Chapter 2

Basic Concepts, Definitions and Classifications

How is an economic system defined? Who are the economic agents, and how is a household distinguished from an enterprise? How is the growth of an economy measured? How is a time series defined, and what is the difference between a trend, a cycle and a seasonal variation? This chapter provides an overview of the main concepts and methodological tools necessary to read and analyse the main economic statistics described in Chapter 4. After introducing the main definitions concerning the economic system, we shall analyse the characteristics of the principal economic agents (households, enterprises and public and private institutions) using examples drawn from OECD countries, and then we shall review the main economic aggregates as defined in national accounts, from gross domestic product to national income. The second part of the chapter will be devoted to a brief presentation of index numbers and time series analysis, which are indispensable tools for "reading" economic statistics.

2 BASIC CONCEPTS, DEFINITIONS AND CLASSIFICATIONS

2.1. The economic system

The term *economic process* refers to to those activities through which goods and services aimed at satisfying human needs are produced, distributed and used. For example: daily consumption of food and other goods and services by individuals; an enterprise's production of machinery used to produce other goods; or even the writing of a book on economic statistics intended to be bought by university students.

One of the fundamental characteristics of activities defined as economic processes is that they involve relations between various agents. The definition of economic agent is therefore absolutely fundamental in determining the nature of the economic processes: *economic agent* refers to a person or legal entity that plays an active role in an economic process. An economic agent can therefore be an individual consumer who purchases goods and services, an enterprise that organises factors of production to generate income, a worker who provides his or her labour in a production process, etc. These individual agents (to which economic theory habitually attributes preferences, objectives, behaviour, etc.) are then normally grouped into *institutional sectors* that represent groupings of institutional units (corporations, households, general government, etc.), each of which:

- is entitled to own goods or assets in its own right; it is therefore able to exchange the ownership of goods or assets in transactions with other institutional units;
- is able to take economic decisions and engage in economic activities for which it is held to be directly responsible and accountable by law;
- is able to incur liabilities on its own behalf, to take on other obligations or future commitments and to enter into contracts;
- has either a complete set of accounts (including a balance sheet of assets and liabilities) or it would be possible and meaningful, from both an economic and legal viewpoint, to compile a complete set of accounts for the unit, if required.

These institutional units are the categories of economic agents normally referred to in the System of National Accounts (SNA).

All the agents within a given territory (a region, country, etc.) and the ways they interact with each other and with other agents outside that territory are defined as an *economic system*. An economic system is not only characterised by the physical or technological factors that determine how its production is oriented (*i.e.* mainly towards agriculture, industry, etc.), but also by cultural and institutional factors that regulate how it functions (laws, regulations, etc.). Thus, there are systems in which the economic relations between individual agents are heavily regulated and systems that freely allow agents to undertake new activities or terminate existing activities. There are economic systems completely open to trade with other systems as well as highly regulated systems; systems in which a few large enterprises produce most of the goods and services and others with a vast number of small enterprises and only a few large ones.

UNDERSTANDING ECONOMIC STATISTICS – ISBN 978-92-64-03312-2 – © OECD 2008

The characteristics of an economic system are also important because they can influence the quality of the statistics describing how the system functions. For example, when economic systems are characterised by the presence of a few large enterprises, it is relatively simple to collect statistics to measure the functioning of the system, but when the economic system is composed of a myriad of small enterprises, it can become extremely difficult and/or time consuming to do so. Similarly, in a system that has a particularly large "underground" (or "non-observed") economy, *i.e.* the economy that is not visible to the tax and administrative authorities, production of accurate economic statistics can be a challenge.

An economic system is generally defined in terms of territorial boundaries. The *economic territory* is the area in which the units reside, operate and pursue their interests. Traditionally, the following types of areas are identified:

- *supranational* economic systems: systems composed of groups of sovereign States that have come together through international treaties that set common standards for the functioning of national economic systems (for example, the group of countries that belong to the European Union);
- *national* economic systems: systems having an economic territory that coincides with the administrative boundaries of a sovereign State (France, Canada, etc.);
- *regional* economic systems: systems defined using the administrative boundaries of sub-national areas (regions, provinces, etc.);
- *local* economic systems: systems not defined on the basis of administrative boundaries, but in terms of specific economic, social or environmental characteristics (for example, "local labour systems" or "industrial districts").

The concept of territory is extremely important in statistics because it allows us to distinguish between *agents residing* in a specific territory and *non-resident agents*, thereby making it possible to measure the contribution of each type of agent to economic variables, such as consumption, production, etc. If an agent has its "centre of economic interest" within the national territory, *i.e.* if that is where it conducts its most important economic transactions for a prolonged period (at least one year), it is defined as "resident". If it does not, it is defined as "non-resident".

In reality, with the development of international trade and the processes of "globalisation", it has become extremely difficult to know the relations existing between resident and non-resident agents and to assess the contribution each makes to the overall economic system. Given the expanding operations of multinationals and the introduction of electronic commerce, it is becoming increasingly difficult to describe in statistical terms the amounts and characteristics of the flow of goods, services and activities among agents residing in different countries and their interrelations.

2.2. Economic agents

As we have just seen, in order to be defined as such under the SNA, institutional units (*i.e.* the basic economic decision-making centres) must have autonomy of decision-making in carrying out their principle function and keep (at least potentially) complete accounting records. Institutional units are then aggregated according to the principle function they perform. In this regard, a distinction is made between three "institutional sectors":

- *households*, which perform three principle functions: 1) consume the goods and services produced by other institutional sectors; 2) produce goods and services that can be sold; and 3) acquire real and financial assets. The household sector also includes *non-profit institutions serving households*, which provide non-market (*i.e.* not intended for sale) services consumed by households;
- *enterprises*, which produce goods and services intended for sale to generate profits and which acquire real and financial assets. This sector also includes *non-profit institutions serving enterprises*, which produce services intended for sale to be consumed by companies;
- *general government*, which, in addition to fulfilling its political responsibilities and role of economic regulation, produces principally non-market services (possibly goods) for individual or collective consumption and redistributes income and wealth.

Transactions between resident and non-resident units are normally measured by aggregating all of the latter into a single institutional sector, defined as *rest of the world*. Let us now examine in detail the characteristics of the main economic agents.

2.2.1. Households

From a statistical standpoint, a *household* consists of a small group of people sharing the same living accommodation, who pool some, or all, their income and wealth and who consume certain types of goods and services collectively, mainly housing and food. This group of people can be bound by ties of marriage, family relationship, affinity, adoption, guardianship or ties of affection, and they habitually reside in the same municipality (even if they are not yet registered by the municipality as residents). A household may consist of a single person. Someone who is temporarily absent remains a member of the household even if he or she is living in a different accommodation, be it in the same municipality or in another municipality within the same country or abroad.

As can be seen, the statistical concept of household is not directly related to that used in legislation, and this is a necessary condition given the need to compare statistics on households over time and space. It should also be pointed out that, as we shall see later, many existing statistical surveys refer to individuals rather than to households. However, some surveys are designed also to provide information

by household, as do the so-called "multi-purpose" surveys or those on household consumption and wealth.

It is important to classify households on the basis of the number of families that constitute them. The *family* within the household is defined as those members of the household who are related, to a specific degree, through blood, adoption or marriage. Consequently, the concept of family is normally more restrictive than that of household; in fact, a household may include one or more families or even no families at all (as in the case of single-person households), or may include several families and isolated family members (for example, a child living at the parental home after being legally separated from his/her spouse) or isolated family members (for example, two siblings whose parents are dead).

For example, according to Italian census data (Table 2.1), in 2001, for nearly 57 million persons residing within the national territory, there were 21.8 million resident households, a sharp increase over 1991 (up by nearly 2 million). Of these, nearly 5.5 million consisted of a single person (mostly widows and widowers) and nearly 6 million of two persons (of which nearly 500,000 were unmarried couples). There were slightly fewer than 6 million households with four or more persons. Some 4.5 million households consisted of couples without children and there were some 9 million couples with children. Lastly, in roughly 2 million cases only one of the two parents was present in a household with children. The average number of members per household was 2.6 (2.8 in 1991), with values ranging from 2.4 in the North-West to 2.9 in the South. The regions with the lowest average were Liguria and Valle d'Aosta (2.2), while Campania had the highest (3.1).

In the case of households with more than one member, a problem of classification arises. Given the individual members' different demographic and social characteristics (employed, unemployed, students, young, elderly, etc.), how can the entity "household" be classified unambiguously on the basis of accurate demographic and socio-economic characteristics? The response normally adopted for household surveys is to refer to the characteristics of the *reference person*. The reference person is the "the person registered in the public records in relation to whom family relationships are defined". Thus, if there is a father in a household, he will usually be the reference person, while the wife will be indicated as the spouse, their offspring as children, etc. If there is no father, then the mother will usually act as the reference person and so on. This means that households can now be classified not only by the *number of members*, but by the *labour market status* of the reference person (economically active or inactive), his or her *employment status* (employed, unemployed, job-seeker, etc.), *age*, *income*, *housing tenure* (owner, renter, etc.) and so on.

Table 2.1 **– Resident households and breakdown by size of household and geographical distribution**
Italy, 2001

Geographical distribution	Households by number of members						Total		
	One person	Two persons	Three persons	Four persons	Five persons	Six or more persons	Households	Members	Average number of members per household
North-West	1 767 208	1 840 037	1 390 009	966 118	207 367	46 461	6 217 200	14 813 530	2.4
North-East	1 116 042	1 208 212	962 636	701 273	184 009	59 838	4 232 010	10 530 285	2.5
Centre	1 061 905	1 188 248	941 315	780 561	208 574	61 596	4 242 199	10 820 324	2.6
South	940 888	1 100 449	935 550	1 150 759	474 806	145 822	4 748 274	13 860 137	2.9
Islands	541 578	568 465	476 696	537 495	191 070	55 689	2 370 993	6 569 745	2.8
Italy	5 427 621	5 905 411	4 706 206	4 136 206	1 265 826	369 406	21 810 676	56 594 021	2.6

StatLink http://dx.doi.org/10.1787/336610548652

Source: Italian Statistical Office.

These classifications are used in many statistical surveys. For example, in the household consumption survey for France, consumer spending can be compared based on the number of household members, the occupational status of the reference person (Table 2.2) or the type of household.

Table 2.2 **– Average monthly household expenditure by occupational status of reference person**
France, 2001

Occupational status of reference person	Food		Non-food		Total
	Value (in euros)	Percentage	Value (in euros)	Percentage	Value (in euros)
Farmer	6 092	19.3	25 553	80.7	31 645
Self-employed	7 165	16.3	36 696	83.7	43 861
Managers and professional workers	8 327	15.5	45 508	84.5	53 835
Employees	4 789	18.6	20 926	81.4	25 715
Manual workers and similar	5 551	19.6	22 782	80.4	28 333
Retired	4 576	20.7	17 529	79.3	22 105
Others not employed	3 201	20.2	12 675	79.8	15 876
Total	5 513	18.4	24 450	81.6	29 963

StatLink http://dx.doi.org/10.1787/336618155337

a) Including expenses for restaurants, bars, cafés and canteens.

Source: INSEE, *Le Budget des Familles en 2001.*

To consume the goods and services produced by companies, general government and the rest of the world, households spend the income they receive from participating in productive activities. They may also engage in productive activities themselves, such as running family businesses (a shop, a restaurant, etc.) or renting accommodation (apartments or houses). In this case, they are defined as *unincorporated enterprises*. More precisely, an unincorporated enterprise is a producer unit not incorporated

as a legal entity separate from its owner (in this case the household, but it could also be the government or a foreign resident); the fixed and other assets used by unincorporated enterprises do not belong to the enterprises but to their owners; the enterprises as such cannot engage in transactions with other economic units nor can they enter into contractual relationships with other units nor incur liabilities on their own behalf; in addition, their owners are personally liable, without limit, for any debts or obligations incurred in the course of production.

Finally, NPISHs (Non-Profit Institutions Serving Households) provide goods and services (usually for free, being financed by regular membership subscriptions or dues) that benefit households, such as assistance to disabled persons, recreational activities for children, etc. NPISHs are created by associations of persons to provide goods or services primarily for the benefit of the members themselves (such as professional or learned societies, political parties, trade unions, consumers' associations, churches or religious societies, and social, cultural, recreational or sports clubs).

