

# **Statistical Agencies and their Users**

**Prepared for the 22nd General Conference  
of the  
International Association for Research in Income and Wealth  
Flims, Switzerland, August 30 - September 5, 1992**

Secondary distribution of any part of this material  
without prior written permission from  
Informetrica Limited is strictly prohibited

**M.C. (Mike) McCracken  
President, Informetrica Limited**



## Statistical Agencies and their Users

This paper offers the views of a user of the Canadian statistical system - one who has used it continually since 1965. Statistics Canada was one of the first statistical agencies to create a databank of economic time series, CANSIM, widely available for all to use (1969). As a result economic research was aided in both the public and private sectors. Publications were widely available, both for research and educational purposes.

Today, as a user, there is a strong signal to minimize the use of statistical information in order to control user charges, to reduce the number of publications kept at hand, and to avoid asking for new information in fear that the incremental data charges will chew up the analytical budget. This seems like a major distortion of choices, one which is counter to the progress made in the technology of statistical analysis and which is counter to the fundamental nature of information - the ability to use it, without denying similar use to others.

For many years, statistical agencies and their users have had a symbiotic relationship, with users supporting the agencies and providing feedback on the data, and the agencies responding to user needs. Both were operating in a benign environment, in which the relatively undemanding user needs could be met with modest increments to the statistical budget or the "dividend" from continuing technological improvements.

But the world is changing. User needs are becoming more sophisticated, technology is allowing them to process much more information, and the pace of decision-making is quickening. At the same time, governments are curbing their discretionary expenditures, seeking new sources of revenue through "user-pay", and focusing on their direct needs, rather than the "public interest". The linkage between the statistical agency and its users is under strain as user fees rise and the capacity of the statistical agency to keep up with user demands lags.

In Canada, user fees have risen for all services. Data previously available on-line at no additional cost, have become the most significant cost in using a computer for analysis. Publication prices have risen by a factor of ten or more. The annual price for the full publication service has risen from \$400 in the early 1970s to \$800 in the late 1970s, then to \$4,000 in the early 1980s and \$20,000 in the mid-1980s. Recent reductions to \$17,500 reflect a reduction in the number of publications. Requests for data are met with quotes of hundreds or thousands of dollars, not with a discussion of the need or utility of such requests.

Some see this as a healthy control of government expenditures and as a way of reflecting need through market signals. Others see this as a hidden tax, a breakdown in the public information system, and a weakening of the evolution of the statistical system through its response to market signals only.

It is possible that this restraint will pass, and a return to a more balanced approach will occur. However, it is also possible that the trend will continue, with further deterioration in the relationship between users and the statistical agencies. This restraint is also likely to reduce the capacity to respond to new data needs, unless they are fully funded by others.

In this paper, the next section describes some alternative scenarios that may influence the capacity of statistical agencies to respond to user needs. Then, the demand, costs, and pricing issues associated with data are explored at some length. Finally, the implications for the statistical system of changing, disappearing, and blurring boundaries are discussed.

It is hoped that a full discussion of these issues by national statisticians from around the world can either suggest alternate strategies or arrest my concerns. Perhaps it will be possible to determine if the trend is continuing or whether the pendulum will swing back towards the public interest. The overall performance of the world economy will likely influence the outcome.

## 1 CAPACITY TO RESPOND

In a recent article [9] three possible scenarios for the evolution of the world economy through 2000 were outlined. The labels used are indicative of the general settings - BLISS, DEJA VU, and COLLAPSE.

In the BLISS scenario, good policy reflected through more sensible macroeconomic settings, a move to full employment, breakout of peace, a rising rate of technological progress, and the emergence of the LDCs, leads to rising real incomes and stability of government finances.

In such an environment, it is likely that government statistical agencies would be urged to "keep up" with technological change, to contribute to a deeper understanding of the economic phenomena, and to make their information widely available.

The DEJA VU macroeconomic scenario is coloured by the continuation of high real interest rates, rising unemployment, capital flight, volatile exchange rates, no productivity growth, static LDCs, restrictive policies and political turmoil.

Consistent with such a setting, statistical agencies would continue to be cut back, with additional user charges imposed whenever possible, staff "rented out", no new statistical programs, no new hardware, and no human capital formation.

Nor is this the worst. The macroeconomic COLLAPSE scenario is one of continuing supply shortages - electricity, food, and other basics. The financial infrastructure disappears, with collapses of private financial institutions, exchange controls, LDC debt repudiation. Business investment is non-existent; government spending is concentrated on defence and "domestic security".

Written in the past tense, the description of the statistical system might be as follows:

After the riots in response to the CPI release, it was decided that the survey would no longer be taken. The National Accounts became a restricted document; the Investment Survey was cancelled due to lack of response. The control of information, public and private, became an important element of government policy. Statistics Canada is now much smaller, with only two publications a year available to

the public. (23-210, Honey Production and Value and 62-532, Funeral Directors.)

Each government department is purported to have a statistical group, but little is known of their affairs, as Ministers closely guard their fiefdoms.

It is of course possible for there to be many other scenarios. Nor is there any necessary link between the macroeconomic scenarios and the paths of the statistical agency. Some just seem to go together!

The point in examining scenarios of any type is to think through the implications of each, and how to avoid the unfavorable and how to achieve the more desirable.

The **macroeconomic environment** can have some influence on the statistical system - both in terms of the resources available to it and the kinds of phenomena that are measured. (Recall the influence of inflation on measurement in the 1970s, the demand for information on the energy sector, and the constraints on expenditure applied first in 1978 and then with subsequent cuts that continue to this day.)

