Part one

The theory of data reconciliation

1 The reconciliation of national accounts

1.1 Introduction

There are, traditionally, three measures of GDP: expenditure, output and income. If these measures are calculated independently, they cannot be expected to be equal, because data are collected by surveys and other disparate sources which cannot be completely accurate. In practice the measures are reduced to two for each of constant and current prices. Factor incomes are not, at least in the United Kingdom, measured independently of the industries in which those incomes are earned. In constant prices output and expenditure are measured separately, but the income measure is calculated by applying the expenditure deflator to current price income, and it does not offer a third independent measure of output. Thus there are discrepancies between income and expenditure measured at current prices and output and expenditure measured at constant prices. This book discusses the question of how these different measures should be reconciled.

The first part of the book presents the statistical theory needed to solve the problem of reconciling inconsistent data. The second part explains how the methods can be applied to the UK national accounts for the period 1920 to 1990 and the third part presents new estimates for the same period, calculated by reconciling the inconsistent estimates provided by Feinstein (1972) for the period 1920 to 1948 and from the official sources from 1948 to 1990. Estimates of the standard deviations of the measurement errors associated with some of the key variables of interest are also presented.

Figure 1.1 shows the discrepancies between expenditure and income at current prices and expenditure and output measured at constant prices as a percentage of the expenditure measure.[1] There are no output measures

[1] Or, more accurately as $100\times\{\log(\text{Income or Output})-\log(\text{Expenditure})\}$.

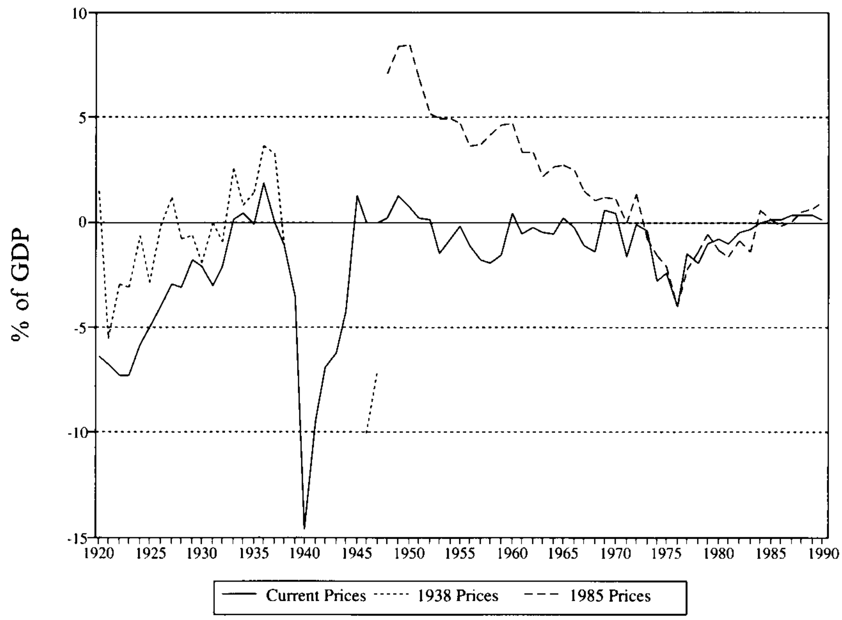

Fig. 1.1 The residual errors in constant and current price GDP

of GDP during the period 1939–45 and therefore no constant price residual can be calculated for this interval. It can be seen that the problems are more pronounced with constant than with current price data, but in neither case are they trivial. The constant price estimates are calculated to 1938 prices for the periods 1920 to 1938 and 1946 to 1947, with none being available for the war years. They are calculated to 1985 prices for the period 1948 to 1990. This means that in 1938 and in 1985 the discrepancies are the same in constant and current prices. In the period after 1948 there is a tendency for the two constant price measures to drift apart – a tendency which is absent from the current price data – while there seems to be a considerable amount of drift in both residuals before 1938. During the Second World War the two current price measures of GDP showed very substantial discrepancies.