The distinction between consuming and producing households is typical of national accounts, while other statistical sources tend to view households above all as entities primarily aimed at consumption (when the survey is of an economic nature) or as the place where interpersonal ties are established and where key functions for society as a whole take place (when the survey is focused on studying demographic or social aspects). In these cases, households are classified by criteria that differ from those adopted in the national accounts, although it is still possible, at least at the conceptual level, to conduct cross-classifications, for example by distinguishing consuming households on the basis of social-demographic characteristics.

2.2.2. Enterprises

The enterprise sector includes various types of entities. The term *enterprise* in the strictest sense refers to the organisation of an economic activity on a professional basis for the purposes of producing goods or providing services intended for sale. An enterprise has a certain autonomy regarding its choices in the field of production, sales and distribution of profits. The entity responsible for the enterprise consists of one or more persons acting individually or in partnership or of one or more legal entities.

Enterprises can operate in all sectors of economic activity (agriculture, industry or services). In the agricultural sector, however, there is a specific type of economic entity, defined as an *agricultural holding*. For agricultural census purposes, an agricultural holding is a techno-economic unit of agricultural production comprising all livestock and all land used wholly, or partly, for agricultural purposes and managed by one person or more, without regard to title, legal form, size or location. In terms of national accounts, it is defined as an economic unit under a single management engaged in agricultural production. The unit may also be engaged in non-agricultural activities, so this concept should not be interpreted too strictly; the aim is rather to value the final production of all agricultural products. Also, establishments or specialised units that provide agricultural services on a fee or contract basis should, in general, be included. The fundamental difference between an enterprise and an agricultural holding is that

the basic requirement for the former is its business-oriented organisation (independent of the sector of activity), while for an agricultural holding, it is the specific activity of agriculture, forestry or animal production that is fundamental.

Enterprises may be classified on the basis of many characteristics. In the SNA, enterprises are aggregated in the institutional sector of "corporations", which comprises corporations and unincorporated enterprises. Moreover, quasi-corporations are unincorporated enterprises that function as if they were corporations, and which have complete sets of accounts, including balance sheets.

Within the institutional sector of corporations two typologies are distinguished:

- *non-financial corporations*: corporations whose principal activity is the production of market goods or non-financial services;
- *financial corporations*: institutional units principally engaged in financial intermediation or in auxiliary financial activities.

Among the various ways of classifying enterprises, classification based on *economic activity* is certainly of key importance. This type of classification is based on the characteristics of the the activity of production units, *i.e.* the characteristics of the goods and services produced, the uses for which these are intended, the factors of production and the production process and technology used. The importance given to these criteria varies from one category to the next. For example, for goods characterised by a particularly complex production process, the final use, the technology and the organisation of production are given priority over the type of goods produced.

The reference classification for economic activities is the *International Statistical Industrial Classification* (ISIC Rev. 3.1) published in 2002. At the European level, the corresponding classification is known as NACE Rev. 1.1, which is totally aligned with ISIC Rev. 3.1. NACE Rev. 1.1 has the following levels:

- Level 1: Categories (one-letter alpha code – A to Q);
- Level 2: Divisions (2-digit codes);
- Level 3: Groups (3-digit codes);
- Level 4: Classes (4-digit codes).

In several countries, the statistical office prepares a national version of the international classifications. These classifications, although roughly consistent with the international ones, take into account the specificity of national production structures and may contain an additional level of detail useful for identifying activities particularly important to that country.

UNDERSTANDING ECONOMIC STATISTICS – ISBN 978-92-64-03312-2 – © OECD 2008

Table 2.3 **– ISIC Rev. 3.1 industrial classification of economic activities – Categories and divisions**

A		Agriculture, hunting and forestry
B		Fishing
C		Mining and quarrying
D		Manufacturing
	15	Manufacture of food products and beverages
	16	Manufacture of tobacco products
	17	Manufacture of textiles
	18	Manufacture of wearing apparel; dressing and dyeing of fur
	19	Tanning and dressing of leather; manufacture of luggage, handbags, saddlery, harness and footwear
	20	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials
	21	Manufacture of paper and paper products
	22	Publishing, printing and reproduction of recorded media
	23	Manufacture of coke, refined petroleum products and nuclear fuel
	24	Manufacture of chemicals and chemical products
	25	Manufacture of rubber and plastics products
	26	Manufacture of other non-metallic mineral products
	27	Manufacture of basic metals
	28	Manufacture of fabricated metal products, except machinery and equipment
	29	Manufacture of machinery and equipment n.e.c.
	30	Manufacture of office, accounting and computing machinery
	31	Manufacture of electrical machinery and apparatus n.e.c.
	32	Manufacture of radio, television and communication equipment and apparatus
	33	Manufacture of medical, precision and optical instruments, watches and clocks
	34	Manufacture of motor vehicles, trailers and semi-trailers
	35	Manufacture of other transport equipment
	36	Manufacture of furniture; manufacturing n.e.c.
	37	Recycling
E		Electricity, gas and water supply
F		Construction
G		Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods
H		Hotels and restaurants
I		Transport, storage and communications
J		Financial intermediation
K		Real estate, renting and business activities
L		Public administration and defence; compulsory social security
M		Education
N		Health and social work
O		Other community, social and personal service activities
P		Activities of private households as employers and undifferentiated production activities of private households
Q		Extraterritorial organizations and bodies

Enterprises may be classified on the basis of other criteria, such as size, organisation, legal form, purpose of products, technological description, etc. The classification by *size of enterprise* is based on the number of employees, with the term *employee* referring to a person who enters an agreement (either formal or informal) to work for an enterprise in return for remuneration in cash or in kind. Employees can have different types of contracts: for example, stable employees are those who have had, and continue to have, an explicit (written or oral) or implicit contract of employment, or a succession of such contracts, with the same employer on a continuous basis. On a continuous basis implies a period of employment longer than a specified minimum determined according to national circumstances. Regular employees are those with stable contracts for whom the employing organisation is responsible for paying taxes and social security contributions and/or where the contractual relationship is subject to national labour legislation.

The size classes can be defined on the basis of various criteria and purposes. For example, the definition adopted at the European level classifies enterprises with fewer than 250 employees as "small and medium-sized enterprises"; this classification is used both for statistical purposes and to identify enterprises that may qualify for special financing granted by the EU. In countries such as Italy, where a myriad of small and very small enterprises exist, the European size classification does not appear to be very useful; therefore, other thresholds are used to analyse these economic systems. Over time, the size distribution of enterprises can change significantly, partly because of legal acts that favour certain groups of enterprises. For example, Table 2.4 shows the evolution of the industrial sector in the United States, based on enterprise size.

Generally, an enterprise produces a variety of products. Furthermore, other activities, known as secondary activities, are often carried out in addition to the principal activity. The term *principal activity* refers to activity whose value added exceeds that of any other activity carried out within the same unit (the output of the principal activity must consist of goods or services capable of being delivered to other units, even though they may be used for its own consumption or own capital formation). An *ancillary activity* is a supporting activity undertaken within an enterprise to create the conditions within which the principal or secondary activities can be carried out; ancillary activities generally produce services commonly found as inputs into almost any kind of productive activity, and the value of an individual ancillary activity's output is likely to be small compared with the other activities of the enterprise (*e.g.* cleaning and maintenance of buildings). Finally, a *unit of homogeneous production* is a producer unit in which only a single (non-ancillary) productive activity is carried out (this unit is not normally observable and is more an abstract or conceptual unit underlying the symmetric input-output tables).

Table 2.4 – **Enterprises and employees by size class in the total industry (excluding construction)**
USA, 1992 and 2002

Size classes (number of employees)	1992		2002		Growth rates (percentages)	
	Enterprises	Employees	Enterprises	Employees	Enterprises	Employees
1-9	214 626	707 793	210 322	693 444	-2.0	-2.0
10-19	55 629	758 189	52 299	711 304	-6.0	-6.2
20-99	68 632	2 809 125	64 467	2 588 453	-6.1	-7.9
100-499	17 203	3 070 757	17 732	2 703 843	3.1	-11.9
500 or more	5 431	12 420 510	10 635	9 600 428	95.8	-22.7
Total	361 521	19 766 374	355 455	16 297 472	-1.7	-17.5

StatLink http://dx.doi.org/10.1787/336664670860

Source: OECD 2006, *Structural and Demographic Business Statistics*, OECD publishing.

Another extremely important distinction is between enterprises situated *in a single location*, *i.e.* enterprises that carry out their activity in a single local unit (establishment), and enterprises with *a number of locations*. Naturally, many small enterprises have a single location, but this is also true of many medium-sized and large enterprises. In addition, enterprises with a number of locations often have local units scattered in many different regions and sometimes in other countries.

The trend towards the globalisation of production, the development of computer networks, the growing international mobility of capital, products and labour and the growing role of groups of enterprises have caused major changes in production units in recent years. Trends towards "relocation" of production have been particularly important, *i.e.* the transfer abroad of entire enterprises or of specific local units. These changes pose difficult problems of classification with regard to the location of enterprises, their ownership structure and the relations between local units operating in different countries. The distinction between individual enterprises and groups of enterprises is important in this regard.

According to the European Council Regulation EEC No. 696/93, an *enterprise group* is an "association of enterprises bound together by legal and/or financial links. A group of enterprises can have more than one decision-making centre (production, sales, etc.) and it may centralise centrain aspects of financial management and taxation". The concept of control is naturally crucial for the definition of a group: control of an enterprise implies the ability to appoint a majority of the board of directors, to run the enterprise, guide its activities and determine its strategy. This ability is exercised by a single direct investor or a group of associated shareholders acting in concert and controlling the majority (+50%) of ordinary shares or voting power. The control of an enterprise may be direct or indirect, immediate or ultimate. However, "effective control" may be exercised when the investor(s) holds a large block of voting stock, even when it is less than 50%, and the remaining shares are widely held by many smaller investors. Control of enterprises may also be exercised through interlocking directorates and inter-corporate ownership links between firms, as in the case of conglomerates.

The notion of *ownership* is different from that of control, since the former corresponds to financial ownership of an enterprise, be it majority or minority ownership. Majority direct ownership (+50%) of ordinary shares, or voting power, by a single investor could imply control of this enterprise by the investor, but minority ownership could also imply indirect control of this enterprise (through another enterprise). In other words, the notion of ownership is not sufficient to indicate if an enterprise is under influence or control.

Within groups, the following main types of enterprises can be identified:

- *holding corporation*: corporation that controls a group of subsidiary corporations and whose principal activity is owning and directing the group. When the holding corporation is resident in a different country from the enterprise controlled, the holding enterprise is considered to be "foreign";
- *the parent (controlled) enterprise*: an enterprise controlled by another institutional unit that controlled directly or indirectly more than 50% of the shares or voting rights of the first enterprise at 31 December of the reference year. When the controlled enterprise is resident within the national territory but is controlled by a non-resident institutional unit, the controlled enterprise is considered to be "foreign".

Using the statistical register ASIA (Table 2.5), the Italian Statistical Institute has estimated that in 2003 there were more than 59 000 groups, which comprised some 138 000 enterprises and employed more than 5.1 million people. Groups involved 3.2% of economically active enterprises and one-third of employed persons. However, this number rose to 20% (and 57% of employment) if it was calculated with reference to the number of corporations only. Some 63.4% of groups are concentrated in enterprises with 1-9 employees (accounting for 4.7% of employment), but those with 500 employees or more, even though they only represent 2% of the total, account for over 3 million employees. There are 649 groups under the direct or indirect control of general government, which account for 12% of the employment of enterprises belonging to groups, while 9.2% of groups are foreign controlled.

Table 2.5 (a) **– Groups, enterprises and employees belonging to groups, by number of employees**

Italy, 2003

Number of employees	Groups		Enterprises		Employees	
	Number	%	Number	%	Number	%
1-19	38 045	63.4	65 155	46.9	239 225	4.7
20-99	15 599	26.0	36 394	26.2	691 950	13.5
100-499	5 016	8.4	18 755	13.5	1 019 649	19.9
500-4999	1 216	2.0	12 513	9.0	1 513 669	29.5
5000 and more	87	0.1	6 115	4.4	1 658 631	32.4

StatLink http://dx.doi.org/10.1787/336677172303

Source: Italian Statistical Office.

