

Compilation and Data-Processing Issues

5.1. This chapter discusses the limitations of reliance on a single data source, the combination of data from multiple sources in compiling all relevant remittance aggregates, and the use of secondary data for estimating missing components such as specific informal remittance flows or adjustments. It outlines a comprehensive approach to improving remittances data and also discusses the problems arising from a misalignment of definitions with data classifications. In most aspects, compiling remittances data is not substantively different from compiling data on other balance of payments items (the more specific problems lie in obtaining adequate data). Therefore, this chapter briefly discusses compilation issues that apply more generally. However, this chapter also draws attention to the database and computing requirements that some approaches may bring.

A. General Compilation Issues

5.2. The tasks of data compilation include the extraction of data from various collections (such as migration data, administrative data, surveys, and reporting systems), processing and compilation of balance of payments accounts and supplementary series, and data management as well as the critical evaluation of data sources and the assessment of data quality. Extracting data requires judgment; compilers have control over some data sources (e.g., an MTO direct reporting system under their control) but also need to evaluate the usefulness of data sources that are not under their control (e.g., administrative data available from government departments). Both the technical tasks of storage and manipulation of data and the critical assessment of data quality are important. Compilation requires sound judgment in the evaluation of data sources and estimation methods. Judgment is particularly important for compiling remittances data because there is no single reliable data source.

5.3. The DQAF sets out good compilation practices applicable to all items, including those related to remittances.⁵³ The selection of adequate data sources and assessment of source data are important aspects of ensuring the accuracy and reliability of balance of payments data. Other important factors are the selection of appropriate statistical techniques, the validation of intermediate results and outputs, and revision studies. Further, the serviceability of statistics depends on factors that include the periodicity and timeliness of data, its consistency, and revision policy and practice.

5.4. Sound source data are measured by different factors, such as the following: data are obtained from a comprehensive data collection program that takes into account country-specific conditions; they reasonably approximate the definitions, scope, classifications, valuation, and time of recording required; and they are timely. Further, the DQAF requires that source data are routinely assessed so that errors, omissions, and other problems can be identified and addressed.

5.5. Statistical techniques cover both compilation procedures and adjustments made to raw data. Compilation procedures should ensure that errors are minimized and that ancillary and benchmark information is used where possible. Sound adjustments must also be employed to make source data consistent with balance of payments requirements. This may include adjustments for definitional differences (e.g., ITRS data capturing net settlement data instead of gross transactions) or estimating missing observations (e.g., under-reporting in a household survey). Also included are procedures used to incorporate unrecorded activities, including remittances through informal channels.

⁵³Readers are advised to consult the DQAF for further details. The RCG will summarize only the most important points. See Appendix 2 for the generic DQAF.

5.6. The validation of intermediate results and final outputs ensures that data are plausible and consistent with other data and relevant information. Compilers who do not use demographic data as the main input are therefore advised to use any available data on migration to verify the plausibility of remittance estimates. Other data sets, including monetary data, can also be used to cross-check remittance estimates derived from other data. Discrepancies in data or counterintuitive results should be investigated.

5.7. From time to time, compilers may need to revise previously released data (including preliminary data). Reviews of revisions should be carried out to ensure that they are not showing a trend. If a trend is discernable (e.g., if preliminary data tend to underreport remittances), estimation techniques should be adjusted.

5.8. The periodicity and timeliness of remittances data should follow that of the overall balance of payments compilation in each country. This implies that the requirements of the applicable dissemination standard are being met.⁵⁴ Compilation systems that rely on household survey data for their remittances data are likely to face problems in obtaining sufficiently timely and frequent data. Estimations may then be used for compiling preliminary data that will be adjusted once household survey data have been processed.

5.9. Remittances data should be consistent with the standard presentation of the balance of payments. This applies to standard items as well as supplementary components, which are “constructed” from elements of standard components. Remittances data should also be reconcilable with national accounts data and consistent with relevant demographic data. Inconsistencies should be investigated.