The inconsistency between the expenditure and the income/output measure of GDP poses problems for the user of economic statistics. The aim of this book is threefold. First of all it presents the statistical theory of methods for resolving these discrepancies. Second, it treats UK national income as a case study for illustrating these methods and produces new estimates of UK national income and expenditure calculated using these methods. Third, it aims to make data users aware of the likely margins of error associated with the data which they use; to that end it presents estimates of standard errors of the data. The data presented in part three are also available on floppy disc (see Preface)

1.2 The average estimate of GDP

Godley and Gillion (1964) suggested resolving the inconsistency by taking an arithmetic average,[2] and for many years the mean of income and expenditure was the recommended measure of GDP at current prices. In constant prices the average of expenditure, income and output was used.

The method of averaging does not, on its own, allocate the discrepancy across the components of national income. If expenditure is above income, then averaging brings down aggregate expenditure, but it offers no indication as to how much of the adjustment should be allocated to consumption, investment, government consumption or trade flows.

There is another problem which arises from the interaction of constant and current price data. It may well be the case that in current prices expenditure is above income, while in constant prices it lies below output and deflated income. If the current price data are adjusted independently of the constant price data there may be large adjustments placed on the implicit deflator. One would be happy with a downward adjustment of current price expenditure if there were a similar adjustment in constant price expenditure. After all, many of the constant price data are simply measured by deflating the current price data. Account should be taken of this in the reconciliation process.

1.3 Allocation in recent estimates

Recently the presentation of official statistics has changed. The discrepancies are allocated in a manner which reflects, one way or another, the views of the statisticians concerned. While their method is nowhere described in detail, it is clear that it does not involve full use of the procedures explained below.

As far as the constant price measures are concerned, the output measure is regarded as defining movements of GDP in the short term, but in measuring output movements between base years (1948, 1958, 1963, 1970, 1975, 1980 and 1985) no weight at all is put on the output index, so that long-term movements in real GDP are estimated from the average of the income and expenditure data deflated by the expenditure

[2] We are grateful to David Savage for drawing their role to our attention.

GDP deflator. The methods by which the output index is elided into the deflated income/expenditure index are described in *Economic Trends*, No. 420 (CSO, 1988).

The current price data are treated slightly differently. The most recent estimates presented in the 1993 *Blue Book* (CSO, 1993a) show no residual at all between 1985 and 1990. This does not mean that the estimates are exact. The notes to the *Blue Book* explain that there were residuals of between £43m and £217m in 1985–8. These residuals were themselves derived after substantial use had been made of the 1989 and 1990 input–output tables (CSO, 1993a, pp. 144–5) to reconcile disparate sources. They have been allocated equally between income and expenditure, but no clues are given as to the components of income/expenditure to which the adjustments have been made. Thus this is, in aggregate, similar to averaging, with the errors allocated across the data in an undisclosed manner. In 1989 and 1990 no residuals are shown because the data are constructed from the input–ouput tables for those years; these tables are constructed to balance. The residuals remain in the pre-1985 data.

The averaging method is essentially arbitrary. The approach based on input–output tables relies on a great deal of disaggregated data and allows consideration of commodity balances as well as discrepancies between income and expenditure. In practice it may be similar to an application of the techniques described in this book, carried out in finer detail. However it is not applied to a long time-series in one go and so the mechanics of the adjustment process remain obscure.

As a separate exercise the CSO applies the techniques described in this book to a span of three years' data at a time. The results are described each November (e.g. CSO, 1993b), but the outcome does not appear in the standard national accounts, and the process of data revision means that, on each occasion, the coherent data set is only three years long.

1.4 Adjustment on the basis of data reliability

1.4.1 Least-squares/maximum-likelihood adjustment

Intuitively it makes a great deal of sense that those adjustments needed to create accounting balance should be borne in some inverse relation to the degree of confidence one has in the data. Any adjustments to expenditure data should be made disproportionately to variables like stockbuilding because it is known that this cannot be measured satisfactorily. The knowledge that the accounting constraints must be satisfied by the true data provides extra information which can be used to refine the accuracy of the data collected through conventional means.