UNDERSTANDING ECONOMIC STATISTICS – ISBN 978-92-64-03312-2 – © OECD 2008

Table 2.5 (b) – **Groups, enterprises and employees belonging to groups, by number of enterprises involved in the group**

Italy, 2003

Number of enterprises involved in the group	Groups		Enterprises		Employees	
	Number	%	Number	%	Number	%
1	22 799	38.0	22 799	16.4	631 833	12.3
2	24 116	40.2	48 232	34.7	854 685	16.9
3-4	9 852	16.4	31 942	23.0	829 308	16.2
5-9	2 472	4.1	14 968	10.8	839 705	16.4
10-49	652	1.1	11 025	7.9	1 133 322	22.1
50 or more	72	0.1	9 966	7.2	824 272	16.1
Total	56 963	100.0	138 932	100.0	5 123 125	100.0

StatLink http://dx.doi.org/10.1787/336677172303

Source: Italian Statistical Office.

Although the activity of multinationals grew dramatically over the last decade, much of international trade is still conducted by enterprises independent of one another and a very substantial number of enterprises produce goods and services for the domestic market alone. Consequently, another important distinction must be made between enterprises that have relations with the rest of the world (*exporting/importing enterprises*) and enterprises only oriented towards the domestic market. As can be seen in Table 2.6, in France there are approximately 19 000 exporting enterprises with more than 20 employees. Naturally, the percentage of exporting enterprises grows as the size of the enterprise increases, reaching a percentage of 40% for enterprises with 2 000 employees or more. It should also be pointed out that the difference between exporting and non-exporting enterprises is extremely interesting from an analytical standpoint, since the economic indicators (productivity, profitability, etc.) for exporters are normally better than for non-exporters, given the same size and activity sector.

Table 2.6 – **Exporting enterprises, employees and exports by number of employees**

France, 2005

Number of employees	Enterprises		Employees	Exports	
	Number	Exports rate (exports/turnover)		Value (million euros)	Share of total exports (percentage)
20-49	10 409	16.5	320 156	8 591.2	3.3
50-99	4 155	22.0	261 576	10 638.9	4.1
100-249	2 872	30.7	406 080	23 802.3	9.2
250-499	1 031	34.6	321 495	25 309.9	9.7
500-999	523	33.1	320 564	35 306.7	13.6
1 000-1999	213	39.1	271 194	33 016.1	12.7
2 000 and more	134	40.3	781 244	123 079.7	47.4
Total	19 337	34.7	2 682 309	259 744.8	100.0

StatLink http://dx.doi.org/10.1787/336765200553

Source: Enquête annuelle d'entreprise 2005, Ministère de l'Économie.

The classification of enterprises by their legal form varies with each country's legislation, making international comparisons rather difficult. For example, Table 2.7 shows the data available for Italy and Norway.

Table 2.7(a) **– Enterprises and employees by legal form**
Norway, December 2004

Legal forms	Enterprises	Employees	Number of employees per enterprise
General partnership	7 739	10 401	1.3
Limited company	128 745	1 096 705	8.5
Public limited company	417	59 203	142.0
General partnership with shared liability	6 205	6 363	1.0
Sole proprietorship	155 505	61 438	0.4
Other legal forms	19 194	159 290	8.3
Total	317 805	1 393 400	4.4

StatLink http://dx.doi.org/10.1787/336776126426

a) Employees in the stock of enterprises per 1 January 2005, employees by December 2004.
Source: Statistics Norway.

Table 2.7(b) **– Enterprises and employees by legal form**
Italy, December 2001

Legal forms	Enterprises	Employees			Number of employees per enterprise
		Self-employed	Dependent	Total	
Sole proprietorships (a)	2 667 160	3 079 521	1 129 363	4 208 884	1.6
Partnerships	824 627	1 548 403	1 426 911	3 011 314	3.6
Corporations	531 590	692 343	6 892 706	7 585 049	14.3
Co-operatives (b)	47 719	114 869	671 218	786 092	16.5
Other forms	12 870	14 686	106 883	121 569	9.4
Total	4 083 966	5 485 822	10 227 081	15 712 908	3.8

StatLink http://dx.doi.org/10.1787/336780633170

a) Also includes self-employed workers and members of liberal professions.
b) Excluding social co-operatives.
Source: Italian Statistical Office.

The *intended use of goods* is another important classification criterion. For many years, several European countries published the production index and other cyclical indicators for industry according to the economic use of the goods (consumption, investment, etc.), aggregating the basic indices according to a scheme defined at the national level. Since 2001, however, cyclical indicators have been aggregated into "Main Industrial Groupings (MIGs)" defined at the Community level. The groupings are as follows: consumer durables, consumer non-durables, capital goods, intermediate goods and energy. The indices concerning groups and/or divisions of economic activity are assigned to each grouping according to the criterion of prevalence, *i.e.* according to whether the goods in that group/division are mainly intended for durable consumption, non-durable consumption, etc.

Classification by the *technology intensity* of goods and of the relevant sectors of economic activity is also interesting. This classification, originally proposed by the OECD and also adopted by Eurostat, distinguishes among four groupings of goods/sectors: high technology, medium-high, medium-low and low technology. The various manufacturing sectors are classified in one of the four groupings based on the specific technological level of the sector, the intensity of R&D spending, and on the technology embodied in the sector's purchases of intermediate and capital goods. This classification is often used to analyse export/import flows and to evaluate the *performance* of the various sectors of economic activity.

For agricultural holdings, the main classification criteria refer to the economic size of the holding, the type of occupancy, the type of farming and the territorial location (mountains, hills, etc.). Two criteria have been adopted to evaluate the economic size of the farm: the amount of labour employed; and the utilised agricultural area (UAA). Given the specific organisation of agricultural production and the fact that the farmer's family members are frequently involved in the farm's activities in various ways and extents, the labour provided is not measured by number of employees but in terms of full-time equivalent units (FTE). A FTE is defined as "total hours worked divided by average annual hours worked in full-time jobs" and makes it possible to define the various types of work actually provided using a standard measurement that can be used to evaluate the size of farms in a way that is analogous to that used for industrial and service enterprises. Typically, farms are subdivided according to whether the labour employed is less than one FTE, between one and 10 FTE, or more than 10 FTE. An alternative measure is the size of the agricultural area used, expressed in hectares and consisting of land used specifically for farming.

Agricultural holdings can also be classified based on the type of occupancy. Of course, the types of occupancy depend on the legal framework, uses and traditions. For example, in Italy (Table 2.8) the following categories are used:

- *those directly farmed by the owner*, with a further distinction depending on whether only family labour is used or also salaried labour. It is this type of occupancy when the owner himself provides manual labour on the agricultural holding;
- *those farmed by employees*: when the owner exclusively employs third parties to perform manual labour, while his work (and that of family members) normally consists of managing the agricultural holding;
- *share farming of complete holdings*: when a natural or legal person entrusts a holding to the head of a family, who undertakes with the aid of his family to carry out all work required on the holding and to bear some of the expenses himself.
- *other forms of occupancy*: this category includes the share farming of individual parcels of land (when the concession does not concern the entire holding, but only several parcels of land and the work is carried out without the aid of family members) and the lease of livestock (a form of stock farming based on an agreement between the owner of the pasture and the owner of the livestock).

Finally, agricultural holdings are classified either as specialised or non-specialised. The former include those specialising in the production of seeds, fruits and vegetables, permanent crops, grazing livestock and grain-fed livestock; the latter are subdivided into farms that practice either field crop combination or livestock combination and those that combine crop and livestock farming.

Table 2.8 **– Farms (including publicly owned holdings) and total area by total area class and type of occupancy (area in hectares)**

Italy, 2003

Type of occupancy	Total area class								
	Less than 1	1 - 2	2 - 5	5 - 10	10 - 20	20 - 50	50 - 100	100 and more	Total
Direct farming by owner	538 187	395 797	459 838	232 384	139 163	99 526	29 408	14 427	1 908 730
• *Only family labour*	455 831	334 230	385 686	190 982	110 226	76 641	19 491	7 159	1 580 246
• *Predominantly family labour*	64 303	47 726	54 055	30 150	20 204	15 780	6 503	4 101	242 822
• *Predominantly labour from outside family*	18 053	13 841	20 097	11 252	8 733	7 105	3 414	3 167	85 662
Farming with employees only	10 209	5 276	10 045	5 879	5 479	5 973	2 775	5 380	51 016
Share farming of complete holdings	-	402	-	142	207	72	47	69	939
Other forms of occupancy	390	221	242	603	244	214	115	549	2 578
Total	548 786	401 696	470 125	239 008	145 093	105 785	32 345	20 425	1 963 263

StatLink http://dx.doi.org/10.1787/336781812148

Source: Italian Statistical Office.

2.2.3. General government

The general government sector comprises:

- all institutional units that produce non-market goods and services intended for collective and individual consumption and financed primarily by compulsory payments by units belonging to other sectors;
- all institutional units whose main function is to redistribute the country's income and wealth.

In particular, the institutional sector of general government is subdivided into the following subsectors:

- *central government*: this includes all the administrative bodies of the central State and the other central bodies whose authority normally extends to the entire economic territory, except for central social security funds;
- *local government*: this includes public bodies (except for social security funds) whose authority extends only over part of the economic territory;
- *social security funds*: this includes all central, state and local institutional units whose main activity consists of granting social benefits funded wholly, or in part, by specific groups of the population, according to legislative or regulatory provisions. This subsector includes government administrations responsible for

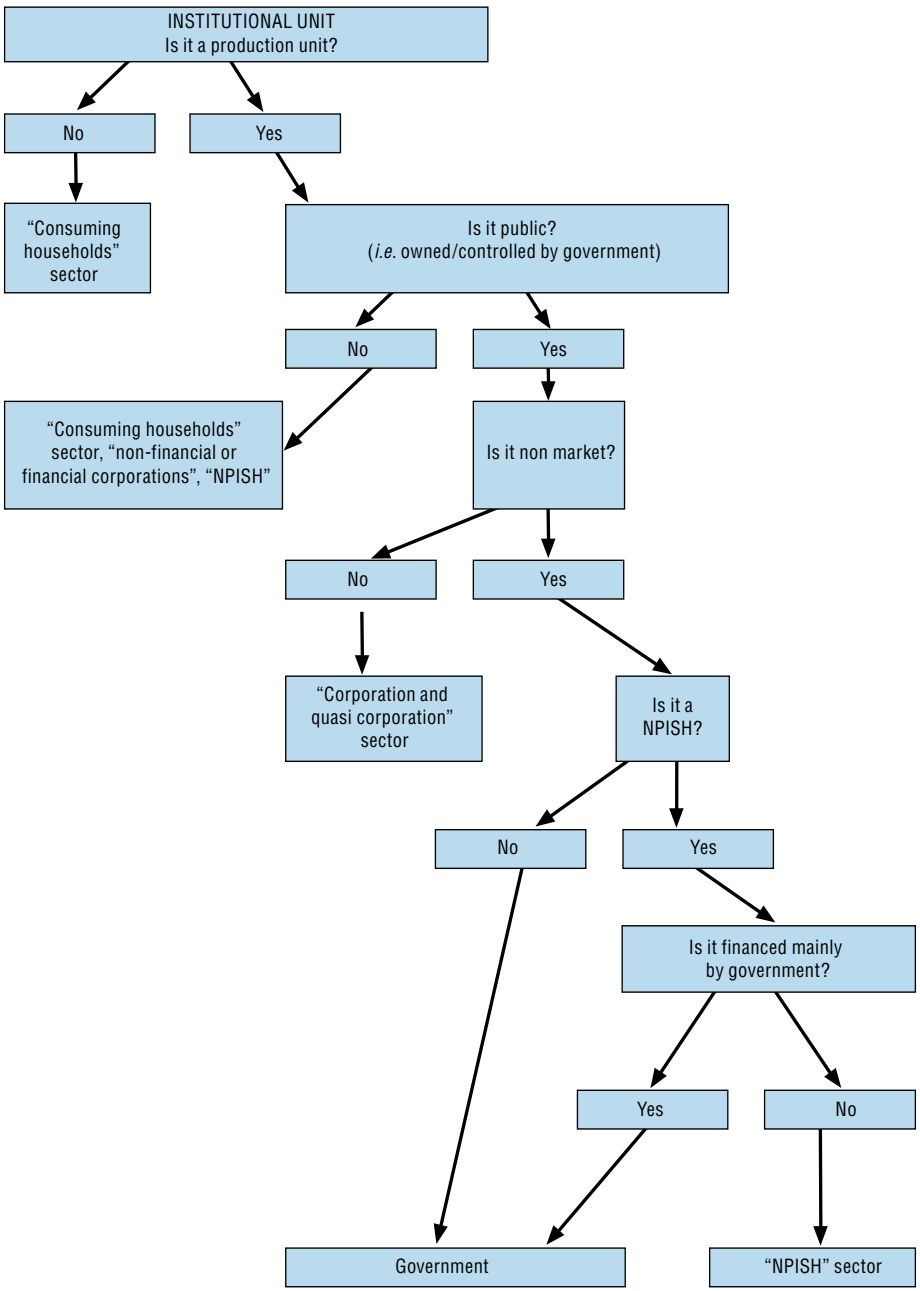
setting, or approving, contributions and benefits, independently from their role as a supervisory or employing body.