## 2 THE STATISTICAL SYSTEM AS INFRASTRUCTURE

In the search for improved competitiveness, "infrastructure" has been identified as a key component. But it should be realized that the statistical system is itself an important type of infrastructure. In fact, it usually gets better with use, rather than wearing out like a road! The information base for private and public decision-makers is crucial for success in advancing living standards, improving productivity, detecting opportunities, and avoiding pitfalls. **The beneficial externalities and the "public good" nature of information should be the fundamental reasons for governments to produce statistics.**

### 2.1 What Part Of Statistics Canada's Output Is A Pure Public Good?

"A pure public good is usually defined in terms of two characteristics:

i) **Supply jointness or nondepletability**; meaning that, at least within a significant range, once the good is supplied to anyone, it requires no additional resources for its benefits to be enjoyed by additional persons.

ii) **Non-Excludability**; meaning that persons cannot be prevented from consuming the good or, rather, from enjoying its benefits. In particular, they can enjoy these benefits without paying anything for them." [1,p.102]

In practice, Baumol notes that most public goods are impure, with some exclusion possible and some marginal costs incurred either by the seller or the buyer. In fact, he says, "It is difficult to think of any good or service for which some degree of excludability cannot be achieved by a supplier who puts his mind to it, and is not hampered in his task by constraints imposed by legislation or regulation." [1,p.103]

The notion of **excludability** is necessary for a price to be charged for the product. Without it, there is little likelihood of it being provided by private industry. But imperfect excludability results in a **beneficial externality**, with gains to those not part of the transaction.

This seems to me to result in two conflicting public policy objectives: reduce excludability to increase the externality, and increase excludability in order to impose user fees or taxes or to "privatize the activity".

As a practical matter I would consider the following "systems" to be "mainly public goods", characterized by beneficial externalities and low marginal costs:

- o Census of Canada
- o National Accounts
- o CPI
- o Labour Force Survey

The externalities come in the form of better public and private decisions, with a scope for "making a difference" in the macroeconomic performance of the Canadian economy. Some parts of these databases are made available widely through the media (e.g., unemployment rate, growth, inflation, population levels and growth) at virtually no marginal cost to the user.

One step removed, are the details behind these statistical systems: the publications, the CANSIM Main Base. They provide the "entrails" required by those advising the decision-makers and interpreting the significance of "the numbers".

At the other extreme, a special tabulation prepared for a monopolist with an eye to preventing competition from abroad (e.g., tabulation of competing imports in preparation for an anti-dumping case) might be viewed as not meeting the notion of "public good" on several grounds.

If the spirit of a "public good" is supply jointness and non-excludability, then the broad statistical system through the point of the first publication, its placement on CANSIM, or its availability in some other form (disk, tape) is perhaps the closest fit.

### 3 THE CLIENTS OF STATISTICAL AGENCIES

Who are the clients of a statistical agency? If the statistical agency is part of a larger department, then the most immediate customer is that department. As a central statistical agency, the Canadian experience is to treat the federal government departments as the primary customer with provincial governments as next most important (in some fields). Private sector interests, nongovernmental organizations (NGOs), and others are viewed as a "residual" market at best.

Such a hierarchy of clients is understandable, given the budgetary linkages to the federal government. As a consequence, however, the government focuses on its priorities in instructing the statistical agencies, and the other users have to live with the consequences of those decisions. The other users usually are described as "always wanting more", but "always receiving less".

This reality is only a "problem" if the other users have no other options, or are unimportant in the broad scheme of things. In market economies, about 80 per cent of value-added occurs in the private business sector, with the balance being provided by governments or non-profit groups. Decision-making in the private sector is surely of some relevance to the overall economy!

What seems to be missing in government circles is any recognition that statistical information used by the private sector could lead to better economic performance, thereby conveying "externalities" for governments through increased revenues, to unemployed through jobs, and to NGOs through improved revenues. Perhaps those in charge of the fiscal and monetary levers feel they control the outcomes.

### 3.1 Typology Of Users

The users of statistical information are not homogeneous in their demands, further complicating the task of a statistical agency interested in serving the broader public. Although within any category there may be a wide range of sophistication, the following typology of users is designed to highlight the nature of the demands on the statistical agency.

- o Media - Their basic interest is in receiving the "current numbers" with enough context to write a column or prepare a (brief) item for the evening news. Release at a specified time each month or quarter, with a well-written press release, and a contact point for quotes or assistance in interpretation are the important elements of service. (Of course, the media never expects to pay for anything!)
- o Governments, Direct - Each department is insatiable in its requirements for information to assist it with its programs or policy formation. In-depth statistics in a particular area are required by line departments, aggregate measures by statistical agencies. There is also a demand for the pre-release of information to allow for the preparation of briefing material for ministers.
- o Governments, Indirect - Politicians react to the media reports and are interested in statistics to support their positions for use in speeches. Some of their political staff may be interested in voter profiles and other information pertinent to their district. If the statistical agency provides such information, it should do so on a similar basis for all political parties.
- o The Financial Community - The security analysts and foreign exchange markets are essentially interested in receiving information in as timely a manner as their competitors, particularly if the information can move markets. The emphasis is on relatively few series, short history, and no need for lengthy documentation.

- o The Econometric Modeller - The economist undertaking to develop a structural model of the economy or some sector, requires a consistent database, long time series, sufficient detail to allow for disaggregation, measures suggested by economic theory (e.g., prices, incomes, quantities, demographics), and with sufficient information about the data to determine if some adjustment to the model is required. Explanations of past "blips" in the data can be very helpful to the econometrician puzzling over residuals.
- o Economic Consultants - Given their project orientation, their need is to assemble information quickly on an industry or topic, taking what is available and analyzing it. Documentation is important, since the area may be new to the analysts.
- o The Corporate Economist - Information about the industry, its principal clients, prices of inputs, and trade data for the commodities it produces are likely to be the key areas. Over time, increasing detail and full exploitation of sectoral information is likely to occur. The analyst will usually "know" the data quite well, and will be frustrated in not having industry information as timely as the corporate data. In some organizations, the economist will also fill out the responses to statistical questionnaires, thereby gaining some insights into the survey.
- o Others - Statistics serve a much broader audience. Everyone forms their expectations about future inflation, output, and other indicators based on some set of statistical reports, usually through the media. Special interest groups utilize statistics in their presentations to the public and governments, picking that information favorable to their cause. Lawyers buttress their case with statistics; otherwise they attempt to discredit statistics used by others.

