users*, because unusual movements will already have been identified; explanations for the expected queries can thus be given immediately. Successful editing enhances both the quality of the data and the confidence of users in the compilation procedures.

5.6. Editing procedures usually rely on *relationships within data* to identify problems and questions. Only rarely will looking at a single number help point to anomalies. The foundation of editing is to compare observations of the same variable in different periods or to compare one variable with other variables that are expected to have some linkage.

5.7. Editing and reconciliation may result in changes in the estimates. It is important that such changes are justified and documented. For example, sometimes mistakes are identified and the correct figure can be used instead. In other instances, a method may have become unsuitable because the assumptions behind it have become obsolete, or the source data may have problems in reporting or coverage. A distinction needs to be made, however, between editing and unacceptable manipulation of data. An unexpected change in a series should lead to checking that there is no error or problem with the data source. Editing may suggest that an alternative source or method is justified; however, data should not be changed just because they are unexpected, as this may lead to charges of manipulation and may undermine the reputation of compilers if it becomes known. Further, in reality, many unexpected developments occur, and the purpose of QNA is to show actual developments in the economy, particularly when they are unexpected. In line with principles of integrity and transparency, QNA estimates should be able to be explained by reference to source data, publicly available compilation methods, and adjustments documented with the supporting evidence.

B. Causes of Data Problems

5.8. There is a range of causes for failure of data to fit expected relationships. When there is a data problem, it is first necessary to confirm that the input data are consistent with those supplied by the data collectors. If the QNA are compiled by computer, as is the usual case, it is necessary to confirm that the computer program is doing what was intended. This check will show whether any anomalies were due to mistakes made in the national accounts compilation system itself. In the interest of good relationships with data suppliers, the possibility of an error in the compilation system should be excluded before pursuing other avenues of inquiry. Causes of data failing to fit expected relationships include the following:

- (a) *Errors in data entry by national accounts compilers.* These include mistyping of numbers, putting numbers in the wrong place, and using old data that should have been updated.
- (b) *Errors in national accounts compilation systems.* At a basic level, these include wrong formulas, which are particularly likely when changes are made to programs, especially in spreadsheets. In addition, the assumptions and indicators may become inappropriate as conditions change; for example, use of a generalized deflator or direct deflation of value added may give acceptable results when there is little relative price change but may become quite misleading under different economic circumstances. Adjustments are required when data sources do not fully meet national accounts requirements and are particularly prone to becoming outdated by economic changes. Examples are adjustments for timing, valuation, and geographic/size/product coverage.
- (c) *Errors in data recording by respondents.* Reporting quality is often a problem, but it can be improved by good questionnaire design, helpful completion instructions, and availability of assistance in completing forms. Timing problems can be particularly important in QNA. Timing problems occur when transactions are not recorded at the time required by the 1993 SNA. The 1993 SNA standard is based on accrual principles and change of ownership; however, many data sources do not meet these requirements. Government data are often recorded on a cash basis. International trade data are typically recorded at the time the goods cross the customs frontier or when the customs authorities process the form. Administrative byproduct data (e.g., value added or payroll tax data) may cover periods that do not coincide with a quarter because the agency is more interested in tax collection than statistical objectives. Businesses may also use different accounting periods that do not exactly match the three-month period used in the QNA, such as weeks, four-week periods, or nonstandard quarters. These problems are also found in annual data but are more significant in QNA because a timing error of the same size is relatively larger in quarterly data.
- (d) *Errors and problems in source collection systems.* Problems can occur in classification, data entry, estimation of missing items or returns, sample design, tabulation, treatment of late response, incomplete business registers, and omitted components. Estimation of nonreporting units is a particularly important issue for QNA because of the higher proportion of missing data owing to earlier deadlines. Early estimates are often based on incomplete response, complemented by estimation processes for the missing respondents. Treatments of outliers may also differ. A systematic difference between early and late estimates suggests that the estimation for the missing components is biased. Large but nonsystematic errors suggest that it would be desirable to put more effort into early follow-up. National accounts compilers need to be sympathetic to the constraints of resources and respondent cooperation faced by their data collection colleagues.