The process for classifying each institutional unit resident in the national territory within the general government sector or in other institutional sectors is relatively complex and conducted in logically successive steps. These steps, shown in Figure 2.1, are based on the legal nature and ownership structure of each unit and on the relevant annual budget data. The process begins with the identification of the unit being analysed (in principle any of the units within the system included in the country's statistical records). The first question is whether this unit engages in a productive activity or not: if it does not, its main function is consumption, and it must be classified in the household sector. If it is a production unit, however, it must be determined whether it is private or public (in the sense that it is owned or controlled, directly or indirectly, by a central or local government body included in the general government sector). If it is private, the institutional unit will be included in one of the other institutional sectors.

If the unit is public, however, it must be determined whether it is market-oriented (market unit) or not (non-market unit), *i.e.* whether the goods or services it produces are traded on the market at economically meaningful prices, or if they are provided to consumers completely or virtually free of charge; only in the latter case, in fact, can the unit be classified in the general government sector. To solve this problem, the ESA considers the unit as non-market if the potential proceeds from the sale of goods and services cover at least 50% of the production costs (expenditure on staff, inputs of goods and services, etc.). This condition must be met for an appropriate number of years to ensure stability in the medium term and avoid having to continually reclassify units from one economic sector to another for purely cyclical reasons.

Once it has been established that the public institutional unit is non-market, it is necessary to determine whether it is a non-profit social institution (foundation, association, etc.). In other words, it must be determined whether, in addition to being controlled by a government (a condition already met in one of the preceding steps), it is also primarily financed by government (through unrecoverable transfers): only if this is the case, in fact, should the social institution be included in the general government sector. If its main sources of income are from capital or transfers of private funds (membership fees, donations, etc.), the unit must be classified in the sector of non-profit institutions serving households (NPISHs).

Figure 2.1 – **Procedure for classifying units in the general government sector**



UNDERSTANDING ECONOMIC STATISTICS – ISBN 978-92-64-03312-2 – © OECD 2008

The rationale behind this admittedly complex procedure for classifying institutional units is based on the importance of defining correctly the boundaries of the general government sector. Firstly, the distinction between market and non-market production units has numerous implications for measuring national income and GDP. In fact, two important economic aggregates – production and value added – are quantified differently, depending on the type of unit: for market services units, these two aggregates are measured using the turnover; for non-market services units, they are based on the sum of production costs. This means that if a non-market unit were classified in the corporate sector, its value added would be undervalued.

Secondly, correctly defining the general government sector is important to ensuring the greatest possible international comparability (or inter-regional comparability if the statistics refer to sub-national territorial levels) and can have consequences not only for economic analysis, but also for policy making and institutional operations. For the European countries, for example, the financial resources levied by Community institutions from member countries and those distributed to these countries through structural funds are determined proportionally based on the measurement of the national (and regional) income produced by the various countries (and regions). The implications are even more obvious if we refer to the measurement of the deficit or public debt, two basic parameters for European economic and monetary policy.

Table 2.9 shows the territorial distribution of the over 15 000 units surveyed and classified as public institutions in the 2001 Italian census, with the relevant employment. It also shows the territorial distribution of the more than 235 000 non-profit institutions, *i.e.* those units, with or without a legal personality, public or private, that produce goods or services and that by law or under their own statutory rules cannot distribute profits or other gains except to compensate persons working on behalf of the institution's founders and its members. Consequently, non-profit institutions include associations, whether they are recognised or not, foundations, voluntary and non-governmental organisations, social co-operatives, political parties, trade unions, etc., which in 2001 had some 500 000 employees and nearly 3.5 million volunteers.

Table 2.9 **– Institutions, local units and employees by geographical distribution**
Italy, 2001

Geographical distribution	Public institutions			Non-profit institutions		
	Legal-economic units	Employees	Other forms*	Legal-economic units	Employees	Other forms*
North	8 710	762 345	71 910	120 884	271 606	1 918 298
Centre	2 343	1 913 539	167 969	48 808	110 186	675 959
South	4 527	533 241	32 582	65 540	106 731	825 338
Italy as a whole	15 580	3 209 125	269 461	235 232	488 523	3 419 595

StatLink http://dx.doi.org/10.1787/336784120254

* Includes project staff, temporary workers and volunteers.

Source: Italian Statistical Office.

2.3. Economic aggregates

As we have already mentioned, in an economic system a virtually infinite number of economic transactions take place between economic agents. These transactions, which may encompass a very wide range of purposes (such as the purchase or sale of goods, services, financial activities, income redistribution, etc.) and be carried out between different institutional sectors, result in changes in the situation of individual agents. Based on the reasons for these changes, individual transactions are classified (according to the rules of the SNA) into homogeneous groups, defined as *economic aggregates*. These aggregates measure the overall outcome of the transactions conducted by all units of the economic system. The aggregates include *private consumption*, consisting of all expenditures by households to purchase goods and services that satisfy their personal needs (food, clothing, etc.); *investment*, represented by expenditures to purchase goods to be used to generate income in one or more successive periods; *production*, *i.e.* the result of the economic activity of production units, etc.

Consequently, a single transaction can contribute to a specific economic aggregate, depending on why it was carried out and the economic agents involved. For example, the cash purchase of a new vehicle can be classified as consumption, if it is carried out by a household to transport the family, or as an investment (or gross fixed capital formation) if it is made by a salesman (*i.e.* an incorporated enterprise) in order to visit customers scattered over the territory. Similarly, the dealership that sells the vehicle will enter the amount received into its turnover, and this will contribute to determining the value of the total production of the economic system.

Economic aggregates (or variables) can be classified on the basis of various criteria, but the most relevant distinction is between *stock variables* and *flow variables.* Flows refer to activities and the effects of events that occur within a certain period of time, while stocks refer to situations at a specific time. To explain in greater detail, flows reflect the creation, transformation, exchange, transfer or extinction of economic value that takes place during a certain period of time; stocks represent the value of *non-financial assets* (also called "real assets", such as dwellings, capital goods, etc.) and/or *financial assets/liabilities* (such as means of payment, securities, etc.) at a particular moment in time. It is important to point out that, by their nature, flows can only be expressed with reference to a certain period of time (a year, a quarter, etc.), while stocks refer to a specific moment in time (last day of the year, of the quarter, etc.). Flows can also be expressed as the total value (amount) of the flows generated during the time period, or as an average value, *i.e.* by dividing the sum of the flows by the amount of time (number of days, months, etc.) included in the period being considered.

Real assets can either be *produced* or *non-produced*: produced assets are non-financial assets that have come into existence as outputs of processes that fall within the production boundary of the SNA (such as the vehicle referred to in the previous example): they consist of fixed assets, inventories and valuables. Non-produced assets are non-financial assets that come into existence other than through processes of production: they include both tangible assets (for example a forest) and intangible assets and also include costs of ownership transfer on, and major improvements to, these assets.

2.3.1. Production, intermediate costs and value added

As we have already seen, production represents the flow of goods and services generated under the control of an institutional unit, using goods, services and inputs of factors of production, such as labour and capital. A good represents a physical object used to satisfy a certain need, that can be traded between various agents and over which property rights can be held. Services are heterogenous produced outputs that change the conditions of the consuming unit who requested the service. In this case, the entity to which the property right refers is not a specific physical good, but the possibility of using the service.

Although production activity can concern an infinite number of goods and services, the SNA clearly defines the *boundaries of production*. Since these boundaries are defined conventionally, they may change over time. For example, the 1993 SNA includes as production some activities that were not contained in the previous version (developed in 1968). According to the 1993 SNA, the boundaries of production encompass, in addition to the products of economic activities in the strict sense, the following types of products:

- the production of all individual or collective goods or services supplied to units other than their producers, or intended to be so supplied, including the production of goods or services used up in the process of producing such goods or services;
- the own-account production of all goods retained by their producers for their own final consumption or gross capital formation;
- the own-account production of housing services by owner-occupiers and of domestic and personal services produced by employing domestic staff.

The 1993 SNA included in the production boundary the creation and use of literary and artistic originals and improvements to historic monuments created through a production process, even if that process took place in antiquity. The SNA also requires that the growth that occurs during the cultivation of crops and the rearing of animals be recorded over time, and not only when the products reach final maturity.

Lastly, it should be pointed out that the boundaries of production also encompass the so-called underground economy, *i.e.* the production of economic agents who evade requirements of an administrative nature. For example, the following are included in production: the activity of craft workers who repair or renovate buildings without declaring this to the administrative and tax authorities; goods produced by an enterprise that is not registered with the authorities and that uses workers not declared to social security agencies. Consequently, the fact that a productive activity is conducted illegally does not mean that it is excluded from the boundaries of production. Even though the statistical quantification of these activities is fraught with difficulties, statistical institutes have developed methods for including the valuation of the "non-observed economy" in official data.

According to the SNA, production should also include illegal activities that produce goods and services paid for by buyers voluntarily, such as smuggling, prostitution, the production and marketing of illegal drugs, etc. Production does not include illegal

activities that involve involuntary payments, such as thefts, which are an involuntary transfer between different agents. However, given the enormous difficulties involved in measuring production derived from illegal activities, only a few statistical institutes include such valuations, while others prefer to publish them in special studies.

Production may be valued in various ways, depending on the prices used. Firstly, it can be measured in *market prices* (or purchase prices), which represent the actual price agreed upon by the transactors. This therefore includes transport costs, trade margins and indirect taxes (such as value added tax), less subsidies for products, *i.e.* minus transfers made by governments (or by EU institutions) to influence, generally in a downward direction, the price charged by the producer/vendor. In the absence of market transactions, valuation is made according to costs incurred (non-market services produced by government) or by reference to market prices for analogous goods or services (services of owner-occupied dwellings).

If we exclude from the market price any tax payable and the transport costs invoiced separately by the producer, and we add any receiveable subsidies, we obtain the price actually received by the producer, *i.e.* the *basic price*. This price is particularly important for analysing the enterprise sector since it is the relevant price for decision-making by an enterprise, *i.e.* it is the price on the basis of which the enterprise's actual income and profitability are measured.

To obtain a certain amount of output, the producing unit must use the various original factors of production (labour and capital) as well as the goods and services produced by other units. The value of the goods and services consumed as inputs during the production process, excluding fixed assets whose use is recorded as consumption of fixed capital, constitutes the aggregate known as *intermediate consumption*. Intermediate consumption is always valued at purchase prices; it does not include *depreciation*, which is the decrease in the value of the fixed capital (such as industrial equipment) used during intermediate consumption.

The difference between production at basic prices and intermediate consumption is called *value added at basic prices*: this represents the value that has been added (output) by the production process to the pre-existing value of the goods and services used (input) during this process.