The solutions to the problem of data inconsistency used hitherto can, of course, be seen in this context. If one believed that income and expenditure data were of equal reliability, then the arithmetic average would be the result produced by adjustment with reference to reliability. But the concept of reliability could be applied to the income/expenditure components and used to allocate the corrections across the series which make up the national accounts. This would lead to a fully consistent set of income/expenditure data.

Such a dataset may lead to important changes in the way economic variables are perceived. For example, in the period 1920 to 1924 a negative personal sector savings ratio is shown by Feinstein (1972). But at the same time the expenditure measure of GDP is well above the income measure. Since personal consumption is the largest component of expenditure and wage income the largest type of factor income, any reconciliation will almost certainly reduce consumption and raise personal income, thereby raising the savings ratio. Adjustments of this type, which do no more than reflect accounting identities, cannot be made unless the adjustments are allocated across the income/expenditure components.

The reliability with which changes in data are measured is of course closely related to that with which levels are measured. The link between the two depends on the degree of serial correlation in the measurement error. If there are large measurement errors in the output index, but they are strongly serially correlated, then it makes sense to put more weight on the output index in the short run than in the long run. But to give no weight to it in the long run, as is current practice, implies that it conveys no information at all about long-run movements in output; this seems inherently unlikely.

The idea of adjusting data with reference to reliability has to be given a precise meaning. The most usual way of doing this is to minimise the sums of the squared changes of the data, weighted by the variances of the measurement errors of the data: this gives balanced data which are a linear combination of the initial data. Thus, if the variance of expenditure were half that of income, two-thirds of the adjustment would be borne by income and one-third by expenditure. As is implicit in this example, it is the relative and not the absolute variances of the data which determine the allocation of the adjustments. This makes considerable sense since the residuals to be allocated are determined by the measurement process. The least-squares estimates which emerge from this process have the attraction that, if the original measurement errors are normally distributed, they are also maximum-likelihood estimators: this is a general property of least-squares estimators.

It is a property of estimators calculated by minimising the sum of squares that they are the linear estimators with the lowest variance (or best linear unbiased estimators). Not only can one calculate the variances of the reconciled data, but one can also be assured that they are lower than those which would be delivered by some other linear combination, such as the arithmetic average.

Obviously, this conclusion is dependent on starting with the correct data variances. If incorrect variances are supplied it is no longer possible to be sure that the reliability of the data is enhanced. Where nothing is known about data reliability, it might be sensible to start from an assumption of equal reliability, but it is still necessary to decide whether this means equal proportionate reliability or equal absolute reliability.

1.4.2 Covariances

The discussion so far has assumed that the measurement errors in one variable are independent of those in other variables. In such a case the variance matrix of the data is diagonal. However, it is not universally true that measurement errors are independent of each other. For example, some of the data used to measure government consumption (on the expenditure side) may also be used to measure output of public administration and defence on the income/output side. In such a case adjustments to government consumption also imply adjustments to public administration and defence output, and will be of relatively little use in reconciling discrepancies between income and expenditure. The covariances between different entries in the data are represented by diagonal elements in the variance matrix of the dataset.

At the level at which we have worked, there is not much evidence of problems of this type, although we have made allowances for the effect described above. However, there are two very important ways in which off-diagonal elements arise. The first is from the presence of constant and current price data. If the constant price data are calculated by deflating current price data, then there will be a covariance between the two equal to the variance of the current price data. The second arises from serial correlation. This means that the error in a variable in one period is correlated with that error in the same variable in a different period, so that an adjustment to a variable in one year would also imply an adjustment to the same variable in neighbouring years. Serial correlation can be taken account of by considering a long time-series of observations of the components of the national accounts simultaneously. The data can be stacked as a single data vector which has to satisfy accounting constraints for each year of the component data. The effect of serial correlation shows up in off-diagonal elements of the variance matrix of this multi-period dataset. The technique is then just the same as when data for a single observation are considered. It is only the size of the problem which is larger. This approach is used in part three of this book.

