# 8 Classifications for economic statistics (Module 8)

## Contents:

### Objectives

This chapter aims to provide an understanding of

* The need for standard classifications in economic statistics
* The international economic classifications
* Key classifications in more detail
* Some practical issues in data collection and analysis

## 8.1 The need for standard classifications

In the context of statistics about economies (and other topics), the various elements need to be grouped according to particular characteristics so that they can be distinguished in any analysis. These elements include kinds of activity, types of products (goods and services), the purpose of expenditures, types of transactions, institutional sectors, employment status and occupation, etc. Standard, systematic classifications are necessary mainly so that statistics from different sources of data can be brought together, compared and analysed. Standard classifications are particularly important for national accounts, price indices and labour market statistics.

The structure of a classification, grouping more detailed items into broader groups, is useful for many purposes, including coding, aggregation and publication. The structure is usually reflected in a system of codes, in which each successive digit (or digits) represents an increasing level of detail. The codes also define the order in which the items are listed.

International agencies encourage countries to use the classifications that have been developed at international level, mainly to enable international comparisons and the compilation of world level statistics. The structure and detail of international classifications tend to reflect the level of development in the largest and most advanced economies (mostly without tropical climates). These may not always suitable for small, less developed ones.

For example, the ***International Standard Industrial Classification of all economic activities*** (ISIC) is the most important and widely used classification in economic statistics. But in a national context, ISIC is sometimes too detailed, and sometimes not detailed enough. The ISIC publication by the United Nations reads (Revision 4, paragraph 159):

*The need for international comparability* ***does not****, however, imply that countries must adopt ISIC as a whole, without modification…*

But of course, a national classification should be compatible with the international standard at some level, for international reporting purposes.

##  8.2 The international economic classifications

#### Classifications maintained by UNSD Classifications Unit

The main international economic statistical classifications maintained by the United Nations Statistics Division are classifications of *economic activities,* of *products* (= goods and services) and of *expenditures* according to purpose. In effect, however, they are all defined at a basic level in terms of the goods or services that are produced, traded or consumed. It is the way in which the products are grouped that differs.

These classifications can be found at and downloaded from:

<https://unstats.un.org/unsd/classifications/unsdclassifications>

They are listed below:

International Standard Industrial Classification of All Economic Activities (ISIC)

Latest version: Revision 4 (2008) – discussed in more detail in Section 2.4 below

Central Product Classification (CPC)

Latest version: 2.1 (2015) – discussed in Section 2.5

Standard International Trade Classification (SITC)

Latest version: Revision 4 (2007) – rarely used now, given the HS (see below)

Classification by Broad Economic Categories (BEC)

Latest version: Revision 4 (2003) – a classification of products by end-use
Revision 5 exists but not published at the time of writing

Classification of Individual Consumption according to Purpose 2018 (COICOP)

Latest version: 2018 – discussed in more detail in Section 2.3 below

Classifications of Expenditure according to Purpose

Published in 2003, this covers four separate classifications:

Classification of the Functions of Government (COFOG)
Classification of Individual Consumption According to Purpose (COICOP 2003, replaced)
Classification of the Purposes of Non-Profit Institutions Serving Households (COPNI)
Classification of the Outlays of Producers According to Purpose (COPP)

COFOG is important in the context of Government Financial Statistics (GFS). The last two are rarely used in practice.

#### Other economic classifications

A full list of international classifications can be found at

<https://unstats.un.org/unsd/classifications/Family/ListByDomain>

Of the economic classifications not listed above, those of most interest for economic accounts are included in the *System of National Accounts 2008*. These include

SNA Classification of Institutional Sectors

SNA Classification of Transactions and other flows

SNA Classification of Assets and Liabilities

In the context of the Balance of Payments and Trade in Services, the following classifications are relevant:

Coding System for Balance of Payments and International Investment Position (CSBPIIP)

Extended Balance of Payments Services Classification 2010 (EBOPS 2010)

In the context of Labour Statistics, the following classifications are used (in addition to ISIC), mainly in relation to individuals

International Classification of Status in Employment 93 (ICSE-93)

International Standard Classification of Occupations 08 (ISCO-08)

In the context of International Trade, the following classification is used by Customs organisations worldwide, especially at a detailed level:

Harmonized Commodity Description and Coding System 2017 (HS 2017)

The HS is updated every 5 years by the Customs Cooperation Council

Finally, the European Union uses the following instead of the CPC

Classification of Products by Activity (CPA)

## 8.3 Key classifications in more detail

### Classification of Individual Consumption by Purpose (COICOP)

COICOP is the international classification recommended by the *System of National Accounts* (SNA) for classifying Household Consumption. It is the basis of the classification of products in the International Comparison Programme (ICP), in which prices are compared worldwide and “purchasing power parities” (PPPs) are calculated. More importantly, at national level, it is often used as a framework for compiling and publishing the Consumer Price Index (CPI).