By totalling the value added at basic prices by individual homogeneous production units, we obtain the value added by industry. A special case is represented by the banking sector, which among its services provides financial intermediation to enterprises, households, general government, etc. Since the costs associated with this service are not normally measurable on the basis of specific commissions, national accounting prefers to measure the total of these costs indirectly, leading to the concept defined as "financial intermediation services indirectly measured" (FISIM).

If we aggregate the value added by individual branches, add to this the value of the indirect taxes on production (less subsidies) and subtract FISIM, we obtain the gross domestic product *at market prices* (GDP), which represents the aggregate of production processes within the national territory. If we subtract from GDP net taxes

UNDERSTANDING ECONOMIC STATISTICS – ISBN 978-92-64-03312-2 – © OECD 2008

on production and imports, and compensation of employees and property income payable to the rest of the world, and we add the corresponding items receivable from the rest of the world (in other words, GDP less primary incomes payable to non-resident units plus primary incomes receivable from non-resident units), we obtain the gross national income (GNI), which represents the final result achieved only by units resident in the country from activity carried out both within the national territory and abroad. The table that presents these aggregates in the SNA is known as the *production account* (Table 2.10).

Table 2.10 **– Production account**
Value in current prices (million euros) – Italy, 2002-2005

Aggregates	2002	2003	2004	2005
RESOURCES				
Production at basic prices	2 514 947	2 587 887	2 690 677	2 752 042
Taxes less subsidies on products	129 807	131 614	139 712	144 479
USES				
Intermediate consumption	1 349 528	1 384 147	1 441 518	1 479 280
Gross domestic product	1 295 226	1 335 354	1 388 870	1 417 241

StatLink http://dx.doi.org/10.1787/336870088018

Source: Italian Statistical Office.

2.3.2. Consumption, capital formation and net foreign demand

The production achieved by an economic system, expressed in terms of GDP at market prices (Y), constitutes, together with imports of goods and services (M), the overall resources available for final consumption (C), gross capital formation (I) and exports of goods and services (X):

$$Y + M = C + I + X \qquad [2.1]$$

This entity is known in national accounts as the *goods and services account* (or also the *resources and uses account*).

Final consumption represents expenditure for goods and services aimed at satisfying human needs. Depending on the type of need met, final consumption can be broken down into *individual consumption*, carried out for the benefit of households, and *collective consumption*, which is done for the benefit of households, enterprises or other institutions. The individual consumption of households (which also includes the consumption of agricultural products they have produced themselves) can also be financed by NPISHs and general government, especially in the areas of health, education, social security and culture. This makes it necessary to distinguish between "final" consumption and "actual" consumption. The former represents the viewpoint of those financing the consumption, while the latter reflects the viewpoint of those benefiting from the consumption. Actual consumption represents the total goods and services used up by individual households; it corresponds to the households' consumption expenditure plus the individual consumption expenditure of general government and NPISHs.

Consumption can also be calculated as *national consumption*, *i.e.* the consumption within the national territory and abroad by the units resident in the national territory, or as *domestic consumption*, *i.e.* the consumption within the national territory both by resident and non-resident units (Table 2.11).

Final consumption is always valued in market prices and is calculated using consumption "functions", *i.e.* by aggregating expenditures for goods and services according to homogeneous categories (food, beverages, transport services, etc.). However, a different classification distinguishes between various types of consumption, based on the type and durability of the good. This leads to the distinction between services and goods, further distinguishing between "non-durable" goods, which are consumed immediately after purchase (food, pharmaceuticals, etc.) or within a limited amount of time (clothing), and "durable" goods, which are purchased to be used over a number of years (furniture, motor vehicles, household appliances, etc.).

Table 2.11 **– National final consumption and domestic final consumption**
Values in current prices (billion yen) – Japan, 2003-2005

	2003	2004	2005
National consumption	370 294	373 896	377 273
Expenditure of general goverment (–)	88 503	89 468	90 684
Expenditure of NPISHs (–)	5 877	6 118	6 396
Expenditure of resident households	275 915	278 310	280 193
Acquisitions abroad by residents (–)	2 783	3 409	3 320
Acquisition in Japan by non-residents (+)	646	804	923
Final domestic household comsumption	273 778	275 706	277 795

StatLink http://dx.doi.org/10.1787/336874201116

Source: National Accounts of OECD Countries, OECD.

Final households' consumption also includes "actual" rent, *i.e.* the rent actually paid by households for the use of dwellings, and the "imputed" rent of homeowner-occupiers. This latter kind of rent is said to be "imputed" since it is estimated based on the rent that owner-occupiers would have to pay to occupy their dwellings, even if the owner-occupiers do not actually make any monetary payments.

Capital formation represents the value of acquisitions less disposals of produced non-financial assets, *i.e.* of assets derived from a production process. It can be broken down into gross fixed capital formation, changes in inventories and acquisitions minus disposals of valuables. *Gross fixed capital formation* consists of the total value of a producer's acquisitions (less disposals) of fixed assets during the accounting period plus certain additions to the value of non-produced assets (such as subsoil assets or major improvements to the quantity, quality or productivity of land) realised by the productive activity of institutional units. Gross fixed capital formation includes physical goods (machinery, equipment, etc.) as well as intangible products (such as software), obtained through a specific production process and used in other production processes for a period longer than one year. The components of gross fixed capital formation are always valued at market prices and can be aggregated by

industry of production, *i.e.* according to the characteristics of the goods (machinery, equipment and various products; buildings; intangible goods), or by owner or user industry, *i.e.* according to the characteristics of the purchaser, in which case the classification by institutional sector is used (households, businesses, general government).

Inventories consist of stocks of outputs still held by the units that produced them prior to their being further processed, sold, delivered to other units or used in other ways, as well as stocks of products acquired from other units intended to be used for intermediate consumption or for resale without further processing. Inventories can be raw materials and intermediate goods, work-in-progress or finished goods. Their purpose is to promote the economically efficient management of the production and distribution processes, making it possible to synchronise better the fluctuations of supply and demand. Consequently, inventories are considered one of the possible uses of the goods produced; the value for the term "I" in the identity [2.1] is the variation in the overall stock of inventories recorded during the reference period. In turn, this variation equals the value of goods intended for inventories minus the value of those that leave inventories to be used in the production process or sold, minus losses of value due to physical deterioration, accidental damage or theft.

Lastly, the term *net acquisitions of valuables* refers to "stores of value" (such as precious stones and metals), acquired mainly by households for purposes other than consumption. The resources and uses account naturally does not enter the amount of the stock of valuables, but its variation during the reference period.

The flow of gross fixed capital formation contributes to determine the evolution of the capital available to the economic system: in particular, if in a given year gross fixed capital formation (GFCF) is higher than depreciation (D), *i.e.* higher than the loss in the value of capital due to its physical deterioration during the production processes, the net fixed capital formation is positive, and this leads to an increase for that year in the "net" capital available to the economic system (NK):

$$NK_t = NK_{t-1} + (GFCF_t - D_t) \qquad [2.2]$$

In other words, depreciation can be considered as that portion of capital formation necessary to preserve unchanged the existing production capacity. What is important for the development of an economic system in the long term is therefore *the trend* of net capital formation. In order to emphasise this aspect, an important measurement of a country's actual economic growth is *net domestic product* (NDP), which is equal to the gross domestic product minus depreciation.

The trend of the NDP may be significantly different from that of the GNP, especially when the average life of capital goods changes significantly over time. This is a trend seen in virtually all industrialised countries since the 1990s when the "computer revolution" caused accelerated depreciation of existing equipment, as it was replaced by new equipment with a higher technology content but a significantly lower average life (*e.g.* the average life of a personal computer is much shorter than that of large industrial machinery). This means that countries that have invested more heavily in the new economy to remain competitive, have seen their gross capital formation (and

thus their GDP) increase, without this being reflected in an overall increase in the country's production capacity (*i.e.* in an equivalent increase in the NDP).

Lastly, with regard to trade in goods and services with the rest of the world, it must be emphasised that in the system of national accounts, both imports and exports are valued *free on board*, *i.e.* without considering the cost of transport and insurance between the border of the exporting country and that of the importing country. The difference between exports and imports is also known as *net foreign demand* and represents the contribution that foreign trade makes to Gross domestic product. In fact, another way of expressing the relationship [2.1] is as follows:

$$Y = C + I + (X - M) \qquad [2.3]$$

The above equation shows how a decrease in the net foreign demand reduces GDP, other conditions being the same.

Exports and imports concern not only goods, but also services such as personal, cultural and recreational services, financial services, communication services, etc. By convention, imports and exports of services comprise respectively the expenditures abroad of residents and the expenditures of foreigners within the national territory.

2.3.3. From gross domestic product to national income

The gross domestic product represents the summary measurement of the income produced over a given period of time (generally one year) by a specific economic system. However, to be used by the various institutional units (households, corporations, general government, etc.) for consumption, capital formation or trade in goods and services with the rest of the world, the income must be transferred from the centres of production to the places where it is used. This is achieved through the remuneration of the two factors of production considered in the SNA: labour and capital.

Labour is remunerated through the payment of compensation to individuals, to which must be added what are known as social charges, which simply represent deferred compensation. The sum of *wages and salaries* (payable in cash or in kind) and the value of the *social contributions* payable by employers (actual social contributions payable by employers to Social Security schemes or to private funded social insurance schemes to secure social benefits for their employees; or imputed social contributions by employers providing unfunded social benefits) constitutes the *compensation of employees*, *i.e.* the portion of income that is paid directly or indirectly to those who have contributed to the production process by providing their labour as employees. The difference between the gross value added and the compensation of employees determines the *gross operating surplus*, *i.e.* the portion of the income produced intended to compensate the labour provided directly by the entrepreneur and the other original factor of production (capital).

As we have seen, the GDP reflects the remuneration of labour and capital used for production purposes: therefore, the *gross national income* (GNI) is also equal to the aggregate value of the balances of gross primary incomes for all sectors. Naturally, it

UNDERSTANDING ECONOMIC STATISTICS – ISBN 978-92-64-03312-2 – © OECD 2008

is also possible in this case to obtain the *net national income* by subtracting the value of depreciation from the gross national income.

Lastly, by deducting from the gross national income the net taxes paid to general government and the net current transfers paid to the rest of the world, the *gross national disposable income* is obtained, which can be used by resident units for the purpose of final consumption. The difference between national disposable income and final consumption (in both cases, after taking into account of an adjustment for pension funds) constitutes *saving*.

2.4. Values at current prices and constant prices

The value of economic transactions is expressed in terms of a specific common currency (dollars, euros, etc.). At the conceptual level, the value (said to be "nominal") of many transactions is determined by multiplying the quantity of the product involved in the transaction by its price, which is in turn defined as the amount of money that must be paid in exchange for a unit of the product. In this way, the value of the transaction is expressed *in current prices*, *i.e.* using the price agreed upon by the parties for that particular transaction at that specific time.

However, as we know, prices vary over time, so economic aggregates expressed in current prices do not enable us to determine to what extent the variations observed over a certain historical period (year, month, etc.) are due to variations in quantities or variations in prices. Consequently, economic statistics focuses not only on aggregates expressed in current prices, but also on the trend of the volumes and the associated price variations. For this purpose, many aggregates are calculated either in current prices or net of the price movements that occurred during the period under consideration. This is done by expressing the value of the various transactions using the prices prevailing during a certain period, taken as a reference. In this case, the resulting economic aggregates are said to be expressed *in constant prices,* and their variations over time are said to be defined "in real terms", in opposition to those "in nominal terms", calculated on the basis of values in current prices.

Let us take the example of the sale of automobiles, and let us assume that in the year 2000 a dealership sold 40 identical cars, with the same qualitative characteristics (model, options, etc.) at a price of 11 000 euros each. Next, let us assume that in 2001 the number of cars (again with the same qualitative characteristics) rose to 45, at a unit price of 11 500 euros. Consequently, the value in current prices of the sales realised by that dealer will be 440 000 euros (40 * 11 000) in 2000 and 517 500 euros (45 * 11 500) in 2001, with a percentage increase of 17.6% [(517 000 – 440 000) /440 000) * 100]. In reality, only part of the increase is due to a variation in the quantities acquired, *i.e.* an increase in sales in real terms, while the remainder is due to a variation in prices. In particular, the increase in sales at constant prices (of 2000) is equal to 12.5% [(45 * 11 000 – 40 * 11 000)/(40 * 11 000) * 100].