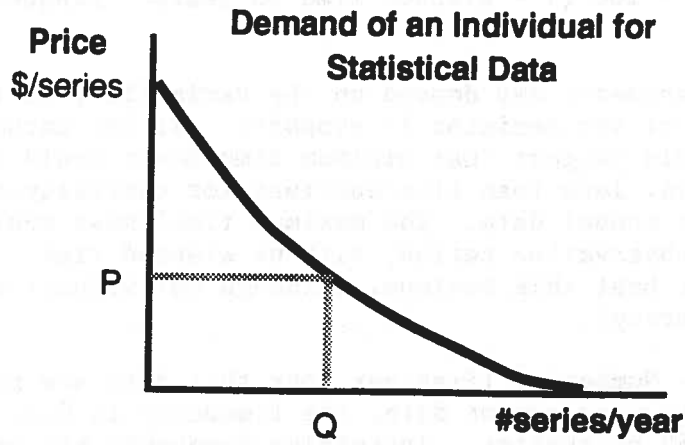
#### 4 THE MARKET FOR STATISTICS

##### 4.1 The Demand For Statistics

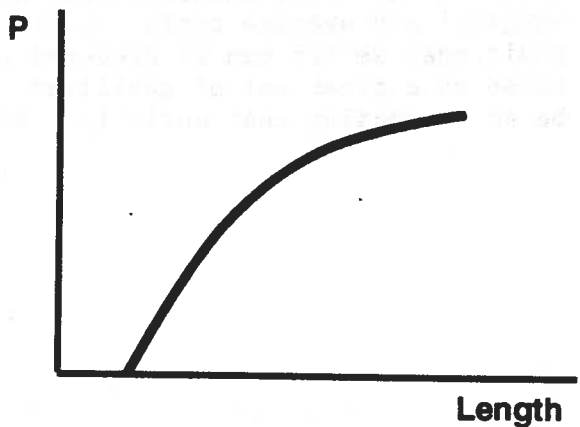
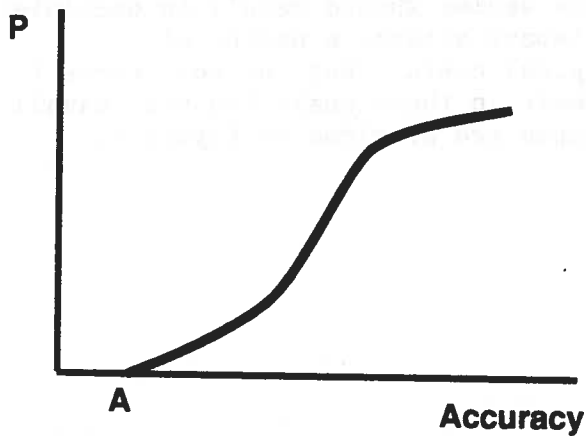
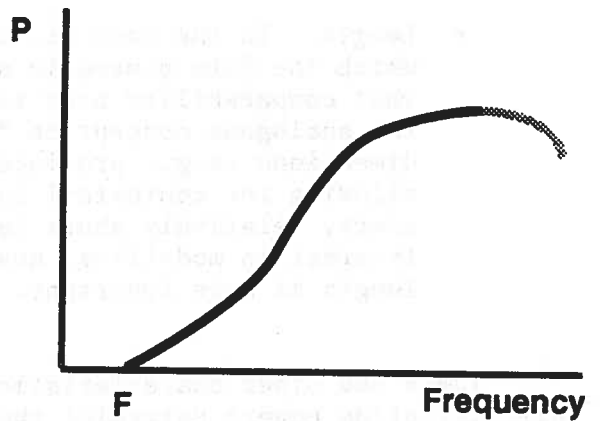
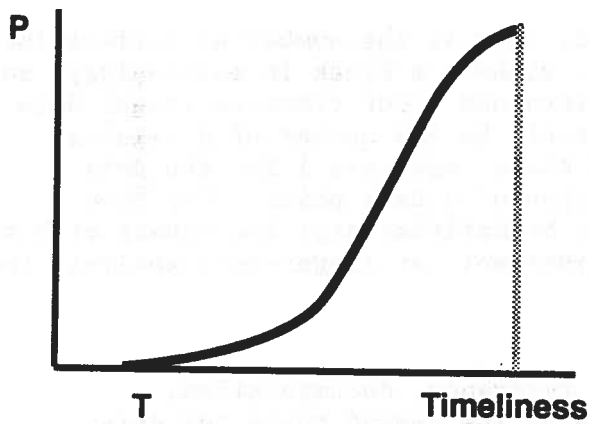
In considering users, statistical agencies might find it useful to have a view about the demand curves they face. There is little reason to suspect that the demand curve looks any different for data than for other goods and services. That is, at higher prices users will "consume" fewer series per year. The "price" is a combination of acquisition costs and the costs to the user of utilizing the data.

But there are other characteristics of data that may be relevant to the user. Four examples are timeliness, frequency, accuracy, and length. More of any of these factors will "shift" the demand curve out, increasing the number of series used at a given price, or allowing a higher price for a given quantity. The relationship between these qualities and price are shown in Figure 1.

Figure 1 Demand Curves



Demand Curves for Characteristics



- o **Timeliness** - Below some level (T) there is no demand for the data, although historians may have more patience. Timeliness can be defined as:

$$\text{Timeliness} = 100 (1 - \text{elapsed time in years} * \text{frequency} / \text{parameter})$$

The parameter may depend on the variability of the data series, or the nature of the decision it supports. If the parameter is set at 5, then it would suggest that minimum timeliness would be five months for monthly data, less than five quarters for quarterly data, and less than 5 years for annual data. The maximum timeliness would be immediately after the observation period, with no elapsed time. Forecasts, of course, can beat this maximum, although not without other problems (e.g., accuracy).