B. Using Data from Diverse Sources

5.10. Remittance transactions are heterogeneous and no single data source can ensure that all transactions are reported adequately. Data obtained from the ITRS and direct reporting from MTOs capture only transactions routed through the respective formal sector institutions. Surveys are subject to underreporting and misclassifications. Also, survey data may be less timely

⁵⁴However, it is feasible to compile the supplementary items (especially “total remittances” and “total remittances plus transfers to NPISHs”) less frequently and timely because these data are less important for macroeconomic management and surveillance.

and frequent than data obtained from other sources. Administrative data are collected for purposes other than compiling remittances data, so neither its coverage nor definitions may be well aligned with balance of payments data needs.

5.11. Given the shortcomings of each individual data source, data from different sources may be combined. Additional data or estimates may be required to account for transactions passing through channels not captured by the main data source (an example, based on combining direct reporting with an ITRS, is presented in Box 5.1). Also, a frequent and timely data source such as an ITRS may be used to compile regular preliminary data, while a less timely but perhaps more accurate source, such as an annual survey, would be useful in revising these data. Compilers therefore have a need for data from numerous sources to ensure sufficient accuracy, coverage, timeliness, and frequency of output data. Compiling data on remittances may be more demanding in this regard than many other balance of payments items would be.

5.12. Secondary data may be used to estimate data items for which no data are readily available (see Chapter 4). These estimates may cover the aggregate of a line item (e.g., personal transfers) or just a subset of transactions not covered by another data source. For example, one country may choose to base its estimates of personal transfers entirely on secondary data because no direct measurement approach is feasible. Another country obtains data on personal transfers from its ITRS, but complements these data by estimates of personal transfers routed through informal channels not covered by the ITRS. A third country uses household surveys as a primary data source but uses estimates to adjust the results for underreporting of remittance receipts by households. An example of using partial data obtained from direct reporting in a data model is outlined in Box 5.2.

5.13. Estimates can therefore either replace or complement direct measurement. In either case, it is important to ensure that using data from two sources for the same item will not lead to double counting. For example, a country could use data from its ITRS as a main source for all balance of payments items but complement these data by direct reporting from MTOs for personal transfers. Although the MTOs will then report all their relevant transactions, the ITRS would still capture net settlement data related to MTOs’ activities. Appropriate adjustments to the

Box 5.1. Combining Direct Reporting with an ITRS

Many disadvantages of the use of an ITRS for compiling remittances data arise from its limited coverage regarding transactions that are settled only on a net basis or that bypass formal settlement structures entirely. Compilers who plan to rely on an ITRS as their main data source should contemplate complementary data sources to compensate for the limitations of the ITRS. Direct reporting by MTOs can provide data on gross remittance transactions, with a reliable geographical breakdown, as opposed to net settlement data. Household surveys can provide data on transactions that an ITRS cannot measure, such as compensation of employees earned abroad or transactions through informal channels. Administrative and other secondary data can be used to estimate any missing components. Models can also be employed as an auxiliary means of controlling the accuracy of data obtained from an ITRS. Econometric forecasts of remittances may be used to check the plausibility of the data obtained from an ITRS. Because ITRS data are very timely, other actual sources are often not available for cross-checking.¹

In combining data obtained from MTOs and data from the ITRS, double counting will occur if MTO settlement

transactions are not omitted from bank transactions. Data reported by MTOs should replace (and not be added to) the transactions initiated by MTOs and reported by banks in the ITRS system. In order to facilitate cross-checks with bank reports, MTOs should separately identify their settlement transactions through resident banks, with sufficient data to allow a reconciliation between the two data sets.

If compilers wish to identify the partner countries of remittance transactions for the compilation of bilateral data, direct reporting can be a useful complement to data obtained from an ITRS or other sources. It has been noted that partner country data on remittances obtained from the ITRS often reflect the settlement country, not the true partner country. From direct reporting, compilers can obtain benchmark data for partner country attribution, which can be applied to data from other sources. For instance, information obtained from MTOs can be used to adjust the country distribution recorded by the ITRS.

¹An example of econometric forecasting is presented in Bank of Spain (2006).

ITRS data must be made to exclude MTOs' settlement transactions (see Box 5.1).