The most recent version of COICOP was published at in December 2018. It replaces the previous version finalised in 1999. It is now more detailed, having 4 levels: Divisions with 2 digits; then Groups, Classes and Sub-classes with 1 extra digit. There is an extra level of detail for Division 01 Food and non-alcoholic beverages. Full details can be found here:

The top level of the classification is as follows:

|  |  |
| --- | --- |
| **Division** | **Description**  |
| **01** | Food and non-alcoholic beverages  |
| **02** | Alcoholic beverages, tobacco and narcotics  |
| **03** | Clothing and footwear  |
| **04** | Housing, water, electricity, gas and other fuels  |
| **05** | Furnishings, household equipment and routine household maintenance  |
| **06** | Health  |
| **07** | Transport  |
| **08** | Information and communication  |
| **09** | Recreation, sport and culture  |
| **10** | Education services  |
| **11** | Restaurants and accommodation services  |
| **12** | Insurance and financial services  |
| **13** | Personal care, social protection and miscellaneous goods and services  |
| **14** | Individual consumption expenditure of non-profit institutions serving households (NPISHS)  |
| **15** | Individual consumption expenditure of general government  |

Divisions 01 to 13 cover expenditures on goods and services by the Household Sector for final use (excluding anything purchased for business purposes). The last two categories concern the purpose of expenditures by the two Sectors mentioned (NPISH and government) on goods and services that are however consumed by individuals. The structure of the codes is shown in the following example:

|  |  |  |
| --- | --- | --- |
| **Code** | **Level** | **Description** |
| 01 | 1 | FOOD AND NON-ALCOHOLIC BEVERAGES |
| 01.1 | 2 | FOOD |
| 01.1.1 | 3 | Cereals and cereal products (ND) |
| 01.1.1.1 | 4 | Cereals (ND) |
| 01.1.1.1.2 | 5 | Rice |

Each Class and Sub-class is also allocated to one of four categories:

* ND Non-durable goods
* SD Semi durable goods
* D Durable goods
* S Services

The table below shows the number of categories in COICOP 1999 and COICOP 2018 by level



At the top level the only difference between the 2018 and 1999 versions is the splitting of former Division 12 into two: Divisions 12 and 13. (Former Divisions 13 and 14 have been renumbered 14 and 15.) Below this level, the structure has been modified slightly and some Groups and Classes have been renumbered. The codes look the same, but they may have a different meaning. For example, code 08.1 was *Postal services* in the 1999 version, but the same code is used for *Information and communication equipment* in the 2018 version. When using the codes, it is therefore crucial to know which version they belong to.

While providing a sound common basis, in practice the structure of the COICOP codes is not necessarily ideal either for data collection, or for data processing or for publication in a national context.

For example, the code for *rice* (and other food items) is 10 characters long (6 digits and 4 dots). Long codes like this can lead to coding and data capture errors. If, in a questionnaire, rows are used listing items of expenditure, it is simpler to use the row numbers (maximum 3 digits) as codes for the items, and to use a “look-up” or correspondence table to assign the corresponding COICOP code to each later, during the analysis.

Another potential problem arises when using COICOP codes in Excel. At Division and Group level the codes look like numbers. Excel may, in certain circumstances, treat these codes as numbers (they should be text), and drop any leading zeros. If this happens, things may go wrong. To be sure no such problem arises, it is recommended that all COICOP codes are given a non-numeric prefix. The dots may then be omitted.

COICOP Divisions (level 1) are often used for publishing the CPI. Although as countries develop the proportion of total expenditure that is spent on food declines, it still predominates. Yet out of 63 Groups (level 2) there is only one for food. This one group contains half of the lowest-level codes. So, to determine what food items may be contributing most to the overall rate of inflation, it is necessary to select the most important items from levels 3, 4 or 5 and ignore others.