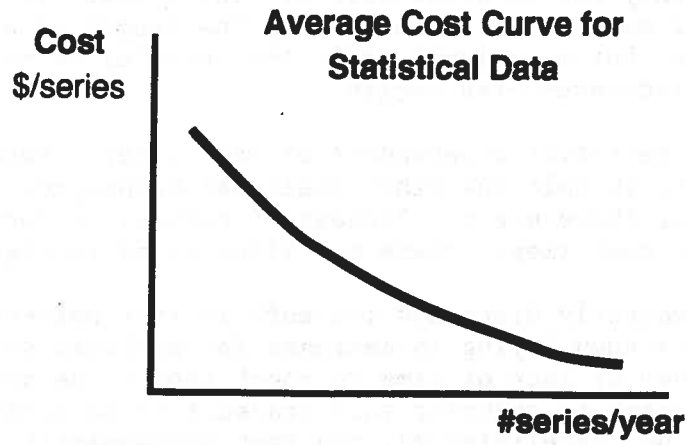
- o **Frequency** - Number of times per year that data are published. In the case of five-year census data, the frequency is 0.2; exchange rates could be 365 or greater. Increasing frequency has some value, at least up to a point.
- o **Accuracy** - If we define accuracy as 100 minus the percentage error, then there is some minimum level of accuracy (A) required, and a maximum of 100. The flatness of the curve at high levels of accuracy suggests that there may be some point at which further accuracy commands little additional price.
- o **Length** - In the case of time series, this is the number of periods for which the data source is available, without a break in methodology, so that comparability over time is maintained. For cross-sectional data the analogous concept to "length" would be the number of available dimensions (e.g., province, income class, age, etc.) for the data, allowing for contextual interpretation of a data point. For some users, relatively short lengths may be satisfactory; for others with an interest in modelling, seasonal adjustment, or longer-term analysis the length is more important.

These and other characteristics (e.g., relevance, documentation, discriminating power) determine the position of the demand curve for data.

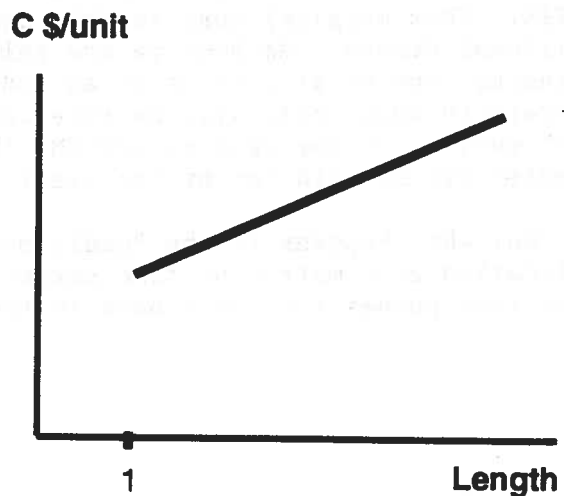
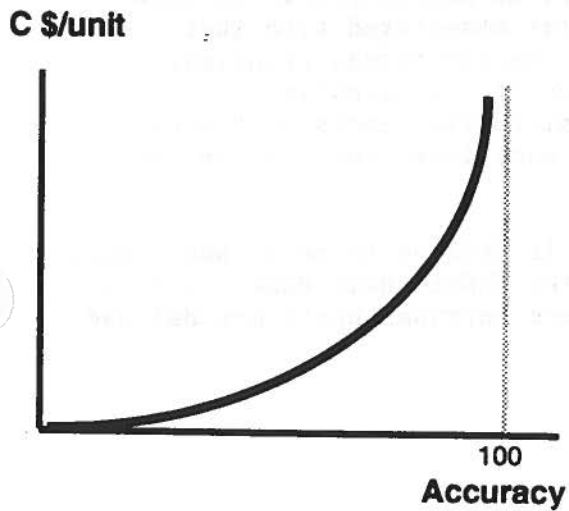
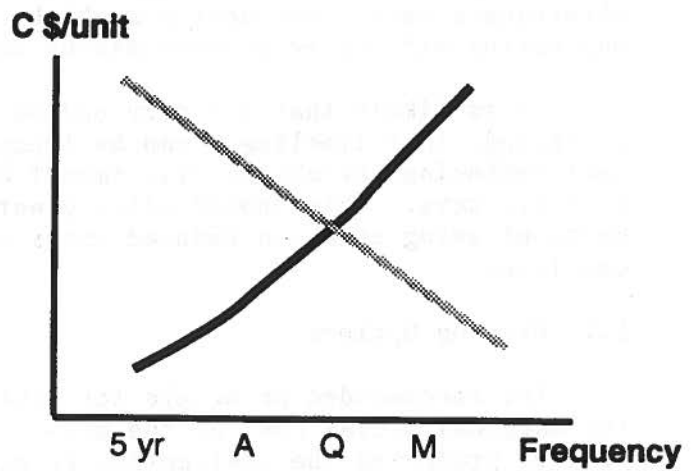
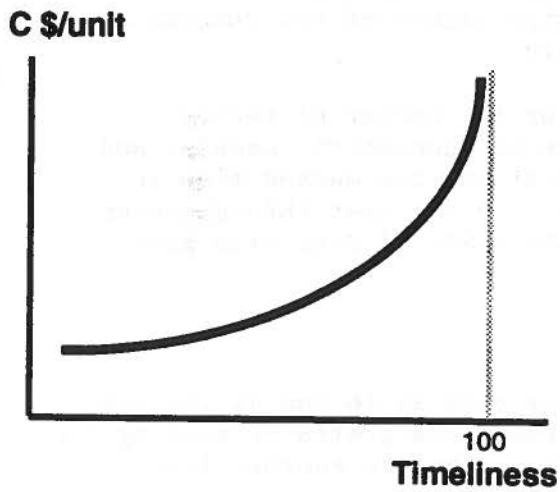
## 5 COSTS FOR SUPPLYING USERS

It is likely that the production of statistical data is a decreasing-average cost activity. The spreading of the overhead of computers, central registers, and professionals over more series should result in declining marginal and average costs. Given that a database exists, a number of additional series can be produced at low marginal costs. But the cost curve is based on a given set of qualities. Improvements in these qualities will usually be an increasing cost activity. Possible shapes are provided in Figure 2.

## Figure 2 Cost Curves



### Cost Curves for Characteristics



Timeliness can be provided at little additional cost up to a point, beyond which costs could escalate rapidly. For frequency, there are additional costs for observing (and publishing) at a more frequent interval. Accuracy is thought to be a continuously increasing cost activity, with rising marginal costs. In some cases complete accuracy can be attainable at finite cost; in others there may be inherent limits of measurement. Increasing the length of a data series incurs some storage costs, but more importantly the costs of maintaining comparability over time increases with length.

