Late Developments in Some Italian Service Surveys: Comparison of Problems and Advantages[1]

by M.R.Ceccarelli[2], R.Gismondi[3], A.P.Mirto[4]

Paper prepared by Italy for the "12th Voorburg Group Meeting" - September 1997
Session 2: *Model Surveys: Presentation of Selected Surveys and Comparison of Response Difficulties and Problems*

1. Tradition and Innovation in the Statistical Surveys of the Year 2000

Prevailing opinion holds that the quality of statistical surveys on enterprises is increasingly dependent on the survey technique utilized rather than on the subject of the survey. The technological level of the units to be observed has for some time set a limit on the degree to which the reference population can be covered. Despite the still scarce presence of information technology devices in households, the share of enterprises possessing basic technologies which can guarantee the transfer of prevalently computerized information is rapidly growing.

On the other hand, the still ample variety of existing situations entails the utilization of a wide array of alternative techniques. The three discussed in this paper refer to important concrete cases and offer valid examples: traditional mailing techniques (in the case of retail businesses, §2), CATI techniques (in the case of medium- and long-distance passenger mobility, §3), and techniques based on data transmission via telecommunications (in the case of tourism supply, §4).

The choice of the surveys is based on both the variety of approaches in capturing data and the current status of the surveys. In this regard, the role played by Eurostat has proven fundamental for at least three reasons:

1. the intense effort to spur the rationalization of pre-existing surveys (the case of domestic trade);
2. the experimental application of modern survey techniques to pre-existing surveys, especially to reduce the burden of response on the enterprises (the case of tourism);
3. the innovative proposal to survey, on a pilot basis, cross cutting phenomena (tourism and transportation) utilizing telephone techniques (passenger mobility).

This summary exposition aims primarily to focus awareness on the real problems underlying the various surveys, in a perspective in which the survey design might not represent the (sole) critical aspect, since there are others of a conceptual nature that are even more serious. In other words, there is no doubt that the utilization of modern survey techniques contributes to reducing the response costs and response times for enterprises, as well as improving, at least to a certain extent, the quality of the data. But, upstream, there must be a clear idea of what is actually to be surveyed and what statistical unit to survey; otherwise, the above-mentioned advantages might be inadequate to guarantee satisfactory results.

In a certain sense, this issue is cross cutting in regard to the three subjects treated in this paper. In section 4 greater emphasis will be placed on the aspects regarding survey techniques. In section 2 the aspects concerning definitions, classifications and computing methods will be stressed. Section 3 will deal with both of these aspects. Finally, section 5 will try to explore more analytically the advantages and disadvantages in the use of one technique or the other, and summary conclusions will be drawn.

To have an idea of the basic penetration of technology in Italian enterprises, we can refer to Table 1.1, which focuses on enterprises with no more than 19 employees. Several estimates based on the preliminary results of the annual survey on their balance sheets for the year 1995 are available for these enterprises. For enterprises with at least 20 employees, we have reason to believe the basic penetration of technology is widespread throughout. On the other hand, the small and medium-size enterprises are precisely those most difficult to reach in the context of a statistical survey characterized by higher response costs.

First, it must be noted that, comprehensively, some 56 out of 100 service enterprises with at least 20 employees possess at least one personal computer, and some 17 out of 100 at least a modem. These shares are higher than the corresponding ones for the industry and construction sectors.

Specifically, the greater presence of modern technologies characterizes financial enterprises as a whole. Over 85% possess at least one personal computer and almost half at least one modem. The high level of integration between industry and the service sector has spurred the strong development of basic technology in services to enterprises. Some three out of four service enterprises possess at least one personal computer and more than one out of five a modem.

In the sphere of services oriented prevalently towards households, the "hotels and restaurants" category is the one most characterized by a lower rate of technological penetration, with 34.2% of sectorial enterprises with at least one personal computer and a scant 6.7% with a modem. In this regard, the pilot study discussed in section 4 focuses on the aspects connected to the statistical representativeness of the structures equipped with this technology compared to the total, and on the problems related to statistical data transmission.

On the other hand, the non-negligible shares characterizing commerce (56.1% and 17.2%, respectively) and transportation (53.4% and 19.9%) must not lead us to hasty conclusions. In fact, they are the result of the aggregation of sectors, such as retail commerce and road cargo transportation, which are very closely linked to traditional organizational schemes (the technological level is only slightly higher than that of hotels and restaurants). And by no accident, these sectors have significant information needs that are now only partially fulfilled.

Overall, the emerging profile confirms that the presence of highly diversified situations will require ISTAT to utilize highly diversified and, eventually mixed, data capture systems for some time to come. The traditional mail survey can be substituted by the telecommunications survey, at least for the share of enterprises able to support it. In this regard, the comparison of the cost and benefits of the various approaches is a priority that cannot be relegated to the whim of the researcher. The same table reveals another indirect indicator of the response costs for enterprises, which are higher the lower the share of enterprises able to keep their own books. This can be assimilated to the share of enterprises which does not have to bear additional response costs because they use outside sources (accountants, sectorial associations, etc.). When sectorial associations are utilized, the additional costs are modest or nil, except for a fixed annual membership fee to belong to the same. In this way, another factor which deters statistical collaboration with the enterprises is highlighted: outsourcing (with the ensuing costs) of the process of filling out the questionnaire, which is very high in the case of hotels and restaurants (66.7% call on an accountant) and retail commerce (almost 60%), and, on the whole, is not negligible for any of the other service activities.

It should be noted that even in the services to enterprises sector a strong technological penetration goes hand in hand with a certain trend to outsource bookkeeping services with accountants (41.7%), given the fragmented nature of these enterprises.

And in this sector the share is the lowest, followed by 45.3% in the transportation enterprises. Finally, there seem to be no significant differences compared to the industry and construction sectors.

Since the response costs and times, as well as the quality of the results, vary with the compiler, we can infer that this variable also is characterized as a strategic determinant with regard to the overall efficiency of a statistical survey system. The application described in section 5 offers an example of how

Table 1.1

Informatic appliances and book-keping procedures in the italian firms - 1995 (*)

ATECO	SETCTOR	The firm has:		Who is the book-keper:		
		Personal computer	Modem	Firm	Accoun--tant	Others
1-2	**Agricolture**	**36,7**	**9,7**	**43,5**	**47,0**	**9,5**
10-45	**Industry and building**	**51,0**	**11,7**	**34,4**	**47,5**	**18,1**
50-93	**Services**	**56,1**	**17,2**	**36,5**	**50,5**	**13,1**
50-52	Trade	57,2	17,9	36,8	47,9	15,4
55	Horeca	34,2	6,7	16,0	66,7	17,3
60-63	Transports	53,4	19,9	38,0	45,3	16,7
64	Posts and telecommunications	50,9	14,0	15,8	63,2	21,1
67	Financial services	86,6	47,3	46,2	47,7	6,1
70	Real estate services	25,4	4,7	25,0	65,0	10,0
71-74	Business services	74,1	23,2	49,4	41,7	9,0
80-93	Other public and personal services	47,2	10,0	27,8	59,0	13,2
1-93	**TOTAL**	**53,5**	**14,5**	**35,5**	**49,1**	**15,4**

Source: Estimates on ISTAT data (1997.1-1997.2).

(*) We refer to enterprises until 19 employees.

the consequences of utilizing one or the other response scheme to a questionnaire can be kept under control.

Finally, if identification of the statistical unit to be observed and creation of a significant typological classification of the statistical units is complicated, then evaluation of the impact of the various survey techniques and of the variability of the entity furnishing enterprise data becomes complicated in its turn. The following section reveals that a possible future modernization of the survey on retail sales may not guarantee, in itself, a tangible qualitative advance in the survey because of still unresolved definition problems.

2. Retail Trade Enterprises: Limits and Validity of the Concept of Specialization

For some time Eurostat has imposed on member states a regulation regarding economic cycle indicators. Although still not ratified, the regulation is to be applied to retail commerce in the field of service enterprises as of January 1998, and entails new common calculation and diffusion schemes of the monthly sales indices (1995 base = 100).

One of the most hotly debated aspects is the uncertainty over the preferred unit of observation. In the choice between enterprise and *kau* there has been a virtually uniform preference for the first, because of the (apparent) greater simplicity of definition and statistical contact with the same. The limits of this choice are confined to points (1) and (2) of section *b* of the draft of the Regulation regarding retail commerce, which states that "*The unit of observation...is the enterprise*", and, "*Furthermore, for enterprises which have a significant (primary or secondary) activity in retail commerce, the national statistical institutes shall also be required to furnish correct results. A report on the corrections performed shall be transmitted to Eurostat*". Exploration of the subject is then termed optional, to be dealt with in a pilot study (Eurostat, 1996).