Alternative structures, national classifications, are allowed. The COICOP publication says

*Countries may choose to either use COICOP directly for their national purposes or to develop their own national classification. Indeed, the need for international comparability* ***does not*** *imply that countries must adopt COICOP as it is, without modification. Countries can use COICOP as a guide in adapting their national classifications to the international standard.*

In the United States, for example, the Bureau of Labor Statistics uses the following main headings when publishing the CPI:

All items

 Food

 Food at home

 Food away from home

 Energy

 Energy commodities

 Gasoline (all types)

 Fuel oil

 Energy services

 Electricity

 Utility (piped) gas service

 All items less food and energy

 Commodities less food and energy commodities

 New vehicles

 Used cars and trucks

 Apparel

 Medical care commodities

 Services less energy services

 Shelter

 Transportation services

 Medical care services

These categories reflect policy concerns and the way of life in America. The underlying detail is structured in a similar way to COICOP and can be transformed into COICOP classes, but it only has eight top-level categories:

|  |  |
| --- | --- |
| **Category** | **Notes** |
| Food and beverages  | Includes alcohol and food and beverages away from home |
| Housing  | Includes hotel accommodation and household insurance |
| Apparel  | Excludes laundry and dry-cleaning services |
| Transportation  | Includes transport insurance |
| Medical care  | Includes medical insurance |
| Recreation  | Includes audio- visual equipment such as televisions |
| Education and communication  | Includes postage and delivery services |
| Other goods and services | Includes tobacco; laundry and dry-cleaning services |

This illustrates the way in which items may be grouped in alternative ways. COICOP has a separate Division for Hotels and restaurants, but in the United States such expenditure is included with Housing and Food and beverages respectively. To reflect developments, information technology might be better grouped with recreation and postage and delivery services with transport.

#### International Standard Industrial Classification (ISIC)

ISIC is used to classify the activity of enterprises, businesses, establishments and units of all kinds that are engaged in producing goods or services. (It is also used to classify people according to the industry they work in.) Activities are largely defined by the products they produce. It is therefore possible to classify most products in the same way.

As mentioned above, ISIC is the most important and widely used classification in economic statistics. Revenue Authorities and other agencies registering businesses may use it to describe their activity, although not necessarily the latest version.

Revision 4 was published in 2008. It comprises 21 sections, represented by the letters A to U. These are further divided into 88 divisions (with 2-digit codes), 238 groups (3 digits) and 419 classes (4 digits). The broad structure is shown below:



It is important to know which version is being used, as the Revision 3 codes look the same with a letter and 4 digits but they can have a different meaning. For example, in Rev.3 the code 1520 was Manufacture of dairy products, in Rev.4 1520 is Manufacture of footwear.

Some of the 4-digit codes have a leading zero, which Excel, in certain circumstances, may drop, thus creating a 3-digit code which will have a different meaning if it is not invalid. As with COICOP, using a non-numeric prefix, such as the Section letter is highly recommended.

As mentioned in the introduction, for analytical and publication purposes the ISIC structure is unlikely to be suitable without some modification. When publishing a summary of GDP by activity, some Sections may need to be combined, while Sections A and C may require splitting. However, the 24 2-digit Divisions in Section C (Manufacturing) of ISIC Rev.4 may be too much detail, so some countries use the following groups:

CA Food manufacturing

CB Alcohol and tobacco

CC Textiles, clothing, footwear and leather

CD Wood, paper and reproduction

CE Refineries, chemicals, rubber and plastic good

CF Non-metallic mineral products (glass, ceramics, cement etc.)

CG Metal products; machinery and equipment

CH Furniture and other manufacturing (including industrial repair and maintenance)

#### Classifications of international trade in goods

The most detailed classification of goods is the ***Harmonised System*** (HS), which is used by Customs authorities worldwide for classifying imports and exports of goods for duty and for statistical purposes. It has legally binding descriptions of each category and well over 5,000 possible codes at the 6-digit level. Additional digits are added by regional organisations (like the European Union) and at national level. The HS is revised every five years at world level (6 digits) but may be revised annually at the more detailed country level. The main structure of the HS at 2-digit level mainly reflects the materials from which the commodities are made. It is not so convenient for economic analysis.

Since 2010 every country reporting its trade flows to the UN COMTRADE database uses the HS codes to do so. Previously many countries used the ***Standard International Trade Classification*** (SITC) to classify international trade, at least for statistical purposes and for reporting trade flows. The SITC has a long history. Its main structure was designed more for economic analysis, grouping commodities mainly according to the extent of processing involved in their production. However, the HS is now accepted as the primary classification system. SITC categories can be derived from the HS, if required.