5.14. Depending on the nature and availability of secondary data, estimation approaches can be tailored to the needs and possibilities of each compiling economy. Estimates of relevant remittance-related transactions can be based on data such as migration records, other demographic variables, labor statistics, social security records, travel or transportation data, or any other data set that is readily available and reliable. Methodologies involving estimation can be used to estimate all relevant remittance items or to fill gaps when data obtained from direct observation are known to be incomplete (e.g., a compiler may decide to use an estimation methodology for all transactions recorded as "personal transfers" while the compiler's colleagues in another country obtain data from MTOs but estimate data on "personal transfers" passing through channels other than MTOs). Both approaches may provide good quality estimates, but compilers should endeavor to use data obtained from direct measurement wherever feasible, at least as a plausibility check of model-based estimates, and to adjust and supplement these data as necessary.

5.15. Work plans for improving remittances data often focus on measuring or estimating compensation of

employees and personal transfers. However, the compilation of "personal remittances" also requires the estimation of travel, transportation, social contributions, and taxes related to short-term employment abroad. In a few cases, direct measurement for these items may be possible. For example, detailed administrative data related to short-term workers in some countries may document the social contributions made by nonresident workers. However, in most cases, direct measurement is not feasible.

5.16. Of course, these data are all subsets of larger items. Travel and transportation related to short-term employment are part of overall travel and transportation services. Social contributions and taxes are part of larger totals. It is difficult to estimate the share of travel, transportation, and transfers attributable to short-term workers. In most cases, a more robust estimate can be derived from estimates of the number of nonresident workers leaving or entering a country, their length of stay, earnings, and subsistence costs. These data are themselves often estimates, but numerous data sources including partner country data are frequently available. Compilers may wish to consult unusual data sources.⁵⁵

⁵⁵The case study from Bulgaria provides some examples. For example, the Bulgarian compilers approximated the living costs of Bulgarian seasonal workers in European Union countries by refer-

Box 5.2. Direct Reporting as Input to a Data Model

Data obtained from MTOs may be useful to estimate a model for remittances based on data from surveys, censuses, and population registers. For example, a simple model for the estimation of remittance outflows is the following:

$$R = \sum_{i=1}^J r_i a_i n_i N_i,$$

where

- R is the estimate of the total remittance outflows;
- r_i is the average amount per remittance sent by the households of a given nationality i ($i = 1, \dots, J$), during a given reference period; this information can be obtained from MTO direct reports (assuming that the average remittance amount is similar regardless of transaction channel);
- a_i is the proportion of households of the given nationality that send remittances; this information can be obtained from a household survey of remittance senders;
- n_i is the average number of remittances sent by households of the given nationality (excluding households that do not send remittances at all); this information can also be obtained through a household survey;

N_i is the size of the population of households of the given nationality living in a country; this information can be obtained from censuses or population registers.

Similarly, the information on the amount transferred from a given country using MTOs may be combined with the following additional information obtainable from a household survey (which together define the “propensity to use the MTO channel”):

- The proportion of persons who send remittances and, within the latter group,
- The proportion of persons using the MTO channel.

This would allow estimating the aggregate amount of remittances sent from a given country through any channel, formal or informal. The “propensity to use the MTO channel” could be used as a benchmark (and would be updated only infrequently) while regular data from MTOs could be grossed up for timely and frequent estimates of total remittances. With appropriate benchmark data from the household survey, the model can also be used to estimate aggregate remittance inflows.

Any estimates should then be compared with the larger item of which they are a part. Other plausibility checks, as well as comparisons with important partner countries, are also advisable.

C. A Comprehensive Approach to Improving Remittances Data

5.17. The large number of remittance transactions and the multitude of channels pose challenges to the compilation of comprehensive statistics. This section deliberates on a practical approach toward building a work program for improving remittances data, with due consideration to reporting burden and compilation cost.

Step I. Understand the current situation

5.18. Compilers wishing to improve remittances data first need to understand the current situation regarding personal remittance transactions, focusing on two factors: identifying possible remitting channels

and recognizing the predominant remitting channels used in the country among those possible channels. If the compiler is not acquainted with the situation, it is necessary to gain an understanding by means of conducting sample surveys; interviewing banks, MTOs, credit card companies engaged in money transfer, and the regulatory authorities; or making estimates based on relevant statistics. Household surveys are often useful for identifying the relative importance of remitting channels. Sometimes, valuable information can be obtained from surveys and investigations carried out by others, such as international organizations or academic researchers.