Nor are these characteristics independent of each other. Thus, for any cost curve it is necessary to hold the other qualities unchanged. For the four qualities considered here, there are six "tradeoff" curves, reflecting the pair-wise linkages of the qualities. These are illustrated in Figure 3.

Perhaps the most frequently discussed tradeoff is that between timeliness and accuracy. This occurs when trying to decrease the publication lag leads to inherent inaccuracy because of lack of time to check inputs, delays in reporting, etc. One approach to reducing this tradeoff is to publish data points that are "timely" but "preliminary", and then subsequently to revise them. But these efforts are not without their costs.

Length and frequency are likely to be positively related, since we measure length by the number of observations and frequency by the number of times per year published and observed. If the length was divided by the frequency, we may find a different shape (see dotted line in chart), since one reason for observing a series frequently might be the changing nature of the phenomenon, suggesting difficulty in maintaining comparability.

It is likely that accuracy can be improved by the number of series collected, that timeliness can be improved by use of econometric models, and that technological change will impact both the cost and the demand side in positive ways. This should allow greater benefits to the user through lower costs of using data and reduced costs in producing a set of data with given qualities.

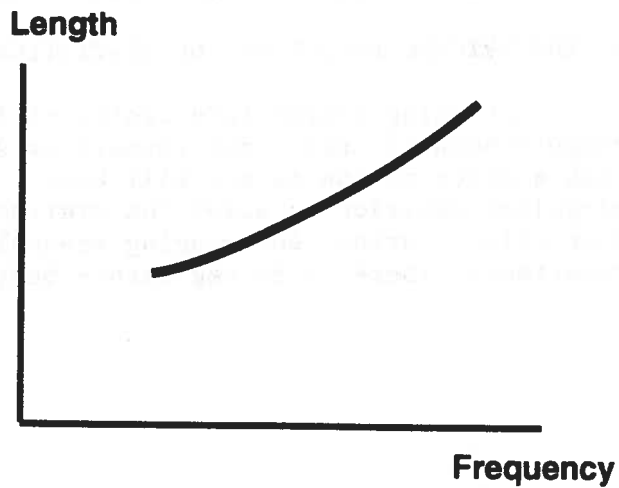
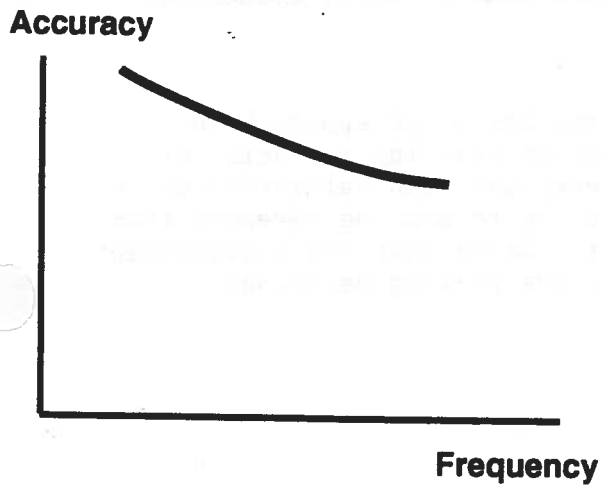
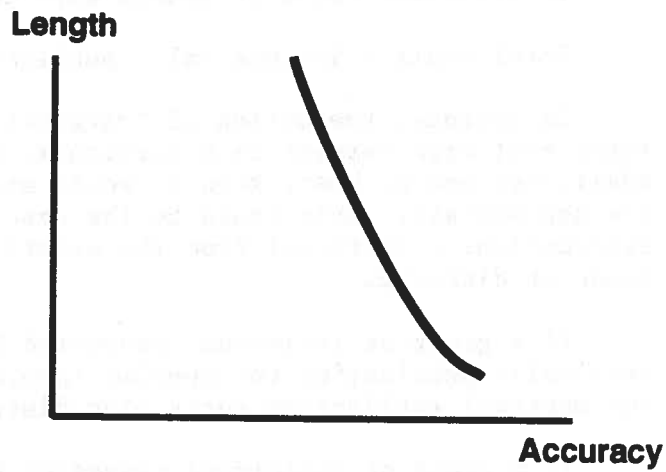
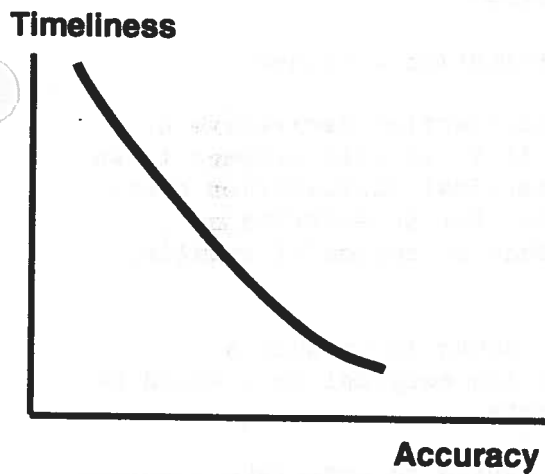
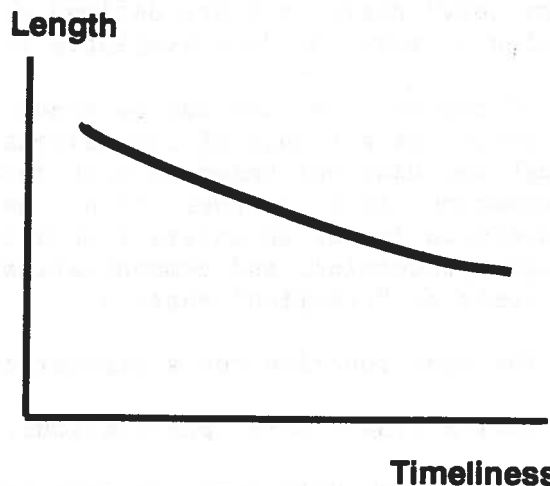
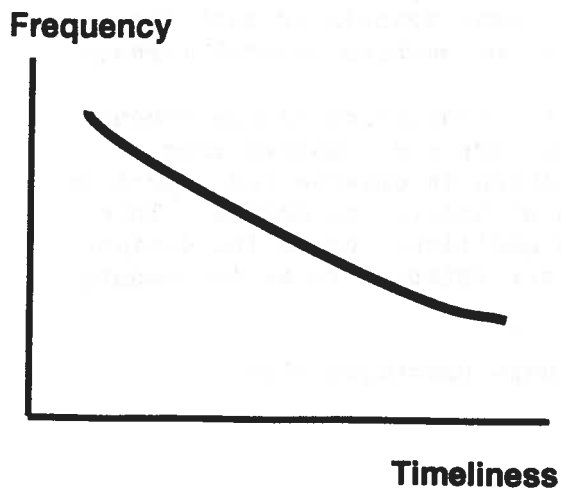
### 5.1 Pricing Options

The recommended principle for statistical agencies is to charge the user the marginal social cost of the data. This has been interpreted as meaning the cost of producing the publication or providing the output in another form.