- (e) *Changes in the structure of economy.* In many instances, it is possible to confirm that there has been a surprising but valid change in the series owing to a known cause, such as a large individual transaction or a business closure. This information helps the national accountant understand the data and deal with queries from users. Some changes in the structure of the economy have the effect of making assumptions used in the national accounts compilation obsolete and so may require changes in methods. For example, the representativeness of an indicator that does not fully match the required coverage may deteriorate.
- (f) *Inexplicable reasons.* There is also likely to be a residue of cases where the movement is surprising, and neither an error nor an actual cause can be found. It is still better to know about such cases, so that a query from a user is not a surprise and in case an explanation subsequently comes to light.

5.9. The causes of some data problems are obvious, while in other cases investigation is needed to identify the cause. Some can be easily resolved, while others involving data collection will take longer to implement; examples of the latter may include problems that require changes in survey coverage or questionnaire design, design of new imputation methods for nonresponse, or revised procedures for incorporation of new businesses in surveys. Even where it is not possible to fix or explain data immediately, it is important that the issues be identified for later investigation and resolution.

C. How To Identify Data Problems

5.10. In this chapter, various ways of identifying data problems are presented. The terminology and classification were developed for this chapter because there is little or no literature about its subject and no standard terminology.

1. Eyeball Testing

5.11. “Eyeball testing”—that is, just looking at the numbers as they will be published, without any additional calculations, tabulations, or charts—is the most basic kind of editing. Even with this limited presentation of data, a number of potential problems will be apparent to the careful eye:

- Different orders of magnitude, different numbers of digits.
- Numbers that change too much—excessive growth or decline.

- Numbers that do not change at all—no change at all may suggest that numbers have been copied into the wrong period.
- Numbers that change too little—a much slower growth than other items may point to a problem.

5.12. Eyeball testing does not use a computer or other tools to pinpoint problems, so it depends solely on the editor’s ability. As a result, many data problems will not be apparent and may be missed. Despite these limitations, such a basic examination can be implemented quickly and is much better than no editing at all. Someone who was not involved in the original calculations is more likely to notice potential problems.

5.13. A slightly more sophisticated check is to present the numbers as charts. Charts of data can be generated readily with spreadsheet and other packages. Unusual movements and inconsistencies stand out in charts to a much greater extent than they do in tables.

2. Analytical Testing

5.14. A more advanced form of editing uses additional calculations or charts to assist in checking data. It is a more sophisticated and time-consuming form of editing but will usually reveal more problems than eyeball testing alone.

a. Logical

5.15. Logical edits are those in which exact relationships must hold, based on mathematical identities or definitions, such as in the following examples:

- Total is equal to the sum of components (e.g., $GDP = \text{Household final consumption} + \text{Government final consumption} + \text{Gross fixed capital formation} + \text{Changes in inventories} + \text{Acquisitions less disposal of valuables [if applicable]} + \text{Exports of goods and services} - \text{Imports of goods and services}$; $\text{Manufacturing} = \text{Food} + \text{Textiles} + \text{Clothing, etc.}$).
- Commodity balances, which are checks of the relationship between supply and use when they have been derived independently. They can best be done as a part of a comprehensive supply and use framework in which balancing and interrelationships between components are dealt with simultaneously. Even without a comprehensive supply and use framework, however, balancing supply and uses of particular products is a useful way to find errors or inconsistencies between data from different data sources. (If the supply and use data are complete, this is a logical edit.)
- Year is equal to the sum of the quarters (in original data; not necessarily true in seasonally adjusted or trend-cycle data).

- Definitions of specific terms (e.g., Implicit price deflator = Current price value/Constant price value; Value added = Output – Intermediate consumption).

5.16. Rounding errors may sometimes disturb these relationships slightly, but they should be relatively minor and not used as an all-purpose excuse for acceptance of inconsistency.

b. Plausibility

5.17. Edits of plausibility rely on expectations of how series should move in relation to past values of the same series and to other series. In contrast to logical edits, there is not an exact requirement that the data must satisfy; rather, data can be seen as being in a spectrum that goes from expected values, to less expected but still believable values, to unusual values, and on to unbelievable values. This assessment requires an understanding of what is a realistic change; that is, the national accountant must have a good grasp of economic developments as well as an understanding of the statistical processes.

5.18. It is important to assess QNA indicators for their ability to track movements in the corresponding annual series. As explained in Chapters II and VI, the annual benchmark-indicator (BI) ratio shows the relationship between the two series. A stable annual BI ratio shows that the indicator is representative. Alternatively, a trend increase or decrease in the BI ratio points to bias in the movements of the indicator series. Volatile changes in the annual BI ratio point to problems that are less easily diagnosed and solved.