5.19. Interviews with banks and money transfer organizations, or other agencies delivering remittances, are useful to identify their role in the intermediation of remittances in the country. Interviews with supervisory authorities provide ways to understand the importance of banks and money transfer organizations in the relevant country from the institutional point of view. Existing statistics that may provide a clue to the scale of the remittance channels should also be used to get a broad understanding.

encing the estimated living costs of students, which are published by universities.

Step 2. Evaluate current data collection system

5.20. Compilers should then evaluate the current data collection system on remittances in the balance of payment statistics. By comparing the major remitting channels and the data collection system identified, compilers can check the match between important channels and data sources covering them, identifying the most important omissions. For the channels that are insufficiently covered in the current data collection system, compilers can assess their relative importance.

Step 3. Prioritize

5.21. A sound compilation strategy needs to prioritize tasks by taking into account data needs, including the coverage, accuracy, timeliness, and periodicity of data, as well as the costs of improving data (this includes an assessment of the need for improved remittances data versus other data needs and the capacity of the compiling agency to implement improvements). The channels identified as being insufficiently covered in step 2 are the targets for the expansion of the coverage.

5.22. Remittances are made through various channels. Improving coverage carries costs that vary depending on the channel and can be large in terms of reporting burden, human resources, and finances. Although full coverage of all remittance channels would be ideal, it is probably unrealistic. Therefore, compilers should identify the important channels and expand data sources strategically. When data cannot be obtained for important channels, estimates should be developed based on the best available benchmarks and indicators.

5.23. Other aspects of the quality of statistics, such as accuracy, timeliness, and periodicity of source data, should also be examined along with the reporting cost. Prior consultation must be held to decide the minimum requirements of data quality. If most of the remitters in the country tend to use a particular channel, the compiler should try to improve the data quality of that channel, whereas improvement of other minor channels may be kept to a modest degree.

Step 4. Improve and expand data by channel

5.24. Different data sources and approaches cover different channels and transactions. Compilers should therefore improve and expand their data source program with specific reference to data gaps and omissions in existing source and priorities identified in step 3.

Different data sources are suited to different channels (as discussed in Chapter 4):

- ITRS and direct reporting from banks may be an effective data source for remittance transactions through banks.
- Direct reporting is a good approach to obtaining data from MTOs provided that a supportive legal environment exists, including a registration and reporting requirement for MTOs.
- Household surveys can be effective in capturing data on earnings abroad of short-term workers and on other remittances through informal channels.
- Other data sources, including administrative and demographic data, are a good basis for estimating transactions that cannot be directly measured.

5.25. The effective ways to capture data depend on the major remitting channels as well as whether the country is the remitting country or the recipient country. For example, it is generally easier to identify remittances in ITRS data in the remitting than in the receiving country, because banks cannot easily determine the transaction purpose when crediting accounts. On the other hand, it may be easier for the recipient country than for the remitting country to conduct household surveys on remittances because of sampling problems related to remittance-sending households. Therefore, compilers should understand the situation of the country, identify their specific data needs, and choose effective ways for improving and expanding their source data.

D. Database Management and Computing Requirements

5.26. Timely, efficient, and reliable data compilation relies on adequate infrastructure for entering, processing, and checking data. Data on remittances may not pose unusual challenges in these areas, but some compilation approaches make higher demands than others.

5.27. If data on remittances are obtained from an ITRS and details on every transaction are individually reported regardless of amount (i.e., if there is no threshold), the amount of data to be entered and processed can be very large. In some countries, the high reporting and processing burden was addressed by applying reporting thresholds. While reducing the data volume, such thresholds also cut out much relevant data

on remittances. ITRS systems that have no thresholds but do have sufficient reporting detail and fast, reliable electronic data transmission, are well suited for compiling remittances data; those without these attributes are less adequate. Detailed direct reporting by MTOs shares many features with a full ITRS. Large amounts of data may be reported, but automation can ensure fast and reliable processing.

5.28. Household surveys can also collect large amounts of data. Coding and entering data are both resource intensive and prone to error. Some compilers

outsource some functions related to enumeration, coding, and data entry. Surveys are, in most cases, demanding in terms of staff, logistics, and other real resources.

5.29. Data models for estimating remittances from secondary data can range from simple to highly complicated and data intensive. Higher complexity does not necessarily result in greater accuracy. Adequate models can be built using simple desktop applications. The reliability of data models is largely determined by the quality of underlying data and their assumptions.