Since ISTAT has already thoroughly reformed the previous survey on retail sales, both for domestic reasons and for those regarding Eurostat, we are already in a position, albeit at a preliminary and experimental level, to furnish certain indications on the impact of the various approaches on the survey system and the nature of the sales indices calculated and diffused. We must note that in this context we will not examine the enterprises that perform retail sales activities *only* in a form different from the primary, but the multiple sales activities performed by enterprises classified as operating in one of the classes of the 52.2 group of the NACE Rev. 1 classification.

As we know, the following definition is deemed valid for the *kau* (kind of activity unit): "The whole of all the parts of an enterprise which contributes to the realization of an activity according to the four-digit class of the NACE Rev. 1 nomenclature and corresponds to one or more operative subdivisions of the enterprise. The enterprise's information system must be able to indicate or to calculate for each *kau* the value of the production, the intermediate consumptions, the cost of labor, the operating surplus, the employment level and the formation of gross fixed capital". The choice of whether to proceed with an approach by *kau* or by enterprise is marked by this peculiarity:

1. it may not be subordinate to an a priori choice of the data capture technique to adopt;
2. alternatively, it may refer to the modes of data capture (choice between enterprise and *kau* as the survey unit) or to the utilization of the same (index calculation technique).

In regard to the first point, an approach by enterprise may involve the type of survey unit to be utilized, which is easier to identify on the basis of the available registers and easier to contact by interview.

In regard to the second, an approach by enterprise does not eliminate the arbitrary element in the modes of utilizing the data captured. In any case, the questionnaire may be designed to capture at least some of the information (including the value of sales) as a function of the individual sales units that can be assimilated, at least roughly, in the *kau*s. The result is the possibility of calculating the sales indices

on the basis of the enterprise's predominant activity, or of utilizing the same enterprise as many times as the typologies of products in relation to which non-void sales have been registered on the form. The choice depends strictly on the degree of correspondence between the enterprise's principal activities[5] resulting from the available extraction list, and on the degree of precision with which the principal activity fully expresses the actual typological array of products sold. *As a rule, the choice leads to different results.*

Nonetheless, it is hard to imagine being able to effectively divide the individual components of a commercial enterprise to the point of capturing data on sales activity with as many questionnaires as there are *kau*s in the enterprise itself. So, beyond the conceptual difference between the two approaches, they cannot always be distinguished from each other, except at the moment one of the following guidelines of calculating average turnovers or sales variation indices must be chosen:

a) all the sales (including those relative to different products) are attributed to the principal activity which characterizes the enterprise;
b) each enterprise is utilized for the calculations as many times as there are different typologies of products in which at least one sale has been made;
c) only the value of sales relative to the typology of products corresponding to the prevalent activity is utilized.

It is easy to grasp how the conceptual binomial between the approach by enterprise or by *kau* and how the more operative of one or the other of the previous guidelines are closely tied to the known limits of the classification guidelines of commercial enterprise activities generally adopted. The current European NACE Rev. 1 classification, equivalent in Italy to ATECO 91 up to four digits, is not wholly satisfactory for retail commerce for two basic reasons:

1. not infrequently we observe the undesirable mix between form of sale and type of product marketed without an absolute definition regarding the concept of specialization of sales which underlies assignment of a sales point to one category or the other;
2. a basic approach prevails that is more oriented to highlighting the "intrinsic characteristics of the product sold" than the "type of utilization of the product", in other words, the consumption function the product is designed to satisfy, a principle which certainly inspires COICOP (Classification of Individual Consumption by Purpose). We must point out that in Italy, when the new survey on retail sales was being designed, the shift from the previous 19-sector classification to a 15-sector classification was motivated by the need to rationalize the internal composition of the various groups of products and to manage more homogeneous groups of products in function of their satisfying different final needs.

On the basis of a classification that is more functional than typological, the same enterprise, when selling highly diverse types of products, should be counted as many times as the consumption function considered varies. In this regard, Eurostat has been trying for some time to promote among the member states application of the CPA classification (based on the "type of sales service" rather than on the type of product marketed), together with the NACE Rev. 1[6] classification. The results have not been altogether satisfactory, due to the difficulty implicit in such an approach.

In Italy this involves designing a matrix in which the lines bear the CPA classes, which up to four digits coincide with the ATECO matrix, and the columns the ATECO classes, evaluating the weight of the non-diagonal terms of this matrix, which represent comprehensively the degree of "imprecision" of a classification of commercial units based solely on the prevalent sales activity. This type of analysis, which, for simplicity's sake, is based only on the marginal terms of the matrix, without considering the internal terms, will be proposed in the second part of this section (Table 2.1).

This is not a trivial problem, also because it may have a direct impact on the correct calculation technique of a sales index. If n_1 stands for the number of sample units which in two successive phases of analysis (for example, two consecutive months) are operative according to the prevalence guideline

in a certain sales area, and n_0 stands for the units that perform the sales activity in both periods, but not in a prevalent form[7], if r_1 and r_0 indicate, respectively, the linear correlation coefficients relative to the two groups of units mentioned above, then the optimal estimator of the difference between average sales in the two periods will be given by a weighted average of the differences between the averages of the units in each of the two distinct groups, and the weight of the units in the second group will show a trend towards 0 (a case in which it would be useless to also consider the non-prevalent activity units) only if:

$$n_1(1-r_0)+n_0(1-r_1)\cong 0,$$

which is possible only if both correlation coefficients show a trend towards convergence.

In practice, we were able to verify that in the period 1991-1995, and with reference only to specialized enterprises, the result was $r_1 >> r_0$, with average values equal, respectively, to 0.867 and 0.513, for an average of 0.743, and with $n_0 \cong 0{,}1n_1$, by which the exclusion of the units of the second type from the calculations would cause a non-negligible loss of efficiency in the estimates and loss of a datum useful in interpreting the degree of development and the dynamic of the vertical integration in the retail commerce sector.

The rest of this section will attempt to explore this aspect, which is more operative than empirical.

The new sales survey was able to cover 53 ATECO categories, corresponding to 17 NACE Rev. 1 classes. In each monthly survey form the enterprises were asked to subdivide the monthly value of sales into 15 groups of products corresponding to particular aggregations of elementary ATECO categories.

Table 2.1 reports, for the specialized enterprises of the sample classified as operating prevalently in each elementary economic category observed, the number of groups (including the one relative to the prevalent activity) in correspondence to which a sale was made in the reference period (in this case the entire year 1996[8]), as well as the percentage of enterprises with sales only corresponding to the group of products relative to the prevalent activity and the position of the individual categories according to a progressive ranking of the previous percentage (in italics and in parentheses). The key to the groups is given in Table 2.3. Certain ineligible enterprises were excluded, for example, those that sell retail but prevalently operate wholesale or in the manufacturing industry. The right side of the table reports analogous data, but referring to the classes of the NACE Rev. 1 classification.

We must first point out that, on the average, some 83 enterprises out of 100 do not sell products classifiable in groups different from the one corresponding to the principal category, and that when this occurs the number of groups different from the main one, in correspondence to which a sales activity is performed, is equal to 3.8.

In particular, the first 14 categories with the lowest shares of enterprises operating in correspondence to the sole prevalent activity are all superior to the NACE 52.43 category. The first five of these categories are, in order, "Lighting articles and electrical material" (38.2%), "Sewing and knitting machines" (54.5%), "Hygiene-health articles" (55.6%), "Thermohydraulic materials" (56.3%) and "Other products" (56.8%).

Overall, we identified 18 categories containing, on the basis of the prevalence guideline, a share equal to not less than 20% of the total enterprises in the category of enterprises also operating in correspondence to categories different from the principal one. The probability of finding enterprises operating in highly diversified sales categories is quite elevated for such categories as "Household articles, glassware and kitchenware" and "Household appliances", into which fall enterprises selling products classifiable in 14 of the 15 different groups mentioned in the survey form.