Another useful way of classifying goods (and potentially services) is according to the end-use of the product. There are three main end-uses: final consumption (by households), intermediate consumption (by producers) and investment goods (mainly machinery and equipment). Several products are of course used both by households and by producers, most notably energy products (mainly refined petrol), which are therefore shown separately. The classification ***Broad Economic Categories*** (BEC) was developed to give effect to this idea. The BEC categories can also be derived directly from the HS and this is done in the COMTRADE database.

Compiling statistics on international trade in services is not as straightforward as it is for goods. But there is an international classification for trade in services called the ***Extended Balance of Payments Services Classification 2010*** (EBOPS 2010). It is an important reference for statisticians when collecting and assembling data on the subject.

#### Central product classification (CPC)

Products (goods and services) are fundamental to production, consumption and trade. As we have seen, there are several ways in which to classify them, depending on the situation. In the context of production, they may be classified according to the activity producing them (ISIC). For household consumption and consumer prices, COICOP is another way. (COICOP covers a more limited range of goods and services, because there are many products that households do not consume.) For international trade in goods, the HS predominates.

However, the main underlying international statistical classification of both goods and services is the *Central Product Classification* (CPC). The structure of the classification is very similar to ISIC, but with some significant differences. The CPC has 5-digit codes: 10 sections, 71 divisions, 329 groups, 1,299 classes and 2,887 subclasses.

The ten Divisions of the CPC are as follows (note the leading zeros):

0 Agriculture, forestry and fishery products

1 Ores and minerals; electricity, gas and water

2 Food products, beverages and tobacco; textiles, apparel and leather products

3 Other transportable goods, except metal products, machinery and equipment

4 Metal products, machinery and equipment

5 Constructions and construction services

6 Distributive trade services; accommodation, food and beverage serving services; transport services; and electricity, gas and water distribution services

7 Financial and related services; real estate services; and rental and leasing services

8 Business and production services

9 Community, social and personal services

The main difference between the CPC and ISIC is that the CPC completely separates goods from services, whereas ISIC does so only partially. For example, in ISIC, agricultural services are included with Agriculture, but in the CPC they are included in Division 8, Business and production services. But every item in the CPC (with very few exceptions) can be allocated to a single ISIC activity. For this reason, the European Union and Malaysia (among others) have developed a *Classification of Products by Activity* (CPA). The first four digits of these CPAs are the same as ISIC and additional digits define the goods or services produced by that activity. This scheme avoids the need for a completely different set of codes for products. But every CPA code has an equivalent 5-digit CPC code. So, by means of a correspondence table, data classified by the CPA could, if necessary, be converted into CPC categories.

In the national accounts, products feature in Commodity Flow Accounts and Supply Use Tables. The SNA recommends the CPC for classifying products. In practice however, very few countries do so. In most published tables, the same classification scheme is used for both products and activities, based on ISIC or its national equivalent. However, in its 2017 publication *Compendium of Supply and Use Tables for Selected Economies in Asia and the Pacific,* the Asian Development Bank used activity groupings based on ISIC, but product groupings based on the CPC*.* The advantages of this different structure are not clear.

The CPC is most useful as a reference to which all the other product classifications can be linked through correspondence (look-up or bridge) tables. Th UNSD maintains a set of correspondence tables on its website.

## 9.4 Some practical issues

#### Data collection

As mentioned above, the structure of the 12-character COICOP 2018 codes is much too cumbersome to use in practice. The number of categories at the most detailed level (538) is also far too many. The level of detail needed for international reporting by the ICP at world level is 155. Although in some regions it may be more, but the full list would be impractical.

Before collecting data from households on their expenditure, it is essential to develop a list of products that is appropriate in the national context. Often in household surveys products are grouped according to the frequency with which they are purchased (or consumed). The COICOP categories would be a useful starting point, but many of then should be deleted or merged, while retaining the COICOP code as a reference for use in the analysis. For example, there is a COICOP code for injera, which is a type of bread consumed in Ethiopia, but unknown in most other countries. This category should be deleted from the list.

In surveys of enterprises, activity is a key classifying variable. Outside of agriculture, ISIC (as modified for local conditions) is the standard used by most countries. Most businesses focus on one particular activity, but larger enterprises may be engaged in more than one. According to international standards, the primary activity is defined as the activity that contributes most to the value added of the entity, as determined by the “top-down” method:

*Step 1*. Identify the ISIC **Section** that has the highest share of the value added.

*Step 2.* Within this section, identify the **Division** that has the highest share of the value added.

*Step 3.* Within this division, identify the **Group** that has the highest share of the valued added

*Step 4.* Within this group, identify the **Class** that has the highest share of value added.