The relationship between an economic aggregate expressed in current prices and the same aggregate expressed in constant prices is know as a *deflator*. Thus, it is possible to have deflators of the gross domestic product, private consumption,

gross fixed capital formation, etc. Valuations in constant prices can also be made either at the prices of a certain base year or at the prices of the previous year. As we shall see in Chapter 4, at the end of 2005 several EU countries changed to a system that deflates national accounting aggregates based on the previous year's prices.

2.5. Index numbers

The deflator reflects a typical statistical ratio, *i.e.* a quotient of two statistical terms, both of which refer to the same phenomenon measured at different times: this type of statistical relationship is called an *index number*. Index numbers may be calculated in reference to either time or space. In official economic statistics, index numbers (or, more briefly, "indices") are more often used for time comparisons. In this case, the index is constituted by the ratio between the measurement of a phenomenon at a given moment in time, and the measurement at another moment taken as a reference (the base of the index). Let us, for example, calculate the ratio between the value of private consumption at current prices for 2000-2005 and its value in 2000 prices. The resulting index (expressed in "base 2000", given that in this year the numerator and denominator are equal) will measure the trend of the implicit prices of private consumption during 2001-2005, in relation to price levels observed in 2000.

We can distinguish between *simple index numbers* and *complex index numbers*: within a time frame, the former make it possible to compare the variation over time of an individual phenomenon in relation to the base; the latter, however, can express the variation of a group of phenomena in relation to their value during the period chosen as the base of the index. A particularly important element is also the distinction between *fixed base* index numbers and *moving base* index numbers. The former always use as a reference (*i.e.* in the denominator) the same quantity (in the preceding example, the value of consumption in 2000), while the latter use in the denominator an element that varies over time. The indices used for time comparisons normally have the value 100 in the reference (or base) period.

The calculation of simple fixed base indices (IF) is relatively easy. Using V_t to indicate the value over time t (for t = 1, 2, ... n) of the relevant economic aggregate, the index expressed in base 0 ($_0IF_n$), for the periods from 0 to n is:

$$_0IF_0 = V_0/V_0 * 100$$
$$_0IF_1 = V_1/V_0 * 100$$
$$_0IF_2 = V_2/V_0 * 100$$
....
$$_0IF_n = V_n/V_0 * 100$$

In the case of a simple moving base index (IM), with previous period used as base ("chain" index), the values of the index for the periods from 1 to n are:

$$IM_1 = V_1/V_0 * 100$$
$$IM_2 = V_2/V_1 * 100$$

UNDERSTANDING ECONOMIC STATISTICS – ISBN 978-92-64-03312-2 – © OECD 2008

$IM_3 = V_3/V_2 * 100$

....

$IM_n = V_n/V_{n-1} * 100$

Given that the percentage variation VP_t is

$$VP_t = [(V_t - V_{t-1})/V_{t-1}] * 100$$

and also that $VP_t = [(V_t/V_{t-1}) - 1] * 100$

and that $VP_t = [(V_t/V_{t-1}) * 100] - 100,$

then $VP_t = IM_t - 100.$

In this way, we see that the simple moving-base indices in each period only express the percentage variations in relation to the previous period.

The fixed-base indices expressed in a given base can easily be converted into another base. To go from one fixed-base index referring to period w to a fixed-base index referring to period z, it suffices to divide the index ${}_wIF_t$ by the value of the index expressed in base w, but referring to a new base period ${}_wIF_z$, and to multiply the result by 100:

$${}_zIF_t = {}_wIF_t/{}_wIF_z * 100$$

It is slightly more complicated to convert a simple fixed-base (expressed in base w) to a moving-base index and vice versa: in the former case, it is necessary to divide each term ${}_wIF_t$ by its preceding ${}_wIF_{t-1}$ and multiply the result by 100. In fact, given that the generic term ${}_wIF_t$ is given by $(V_t/V_w * 100)$, it follows that:

$$IM_t = (V_t/V_w * 100)/(V_{t-1}/V_w * 100) * 100$$

which becomes:

$$IM_t = (V_t/V_w * 100) * (V_w/V_{t-1} * 100) * 100$$

that is:

$$IM_t = (V_t/V_{t-1}) * 100$$

When converting from a moving-base index IM_t into a fixed-base index ${}_0IF_t$ (where 0 refers to the first term of the series being considered), each term IM_t must be multiplied by the product of all the moving base indices included between the time 0 and the time t-1. For example, to calculate ${}_0IF_5$, the following operation must be performed:

$${}_0IF_5 = (IM_1 * IM_2 * IM_3 * IM_4 * IM_5) * 100$$

If we now wish to express the fixed-base index with reference to a term different from 0, such as z, after calculating the series ${}_0IF_t$ we can, as previously described, change the base by dividing all the terms ${}_0IF_t$ (for t = 1, 2, ... n) by the value ${}_0IF_z$ and multiply the result by 100.

As already indicated, complex index numbers are able to express, in a single synthetic value, the variation of many phenomena with respect to a period chosen as the base of the index. The synthesis of the basic information is made by calculating its weighted average using weighting schemes selected in accordance with the objective being pursued. Consequently, the construction of a complex index number requires that the following elements be determined: the statistical data on the phenomena to be considered in the index; the base of the index; and the weighting scheme.

There is an extraordinary variety of indices, which differ mainly in the weighting scheme used: there are indices that use arithmetic, geometric, harmonic means, etc., and weight schemes that give greater importance to the base period or to the period for which the index is being calculated. The following index numbers are widely used in the field of economic statistics: Laspeyres (IL), Paasche (IP) and Fisher (IF) indices: the first two use arithmetic means and differ in the schemes of weights adopted, while the Fisher index is equal to $\sqrt{IL*IP}$, *i.e.* the geometric mean of the Laspeyres and Paasche indices.

Let us now consider the Laspeyres index, assuming that we wish to construct a synthetic index of the prices of a given set of consumer goods for the period from 1 to t. To calculate a synthetic index that expresses the variation in the general level of prices observed, we must naturally attribute to each price a weight that expresses the importance of the relevant good in the overall consumption expenditure. Let us assume that we choose as a weight the value of the consumption of the good i during the time t (v_{it}) equal to:

$$v_{it} = p_{it}\, q_{it}$$

where p_{it} is the price of that particular good and q_{it} is the relevant quantity consumed. Let us also assume that we choose the period 0 as a reference base: thus, for the period 0 we shall have the price of the individual good p_{i0} and the relative value consumed v_{i0}. The Laspeyres index, which represents a weighted mean of basic price indices for the period t in base 0 is therefore:

$$ {}_0IL_t = \frac{\sum_{i=1}^{n} \frac{p_{it}}{p_{i0}}\, p_{i0}\, q_{i0}}{\sum_{i=1}^{n} p_{i0}\, q_{i0}} = \frac{\sum_{i=1}^{n} p_{it}\, q_{i0}}{\sum_{i=1}^{n} p_{i0}\, q_{i0}} \qquad [2.4] $$

As can be seen, both the numerator and the denominator contain the quantity consumed during the time 0 (q_{i0}): consequently, the index is not affected by the changes that occurred in the pattern of consumption during the periods after the base period; it is affected only by the basic price variations (p_{it}/p_{i0}).

UNDERSTANDING ECONOMIC STATISTICS – ISBN 978-92-64-03312-2 – © OECD 2008

However, the Paasche index for the period t in base 0 is expressed by:

$$ {}_0IP_t = \frac{\sum_{i=1}^{n} \frac{p_{it}}{p_{i0}} p_{i0} q_{it}}{\sum_{i=1}^{n} p_{i0} q_{it}} = \frac{\sum_{i=1}^{n} p_{it} q_{it}}{\sum_{i=1}^{n} p_{i0} q_{it}} \qquad [2.5] $$

In this case, the weighting structure is provided by the quantities consumed q_{it} during period t, and not during the base period, as for the Laspeyres index. Consequently, the Paasche index is affected by the changes due both to the variation in prices (p_{it}/p_{i0}) and in the pattern of consumption between the period chosen as the base of the index and period t.

Laspeyres and Paasche indices are widely used to construct price, volume and value indicators. When Laspeyres and Paasche formulae are used to construct price indices, the results obtained show a systematic bias. If, as is shown by economic theory, the quantity consumed falls as the price rises (all other conditions being the same), then use of the Laspeyres index tends to overestimate the general price trend of the products included in the index: in fact, if the pattern of consumption is kept unchanged during the base period, the substitution effect for products whose prices increase more than average is not taken into account. In other words, the index measures the variation in the general level of prices that would have occurred if consumers had not "adjusted" their own behaviour to price increases, thereby reducing the quantities consumed of those products whose prices rose more than others. Consequently, the Laspeyres index will always indicate a variation in the average level of prices that is greater than that measured with the Paasche index. On the other hand, the Paasche index will have a negative "bias", giving less-than-average weight to those products whose prices increased the most during the period considered. The Fisher index, as a geometric mean of the two indices, does not have these biases.

It should also be pointed out that, in ordinary practice, the Paasche index can be difficult to use because of the lag in the availability of information on the structure ofweights. This is the case, for example, when calculating a monthly consumer price index for the 2000-2005 period, in base 2000 = 100. While the prices p_{it} can be observed monthly, the quantities consumed q_{i2005} can only be evaluated at the end of the year: consequently, the Paasche index (and the Fisher index) for 2005 can only be calculated in 2006, a problem that does not arise for a Laspeyres index calculated for 2000-2005 and expressed in base 2000.

2.6. Time series

A substantial portion of economic statistics are finalised by comparing economic aggregates over time, and these comparisons are generally conducted using time series. In intuitive terms, a *time series*

$$Y_t = y_1, y_2, y_3, \ldots y_n \qquad [2.6]$$

is a succession of numerical data arranged on the basis of time t. Given that time is a continuous variable, but that statistical observations of economic phenomena are normally carried out with a certain periodicity (year, quarter, month, etc.), economic time series are generally "discrete", *i.e.* composed of a finite number of observations made with reference to conventionally defined time periods.

The trend of time series is naturally influenced by many factors: some of these cause temporary movements in time series, while others have an underlying influence on trends. Very schematically, we can imagine a time series as being composed of three basic "components" (Figure 2.2):

- the *trend-cycle* (CT_t) represents the underlying trend of the series, which is typically determined by intrinsically economic factors. In turn, the trend-cycle can be broken down into the long-term trend (the trend) and the fluctuations, which are not necessarily regular, observed in relation to the trend over a multi-year period (the cycle);
- *seasonality* (S_t), determined by climatic, cultural and organisational factors that cause movements in the time series, which are repeated with some regularity from year to year;
- *irregularity* (A_t), determined by an infinite series of temporary and random phenomena that cannot be identified either with the trend-cycle or seasonal variation.