On the other hand, among the categories best identified by the prevalence guideline are those relative to foods and beverages. Except for "Pastries and sweets" (77.8%) and "Coffee" (81.8%), these

Table 2.1

The new retail trade sales survey: specialised enterprises classified per ATECO and per NACE Rev.1 - Year 1996

ATECO	CATEGORY	Number of groups, % share of firms in one group only and decreasing ranking			NACE Rev.1	CLASS	Number of groups, % share of firms in one group only and decreasing ranking		
52210	Fruit and vegetables	3	95,3	*(40)*	5221	Fruit and vegetables	3	95,4	*(13)*
52221	Cattle, pork, ovine, horse meat	2	96,3	*(42)*	5222	Meat	2	96,6	*(15)*
52222	Poultry, game	1	100,0	*(46)*					
52230	Fish	1	100,0	*(46)*	5223	Fishes	1	100,0	*(17)*
52241	Bread	2	97,6	*(45)*	5224	Bread and cakes	7	87,2	*(7)*
52242	Cakes	7	77,8	*(18)*					
52250	Beverages	2	97,2	*(44)*	5225	Beverages	2	97,2	*(16)*
52271	Milk	1	100,0	*(46)*	5227	Other food products	6	93,8	*(11)*
52272	Groceries	2	93,5	*(36)*					
52273	Coffee	5	81,8	*(19)*					
52274	Other food and beverages	3	94,9	*(39)*					
52310	Pharmaceutical products	4	90,3	*(29)*	5231	Pharmaceutical products	4	90,3	*(9)*
52320	Otrhopaedic appliances	3	84,6	*(22)*	5232	Otrhopaedic appliances	3	84,6	*(6)*
52331	Herbalists	3	85,4	*(23)*	5233	Cosmetics	9	89,7	*(8)*
52332	Perfumeries, cosmetics and personal goods	9	90,8	*(30)*					
52411	Textiles for dressing	1	100,0	*(46)*	5241	Textiles	3	94,2	*(12)*
52412	Textiles for the house	2	97,1	*(43)*					
52413	Linen	2	75,0	*(15)*					
52414	Hosieries and knitwears	2	75,0	*(16)*					
52421	Adults clothes	5	91,5	*(32)*	5242	Clothes and shoes	6	93,6	*(10)*
52422	Children and babies clothes	1	100,0	*(46)*					
52423	Lingeries, shirts	1	100,0	*(46)*					
52424	Haberdasheries	4	89,5	*(28)*					
52425	Fur-coats	1	100,0	*(46)*					
52426	Hats, umbrellas, gloves, ties	1	100,0	*(46)*		SHOES AND LEATHER ARTICLES			
52431	Shoes	5	96,2	*(41)*	5243	~~Calzature ed articoli in cuoio~~	5	95,4	*(13)*
52432	Travel articles	3	91,7	*(34)*					
52441	Furnitures	4	93,7	*(37)*	5244	~~Mobili, articoli d'illuminazione~~	14	80,5	*(3)*
52442	Glasswares and crockeries	14	83,3	*(21)*		FURNITURES AND LIGHTING ARTICLES			
52443	Lighting articles and electric materials	9	38,2	*(1)*					
52444	Wood, wicker, plastic articles	3	66,7	*(11)*					
52445	Various articles for domestic use	13	64,7	*(8)*					
52451	Electrical household appliances	14	65,1	*(9)*	5245	Electrical household appliances	14	71,0	*(1)*
52452	Radio, tv, personal computers	12	72,3	*(13)*					
52453	Records and tapes	5	89,2	*(27)*					
52454	Musical instruments	4	92,1	*(35)*					
52455	Sewing machines	4	54,5	*(2)*					
52461	Ironmonger's shops	6	91,5	*(33)*	5246	Ironmonger's shops	11	71,8	*(2)*
52462	Paints	5	87,9	*(26)*					
52463	Hygienic and health articles	10	55,6	*(3)*					
52464	Building appliances	4	76,7	*(17)*					
52465	Hydraulic appliances	5	56,3	*(4)*					
52466	Agricolture machines	4	62,7	*(7)*					
52471	New books	3	86,7	*(25)*	5247	Stationery, books, newspapers	8	82,9	*(5)*
52472	Nespapers, magazines	8	72,4	*(14)*					
52473	Stationery and office stuffs	7	86,3	*(24)*					
52481	Office appliances	5	82,6	*(20)*	5248	Other products	12	81,9	*(4)*
52482	Opticd, photography, precision instruments	7	94,8	*(38)*					
52483	Watches, silver articles and jewelry	7	91,4	*(31)*					
52484	Toys and games	11	61,2	*(6)*					
52485	Sport, bykes, weapons	11	66,7	*(12)*					
52486	Art and cult	6	65,8	*(10)*					
52489	Other products	7	56,8	*(5)*					
	Average	4,8	83,1						

Source: elaborations on ISTAT data.

Thel 15 groups of products have been specified in the table 2.3.

categories show shares of enterprises operating solely in the prevalent category oscillating between 95% and 100% of the total.

Finally, the most summary classification based on NACE Rev. 1 shows even more clearly how the prevalence guideline is less effective for all the classes from 52.44 to 52.48, as well as for 52.32, relative to medical and orthopedic articles and, to a lesser extent, 52.24, relative to "Bread, pastries and sweets". For all the other classes the guideline can be considered quite satisfactory.

A first summary conclusion reveals that, in reference to a particular subset of the retail distribution sector photographed by the monthly sales survey, for every five enterprises there is, on the average, one that sells in correspondence to a group of products different from the principal one. So, in defining the sample size it might suffice to choose five-sixths of the units effectively necessary to satisfy the pre-established quality limits of the survey, at least in terms of sampling error.

Table 2.2
Retail trade sale survey: not specialised firms and outlets per ATECO and kind of product sold - Year 1996

ATECO	Type of sale	OUTLETS BELONGING TO NOT SPECIALISED FIRMS			SHARE OF SALE		
		Food	Not food	Total	Food	Not food	Total
52111	Hypermarkets	78,0	22,0	100,0	67,2	32,8	100,0
52112	Supermarkets	53,1	46,9	100,0	93,8	6,2	100,0
52113	Minimarkets	81,6	18,4	100,0	96,3	3,7	100,0
52114	Other food shops	90,3	9,7	100,0	69,1	30,9	100,0
52121	Department stores	4,9	95,1	100,0	26,1	73,9	100,0
52122	Other not specialised stores	27,1	72,9	100,0	57,3	42,7	100,0
5211	Not specialised with a prevalence of food sales	66,5	33,5	100,0	87,9	12,1	100,0
5212	Not specialised with a prevalence of not food sales	18,0	82,0	100,0	38,0	62,0	100,0
521	Not specialised	63,1	36,9	100,0	86,4	13,6	100,0

Source: elaborations on ISTAT data.

On the other hand, the occurrence of sales classifiable in different groups always entails a careful preliminary verification of the precision of the classification derived from the list from which the sample units were drawn. This type of verification may also be necessary even for the single-product enterprises which list on the survey form possible classification errors due to failure to update the register.

For simplicity's sake, in the previous table the relative composition of the value of sales by product typology of goods sold was not studied, as, vice versa, is done in Table 2.2, which refers only to the unspecialized enterprises inserted into the sales survey sample.

Of 100 sales points belonging to non-specialized enterprises observed, on the average some 63 sell only food products and the remaining 37 only non-food products. These shares correspond to a retail spending portfolio of which 86.4% is spent on food products and the remaining 13.6% on non-food

Table 2.3

Retail trade sales indexes for specialised enterprises calculated with two different criteria - Average 1996

Cod.	GROUP	(1) Share of firms with sales in one group only	(2) Sales index including firms (1) only	(3) Sales index including all firms	(4)=(3)-(2) Difference
1	Food and beverages	94,8	104,4	106,9	+2,5
2	Pharmaceutical products	89,8	104,1	105,3	+1,2
3	Clothes and shoes	93,2	108,4	108,0	-0,4
4	Leather products	95,4	113,9	113,7	-0,2
5	Furnitures and textiles	84,9	105,7	109,5	+3,7
6	Electrical household appliances	65,1	105,2	105,5	+0,4
7	Radio, tv, personal computers	74,3	93,6	98,6	+5,0
8	Photographic goods	94,8	101,6	102,1	+0,5
9	Small products for the house	80,2	89,4	93,3	+3,9
10	Ironmonger's shops	79,0	106,4	107,4	+1,0
11	Cosmetics and personal goods	89,7	102,8	101,5	-1,3
12	Stationery, books, newspapers	82,9	118,7	118,0	-0,6
13	Cd, tapes, musical instruments	90,0	97,8	97,1	-0,7
14	Toys, Games, Sport, Camping	64,8	108,5	110,6	+2,1
15	Other products	82,4	104,8	108,1	+3,3
E1	Food and beverages	94,8	104,4	106,9	+2,5
E2	Pharmaceutical and cosmetic products	89,7	103,5	103,4	+0,0
E3	Clothes and shoes	94,3	111,3	111,0	-0,3
E4	Furnitures and articles for the house	72,2	101,8	104,4	+2,6
E5	Other products	81,2	106,7	108,7	+1,9
	TOTAL	83,0	104,2	105,8	+1,6

Source: elaborations on ISTAT data.
The indexes have the base 1995=100.

products, a fact which testifies to the persistent lag of big distribution in Italy in the market of non-primary goods.