1.4.3 Validation of the variance matrix

The approach assumes that a data variance matrix can be inferred from available sources. The researcher is, nevertheless, not left completely at the mercy of those providing the data variance matrix. It is possible to carry out two tests to see whether the variance matrix is consistent with the data. From the variance matrix it is possible to calculate the variance matrix of the accounting residuals. This is simply a linear transformation of the data variance matrix. In the simplest case, in which there is only one accounting residual, the variance matrix is a scalar equal to the sum of the income and expenditure variances. One can then test the actual variance of the residual against the hypothesised variance calculated from the statisticians' estimates of data reliability, and adjust relevant parameters in order to ensure that well-defined criteria are satisfied.

1.4.4 Hypothesis testing

The structure offered by least-squares/maximum-likelihood adjustment provides opportunities for testing behavioural hypotheses. There is no doubt about the validity of the accounting restrictions; they must be true. But one may wish to test behavioural restrictions to see if they are true. The algebra of behavioural restrictions is just like that of accounting restrictions, but since the introduction of restrictions is achieved by maximum-likelihood estimation, successive restrictions can be tested by means of likelihood-ratio tests. One may, for example, wish to test whether the measurement error in the data is such that they would accept the hypothesis that imports equals exports. It is not enough simply to look at the measurement errors in imports and exports in order to assess this. A reduction in imports would require an increase in factor incomes or a reduction in domestic demand if the GDP identity is to be maintained. The test must therefore be carried out with reference to all the data in the accounts concerned, and not just with reference to external flows.

1.4.5 Unobserved components

It might be thought that, in the national accounts, there would be accounting constraints applying to the sectoral data as well as to the components of GDP. For example, the current account of the balance of payments, identified as the gap between exports and imports of goods and services (including property income), should be offset exactly by the capital flows identified in the financial accounts. This would suggest that the observed gap between the two is a residual like the residual in the GDP account.

Unfortunately this is not correct. Observed financial transactions do not cover all financial transactions. Large items such as flows of trade credit are not recorded and it is unlikely that it will ever be possible to record them in full sectoral detail (Begg and Weale, 1992). This means that the sectoral financial balances do not offer extra usable constraints. They must be calculated as residual items; the fact that balancing ensures that the GDP identity is satisfied means, nevertheless, that these residual balances will satisfy the constraint of adding to zero across all the sectors. More generally, one must be careful to distinguish genuine accounting residuals from figures which combine them with unobserved data.

1.4.6 Unknown variances

The discussion so far has focused on the need for some estimates of data reliability in order to make the least-squares adjustment possible. Is it possible to make the adjustments without knowledge of data reliability? The answer turns out to be yes. If the residuals are regressed on the data, then the regression coefficients converge asymptotically to the weights which would be used to allocate the residuals if the true error variance matrix were known. Furthermore, it is possible to carry out statistical tests on the weights to test, for example, the hypothesis that two different observations of the same variable should be given equal weight.

The analysis can be extended to the case where the measurement errors are proportional to some linear combination of the data. One would not expect the absolute standard deviation of money consumption to be the same now as it was in 1948, but it may well be a constant proportion of nominal GDP. Thus problems of heteroscedasticity can be avoided.

Nevertheless, this result is not as helpful as might be thought, at least with the sort of dataset we consider in part two. It turns out that, with the number of observations available, no degree of precision can be given to the estimated weights. This problem could have been resolved by making more assumptions about the underlying structure of the data. The methods suggested by Harvey (1989) would be suitable means of doing this. Instead, however, we have used assumptions about the data variance. These are derived from published information on data reliability and are then assessed using the tests of the data variances summarised above.