There are many ways of representing how these components determine the evolution of a time series, although it should be emphasised that in all cases these representations are only a statistical abstraction constructed for the purpose of measurement and analysis. For example, we can imagine that the time series is the result of the sum of the three components just described (additive model):

$$Y_t = CT_t + S_t + A_t \qquad [2.7]$$

or of their product (multiplicative model):

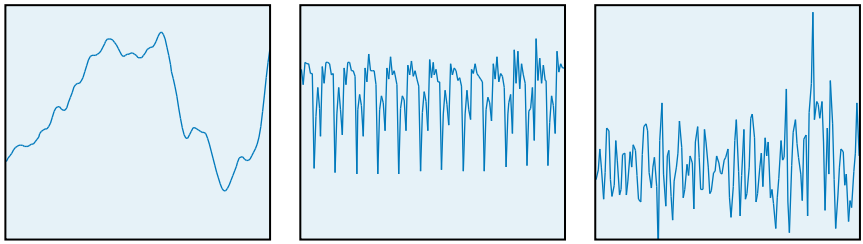
$$Y_t = CT_t * S_t * A_t \qquad [2.8]$$

or of a mixed scheme:

$$Y_t = CT_t * S_t + A_t \qquad [2.9]$$

UNDERSTANDING ECONOMIC STATISTICS – ISBN 978-92-64-03312-2 – © OECD 2008

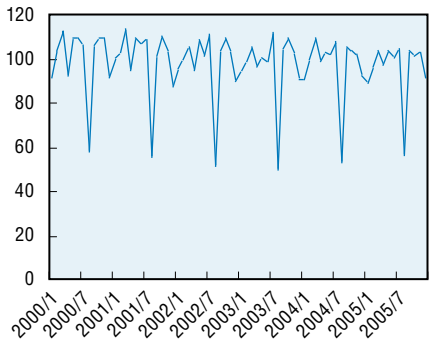
Figure 2.2 **– Components of a time series: trend-cycle, seasonal variation, irregular variation**



The decomposition of a time series into basic components can be extremely useful for understanding some economic phenomena better. For example, let us look at the index that measures industrial production in Germany (Figure 2.3). As can easily be seen, the movements of the index are dominated by fluctuations that are repeated annually because of the closing of industrial establishments during the summer months and the Christmas holidays, while the trend-cycle is difficult to read. In the case of the consumer price index of Spain (Figure 2.4), the dominant component is the trend, making it difficult to assess the role played by seasonality, which is nevertheless important for the prices of many products (such as fresh food products), or by the cycle.

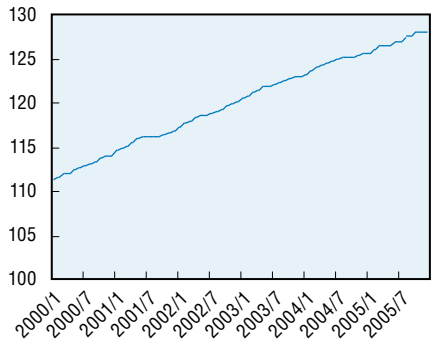
To isolate the most significant components for the purposes of economic analysis, many methods for decomposing a time series have been developed; of particular importance are "seasonal-adjustment" methods, which make it possible to isolate the seasonal component and produce series "net" of this component and/or the irregular component. For example, Figure 2.5 shows the seasonally adjusted series of the Italian industrial production index obtained by applying the TRAMO-SEATS procedure, which together with the X12-ARIMA procedure is most frequently used by statistical institutes to produce seasonally adjusted economic time series. The analysis of Figure 2.5 clearly shows a phase of expansion between the beginning of 2000 and the initial months of 2001, followed by a phase of recession (until the end of 2001), and then substantial stagnation that continued throughout 2002, and later a slow slippage in production following a moderate negative trend.

Figure 2.3 **– General industrial production index – Raw data**
Germany, 2000-2005



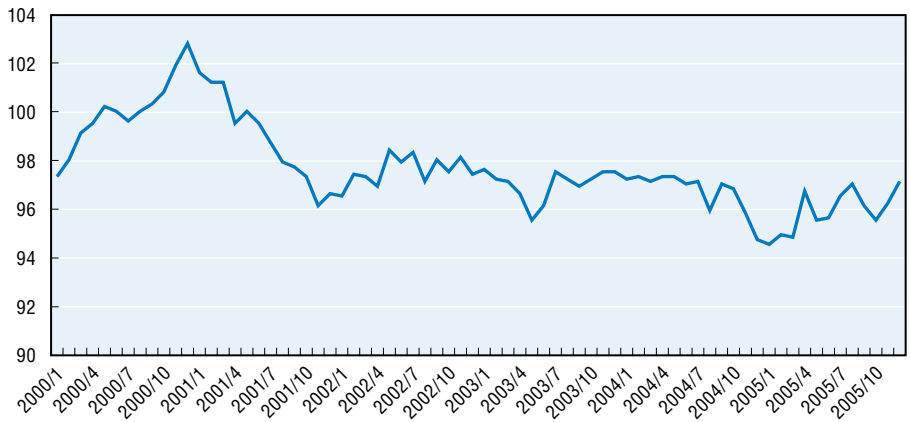
Source: OECD, *Main Economic Indicators.*

Figure 2.4 **– Consumer price index – All items**
Spain, 2000-2005



Source: OECD, *Main Economic Indicators.*

Figure 2.5 **– General industrial production index – Seasonally adjusted data**
Italy, 2000-2005



Source: OECD, *Main Economic Indicators.*

2.6.1. Measuring the movements of a time series

The graphic analysis of a time series, even though it provides useful information for understanding the evolution of the phenomenon being analysed, cannot by itself estimate the intensity of its movements. To do so, we must use statistical measurements capable of comparing the values of the phenomenon at various moments in time; but to make a "significant" time comparison, *i.e.* one which is able to provide the user with a sensible result in economic terms, it is necessary

UNDERSTANDING ECONOMIC STATISTICS – ISBN 978-92-64-03312-2 – © OECD 2008

to make some choices. Firstly, we must specify the period of time to which we are referring, *i.e.* the *time interval (or horizon)* of the analysis. The time interval is typically defined by two specific observations (the beginning and end of the interval), and the time comparison can be made by comparing contiguous periods (years, quarters, months, days, weeks, etc.) or periods that are separated from each other in the time frame. Secondly, we must identify the *frequency* at which we wish to measure the phenomenon over time, a choice closely linked to the characteristics of the economic variable being analysed and to the cognitive objectives of this analysis. If, for example, an analyst wishes to understand the current or future economic situation, he/she will tend to observe the trend of economic aggregates over the most recent quarters or months, while if the objective is to evaluate the changes that have occurred in an economic system following major legislative changes, the comparison will presumably be based on annual data referring to a long time interval. Furthermore, the choice of comparable time units is a necessary but not sufficient condition for making meaningful comparisons. In fact, as we have already mentioned, economic time series are influenced by climatic, cultural and organisational factors that tend to introduce into these series movements that are of little interest to economic analysis.

To measure the intensity of the movements of a time series, various indicators can be used. Percentage variations are very frequently used. Given a time series I_t, where t indicates the generic period of observation, the following percentage variations can be defined:

$$\text{month-on-month growth rate: } M_t = \frac{I_t - I_{t-1}}{I_{t-1}} * 100 \qquad [2.10]$$

$$\text{year-on-year growth rate: } Y_t = \frac{I_t - I_{t-z}}{I_{t-z}} * 100 \qquad [2.11]$$

Where z = 1 for annual series, z = 4 for quarterly series and z = 12 for monthly series. For monthly and quarterly series the following relation is then valid:

$$M_t \gtreqless M_{t-z} \Leftrightarrow Y_t \gtreqless Y_{t-1} \quad \forall\, t \qquad [2.12]$$

Therefore, if for monthly series the cyclical variation for the month t is higher (lower) than that recorded 12 months earlier, the trend variation for the month t will necessarily be higher (lower) than that recorded in month t-1. With a few changes, it can also be shown that:

$$Y_t - Y_{t-1} = (M_t - M_{t-12}) * (I_{t-1}/I_{t-12}) \qquad [2.13]$$

and that:

$$Y_t = M_t\,(I_{t-1}/I_{t-12}) + M_{t-1}\,(I_{t-2}/I_{t-12}) + \ldots + M_{t-11}\,(I_{t-12}/I_{t-12}) \qquad [2.14]$$

i.e. that the year-on-year variation at time t is equal to the weighted sum of the last 12 month-on-month variations.

If we define as A_T the simple annual average for year T of the monthly values I_t for year T:

$$A_T = (I_{tT} + I_{t+1T} + ... + I_{t+11T})/12 \quad [2.15]$$

we can calculate the average annual growth (AA_T) as:

$$AA_T = (A_T - A_{T-1})/A_{T-1}) * 100 \quad [2.16]$$

that is:

$$AA_T = (\{[(I_{tT} + I_{t+1T} + ... + I_{t+11T})/12] - A_{T-1}\}/A_{T-1}) * 100 \quad [2.17]$$

This equation shows us how the average annual variation for the year T depends on the relationship that is established between each month of year T and the average for year T-1. In the case of time series that increase from month to month (*i.e.* positive monotonic series, such as nominal contractual wages), the level of the average of year T will be located approximately around the value recorded during the central months of that year, and the average annual growth AA_T will depend on the dynamic of the time series during the second part of year T-1 and the first part of year T.

Lastly, let us define as *cyclical gain* at month t of year T (AC_{tT}) the relationship:

$$AC_{tT} = [(S_{tT} - A_{T-1})/A_{T-1}] * 100 \quad [2.18]$$

where the quantity S_{tT} expresses the average value of year T that would result if for the rest of the year the time series remained at the level recorded during month t:

$$S_{tT} = [\sum_{i=1}^{t} I_{iT} + (12-t)\, I_{tT})] /2 \quad [2.19]$$

The quantity S_{tT} expresses the average of two terms: the values of the time series actually observed in the first t months of year T and their extrapolation to the remaining part of the year obtained, assuming that the level reached in month t remains unchanged for the months from t+1 to 12. The cyclical gain at March of year T will then express the average annual variation that would be seen in year T if the series remained unchanged for the months April to December at the level reached in March. Similarly, the cyclical gain calculated from December of year T-1 expresses the variation that would occur in year T if the time series remained unchanged at the level reached in December of the previous year, *i.e.* the part of the growth that would be attributed "arithmetically" to year T, but which in reality occurred during the previous year.

It can also be shown that:

$$AA_T = AC_{12T-1} + [M_1\, (I_1 + AA_{T-1}) * 12/12] + [M_2\, (I_2 + AA_{T-1}) * 11/12] + \ldots +$$
$$+ [M_{12}\, (I_{12} + AA_{T-1}) * 1/12] \quad [2.20]$$

i.e. that the average annual growth in year T is equal to the cyclical gain of December of year T-1, plus the weighted sum of the 12 month-on-month variations (from January to December) of year T.

UNDERSTANDING ECONOMIC STATISTICS – ISBN 978-92-64-03312-2 – © OECD 2008

Lastly, let us point out that the average annual growth cannot simply be obtained as the average of 12 monthly trend variations. In fact, the relationship between these two measurements is as follows:

$$AA_T = [Y_{1,T} * (I_{1,T-1}/A_{T-1})] + [Y_{2,T} * (I_{2,T-1}/A_{T-1})] + \ldots + \\ + [Y_{12,T} * (I_{12,T-1}/A_{T-1})] \quad [2.21]$$

From above formula,it is seen that AA_T is a weighted (and not a simple) average of the year-on-year variations, with weights provided by the level of the time series reached at time t-12 with respect to the annual average T-1.

2.6.2. Seasonal adjustment procedures

Although many seasonal adjustment procedures based on different methodological approaches have been proposed over time, two procedures are currently most widely used: X12-ARIMA and TRAMO-SEATS. The former performs the decomposition between the trend-cycle (CT_t), seasonality (S_t) and irregularity (A_t) through the successive application of different-sized moving averages, while the latter does so using the ARIMA (Auto Regressive Integrated Moving Average) models.

Given a time series y_t, the series y_t^* obtained by applying a moving average of "order" m (=n1+n2+1) is given by:

$$y_t^* = \sum_{i=-n1}^{n2} \vartheta_i y_{it} \quad [2.22]$$

In other words, a moving average represents an operator that transforms the original series into a linear combination, with weights ϑ_i. There is a potentially infinite variety of moving averages. For example, if all weights ϑ_i are equal, the moving average is said to be "simple", but if they are different, the average is said to be "weighted". If n1=n2, the moving average is said to be "centred", and when the order of a centred moving average is uneven, the resulting individual values of the series y_t^* necessarily refer to the same time periods of y_t. If, however, the order is even, this correspondence is lost. If we assume, for example, that we have a monthly series y_t, and we apply a moving average of order 4 (*i.e.*, considering y_t, y_{t+1}, y_{t+2}, y_{t+3}), the result obtained can no longer refer to the months to which the original series referred, but to theoretical periods that "straddle" them. To obtain a result referring to the original months, we must calculate the arithmetic mean of two successive terms of the moving average obtained in the first step.

Lastly, a moving average is said to be "symmetric" if its order is uneven, then n1 = n2 and the value of the weights ϑ_i are identical for each of the terms t+i and t-i, $\forall i$. However, if the weights are not identical, or if n1 is different from n2, the moving average is said to be "asymmetric".