The presence of conspicuous classification problems is also revealed by another fact: approximately one-third of the sales points belonging to enterprises operating on the basis of the classification of the original list in the form of "non-specialized enterprises selling prevalently food products" sell non-food products only. This incongruence can be found, in a more reduced form, in the sales points of enterprises which sell prevalently non-food products (18.0% sell only food products), even though the previous shares referring to the number of sales points is reversed if we consider the value of sales. For the prevalently food and non-food enterprises the figures are 12.1% for non-food sales and 38.0% for food sales, respectively.

In greater detail, the department stores seem to be the non-specialized distribution typology best identified by the current classification (only 4.9% of the relative sales points sell food products, but these products account for 26.1% of sales). However, the management structure of supermarkets seems to be heavily concentrated in enterprises operating in both the food and non-food sectors, to the point that the two typologies of sales points are equivalent.

To avoid inexact conclusions on the precision of the basic classifications, in this case as well the composition of the sales value plays the most significant role. Consequently:

- approximately one-third of the receipts of hypermarkets come from non-food products;
- almost all the sales in supermarkets and in minimarkets are in the form of food products, in full conformity with the basic classification;
- less than one-third of the sales in department stores are in the form of food products;
- the structure of "Other food enterprises" and "Other non-specialized enterprises" seems, at least to a certain extent, ambiguous. For the latter group, in particular, a specialized ad hoc classification would seem more appropriate, also because it is difficult to imagine large unspecialized sales areas that do not take the form of a hypermarket, supermarket, minimarket, department store or other food enterprise (for example, a discount store).

Aside from the previous considerations, Table 2.3 allows us to infer the effective impact of the prevalent activity classification guideline on the calculation of the sales indices.

If we leave aside the original NACE classes and consider directly the 15 national groups and the five established for Eurostat in accordance with the economic cycle Regulation (excluding non-specialized commerce and mail-order sales), it is evident that the four national groups with a share of enterprises with sales only in correspondence to the pertinent group below 80% are (column 1): "Games, toys, sports, camping articles" (64.8%), "Household appliances" (65.1%), "Radio, television, recorders, information technology products" (74.3%), and "Household utensils and hardware" (79.0%). With reference to the Eurostat classification, the foregoing circumstance occurs only for "Furnishings and household articles" (72.2%), while the product groups most clearly specified by prevalent activity are: "Footwear and leather goods" (95.4%) and "Photo-optical products and film" (94.8%) for the national grouping, and "Food and beverages" (also 94.8%).

The most significant part of the table is certainly the comparison between the sales indices calculated considering only the enterprises definable as "single-product" (column 2) and those with sales in at least two distinct groups (column 3). These reflect the type c) and b) calculation approaches mentioned in the first part of this section.

In this regard, we must clarify that the calculation of the second type of index examines the same enterprise as many times as there are different product groups in which a sale has been made in the reference period. For example, if there were only two enterprises, of which one sold exclusively food products and the other food products and household articles, the index of group 9 (household products) would utilize the sales value of household products of the second enterprise, while the index of group 1 (food and beverages) would utilize the *values of food sales of both enterprises.*

It is clear then, that the second approach is actually closer to the approach of a survey in which the survey unit is the *kau* and not the enterprise, as proposed by the preliminary drafts of the Eurostat Regulation. While we are far from an approach strictly similar to the one by *kau*, we can clearly conceive that every retail commercial enterprise can be composed of as many entities similar to *kau*s as there are distinct product groups in which it offers a sales service. The real problem lies in the fact that a classification based only on the prevalent activity of the enterprise has serious limitations, because it does not allow the information at the level of the unit of economic activity to be processed correctly. The unit of economic activity should be the true unit of analysis, rather than the enterprise, even with all the conceptual and operative problems such a choice entails.

The results reveal that, overall, the second index is higher than the first by 1.6 points. This indicates that the sales points which diversify their offer by adopting a sales strategy on multiple groups of products are able to distinguish themselves by a dynamic trend that is effectively better than the trend of single-product enterprises.

This regularity marks the majority of national groups (10 out of 15) and the Eurostat groups (4 out of 5). Excluded are the sales activities related to products for which the diversification of the array offered seems to serve the purpose of directing the clients' attention more towards the purchase of effectively prevalent goods rather than towards guaranteeing a comprehensively better sales performance. This evidence is typical in clothing, footwear, stationery, music stores (in the broad sense) and perfumeries. In these shops there is an often vital need to offer products collateral to the prevalent ones to raise the overall sales volume but not necessarily the longitudinal dynamic of these enterprises.

Since ISTAT currently calculates indices of the second type[9], which objectively seem more compatible with consumers' requests, it becomes necessary, consequently, to verify their methodological compatibility with the calculations of the other countries. It is also necessary to follow its comparative dynamic with those of the first type, in order to maintain control over at least the parallelism in the longitudinal trends of the two historical series.

- The studies we have dealt with in a preliminary way in this report lead to two considerations:
- the approach by enterprise or by *kau* does not actually concern the unit effectively surveyed, which in any case is the enterprise, since the *kau* cannot be tangibly identified, but the type of information requested of the enterprise. In turn, the information derives strictly from the type of utilization to be made of the data gathered. In other words, the choice of the unit of analysis does not seem to derive so much from an a priori reasoning, but from a process based, in order, on the nature of the information to be diffused, on the technique of generating meta-data and on the classifications adopted to satisfy the requirements of consumers, which may at least in part diverge from the "official" requirements.
- To ask the specialized enterprises to itemize the sum of monthly sales by type of product sold may, in certain instances, represent an additional burden for the statistical institutes and for the enterprises themselves. When this occurs, it might be necessary to revise the classifications or modify the concept of "specialized enterprise" (this aspect cfr. Puglisi, 1995).

The issues raised here undoubtedly warrant more specific study in Eurostat, as well. For an economic cycle Regulation of such scope ad hoc meetings should explore in depth the problems regarding classification of enterprises and calculation of the indices, while freeing themselves from the constraints set by the Structural Regulation, regarding which, for the time being, the question of longitudinal studies is less relevant.

3. The Pilot Survey on Long-Distance Mobility[10]

Economic integration has wrought immense changes in the traffic and transport system in Europe. Abolition of the frontiers within the European Union has led to an increased flow of transportation. In turn, this has created greater bottlenecks in existing transport structures, which have proven inadequate to support the new mobility requirements. The increased demand for transportation infrastructures has led the Community to launch a series of initiatives to develop the information system of European transportation, which was seriously depleted by the elimination of data capture at customs points. These initiatives are vital for any type of choice on questions of transportation policy, environmental policy, territorial development and management of energy resources.

The massive growth of long-distance and international transportation has augmented, first, the request for statistical information on passenger mobility. Above all, there has been expanded interest in the variables, still unavailable, regarding the passenger profile, the reason for travelling, the "chain" of transportation means utilized, the reasons underlying the choice of the means utilized, the O/D (origin/destination) matrix (on the regional NUTS II basis), etc.

To contend with these new requirements and to develop methods and a classification which will permit comparable data to be gathered on the European level, the European Union has requested that its member states implement pilot surveys on long-distance (over 100 km) mobility. The aim is to perfect a methodology valid for surveys of broader dimensions. The countries that have joined this initiative, which is coordinated by Eurostat and the DG VII of the Commission are: Austria, France, Germany, Italy, Portugal, Spain and Sweden.

In the sphere of this project ISTAT has developed two pilot surveys, which were conducted parallel to each other. One was conducted via mail, the other via CATI (Computer Assisted Telephone Interviewing). The two surveys aimed to test the survey methodology, not only as regards the typical survey phases (sample design, questionnaire), but especially to evaluate, for such a complex survey, the most suitable technique to adopt in administering the questionnaire.

The survey proved to be especially ambitious. Besides capturing the O/D matrix of the comprehensive trip, it sought to reconstruct the itinerary followed by the interviewee, dividing it into legs. For each leg the interviewee was asked which means of transportation was/were utilized, the purpose of each leg, the number of people with whom he travelled and, a very difficult fact to recall, the day and hour of departure and arrival.

The questionnaire was divided into four sections:

1. information on the household (composition of the household, working status, automobiles utilized by the household);
2. information on the respondent (sex, marital status, educational level, region in which he was employed in the previous month, ownership of subscription to public transportation and of driver's license);
3. information on the trips (trips made in the month of April, O/D, prevalent means of transportation, principal reason for travelling);
4. information on the individual legs made during the trip (place, date and hour of departure, place, date and hour of arrival, purpose of the leg, means of transportation).

The two surveys, conducted in the month of May, were structured according to the methodology summarized in the following table:

Table 3.1 - Comparison between two survey techniques

Analysed character	Way of submission of the questionnaire	
	Telephonic interview with the aid of computer (CATI)	Mail interview, using the Electronic Post System (POSTEL)
• Universe	• Residents.	• Residents.