Where does the threshold of marginal cost lie? The simplest answer is to observe that the additional output is to be measured as an incremental copy of a publication, an additional diskette, or an additional series retrieved from CANSIM. Thus marginal cost is that additional cost associated with that additional demand. As such we are talking about the additional printing, packaging, and mailing costs of an additional copy of a publication. (Henceforth such costs will be referred to as distribution costs or "second copy" costs.) In the case of CANSIM, there is no additional cost, except for computer cycles paid for by the user!

But what happens if the "additional output" is defined to be an additional publication or a matrix of time series added to the CANSIM Main Base. Such a definition pushes the costs back to the point where marginal costs are defined

### Figure 3 Tradeoffs between Characteristics



differently; namely to include those additional costs associated with the preparation of a publication as well as the distribution costs. In the case of CANSIM, additional costs will be incurred for the initial conversion to machine-readable form, cataloging, checking, updating, and storage on-line. For purposes of contrast, these costs will be referred to as **publication costs** or **"first copy" costs**, and are defined as additional costs associated with the decision to make the data available in some format, or perhaps several formats.

Of course, the line can be drawn further back. Statistics Canada produces many series as a result of tabulations of censuses, surveys, observations or through calculations based on such data. The decision to observe and report on a phenomenon, in most cases, is at the discretion of Statistics Canada. Thus the decision to add an activity entails incurring additional costs for design, surveys, processing, and communications. (These are referred to as **incremental data costs** or **"creation" costs**.)

The cost function for a statistical agency looks something like:

$COST = f(\text{data sets, publications, users}), \text{ or}$

$\text{Incremental data costs or creation costs} = f(\text{data sets})$

$\text{Publication costs or first copy costs} = f(\text{publications})$

$\text{Distribution costs or second copy costs} = f(\text{users})$

$\text{Total costs} = \text{incremental} + \text{publication} + \text{distribution} + \text{"fixed"}$

In essence, the notion of **"marginal cost"** is the **partial derivative** of total cost with respect to a particular variable. If it is with respect to an additional use or user, then it would appear that **marginal distribution costs** are appropriate. This would be the case for an order for an existing publication, a retrieval from the existing CANSIM base or copies of existing tapes or diskettes.

If a group or individual requested Statistics Canada to produce a particular publication (or special tabulation) then the marginal cost would be the **marginal publication costs plus distribution costs**.

If a group or individual requested Statistics Canada to undertake a survey, census, or other calculations, then the marginal costs would be **incremental plus publication plus distribution**.

**Where the line is drawn depends on the product that is being requested.**

## 6 THE PRICING POLICY OF THE STATISTICAL AGENCY

Increasing expenditure restraint in Canada has led to an approach to nongovernmental users that emphasizes selling them on existing products, at as high a price as the market will bear. This tendency has been reinforced by the misguided decision to allow the statistical agency to retain the revenues from such sales, further encouraging monopoly behavior. Since they are a government department, there is no regulatory body reviewing the pricing decisions.

What has been the private sector reaction to this approach? Although "market" data are hard to come by, many economic research organizations in the private sector have reduced the number of publications they subscribe to, reduced the number and frequency of machine-readable retrievals, and forced an increasing reliance on other sources of information (newspapers, newsletters, opinions) for their reading of the economy.

Companies with on-line management information systems rely on their own sales, inventory, back-order information and other market data. Surveys and market research fill in the holes. Unfortunately, these tend to be done more by the larger companies, leaving the smaller companies in the fog.

Perhaps this is all for the better, leaving the private sector to generate the data it needs independent of governments, and for governments to collect only what they need. However, it does seem to be leaving some economies of scale unexploited and reduces the net social benefits associated with the dissemination of information. As well, there is the risk that governments and the private sector could be working at cross-purposes, because of different perceptions of the state of the economy.

If the private sector interests are ignored, then voluntary response to statistical agency surveys may also deteriorate. This can be overcome by requiring firms and individuals to file, or further exploiting administrative data. At the same time, users of the output from the statistical agency may become less supportive, resulting in having no "friends" when the inevitable cuts occur.

## 6.1 Back To The Basics?

**Should Statistics Canada return to the basics? What about monopoly rents forgone?**

There is little guidance to suggest what Statistics Canada should or should not do. What is a "basic" to one person may be a "frill" to someone else. But the notion of creating a "monopoly" for purposes of maximizing the rents obtained, strikes me as perverse public policy, whether the entity is in the public or private sector.

Statistics Canada is a "natural monopoly" as a result of economies of scale in a statistical system and a "legal monopoly" because of the Statistics Act which, in some cases, compels responses in return for protection of privacy. It is precisely because of its position that it is in the public sector, with a clear intent that it not behave as a profit-maximizing monopolist, but rather as a benign generator of public benefits. This suggests that the public servants should focus on the pursuit of the public interest, not entrepreneurial activities.

Statistics Canada should also be concerned about the industrial organization of the "private statistics processing industry". It has a role to ensure that entry into this industry is not restricted by barriers, thereby preventing monopoly rents downstream. Such historical actions as the general availability of publications, the broad distribution of CANSIM tapes, the publication of a directory of survey organizations, and the open assistance to

groups are consistent with this "caretaker" role, and have ensured a reasonably competitive private sector.