5.19. The following are some other editing calculations that can be made to assess the plausibility of data:

- Percentage changes (e.g., for quarterly estimates, compared with one quarter or four quarters earlier) can be calculated. These can help identify cases where rates of growth or decline are excessive, or where one component is moving in a different way from a related series. It may be feasible to develop thresholds to identify unusual changes on the basis of past behavior. As well as being useful in editing, percentage change tables are a useful supplementary way of presenting data.
- Contributions to growth,¹ which show the factors behind growth in aggregates (rather than just

growth of series in their own right), can be calculated.

- Commodity balances can be made. (These are already discussed under logical edits. If the supply and use data are incomplete, this is more a test of plausibility.)
- Ratios of various kinds can be calculated (particularly where series have independent sources):
 - ▶ Implicit price deflators—that is, the ratio of current price values to constant price values, are a kind of price index.
 - At a detailed level, if the value and volume measures have been obtained independently, a peculiar implicit price deflator movement will indicate incompatible trends.
 - At an aggregated level, it is useful to calculate the corresponding Laspeyres price indices. Comparison between the Laspeyres price indices and implicit price deflators points to the effect of compositional changes on the implicit price deflators. No extra data are required to calculate the Laspeyres price indices, and they are of analytical interest in their own right.
 - ▶ Productivity measures show the relationship between inputs and output/value added and, hence, may point to problems in input or output data. The most common and simple measure is labor productivity, that is, output or value added at constant prices per employee or hour worked. For example, the output, value added, and employment series may not look unreasonable individually, but they could be moving in incompatible ways. In this case, the productivity measure will highlight the inconsistency in the trends by the implausible movement. Some countries publish labor or total factor productivity estimates; again, these are of analytical interest.
 - ▶ Ratios between other closely related series (e.g., construction in gross fixed capital formation and construction output in production estimates; value added and output for the same industry; components to total ratios, such as manufacturing/total; inventories/sales).
 - ▶ Other ratios between series. Less stable ratios will occur for series that are linked by behavioral relationships, for example, consumption and saving to income, current account deficit to saving. However, changes in these ratios can point to data problems and also help national accounts compilers advise data users.
- Implicitly derived series should be examined closely, as they may highlight data problems, for example, intermediate consumption when value added has been derived with an output indicator.

¹Calculated as $(x_t - x_{t-1}) / A_{t-1}$ where x is the component series and A is an aggregate. For example, if household consumption has increased by 5 since the previous period, and GDP was 1000 in the previous period, the change in household consumption makes a contribution to GDP growth of 0.5 percentage point.

- Revisions (since the previous publication or several publications earlier) should be examined. Newly introduced mistakes will show up as revisions. Consistent patterns of revisions (i.e., consistently upward or downward) suggest a biased indicator. Large, erratic revisions may indicate a problem with early data that can be investigated. The incorporation of annual benchmarks into quarterly estimates will cause revisions and could reflect problems in the sources or methods for either annual or quarterly data. To calculate and track down the causes of revisions, it is necessary to archive data from previous releases, by keeping printouts and copies of computer files or by saving earlier data in the computing system under separate identifiers.

5.20. It is not a coincidence that many of these tools for plausibility editing are also of interest to users of the statistics. Both editors and analysts are performing similar tasks of looking at how the data are moving and why.

5.21. Analytical editing can be done with charts or tables. Usually, the interest in this case is in big changes rather than precise relationships. Charts are particularly suitable in this task because they can be read by glancing, especially to identify outliers. Line charts and bar charts are alternative presentations that give different emphases. Charts may sometimes take more time to set up than tables but are worth it because of their usefulness. Tables allow errors to be traced more easily because an exact number is known, so they might be used to investigate a problem detected by a chart. Choices between charts and tables are often influenced by the capacities of the computer processing system being used. Different formats each have their own uses, so it is desirable to have a range of presentations.

5.22. In general, editing and reconciliation are best done at both detailed and aggregate levels. In aggregate form, problems can be hidden by large values of data or by errors in offsetting directions canceling each other out. With more specific identification of the affected components, it is possible to focus on the cause of the problem. Some problems are only apparent at a detailed level, because they get swamped at a higher level of aggregation. In other cases, the level of “noise” or irregular movements in the series is high at a micro level, so problems may become more obvious at a higher level, as the noise in the series becomes relatively smaller.