• Population from which the sample has been drawn	• Subscribers found in the telephonic list SEAT (not commercial users).	• Subscribers found in the telephonic list SEAT (not commercial users).
• Unit or relevation	• Who belongs to the household with at least 18 years, better if with one journey at least.	• Who belongs to the household with at least 18 years, better if with one journey at least.
• Size of sample	• 7.000 households for obtaining 2.000 valid interviews.	• 300 households.
• Sampling design	• Random regional stratification (20 regions).	• Random selection of 300 households equally shared among three towns: Milano, Roma, Napoli.
• Survey presentation	• Lecter of presentation, signed by the President, sent 15 days before the telephonic contact.	• Lecter of presentation, signed by the President, joined to the questionnaire
• *Auditing*	• *Random* control of the interviewres' activity.	• Telephonic reply and telephonic reiterview.
• Period of reference	• The month before the week of interviews (april).	• The month before receiving the questionnaire (april).

Source: ISTAT - Pilot survey on long distance mobility - 1997.

Of the two, the survey which yielded the best results, in terms of quality and time, was clearly the CATI telephone survey. Although the costs seemed higher at the outset, in the end they did not differ significantly from the mail survey. In fact, we must take into consideration that, because of an almost total drop off in responses to the mail survey, we had to conduct telephone solicitations, reinterviewing the subjects at their disposal. In addition, we must also consider the costs of the operations implicitly successive to sending the mail questionnaire: data control and editing, insertion of the data on information supports and subsequent control of the imputed data.

We can now analyze the two surveys in greater detail. As we mentioned above, the interviews for the telephone survey were performed with the aid of a CATI system. To capture the information we had to transpose the paper questionnaire onto an information support. Although this operation involved five days of work by Metron R&C (the company which contracted the CATI interviews), it resulted in better quality telephone interviews compared to the mail interviews. The questionnaire on the information support automatically managed:

• skipping and moving back and forth from one question to another;
• exact keying in of the data (acceptance only of codes provided by the question);
• verification of compatibility in the information provided by the interviewee.

Besides guaranteeing a flexible and dynamic management of the questionnaire (moving back and forth in the questionnaire, development of the trip in legs, etc.), the CATI methodology proved invaluable in following the chronological compatibility of the development of the trip in the individual legs composing it.

To carry out the telephone interviews a staff of 20 female telephone interviewers and one supervisor was assembled. After the first day of briefing, the interviewers conducted the interviews over ten days, from Monday to Friday from 5:00 pm to 9:00 pm, and Saturday from 9:00 am to 6:00 pm. In addition, the ISTAT project head and his delegate performed a constant auditing of the work of the interviewers by making random checks.

On the whole, administration of the questionnaire went smoothly, also thanks to the strong interaction that was often created between interviewer and interviewee. In numerous cases the interview took place serenely, and the trip was able to be reconstructed in detail.

The section involving the introduction and screening was simple and fast, allowing the interviewee to get immediately to the subject of the survey through a series of data that was easy to capture and in no way confidential (habits tied to mobility, ownership of means of transportation, etc.).

The only problems encountered during the interviews regarded the section on the trips. There was no difficulty in recalling the events to be captured, but problems did arise in reconstructing the trip in "legs". The definition of "leg" established by Eurostat[11] was too complex and difficult to comprehend by the interviewee. It proved possible to subdivide the trip correctly, especially in the most fragmented situations, thanks only to the sensitivity and savvy of the interviewer. In fact, in a number of cases, the interviewer sensed the need to explore the dynamics of the trip, asking for clarifications and explanations on the itinerary which could not be inferred from the questions already posed. This is a highly unusual situation for a telephone capture technique, since the operator normally has the sole task of transcribing the answers received, without influencing in any way the interviewee, in order to avoid biased answers. This aspect entailed careful training of the interviewer and an in-depth control of the interviews in the room in which they were being conducted. In fact, several interviews performed during the first days were incorrect in content and had to be substituted.

To perform 2,000 interviews, 4,507 of the 7,000 households available (some 65%) were utilized. Overall, 6,245 telephone contacts (1.4 for each household) were made. Participation in the survey was good and in line with that observed in other telephone surveys of the population. There was a total of 1,048 refusals, in other words, 23.3% of the households contacted. From the information gathered on the refusal form, we observe that the individuals who did not participate in the survey declined because they felt they had been disturbed, called at an inopportune moment, or because they declared they had no time to devote to an interview (27% gave this last reason). Others declared they had no interest in the subject of the survey (18%), while a good 16% of the refusals were due to a skeptical or suspicious attitude to the caller. Finally, we noted there is less collaboration during the telephone contact when a woman, particularly an older woman, responds.

On the average, the interview lasted 12 minutes. Just under 10 minutes for the individual who made no trips, and approximately 16 minutes for the individual who took at least one trip with several legs. No interview took more than 30 minutes to complete. The costs of the telephone survey averaged approximately Lire 15,000[12] per interview, including sending the letter of introduction to the entire sample, set up of the electronic questionnaire and supply of the file with the results, previously controlled on magnetic support.

With regard to the mail survey, as we already noted, the response rate was approximately 16%. Although the questionnaire was sent via the "Postel" (electronic mail) system, the time it took to send and return the questionnaire was rather long, on the average around 45 days. If we exclude the formulation of the paper questionnaire and the design and drawing of the sample, it took approximately 70 days to perform the mail interview (15 days to transpose the questionnaire into electronic mail graphics, 45 to send and return the questionnaire, 10 for re-interviews, keying in data and other operative phases).

Because of the low level of participation, telephone solicitations had to be made to verify whether the form had been received and the possible willingness to grant the interview via telephone. The majority of people contacted declared they did not respond to the mail questionnaire for three basic reasons:

1. lack of time;
2. questionnaire too complicated;
3. lack of interest in the subject.

Some 65% of the people reinterviewed declared their willingness to give the interview via telephone; 27% said they were uninterested or intruded upon; and 8% said they had no time.

If we consider the results, the time involved and the quality of the data obtained, the sole advantage of the mail interview was the relatively low cost (approximately Lire 4,700 per interview[13]). To these

costs we must add the cost of personnel for solicitations and telephone re-interviews, data control and editing, insertion of the data on information support, and the subsequent control of the imputed data (approximately Lire 6,700 per interview[14]).

If we compare the two surveys we can summarize the principal differences as follows:

Table 3.2 - Comparison of advantages and disadvantages

Analysed character	Way of submission of the questionnaire	
	Telephonic interview with the aid of computer (CATI)	Mail interview, using the Electronic Post System (POSTEL)
• Saving of time and costs	• It avoids the needs of mail replies and the informatic transfer of gathered data, it reduces the *ex-post* quality controls and data editing.	• Saving of costs of about 3.600 lire per interview.
• Data quality ad reliability	• More controls (incompatibility rules, a more complex electronic questionnaire), increasing quality and reliability.	• Higher flexibility in terms of questionnaire design and a better overview of the questionnaire.

Source: ISTAT - Pilot survey on long distance mobility - 1997.

In Table 3.2 a comparison of the two techniques shows that the telephone survey offers greater guarantees of quality and speed. The costs savings gained with the mail interview are offset by the advantages in time saved and better quality and reliability of data obtained with the telephone survey.

We must observe that in Italy, unlike in other European countries (especially the northern European countries), the results of the telephone surveys are not heavily biased by the respondents' use of answering machines, which automatically reveal the telephone number of the caller, or by an elevated percentage of unlisted numbers. In fact, Italians have shown they are favorably disposed to respond to telephone interviews, while they tend not to take into consideration questionnaires received in the mail, especially when the response is not compulsory.

Another method that could respond well to this type of survey is undoubtedly the direct interview, especially if aided by a computer (CAPI - Computer Assisted Personal Interviewing). In this case, however, the costs of the survey are considerably higher and make it impossible to utilize this technique for surveys conducted over the entire national territory and characterized by a numerous sample.

Again in the sphere of computer-aided techniques, an experiment in utilizing CASI (Computer Assisted Self-Interviewing) was designed by the Italian State Railways, also to survey passenger mobility.

The survey conducted on a panel of 1,800 households (4,112 individuals), with a programmed 15% rotation rate (with an average duration of 2.5 years on the panel), requires the households to insert data on trips longer than 10 km made by each household member outside the municipality of residence. The data are inserted on a portable computer made available to the household. The data are automatically downloaded via modem on a weekly basis to the central computer, to which the portables are connected. While this survey technique presents undoubted advantages, its very elevated costs do not allow widespread utilization.