If further tailoring, packaging, etc. is required by a client then charges to cover these costs can be imposed. If it is done by Statistics Canada that should be because Statistics Canada can do it better. But competition should be allowed to ensure that this is the case. With an existing private sector, it would be inconsistent with government policy for Statistics Canada to pursue the development of this competing line of business in an aggressive way.

## 6.2 What Pricing Options Should Be Considered?

The first step should be a review of the desirability of raising net revenue from the sale of data. Any pricing that deviates from the appropriate marginal social cost will be socially distorting. If some net revenue must be raised, then the issue becomes one of minimizing the distortions. This is consistent with governments covering the fixed costs up to and including the first copy. If some contribution to fixed costs must come from users directly, the answer is "Ramsey pricing". In essence, it suggests a markup for declining average cost activities that is inversely proportional to the demand elasticity. (If of course marginal costs are not declining then pricing at marginal costs does result in a producer's surplus.)

The logic is apparent. If there are minimum deviations from marginal cost for those activities that are elastic, then there will be minimum distortion of their consumption; conversely, if inelastic items are marked up much more, it will matter little to the quantities of these items consumed, and again minimize the distortion. (This does not mean that those with inelastic demands will be pleased, since clearly they will pay a larger shot!) In an extreme case, if one item with a completely inelastic demand can be found, then all of the additional revenue requirement can be obtained from a single source with no distortion of consumption patterns.

Where might such demands originate? A first place to look would be to the demands of government, particularly where there are legislated requirements for data. Such demands would be highly inelastic, implying a major part of the burden (all?) should be met there. At the extreme, we are back where it all started, with governments carrying the budget of the statistical agency and publications being priced at marginal distribution costs.

A straightforward approach that would have some appeal as self-regulating, would be to allow Statistics Canada to sell any output in any form at any price, but also to allow anyone else to sell a copy of such output or other derived outputs at any price they choose. In essence, the copyright protection that currently is in place, would be unilaterally dropped for Statistics Canada outputs. (It should be recalled that this is the situation in the U.S. for all government publications.)

The key principle is that pricing of Statistics Canada outputs should not be viewed as a net source of additional revenue over and above the marginal distribution costs. Thus the prices for the various options will reflect the marginal resource costs of making them available, and the users will be led to rational decisions.

### 6.3 What About CANSIM? Alternative Delivery Vehicles?

One view starts with the assumption that the placement of a series on the CANSIM Main Base is an "initial" publication cost and considered in the interests of Statistics Canada for purposes of generating subsequent publications or internal analysis, or in the interests of the government for its analysis and to ensure that consistent data series are used throughout the government. In such a case, the marginal cost becomes the distribution costs or "second" copy costs.

Then the question becomes an issue of pricing the various distribution formats from the Main Base. One option is the Mini-Base, another is diskettes, another CDs, etc. The pricing of such distribution devices should reflect their marginal distribution costs.

If some new technique (e.g., diskettes) represents a cheaper method with all of the same characteristics as an older method (e.g., retrieval from the Minibase through a timesharing system) then surely such an innovation should be promoted, not discouraged. If one method has advantages and disadvantages compared to other methods, then there will likely remain a number of such methods.

If one starts with the view that CANSIM is not necessary for Statistics Canada or the federal government, then the marginal costs of CANSIM become the publication or "first copy" costs plus the distribution costs. But if this is the case, then a more basic question must be asked. Why should CANSIM exist or continue as a Statistics Canada service?

### 6.4 Related Issues

If it is decided to run Statistics Canada as a profit-maximizing monopolist or independent taxing authority, then the fundamental issue becomes: **How to regulate this activity, and who should do it?** Answering such a question is of course not to be the responsibility of the monopolist!

Additional problems arise if costs differ widely between otherwise similar users of the statistical system. How does Statistics Canada "police" the use of data by secondary distributors for their own purposes?

If Statistics Canada is "pricing to what the market will bear", how about looking at selling numbers before their release date to selected brokers? Or perhaps an auction? Is there a market for "favorable" numbers, particularly around an election? I believe of course that professional national statisticians would find such activities abhorrent, and that Statistics Canada would not go so far. But when the signal is to "wring out a buck", then I am concerned about the focus of the institution.

Nor is Statistics Canada alone. It is often thought of as an example or precedent for the other two-thirds of the information sector that originates through other federal government departments and the large provincial component. If proper pricing rules can be evolved at Statistics Canada, then its positive impact can be much greater than Statistics Canada alone (and conversely).

## 7 CHANGING, DISAPPEARING AND BLURRING BOUNDARIES

Lack of resources and pricing issues are not the only problems of national statisticians. Even with adequate resources there will be numerous problems and challenges remaining for both statistical agencies and their users. One of the ongoing, and perhaps accelerating, issues is that of boundaries.

At the same time that Statistics Canada is working to improve the industrial classification process, as described by John Crysdale [4], the business community is developing a new industry - general-purpose manufacturing or computer-integrated manufacturing (CIM), which aspires to produce any thing at any time at any place.

Deregulation of the financial industry is underway, with the four pillars of banking, insurance, trusts, and brokerage activities, becoming "financial and other services". (Some institutions are already offering payroll services, economic forecasts, and soon leasing of automobiles. It won't be long before travel services, auto repair, and video rental will be added.) Similar changes in transportation and communications will further blur the boundaries.

Consider the boundary issues from a statistical viewpoint of the following items:

1. Municipal reorganizations into regional government.
2. The continuing movements of enterprise-establishment relationships.
3. Shifts of responsibility between federal-provincial-municipal levels of government.
4. The privatization of Crown corporations through municipal corporations (airports), public share sales (Air Canada), outright sale to existing businesses (Boeing "purchase" of deHavilland and Canadair with a subsequent sale to Bombardier and the Ontario government), or the creation of special operating agencies, with a foot in both the public and private sector.
5. The effect of human rights legislation prohibiting mandatory retirement at age 65. What is "old" - "retired"?
6. What is licit or illicit? Gambling if state-run is legal, otherwise not. In Manitoba, regular casinos are acceptable, in Alberta only if "charity-run". Alcohol is legal, unless in the blood in sufficient concentration and while operating a vehicle. Some issues of some magazines are pornographic. Are their sales recorded in that month? Recent changes in legislation legalized video-gambling machines, essentially slot machines in the Maritime provinces. Should GDP be increased as a result? What about the years before legalization, when the same machines were in operation?
7. Boundaries between corporations, non-profit entities, non-governmental organizations (NGOs), associations of individuals, cooperatives, etc.