Moving averages are particularly important because they make it possible to approximate complex mathematical functions through which some of the

components of time series can be represented. For example, the trend-cycle (CTt) can be represented by a polynomial function f(t) of time t:

$$CT_t = a_0 + a_1\, t + a_2\, t^2 + \ldots + a_n\, t^n \qquad [2.23]$$

which can be approximated by a centred moving average with weights appropriately selected on the basis of the order of the polynomial to be approximated. A third-grade polynomial can in fact be approximated by a five-term moving average with weights [-0.086, 0.343, 0.486, 0.343, -0.086], or by a seven-term average with weights [-0.095, 0.143, 0.286, 0.332, 0.286, 0.143, -0.095].

It is not only the trend-cycle that can be approximated through moving averages but also the seasonal component. This means that by applying moving averages, it is possible to "filter" the various components of a time series and this is exactly what the X12-ARIMA procedure makes possible. In very schematic terms, the X12-ARIMA procedure (derived from X11-ARIMA developed by Statistics Canada on the basis of the procedure originally proposed by the U.S. Census Bureau) performs the seasonal adjustment of a monthly time series $Y_t = CT_t \cdot S_t \cdot A_t$ through the following steps:

1. Elimination of outliers from the time series that are due to purely accidental factors (strikes, floods, etc.) and preliminary processing of data for corrections due to holidays, different numbers of working days in each month, etc.
2. Calculation of the preliminary estimate of the trend-cycle (CT_t') through the use of a centred 25th-order moving average. The relationship (S_tA_t)'= Y_t/CT_t' then provides a preliminary estimate of the product of the seasonal and irregular components.
3. For each term of the series (S_tA_t)' a five-term weighted moving average is calculated (with weights 1/9, 2/9, 3/9, 1/9, 2/9) that is "vertical", *i.e.* the datum of each month is "mediated" with the data on the same month of the two previous years and the two following years. In this way a preliminary estimate of the seasonal component is obtained (S_t'), which is later completed through a 13-term weighted moving average. By dividing (S_tA_t)' by the preliminary estimate S_t', the preliminary estimate of the component A_t' is obtained and thus the preliminary estimate of the seasonally adjusted series D_t'= Y_t/S_t'.
4. Once the first phase has been completed, *i.e.* when the preliminary estimates of the various components have been obtained, an "intermediate" estimate of the trend-cycle (CT_t") is calculated by applying to the preliminary seasonally adjusted series D_t' the five-term Henderson moving average with weights -21/286, 84/286, 160/286, 84/286, -21/286, thereby obtaining an intermediate estimate of the series (S_tA_t)" = Y_t/CT_t".
5. Calculation of the final estimate S_t^* of the seasonal component by applying a "vertical" seventh-order moving average to the series (S_tA_t)", with weights 1/15, 2/15, 3/15, 3/15, 3/15, 2/15, 1/15, later adjusted through a 13-term moving average, making it possible to obtain the intermediate seasonally adjusted series D_t" = Y_t/S_t^*.

UNDERSTANDING ECONOMIC STATISTICS – ISBN 978-92-64-03312-2 – © OECD 2008

6. Final calculation of the trend-cycle (CT_t^*) by applying to the intermediate seasonally adjusted series D_t" a Henderson moving average of variable size (9, 13 or 23 terms), selected on the basis of the size of the irregular component in the specific time series (the greater the irregularity, the longer the moving average recommended).
7. Calculation of the final estimate of the irregular component $A_t^* = D_t"/CT_t^*$.
8. Evaluation of the quality of the seasonal adjustment by calculating statistical tests.

Stochastic processes and ARIMA models

As we have seen, although a moving average can approximate an n-grade polynomial, *i.e.* a deterministic function, the time series x_t may also be written as the sum of two components:

$$x_t = f(t) + u_t \qquad [2.25]$$

where f(t) represents a generic function of time and u_t represents a random component. In reality, the separation between the deterministic and the random component, as well as their relative importance, is much less clear than what one might imagine intuitively. In 1921, the English statistician Yule observed that the successive application of moving averages to a series of purely random values produced a result characterised by pseudo-periodic movements, while a few years later the Russian statistician Slutsky observed that as the number of moving averages applied increased, the pseudo-periodic movements tended to assume a clearly sinusoidal form. If then, by applying moving averages, we go from a purely random process to one of a pseudo-deterministic type, then we can also imagine phenomena, such as the trend-cycle or seasonality, as the linear combination of purely random elements.

Let us define the "stochastic process" X_t as a family of random variables arranged according to a parameter t (*i.e.* time) belonging to a parametric set T. In other words, let us define X_t as a process that is evolving over time following probabilistic laws. Let us now assume that, as t varies, the process generates infinite realisations, which are distributed according to the characteristics of the random variables that make up the process. This means that a time series x_t can be imagined as the succession over time of specific realisations of a stochastic process observed at time t = 1, 2, ..., n. The time series x_t therefore represents a sample drawn from the infinite probability distributions that make up the process, and by studying its characteristics we can try to identify the characteristics of these distributions, *i.e.* the characteristics of the stochastic process X_t that generated the observed time series.

While in theory there is an infinite quantity of stochastic processes, in practice the modern analysis of time series is concentrated on stationary processes (in the weak sense), *i.e.* those processes that have a mean, variance and covariance that are not dependent on time t:

Stochastic processes and ARIMA models *(cont.)*

$\mu(t) = E(X_t) = \mu \qquad \sigma^2(t) = E(X_t - \mu)^2 = \sigma^2 \qquad \gamma(t, \tau) = E(X_t - \mu)(X_{t-}\tau - \mu) = \gamma(\tau)$ [2.26]

If a stochastic process ε_t is stationary and:

$\mu(t) = E(\varepsilon_t) = 0 \qquad \sigma^2(t) = E(\varepsilon_t - \mu)^2 = \sigma^2 \qquad \gamma(t, \tau) = E(\varepsilon_t - \mu)(\varepsilon_{t-}\tau - \mu) = 0$ [2.27]

i.e. if the process has zero mean, constant variance and covariance equal to zero, it is said to be a *white noise*.

Any stationary stochastic process can be decomposed into two parts:

$$X_t = D_t + ND_t \qquad [2.28]$$

a deterministic component (D_t); and a non-deterministic component (ND_t). And it can be represented as follows (Wold's decomposition):

$$X_t = D_t + \sum_{i=0}^{\infty} \vartheta_i \varepsilon_{t-i} \qquad [2.29]$$

in which ε_t represents a white noise process, with $\vartheta_0 = 1$ and $\sum_{i=0}^{\infty} \vartheta_i^2 < +\infty$.

Consequently,[2.28] indicates that any stationary stochastic process can be decomposed into a deterministic component and into a weighted moving sum (or average) of white noises in relation to "past" times. If we leave aside the deterministic component, under specific conditions it can be shown that an infinite moving sum (average)

$$\sum_{i=0}^{\infty} \vartheta_i \varepsilon_{t-i}$$

can be approximated by a process of the type $\sum_{i=0}^{\infty} \phi_i X_{t-i}$.

This means that a time series x_t, the finite realisation of a stationary stochastic process X_t , can be represented as follows:

$$\sum_{i=0}^{\infty} \phi_i x_{t-i} = \mu + \sum_{i=0}^{\infty} \vartheta_i \varepsilon_{t-i} \qquad [2.30]$$

i.e.:

$$x_t + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \phi_3 x_{t-3} + \ldots = \mu + \varepsilon_t + \vartheta_1 \varepsilon_{t-1} + \vartheta_2 \varepsilon_{t-2} + \vartheta_3 \varepsilon_{t-3} + \ldots \qquad [2.31]$$

UNDERSTANDING ECONOMIC STATISTICS – ISBN 978-92-64-03312-2 – © OECD 2008

Stochastic processes and ARIMA models *(cont.)*

said to be an "autoregressive moving average" in that it is a combination of two components: the first in which the variable x_t is regressed upon itself lagged in time, and the second composed of an infinite moving sum (average) of white noise processes. By introducing the "lag" operator B, so that $x_t - x_{t-h} = (1 - B^h)\, x_t$, then [2.31] can be written as:

$$x_t + \phi_1 B\, x_t + \phi_2 B^2 x_t + \phi_3 B^3 x_t + \ldots = \mu + \varepsilon_t + \vartheta_1 B\, \varepsilon_t + \vartheta_2 B^2 \varepsilon_t + \vartheta_3 B^3 \varepsilon_t + \ldots \quad [2.32]$$

i.e.

$$(1+ \phi_1 B^1 + \phi_2 B^2 + \phi_3 B^3 + \ldots)\, x_t = \mu + (1+ \vartheta_1 B^1 + \vartheta_2 B^2 + \vartheta_3 B^3 + \ldots)\, \varepsilon_t \quad [2.33]$$

which, in compact form, can be written as:

$$\phi\,(B)\, x_t = \mu + \vartheta\,(B)\, \varepsilon_t \quad [2.34]$$

with ϕ (B) and ϑ (B) polynomials in lag operator B. As we said earlier, all this requires that the time series (as well as the stochastic process from which it was generated) be stationary. When the series is non-stationary in variance, the use of logarithmic transformation (or other similar transformations) makes it possible to overcome this problem. However, when the time series x_t is non-stationary in mean, this stationarity can be obtained by applying differences of the type $\nabla = (x_t - x_{t-1})$. This means that the general form of an ARIMA model is as follows:

$$\phi\,(B)\, \nabla\, x_t = \mu + \vartheta\,(B)\, \varepsilon_t \quad [2.35]$$

Estimating opportunely the parameters ϕ' and ϑ' and resolving the model with regard to the variable x, it is possible to obtain forecasts x_{t+1}, x_{t+2}, ... to lengthen the historical series and then apply the centred moving averages used in the X12-ARIMA procedure.

In conclusion, the decomposition of the time series is conducted by the X12-ARIMA procedure according to the following expression:

$$Y_t = CT_t^* \cdot S_t^* \cdot A_t^* \quad [2.24]$$

i.e. by applying the results obtained in steps 5, 6 and 7.

As was indicated earlier, the application of moving averages is a relatively easy method for approximating rather complex functions. Unfortunately, the calculation of moving averages involves the loss of important data at both ends of the time series, with the data being increasingly numerous the higher the order of the moving average (a 13-term centred moving average, for example, involves the loss of six data at the beginning and six data at the end of a time series). Although the loss of data at the beginning of the series can be a relatively minor drawback, the loss of final data, *i.e.* those that show the most recent trend of the phenomenon being studied, can be particularly problematic, as this is normally the subject that interests economic analysts. Two solutions are most frequently used to overcome this problem: to adopt for the final part of the series asymmetric moving averages, *i.e.* based only on

past data, or to lengthen the time series by forecasting future data. However, both methods present risks of distortion: in fact, the application of asymmetric moving averages tends to produce phase shifts in the estimation of the trend-cycle, with negative effects on the accuracy of the estimate of the seasonally adjusted series. In addition, the quality of the forecasting of future data strongly affects the accuracy of the estimate of the seasonally adjusted series, which may require substantial revisions if there are significant forecasting errors. The X12-ARIMA procedure makes it possible to adopt both approaches, using ARIMA models to forecast the missing data (see box).

On the other hand, the TRAMO-SEATS procedure is based on the following steps:

1. Estimation of the ARIMA model, elimination of outliers and other preliminary processing.
2. Identification of the ARIMA models for the trend-cycle and seasonal variation components, having hypothesised orthogonals between them and with regard to the irregular component.
3. Estimation, through what is known as "canonical decomposition" (which tends to maximise the variance of the irregular component), of the trend-cycle, seasonality and irregularity, and calculation, through appropriate transformations, of the seasonally adjusted series.
4. Evaluation of the quality of the decomposition by calculating statistical tests.

Although they are based on rather different theoretical approaches, the two procedures tend, for many time series, to produce fairly similar results. However, if the time series has special characteristics (high irregularity, changes of level, etc.) the results may differ significantly.

UNDERSTANDING ECONOMIC STATISTICS – ISBN 978-92-64-03312-2 – © OECD 2008