4. Innovative Aspects in Statistical Surveys on Tourism Supply

ISTAT's surveys on tourism supply were radically overhauled in 1997. The main reasons for the change were:

1. to design an information capture system better able to satisfy the information needs of consumers;
2. to improve the quality and velocity of statistics on tourism enterprises.

To achieve the first objective ISTAT had not only to identify the demand for information by consumers of sectorial statistics, but also to study the administrative and statistical sources, the demands of Italian legal and fiscal regulations and of the enterprise information system, the latter in compliance with Community Regulations and Directives. ISTAT completed this phase of study and analysis performing market research designed to analyze the principal accounting and data management methods of tourism enterprises (hotels and travel agencies).

Improvement in the quality and timeliness of tourism statistics gave rise to a new capture system, implemented by designing and creating a telecommunications network to transmit enterprise data, now sent via an electronic questionnaire instead of the paper one. In particular, the goal of the network was to connect the public sector (ISTAT) and the private one (the tourism enterprises) utilizing a common EDI (Electronic Data Interchange) transmission standard.

In this perspective several procedures were developed to control and return the information to the enterprises. An electronic questionnaire was created, along with a data-set of enterprise information and software to manage and transmit this information.

To study the demand for information, a distinction was made between national demand (enterprises, sectorial associations, policymakers) and international demand (Eurostat and its Regulations and Directives, and other organisms such as the OECD and the WTO) based on their relative consumers.

In order to analyze the requirements of national consumers of sectorial statistics, ISTAT undertook a series of initiatives to establish collaborative ties with sectorial associations and public organisms competent in tourism affairs (ENIT - Italian State Tourist Office; the Tourism Department of the Ministry of Industry; the various APTs - Tourism Promotion Agencies; and the EPTs -the various Provincial Tourism Departments). These contacts led us to make the following considerations:

- the structural characteristics of the enterprises and of the local units (dimensions, economic activity/-ies exercised, legal form, period of activity, number of stars, location, etc.) are given very serious attention but are not of primary importance;
- the economic characteristics (total turnover, acquisition of goods and services, employment) are deemed to be of equal importance to the structural characteristics since they are required by Community Regulations and Directives and allow, at the same time, the enterprise to evaluate its competitive position vis-à-vis others in the sector;
- the characteristics of marketing strategies (sectorial potential, utilization of training techniques, promotional activities, etc.) are of prime importance to the consumer and related to those of the first two groups.

At the international level, the Regulations and Directives lead to standardizing information gathered on the national level in terms of definitions, classifications, field of observation, units of observation and timing.

Study of the information demand was followed by study of the supply. Analysis of the administrative sources and, in particular, the fiscal and legal norms and the enterprise information system allowed us to verify whether the aggregates which potentially can be surveyed in the system of tourism enterprises are able to satisfy the information needs of sectorial users. The current and potential supply of information consists of:

1. statistical sources: surveys conducted by ISTAT, by other public organisms, by private research bodies and by sectorial associations;
2. administrative sources: primarily data available from the public organisms belonging to the National Statistical System (Sistan);

3. enterprise information system: composed of the general accounting (balance sheets), industrial accounting (referring to the local units), and service accounting (typical of hotel enterprises) which, in addition to satisfying administrative requirements, allows the enterprise to monitor and set the policy for its own management.

In regard to the second objective, upstream of the design of a telecommunications network to transmit data, a feasibility study was carried out on the basis of the results of the "Pilot Survey on Hotels and Travel Agencies - Year 1995". Besides investigating the general, industrial and service accounting, the accounting system and the type of VAT liquidation for hotel enterprises and travel agencies, the feasibility study also analyzed the degree of information technology the businesses possessed. This survey brought to light the following aspects:

- the high percentage of tourism enterprises using an ordinary accounting system (approximately 90%) should permit us to capture the data required for national accounting purposes (SEC95) even for the small and medium enterprises.
- The low percentage of hotel enterprises that keeps its own books either in-house or by utilizing sectorial associations could prove an obstacle to direct data capturing from the enterprise. This problem is less evident in the case of travel agencies, since 57% of these agencies manage their own books in-house (cfr. Table 4.1).
- Approximately half the hotel enterprises manage their own books and, consequently, valuable information on the performances of the individual local units could be captured.
- Some 56% of hotel enterprises are equipped with computer technology, as are 95% of all travel agencies.
- The degree of information technology possessed by hotel enterprises varies according to the size of the hotel, and a wide array of software is used.

Tabella 4.1 - Hotels and travel agencies by accounting management and classes of employees - % shares

	TRAVEL AGENCIES				HOTEL			
ACCOUNTING MANAGEMENT	EMPLOYMENT				EMPLOYMENT			
	1-5	6-19	>20	Total	1-5	6-19	>20	Total
enterprise itself	40,00	80,00	93,33	58,46	9,39	15,31	50,94	18,46
accountants	50,66	15,00	6,67	34,62	75,81	69,05	44,34	67,95
enterprise associations	2,67	2,50	0,00	2,31	12,27	13,27	1,89	11,08
CAAF	0,00	0,00	0,00	0,00	0,36	0,00	0,00	0,15
others	6,67	2,50	0,00	4,61	2,17	2,37	2,83	2,36
TOTAL	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Source: “Pilot survey on hotels and travel agencies - Year 1995”. Eurostat project, April 1997.

Table 4.2 - Hotels and travel agencies per used softwares and class of employees - % shares

	TRAVEL AGENCIES				HOTEL			
ACCOUNTING SOFTWARES	EMPLOYMENT				EMPLOYMENT			
	1-5	6-19	>20	Total	1-5	6-19	>20	Total
market leaders *	94,37	98,04	100,00	96,45	14,28	18,74	25,00	20,00
self-developed	0,00	0,00	0,00	0,00	42,86	34,38	28,41	33,96
others	5,63	1,96	0,00	3,55	42,86	46,88	46,59	46,04
TOTAL	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

* Hotel software market leaders: CDS, Sysdata, Fidelio, Nyce, Unif. system of account.

Travel agencies automated bookings market leaders: Amadeus Worldspan Sabre, Sigma.
Source: "Pilot survey on hotels and travel agencies. Year 1995". Eurostat project, April 1997.

- Travel agencies employ a number of telephone reservation systems that are generally widespread throughout the entire sector.
- The utilization of the modem and of e-mail is still rather circumscribed (37% and 17%, respectively), especially by hotel enterprises.

In light of the results of the previous survey a telecommunications network was designed to link a limited sample of hotels with ISTAT by means of the corresponding regions. The objective of this experiment was to test the transmission of both economic and structural data through a telecommunications link of the different information junctions.

The network operated in accordance with the following mechanism: an initial telecommunications link was set up between the individual hotels and their region. This connection served to capture disaggregated information from the hotels which the regional administration otherwise had no possibility to access. A second link, between the region and ISTAT, allowed all the information of interest to be captured and to serve as a first elementary data base. BBS ISTAT is the capture point for this information.

The individual data are subsequently aggregated and made available for consultation, through BBS, to the individual hotel structures interested (return of information to the enterprise). A more disaggregated data-set is defined for the Sistan administrations and made available through BBS.

The variables object of the electronic questionnaire regard, first, the arrival and presence of tourists, distinguished by nationality (according to the scheme of the paper form currently used for the statistical survey on the movement of clients), and, second, a number of relevant economic variables (total turnover, purchase of goods and services, average prices charged, cost of labor, number of employees and self-employed workers). The first set of variables is transmitted on a daily or, at most, a weekly basis. The economic variables are transmitted monthly.

Thus, the hotels fill out an electronic questionnaire that captures information first filed in the individual hotel structures by the software simulating the questionnaire. Subsequently, a transmission software automatically sends, at a pre-established temporal interval, the individual data-sets to the respective regional office. The latter gathers all the information from the relative hotels and, utilizing the same communications software, transmits the data to ISTAT.

To implement the network each administration (ISTAT, regions) and statistical unit must have a PC and modem dedicated to the capture and transmission of data via cable. The transmission of data from each individual structure complies with the EDI transmission standard adopted by Eurostat. The files are coded according to RDRMES (Raw Data Reporting Message) type messages utilized precisely in correspondence to data originating from the questionnaires. The following software procedures have been developed:

- a software reproducing in electronic format the paper questionnaire of the ISTAT survey, placed in every participating hotel;
- a communications program with BBS ISTAT to transmit information from the unit of observation;
- an ISTAT software for recodification of the data-sets of the individual hotels.

As a result of this experience we are able to express several considerations on the utilization of this data capture technique. This technique pertains to the CASI (Computer Assisted Self-Interviewing) category and, in particular, to the PDE (Prepared Data Entry) system. The cornerstone of the experiment was the adoption of an electronic questionnaire in place of the paper one, with the consequent benefits and drawbacks.