## 7.1 Globalization

Another area in which some see the "boundaries" becoming "blurred" is between nations. The general term is "globalization", usually with some awe thrown in. Yet this is hardly a new phenomenon. Trade between nations has been carried on for centuries. Distances between countries measured by travel time have been shrinking over the same period, with substantial gains through air travel in the post-war period. Communications between nations is not new, although the "band-width" has been broadened with satellites in the last twenty years. What is perhaps new is the "involvement" of many more people through travel, through television, through immigration, and through the workplace in activities with a link to other countries.

The other major force must be the "international", "multinational", or "transnational" corporation, with operations in many countries, with personnel moving throughout the company, and with goods and services being exchanged within the organization, as well as outside. (In fact, the "globalization theme" may originate from the song-books of the large corporations! For a controversial view of the effects of such organizational structures on the statistical system and economic policy, see Kenichi Ohmae, particularly the chapter, "Lies, Damned Lies, and Statistics" [10,p.137ff].

## 7.2 The Corporation And The Individual - Different Rules?

With the myriad of activities of corporations and other forms of organization, users increasingly want information in a number of different ways. But reports by region, by industry, by ownership class, and by size for a group of firms might, through residual disclosure, reveal something about a particular corporation. Current practices are to avoid such a possibility, with a result that confidential cells abound in statistical publications and new tabulations are severely limited by what has been prepared previously.

With the increasing emphasis on markets, the growth in firm size, and the "corporate" state dominating international transactions, the time has come to re-think our attitudes towards corporate information.

Corporations and other business organizations are creatures of the state, licenced to operate, granted certain privileges of limited liability and the possibility of perpetual life, and required to file certain reports and to collect certain taxes. There is no reason that this licence cannot require the filing of certain data for statistical purposes, and also provide for unfettered release of all such information to the public at large. The notion of confidentiality and privacy as matters of state concern would cease to exist for organizations in our society.

With this change, then the task would be to work to fulfil the necessary condition assumed in economics for perfect competition, namely that all parties have full information. If the information filed with Statistics Canada was also provided to all of the shareholders, then there would be a net improvement in the base for both public and private decision-making. Issues of residual disclosure would disappear, and the activities, both positive and negative, of corporations in Canada would be more apparent.

The treatment of the individual in the statistical system is a separable issue. Privacy is perhaps best treated as a "right", with any compromise subject to checks and balances. In fact, the separation of business and persons might help clarify the nature of these rights. If a group of people choose to organize a business, then their income from that business could be a matter of public record. Their total income, or health records, or recreational propensities would remain a matter of privacy, with any questions for statistical purposes carefully designed to avoid disclosure (as is presently done).

### 7.3 Implications

What does this do for the national statistician? In some cases, boundaries may be impossible to establish, although I suspect that this need not be the case. Rather, the problems may be difficulty in measurement, the magnitude and number of transactions, the comparability of measures, and the legal right to require reporting of the various dimensions of a transaction. This may lead to international agreements for the exchange of information, not unlike the international tax agreements currently in place.

At the same time, with a heightened awareness of other countries, there is a demand for increased comparability of derived statistics. Producing data for one country is not easy; doing so in a fashion that is comparable with many other countries at different stages of development, with different historical processes, and different cultures is challenging indeed.

## 8 RESOLUTION?

If statistical agencies are to receive adequate resources in order to pursue the public interest and to cope with the many emerging statistical challenges, then the strong support of the user community is likely to be a necessary condition. Certainly, national statisticians should develop and maintain good communications with their users and suppliers of information.

At the same time, users of data should become more familiar with what is being done, the process by which data are calculated and assembled, and the constraints that apply to statistical agencies.

Advisory councils can also serve a positive role, both in improving the quality of data submitted and identifying the user needs. But reaching the users directly is important, both for education of the user and to identify their specific requirements.

But such activities are likely to be insufficient alone to reverse the trend. Political leadership and strong lobbying are more likely to produce results. This will require the development of the infrastructure concept as applicable to statistics, and the fostering of a stronger desire to use quantitative information as the basis for public policy formulation.

### References

- [1] William J. Baumol, **Microtheory: Applications and Origins**, (MIT Press, Cambridge), 1986.
- [2] Carlo M. Cipolla, **Fighting the Plague in Seventeenth-Century Italy**, University of Wisconsin Press, 1981.
- [3] Carlo M. Cipolla, **The Economic History of World Population - 7th ed.**, Barnes & Noble, 1978.
- [4] John S. Crysdale, "Industrial classification in the Canadian Census of Manufactures: Towards less art and more science", **Statistical Journal of the United Nations ECE 5** (1988) 377-92.
- [5] Robert Eisner, "Extended Accounts for National Income and Product", **The Journal of Economic Literature**, Volume XXVI-4 (December, 1988), p.1611-1683.
- [6] I.P. Fellegi and M.B. Wilk, "Is statistics singular or plural?", **The Canadian Journal of Statistics**, Vol.16, Supplement (1988), p.1-8.
- [7] Martin Lefkowitz, "The Statistics Corner: Guest Editorial on needed improvements in U.S. economic statistics", **Business Economics**, Volume XXIV-1 (January, 1989), p.48-50.
- [8] M.C. McCracken, "Data for Dollars: Charges for Public Goods", **Monthly Economic Review**, Volume 4 No.6 (June, 1985), p.1-3.
- [9] M.C. McCracken, "Looking Back From the Year 2000", **Conference Board Research Bulletin No.198**, New York, 1986.
- [10] Kenichi Ohmae, **The Borderless World: Power and Strategy in the Interlinked Economy**, Harper Perennial, 1991.
- [11] Charles A. Waite, "Statistics Corner: Guest Editorial on Recent Developments in the Census Bureau's Economic Statistics Programs", **Business Economics**, Volume XXIII-3, (July, 1988), p.55-59.