The relative value added of each activity is the ideal. Alternatively, the value of sales (less purchases in the case of traded goods) or the numbers of employees may be used. Best practice would be to ask for the details of sales by product (less the cost of the goods sold in the case of wholesale and retail trade) to establish which is the most important. In some countries, this is done annually for the very largest enterprises, but only every five years for medium sized ones.

In practice, the business’s own description of its activity may be sufficient in the absence of more information. Those involved in coding should consider first which Section a respondent belongs to. This is the most important step. The treatment of multi-nationals requires special care. For example, at international level, petroleum companies may be engaged primarily in refining (manufacturing), but they must be classified according to their activity in the domestic economy.

Often it is not feasible to collect data in the full detail ideally required. For example, the following categories are recommended for collecting information on visitors’ expenditure in the domestic economy:

i. Package travel, package holidays and package tours

ii. Accommodation

iii. Food and drink

iv. Local transport

v. International transport

vi. Recreation, culture and sporting activities

vii. Shopping

viii. Others

Some assumptions will be needed to allocate the expenditures to more specific products used in compiling a Supply Use Table.

Similarly, it may be difficult for businesses to provide full details of their expenditures on goods and services. Better quality data is likely to be obtained using a few broad categories which they already use for accounting purposes.

## 8.5 Summary statistics

It is often the case that the source data is more detailed than is required for analytical or publication purposes. For example, Customs data is usually available according to a classification based on the Harmonised System (HS) which has well over 5,000 possible codes at the 6-digit level. Although the detail may be useful when you need information on a specific item, for most statistical purposes they need to be aggregated into a much small number of broader categories. In the context of consumer prices, the BLS has established categories of interest and summarised the statistics accordingly.

Compiling a Supply Use Table (SUT) is the best way to establish the level of GDP. For this purpose, a common standard classification is needed for activities and products. As mentioned above, most countries use the same classification for both. Some countries have developed a detailed CPA which can be used for this purpose. Some developing countries have used a more specific national CPA with fewer categories, say 60-80 activities and 150-200 products. Having decided on the appropriate categories (some of which may usefully distinguish the end-use of certain products), correspondence tables need to be established between the codes used in the source data and the CPA used in the SUT. The UNSD correspondence tables can be used to help achieve good quality links. In some cases, however, a source category may have more than one corresponding destination category. When this occurs, the simplest solution is to determine which destination code is the most likely or most appropriate.

Once a correspondence table is established with a single destination category for each source category, the source data is easily converted into the new structure. In Excel, the VLOOKUP function can be used to assign a destination code to each record with a source code. Then, given a list of the destination categories and codes, the SUMIF function is often the best way to obtain the results according to the destination categories. Alternatively, a Pivot Table can be set up, although this method ignores destination codes that do not feature in the dataset. The SUMIF function is also useful when aggregating structured codes to higher levels. The LEFT function can be used to establish which detailed codes belong to each higher-level category.

## 8.6 Exercise 8.1

*Tick one box to indicate whether each statement is true or false*

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Statement | True | False |
| 1 | UN Statistics Division (UNSD) maintains all economic classifications |  |  |
| 2 | Information about most international classifications is available on the UNSD website |  |  |
| 3 | There are many ways to classify products |  |  |
| 4 | COICOP 2018 has 15 Divisions of which 13 are for household expenditure |  |  |
| 5 | COICOP 2018 codes are ideal for collecting data in the field |  |  |
| 6 | Out of 63 Groups in COICOP 2018, only one is for food |  |  |
| 7 | Half of the most detailed COICOP codes are in Group 01.1 |  |  |
| 8 | The USA uses COICOP to classify its CPI |  |  |
| 9 | The International Comparison Program (ICP) uses COICOP |  |  |
| 10 | ISIC is the most widely used classification of economic activities |  |  |
| 11 | In ISIC, some Divisions include the production of both goods and services |  |  |
| 12 | The same ISIC 4-digit code can have different meanings in Rev.3 and in Rev.4 |  |  |
| 13 | No ISIC 4-digit code has a leading zero |  |  |
| 14 | Countries are not allowed to modify ISIC for national purposes |  |  |
| 15 | Most countries use SITC to classify imported goods |  |  |
| 16 | The HS has more than 5,000 categories at 6-digit level |  |  |
| 17 | The HS covers both goods and services |  |  |
| 18 | The CPC and ISIC have the same top-level structure |  |  |
| 19 | Many countries use the same classification for both activities and products in their published national accounts |  |  |
| 20 | The European and Malaysian CPAs have the same structure as ISIC |  |  |