We can point out the characteristics and highlight the virtues of this methodology:

- the strength of this technique is clearly the greater transmission speed and, therefore, the diffusion of data at different levels of territorial disaggregation;
- consequently, the reduced data transmission times reduce the response costs to the enterprises, which are naturally relieved from the onus of transmitting the data by fax or even by mail;
- it is particularly necessary to insert "help on-line" so that, through user interface windows, the user can be guided through navigation of the management program;
- the possibility of capturing data of diverse nature (economic, structural) through a single questionnaire allows integration procedures regarding the following aspects to be activated:
  - immediate availability of diverse types of information;
  - comparability between economic and demographic information;
  - comparability spatially and territorially;
  - comparability in terms of classifications, definitions and methodologies utilized;
- as a rule, the present technique allows non-sample errors to be minimized by introducing, for example, controls in the phase of initial insertion of the information (cfr. Table 5.1);
- finally, the feature of greatest impact on the capture units is undoubtedly the return of information to the enterprises, with all the prospects this affords, including the possibility of confronting the performance of the individual hotel with those of the province/region to which it pertains.

As we know, the foregoing advantages are offset by the following problems:

- difficulties in designing generalized interface software corresponding to the vast spectrum of hotel software employed. This problem is certainly a hindrance to automated data capture;
- presence of economic variables which do not reside within the business and, therefore, difficulty in capturing them automatically since they are in the hands of brokers (sectorial associations, accounting firms, etc.);
- qualities of reliability and confidentiality of the public transmission network. This problem has been raised in connection with the need to capture economic-accounting data. Certain hotels contacted are unwilling to furnish information on given variables without a statistical obligation to participate in the survey and, above all, without adequate guarantees of the confidentiality of the information.

The principal advantages and disadvantages can be summarized in the following table:

Table 4.3
Advantages and Disadvantages in the use of telematic relevation systems in the field of tourism statistics

Advantages	Disadvantages
1. Higher timeliness in data transmission. 2. Lower burden of response for enterprises. 3. Reduction of measurement: automatic controls during data registration. 4. Aggregate back flow of information for respondent enterprises. 5. Multipurpose questionnaire. 6. Uniform standard for data transmission (EDI message).	1. Heterogeneous situations, varying from hotel to hotel, for what concerns the availability of hardware and software appliances. 2. Starting difficulty in the transmission of economic data because of problems of confidentiality in the network. 3. When economic data are not available in the enterprise.

In conclusion, the process of innovation in tourism statistics is still in embryo and is closely related to the diffusion of computer culture at the enterprise level. We must also stress that this process is necessary only if there is a substantial improvement in the quality of the final data.

5. Quality, Costs and Timing

In the attempt to make summary evaluations of the efficiency and effectiveness of the different capture techniques utilized in the experiments, Table 5.1 sets forth a simple analytical scheme in which the "traditional" mail surveys, surveys utilizing telecommunications transmission of data and telephone surveys are compared. Although the evaluations presented here refer to the three surveys just described, we have grounds to believe that, since these surveys are opportunely selected models representative of a wider array of concrete cases, the way they have been approached preserves a certain autonomy of information and a broad adaptability to a wide variety of contexts. The evaluations are the fruit of experience matured in the sphere of each of the operative contexts considered, and they were supplied by the heads of the respective surveys.

Hypothetically, the problem involves choosing one or the other of the techniques on the basis of different indicators of opportunity.

The effectiveness of the capture technique is evaluated on the basis of the "quality" of the same. "Quality" is understood in the strict sense and in reference to the different components into which the quadratic average of the error of the estimator given by the sample average can be divided (the sample average is here used as a reference because, in practice, it is the most frequent). Efficiency is evaluated considering the response costs per enterprise and the time required to transmit and process the data. This approach simplifies the far more complex system of quality control of the production process. At the same time, characteristics of the capture such as pertinence, compatibility, diffusion and comparability (cfr. Gismondi, 1996) are excluded, since we suppose they are guaranteed in any case, independent of the capture technique utilized.

The estimate of the share of enterprises which can be interviewed with the three above techniques was reached on the basis of the preliminary results of an ISTAT survey referring to 1995, a summary of which is reported in Table 1.1 (a more detailed discussion can be found in Anitori, Gismondi, Riccardini and Trovato, 1997). We believe that all enterprises with at least one telephone line can be interviewed with a CATI-type technique. To comprehend the heuristic significance of the components of the average quadratic error ("Quality" section of the table), we must bear in mind the following (ISTAT, 1989, p. 22):

- the term VCA/*n* expresses the sample variance and, therefore, the component of the estimated error due to the non-exhaustive nature of the capture made; it depends on chance and tends to annul itself if the sample number *n* is very high;
- the term B^2 (square of the non-sample bias) expresses the "systematic" component of error, fixed with the variance of the unit interviewed and dependent on "external" causes such as, for example, the errant specification of a question. We must point out that the sample bias is nullified using the sample average with a simple random sample design;
- the term VSR/*n* expresses the simple response variance, in other words, the component of comprehensive error due to the occurrence of individual response errors, which vary in entity and by cause with the variance of the unit interviewed. This may depend on the inexact comprehension of a question, on the use of *proxy* variables rather than the one required (for example, turnover net of taxes rather than gross) or on memory factors in the case of retrospective questions. This term also decreases to zero as *n* increases;
- the term VCR expresses the correlated response variance, depending on the possible bias effect induced by the survey taker (obviously, when he is present); it does not depend on the size of the sample.

The points for every type of capture and every item considered range from 1 to 3, and have a meaning ranging from "low" to "medium" to "high", with an exclusively hierarchical significance. In other words, they are related to one another, so that a "3" could equal a maximum component error

but, in theory, could be low in absolute value. The points must be interpreted inversely to their numerical size: a higher number of points will correspond to greater costs, and vice versa.

The points for the mail questionnaire (which can be returned via telefax) are obtained from the weighted average of the corresponding points assigned in the case in which the enterprise itself, the accountant or others fill out the form, with weights equal to the relative shares with which these situations actually occur. On the basis of operative experiences, we have been able to verify the following:

a) with reference to quality, the sample variance is virtually constant and not elevated (1), reaching an average level (2) for the questionnaires filled out by others, only because of the effect of the low number *n* of cases in which this actually occurs.
b) The systematic error is reduced (2) if filling out the form is not delegated to the enterprise (3), and it should be wholly uninfluential in the case of CAPI (1).
c) The individual response error remains low (1) unless the enterprise itself responds (2), whether or not CATI is utilized.
d) The bias effect of the survey taker is obviously significant only in the case of the telephone interview (2).

The preceding joint effects create a ranking on the basis of which utilization of telecommunications data exchanges is clearly preferable (average points = 1.25), while the use of a mail questionnaire or telephone interview is equivalent, unless the accountant responds (1.25), in which case the first method is preferable.

With regard to costs, the fixed component, independent of the number of forms sent out, remains low (1), unless the installation on a one-time basis of the necessary hardware, as in the case of the utilization of a modem (2), is required, or unless the enterprise utilizes its membership in a sectorial association to fill out, at the same annual cost, an unlimited number of survey forms (questionnaires filled out by others). The variable component of costs is more heterogeneous, reaching the maximum if the accountant is utilized (3), since payment is as a rule required for each questionnaire answered and sent. Overall, we must point out that the response cost for enterprises is higher if questionnaires are sent by mail.

With regard to timing, return of the completed questionnaire is almost always belated compared to the established deadline (3), but the time is reduced downloading the data on a personal computer if the telephone interview is utilized (2), or if the data are systematically transmitted via telecommunications (1). These features have a direct impact on processing times. Overall, the use of the traditional mail questionnaire is clearly the most costly approach in terms of response and processing times (average points = 2.68), while the costs are minimum when the modem is used (1).

The final ranking obtained with the summary of the points acquired on the basis of quality, response costs, return and processing time favors telecommunications (1.25), followed by the use of CATI (1.50) and, last, the traditional mail questionnaire (2.68, a signal that the technique is becoming obsolete). Among the responses returned by mail, the most preferable were the enterprises that delegated to others the task of answering the questionnaire. Since a good capture system should guarantee an average levelling of the general summary index below the average value of 1.5, the value of 1.67 actually attained (the last term in the lower right of the table) suggests that a greater effort is needed to orient the current capture systems towards the use of more modern survey techniques, or techniques that are alternative to a capture in which the component of self-answering prevails.