5.23. Problems are sometimes more apparent in constant-price or seasonally adjusted data. These presen-

tations remove some sources of volatility and hence isolate remaining fluctuations. For example, an unadjusted series may have a strong seasonal pattern, with quarter-to-quarter changes so large that trends and irregularities are hidden.

5.24. Discrepancies and residual items should receive particular attention because they are not derived directly, and problems in certain components are often highlighted by the balancing item.

D. Reconciliation

5.25. When there are two or more independent measures of an item, inconsistencies inevitably will arise. The inconsistencies could be between two measures of GDP estimated by different approaches or, in a detailed system, between the supply and use of a particular product. Reconciliation is the process of dealing with these inconsistencies. This section discusses different options for reconciliation and the considerations that need to be taken into account in choosing among them. Reconciliation issues arise in both annual and quarterly estimates. The approach to ANA reconciliation will typically be the starting point for QNA, although some different approaches may emerge because of the quarterly emphasis on speed and time-series maintenance. In addition, the QNA data will be strongly influenced by the reconciliation carried out in the annual data because the annual balances (or imbalances) will be passed to QNA through the benchmarking process. The options available are reconciliation by detailed investigation, reconciliation by mathematical methods, or publication of discrepancies in varying ways.

5.26. One important type of reconciliation is the process of balancing data at a detailed level within a full *supply and use* (or *input-output*) table framework or through commodity balances for key products. Supply and use tables provide a coherent framework to identify inconsistencies at the detailed product level. Supply and use balancing is at its most useful when investigations are used to identify the cause of discrepancies. Even if supply and use data are not available in a comprehensive framework, a partial version in the form of commodity balances for particular products can provide some of the benefits of supply and use tables for reconciliation. A few countries use a supply and use framework on a quarterly basis, typically at a less detailed level than annually and as a compilation tool that is not intended for publication.

5.27. Another type of reconciliation occurs when there are *independent estimates of GDP* by two or more approaches but without the details of a supply and use framework.² In such cases, discrepancies become apparent only when the data are aggregated, making well-based reconciliation difficult or impossible because the aggregate discrepancies provide no indications of which components are causing the discrepancies. Investigations may still prove useful, however, as patterns in the discrepancies may point to specific problems (e.g., reversed fluctuations point to timing problems, persistent differences of a similar size point to a bias in a major source, and procyclical differences may point to problems in measuring new businesses).

5.28. Some countries have a mix of methods in which supply and use balancing occurs on an annual or less frequent basis, while independent estimates are made quarterly. In these cases, the quarterly discrepancies will cancel out within the quarters of balanced years and generally tend to be smaller because of the benchmarking process.

5.29. A number of countries do not have an apparent problem of reconciliation because they do not have supply and use tables; they have only one approach to measuring GDP; or they have two or more approaches, but only one is derived independently, with one component in the other(s) derived as a residual. Besides the analytical interest of having different approaches, however, discrepancies can be useful pointers to data problems that would otherwise be undiagnosed.

5.30. For both supply-use and independent measures of GDP, investigation and resolution of the problems is the ideal method of reconciliation. The processes of confrontation and reconciliation at a detailed level can identify many issues and are highly regarded by national accounts compilers. The extent of adjustment that can be made should depend on the expertise of the statistical compilers. Adjustment should not be made lightly but should be based on evidence and be well documented. There is potential for concern if uninformed guesses are made or adjustments are made with a view to meeting some political objective (or that accusations could be made that politically-motivated manipulation has occurred). Adjustments should be monitored to see if they later need to be reversed.

5.31. For cases in which there is insufficient time, expertise, or information for investigation to achieve

complete reconciliation, there are a number of alternatives for treating the discrepancies. There is no international consensus, however, and treatments must account for national circumstances.

5.32. One technique to remove discrepancies is the allocation of discrepancies to a single category by convention. The discrepancy is, then, no longer apparent. Usually the chosen category is large (such as household consumption) or poorly measured (such as changes in inventories). In effect, the estimates are no longer independent, and one source is forced to equal the other. As a consequence, the information content of the chosen component is reduced or even lost. And although the discrepancy is hidden in this way, it is not solved. At least, the component should be properly labeled, for example as “changes in inventories plus net errors and omissions.”

5.33. A related option for removing the remaining discrepancies is to allocate them by mathematical or mechanical techniques across a number of categories. The chosen categories could be a selected group or all categories. Methods may involve simple or iterative prorating; for example, the RAS method is an iterative prorating method used for supply and use tables and other multidimensional reconciliation situations. The selection of which categories to adjust by prorating and which categories to leave unchanged should be based on explicit assessments of which estimates were better. Like allocation to a single category, the problem with allocation across several categories is that the process removes some of the information content of the original data. As a result, balance may be achieved at the expense of damaging the time-series quality of the individual components. If an error that belongs in one component is distributed across a number of components, all the components will be less accurate. If the discrepancies are trivial, this may not be of concern. But if they are significant, these techniques merely hide the problem rather than solving it. It is a disservice to users to leave them unaware of the actual extent of uncertainty. Minimizing problems in data sources can also undermine the attempts of national accountants to highlight those problems and reduce the chance of bringing about improvements. Because of the greater significance of timing problems in source data and the reduced time for investigation of the causes of inconsistencies, the limitations of reconciliation are more serious in QNA than in ANA. As a result, some countries that have balanced ANA allow imbalances in QNA.

²These issues are dealt with in Bloem et. al. (1997).

5.34. The alternative to reconciliation by investigation, allocation to a single component, or mathematical removal is to present the remaining discrepancies openly. Within that alternative, one presentation is to publish more than one measure of GDP or supply and use of a product. Alternatively, a single measure can be identified as preferred on the basis of a qualitative assessment of data sources or mathematical testing of the properties of the alternative measures (or a mixture of them). Explicit statistical discrepancy items would then be needed (in aggregate for independent measures of GDP; at the product level for supply and use), so that the sum of the items equals the preferred total.

5.35. The main concerns about showing explicit discrepancies are that they may cause confusion among users and criticism or embarrassment to the compilers. To the extent that the discrepancies represent problems that have identifiable causes and can be solved, the criticism is justified and investigations should have been carried out to make appropriate adjustments. To the extent that the discrepancies are trivial, mechanical techniques would be justified to remove them. In the remaining cases where the differences are significant and the causes unknown, however, it is better to admit the limitations of the data because the uncertainty is genuine. The ultimate objective must be to solve the problem, and being transparent to users about shortcomings is more likely to help bring about the required changes in data collection or compilation resources. While it is understandable that some compilers might be inclined to “sweep problems under the carpet,” in the longer term, being open will avoid even more serious—and valid—criticism about secretiveness and covering up important problems.

5.36. The objective of soundly based reconciliation is the same in both ANA and QNA. Similarly, the options and considerations to be taken into account in choosing between them apply in both situations. There are, however, some procedural and practical differences. Procedurally, QNA reconciliation problems are likely to be most severe for the most recent quarters, because for earlier quarters the same issues would already have been identified in the ANA. Benchmarking brings the benefits of annual reconciliation to QNA, so that additional quarterly reconciliation may be a lower priority. There are also practical considerations, because there is less opportunity to investigate discrepancies during quarterly compilation.

5.37. Benchmarking means that QNA will benefit indirectly from the reconciliation carried out on the

annual data, so that discrepancies may be smaller and reconciliation less urgent. If the ANA are already balanced and the QNA are benchmarked, the need for separate reconciliation is reduced. For the balanced years, discrepancies within quarters will cancel out over the whole year and tend to be small. For quarters outside the annually reconciled period, the discrepancies will tend to be smaller close to the benchmark years. For the most recent quarters that have no annual benchmark, if the indicators correctly track their benchmarks, previously identified causes of inconsistencies will already have resulted in adjustments that are carried forward. Accordingly, the QNA discrepancies will tend to be limited to those caused by noise, divergence between benchmarks and indicators, or data problems that have emerged since the last benchmark. Of course, if the annual data contain unreconciled inconsistencies, they will also be carried forward to the QNA, which will be at least as imbalanced as their ANA equivalents. The implications of benchmarking for reconciliation are discussed further in Chapter VI.