Table 5.1

Quality, costs and response timeliness: an overall comparison among efficiency estimates using different survey techniques (estimator: sample mean)

SURVEY TECHNIQUE	Share of firms	QUALITY					COSTS FOR FIRMS			TIMES			TOTAL Q+C+T
		VCA/n	VSR/n	VCR	B^2	Total Q	Fixed costs	Variable costs	Total C	mailing	Calcula--tions	Total T	
1 POSTAL QUESTIONNAIRE	**82,8**	**1,1**	**1,4**	**1,0**	**2,4**	**1,47**	**1,1**	**2,4**	**1,75**	**3,0**	**2,4**	**2,68**	**1,97**
Filled by enterprises	30,2	1,0	2,0	1,0	3,0	1,75	1,0	2,0	1,50	3,0	3,0	3,00	2,08
Filled by an accountant	41,8	1,0	1,0	1,0	2,0	1,25	1,0	3,0	2,00	3,0	2,0	2,50	1,92
Filled by others (a)	10,8	2,0	1,0	1,0	2,0	1,50	2,0	1,0	1,50	3,0	2,0	2,50	1,83
2 MODEM	**17,2**	**2,0**	**1,0**	**1,0**	**1,0**	**1,25**	**2,0**	**1,0**	**1,50**	**1,0**	**1,0**	**1,00**	**1,25**
3 CATI INTERVIEW (b)	**100,0**	**1,0**	**2,0**	**2,0**	**1,0**	**1,50**	**1,0**	**2,0**	**1,50**	**2,0**	**1,0**	**1,50**	**1,50**
AVERAGE (1-2-3) (c)	**200,0**	**1,1**	**1,7**	**1,5**	**1,6**	**1,46**	**1,1**	**2,1**	**1,60**	**2,3**	**1,6**	**1,95**	**1,67**

Source: ISTAT estimates..

Legenda: 1=Low - 2=Medium - 3=High.

(a) Mainly, cathegory associations.

(b) IT has been considered apart and it has been supposed suitable for the 100% of firms, because all firms should be furnished with one telephone at least.

(c) The 200,0 value is simply given by the sum of weight with which techniques 1,2 and 3 have been included in the mean. Other weights derive from table 1.1.

VCA/n = global sample variance (it depends from n).

VSR/n = simple response variance or individual response error (it depends from n).

VCR = correlated response variance or error induced by the interviewer (it doesn't depend from n).

B = Non sampling bias or systematic response error.

References

P.ANITORI-R.GISMONDI-F.RICCARDINI-G.TROVATO (1997), *"New" and "Old" Services: An Overview for Some Italian Industries*, paper prepared for the "12th Voorburg Group Meeting", Session 4, Copenaghen.

M.R.CECCARELLI (1997*), L'indagine pilota sulla mobilità passeggeri di lunga distanza - rapporto finale*, documento interno, Istat, Roma.

M.CHIARO (1996), "I sondaggi telefonici", CISU, Roma.

CONSORZIO "MEST" (1996), "Long Distance Diaries Today: Initial Review and Critique", deliverable D1 of the 4th Framework project "Methods for European Surveys of Travel Behaviour", Innsbruck.

D.ETTEMA-H.TIMMERMANS-L.V.VEGHEL (1995), "Effects of Data Collection Methods in Travel and Activity Research", Eirass.

EUROSTAT (1995.1), "Methodological Manual on Business Statistics - Chapter General Framework", Luxembourg.

EUROSTAT (1995.2), "Methodological Manual on Services Statistics - Chapter Horeca and travel agencies", Luxembourg.

EUROSTAT (1995.3), "Besoins en donnés de la Commission", documento discusso nel "Gruppo di lavoro sulle statistiche dei trasporti di viaggiatori e sulla mobilità", 23-24 ottobre, Luxembourg.

EUROSTAT (1996), *The Draft Regulation on Short-Term Indicators (ver.10.0e)*, Luxembourg.

F.FRANCESCHINI-M.TERZAGO (1997), *Quality Function Deployment: requisiti per l'uso*, in "De Qualitate", aprile 1997, pp.13-29, Roma.

R.GISMONDI (1993), "Le esigenze degli utilizzatori di statistiche sulle attività Horeca - Indagine pilota", Aprile-Maggio 1993, documento interno Istat.

R.GISMONDI (1996), *Qualità dell'informazione statistica, nuove tecnologie ed evoluzione dei mercati nel settore dei servizi in Italia: problemi e prospettive*, paper presented at the "11th Voorburg Group Meeting", Newport.

ISTAT (1989), "Manuale di tecniche d'indagine - vol.6", Istat, Roma.

ISTAT (1997.1), "Conti economici delle imprese da 1 a 19 addetti - Anno 1994", mimeo, Istat, Roma.

ISTAT (1997.2), "Conti economici delle imprese da 20 addetti in poi - Anno 1994", mimeo, Istat, Roma.

ISTAT (1997.3), "Il progetto SERT-turismo: rapporto finale", Istat, Roma.

ISTAT (1997.4), Rapporto finale "Indagine pilota relativa agli alberghi ed alle agenzie di viaggio. Anno 1995", project leader Anna Pia Mirto, Istat, Roma.

P.MILGROM-J.ROBERTS (1992), "Economia, organizzazione e management", Il Mulino, Bologna.

G.PUGLISI (1995), *La nuova classificazione delle attività nel contesto di un sistema statistico integrato delle imprese in Europa*, in "Il registro statistico europeo delle imprese" (a cura di S.Biffignardi e M.Martini), pp.411-426, Franco Angeli, Milano.

TRANSPORT INFRASTRUCTURE COMMITTEE (1994), "Mobility on the Trans-European Road Network - Action NEMO II - Final Report", Motorway Working Group, Bruxelles.

E.VEIL (1994), *Selected OECD Countries' Strategies for the Collection of Service Statistics*, paper presented at the "9th Voorburg Group Meeting".

NOTES

1 The work was coordinated by R. Gismondi and is the fruit of shared reflections by the authors. The opinions erein expressed are not binding on ISTAT. In particular, R. Gismondi wrote sections 1, 2 and 5, M.R. Ceccarelli section 3 and A.P. Mirto section 4.

2 Istat - Researcher in Transport Statistics.

3 Istat - Head of Unit Tourism, Transport and Internal Trade.

4 Istat - Researcher in Tourism Statistics.

5 For a close examination of the retail trade classification guidelines, a number of critical proposals were made in a paper presented by Italy in session 3 of the present Voorburg Group Meeting.

6 In accordance with this nomenclature, the general classification procedure is as follows. Units whose principal activity in terms of added value (or other indicator) clearly consists of retail commerce are classified in one of the classes from 52.11 to 52.50. Assignment is made on the basis of the following rules:

- if the products sold fall exclusively within a single NACE Rev. 1 class, assignment is obviously made to that class.
- If the products sold include commodities of several NACE Rev. 1 classes (lower limit of the percentage by class equal to 5% of the total), we must determine whether one of these classes represents a share of 50% or more in terms of added value. In this case the unit would be assigned to that class.

If, instead, no NACE Rev. 1 class reaches 50% of added value, another path must be chosen. If the products sold are assignable to five or more NACE classes, each above 5% but below 50%, the unit is considered a department store; but if food, beverage or tobacco products account for at least 35% of the chosen indicator, assignment is made to class 52.11 (sale in non-specialized, prevalently food business). If, instead, the products do not cover more than four classes, the principal activity is determined by first selecting the principal group (three digits) and, subsequently, the class within this group. For example, if the choice was between three classes - 52.44 (45%), 52.25 (30%) and 52.27 (25%) - the assignment would be made to class 52.25 (and not 52.44) because the largest relative share corresponds to group 52.2 (equal overall to 30% + 25% = 55%).

7 For simplicity's sake, we ignore the units that perform the sale as prevalent or secondary only in one of the two reference periods.

8 The data were observed with a single retrospective model. On the basis of this model the information was captured for each month of 1996. The occurrence of sales in a group different from the prevalent one was, therefore, a feature of *all* the months of that year.

9 The indices proposed in this context, without considering the weighted coefficients, are different from the ones diffused officially, although the results of the simulations proposed here are analogous to those that can be obtained using the correct weight.

10 To better comprehend the monetary evaluations cited in this section we note that the Lira-ECU exchange is approximately 2,000 Lire per ECU.

11 "Leg" signifies every stop made during the trip to change means of transport, because the reason for the trip has changed, or to spend the night. A number of difficulties arose because of linguistic difficulties, since the distinction made by Eurostat was between *journey* (change in motivation) and *stage* (change in means). The Italian word "tappa" combines both English terms.

12 Estimated costs.

13 This figure includes Lire 510,000 in fixed costs, which, in the case of a more numerous sample, reduce the unit cost indicated. In this case, as well, the costs were estimated.

14 The costs are estimated on a sample number of 300 units.