5.38. QNA are typically compiled with less time, information, and detail than ANA. The reduced time and information tend to restrict the capacity to investigate problems that have emerged in the most recent quarters. Timing errors and statistical noise may be difficult to resolve by investigation. These issues are more significant in QNA because they tend to cancel out over a whole year. In terms of user interests, analysis of QNA tends to strongly emphasize the time-series aspects of QNA data rather than structural relationships. Also, in a quarterly supply and use system, the tables are compilation tools and are not generally published in their own right, so that time-series consistency is given more weight than structural balance. Therefore, there is likely to be less investigation and more acceptance of unresolved discrepancies in a QNA system than an ANA system.

E. Editing as Part of the Compilation Process

5.39. Editing can occur at all stages of data processing:

- (a) before receipt by the national accounts compilers,
- (b) during data input (i.e., the data as supplied to the national accounts compilers),
- (c) during data output (i.e., the data as planned to be published), and
- (d) during intermediate stages:

- (i) before and after benchmarking,
- (ii) before and after deflation,
- (iii) before and after reconciliation, and
- (iv) before and after seasonal adjustment.

5.40. Good editing practices should be applied by all compilers of statistical data. Those who collect the data need to monitor the results and anticipate queries for their own purposes. In some countries, the national accounts compilers have contributed toward educating the data collection staff through the perspective that comes from seeing macroeconomic links, from undertaking deflation and seasonal adjustment, and from maintaining consistent time series. In addition, national accounts compilers may have meetings or standardized data supply forms to allow the data collectors to notify them of major movements in the data, known economic developments, response rates, standard errors, changes to questionnaires, and other changes in methods. Good procedures or structures for interaction between data collection staff and national accounts compilers help maintain effective cooperation and avoid conflicts.

5.41. Editing at each stage through the compilation process is desirable. Each stage of processing and adjustment can introduce new errors or hide earlier ones. Earlier identification of problems and errors is generally preferable.

5.42. Original estimates, adjustments, and reasons should be documented along with supporting evidence. As a good practice, when national accounts data are changed during the editing process, the source data, original estimates, and adjusted estimates should be stored. Although only the adjusted data will be published, it is important to be able to document how the source data were amended and the cause of the problem. Documentation is necessary so that the reasons may be understood and verified later. While it is tempting to put off documentation work, memories are not a good substitute, because people move on to other jobs, forget, are on leave at a crucial time, or have conflicting recollections. Documentation is a defense against accusations of manipulation. As later data become available, patterns may be more apparent from a consistent series of original data, or alternative adjustments may be developed. Later information may lead to the conclusion that some adjustments were ill-advised and should be revised. Documentation could be on paper files or, better still, on the computer system if it allows different versions of a series to be saved and associated metadata to be linked to a series.

5.43. The ability of the national accounts compiler to make adjustments is limited if consistency with some or all published source data is a constraint. In some countries, particular data are regarded as binding for QNA compilation because of their relatively high quality or need for consistency. While some sources may not be published, making overt inconsistency not an issue, the basic criterion for adjustments should be their justifiableness. In some countries, data that are known to be particularly poor are identified as being subject to adjustments (e.g., consistency between the production and expenditure estimates being achieved by adjustments to changes in inventories because that component is known to be of poor quality).

5.44. Deciding how much editing work to do depends on staffing, deadlines, and knowledge of the kinds of problems that typically arise. In the abstract, more editing is always better. In practice, the extra work and time required to establish editing systems and then check the data mean that edits must be limited to the types that are most likely to be useful.

5.45. Computers have greatly increased the capacity for editing. At the first stage of computerizing the national accounts, the tasks from clerical systems are often transferred directly to computers without changes. However, this does not fully use the capacity of computers to do additional tasks. The next stage in the evolution of processing is to use the strengths of the computer to implement new tasks, especially editing. Calculations for editing (such as percentage changes and ratios) that would be time consuming in a clerical system involve very little cost in a computerized system and so are much more feasible. At the same time, computerized systems may need more checking because the data processing itself involves less human observation.

5.46. The compilation schedule needs to allow time for editing and subsequent investigation and revision of data. If time is only allocated to carry out basic data entry and calculation tasks, it will not be possible to make any changes before the publication deadline.

5.47. More complicated estimation methods for particular components are at more risk of mistakes. Similarly, the need for editing is stronger when data or methods are weak because the risk of inappropriate results is greater. Because numbers in a computer are all treated as numbers regardless of their origin, it is important for the compiler to bear in mind the link between the quality of data input and the quality of data output: “garbage in, garbage out.”