INTRASTAT AND PRODCOM - NEW EUROPEAN SURVEY SYSTEMS

Graham Jenkinson, UK Central Statistical Office,
Government Offices, Great George Street, London, SW1P 3AQ

KEY WORDS: Trade, production, European, nomenclatures

INTRODUCTION

The European Community Statistical Office (Eurostat) has for some time pursued a policy of establishing common statistical systems across all of the Community's 12 Member States. This is becoming more structured with its proposal for a "European Statistical System" likely to go to the Council of Ministers in the relatively near future.

Two of the earlier initiatives came into force at the beginning of 1993. One is INTRASTAT which is a new system to measure visible trade between Member States. It replaces information from Customs documents which are no longer required at European Community (EC) internal frontiers following the completion of the Single European Market at the end of 1992. The other is PRODCOM, a system to collect detailed data on output by product which has to operate first in respect of 1993 and which in the UK replaces a set of less comprehensive sales inquiries.

Both sets of data have to be collected direct from businesses and they use related commodity coding systems which can be linked together. The systems generate a number of clear advantages:

- consistent systems across the EC;
- clearly specified quality requirements;
- good and consistent information for business decisions;
- good and consistent information for EC and national political decisions;
- reduced costs of collecting trade data.

However, by far the most important advantage stems from the use of a common nomenclature basis so that it is possible to look at detailed product statistics for each country in the EC on a completely harmonised basis for imports, exports and production for home sale.

This paper describes the systems, discusses their similarities and differences and finally considers some of the implementation issues. While describing some of the EC issues it is written mainly from a UK perspective.

INTRASTAT

INTRASTAT replaces for intra-EC trade the former system based on the use of Customs documents. The basis of INTRASTAT is a link with value added tax (VAT) procedures on goods traded between Member States. This link with an administrative revenue collecting system will provide a mechanism for ensuring that the statistics are as complete and accurate as possible.

INTRASTAT is a theoretically simple system. Its main features are:-

- All businesses registered for VAT are required to complete two additional boxes on their VAT returns. These record the total values of goods exported to and imported from other Member States. (In the UK, VAT returns usually cover three month periods. Also in the UK the new boxes have been completed since 1992 to ease the transition between old and new systems).
- Larger traders above thresholds complete detailed returns at least monthly.
- Smaller traders below these thresholds do not provide any further statistical information: estimates at a detailed level are made from the total values reported in the boxes on their VAT returns.
- The smallest businesses falling below the turnover limit requiring registration for VAT are completely excluded from the system.

The EC legislation only allows a short list of data items to be reported on the detailed monthly returns. Eight data items have to be collected, Member States may collect up to five other specified items but no others are allowed (at least on a statutory basis). Eight digit commodity codes must be used comprising the 6-digit Harmonised System (HS) codes plus 2 extra digits from the EC combined nomenclature (CN).

The totals of the values reported on these monthly returns should be reconcilable with the quarterly

amounts reported in the new boxes on the VAT returns. These checks will be simple and straightforward for most traders and most trade although there are exceptions. The checks will be an important control on the quality of data provided by traders.

PRODCOM

PRODCOM is literally a list of product descriptions and associated codes. (This will be referred to as the PRODCOM list.) PRODCOM can also be taken to mean the inquiry into the value and volume of sales of products, classified according to the PRODCOM list. (This will be referred to as the PRODCOM inquiry.) PRODCOM in general refers to the system to collect and publish the sales values classified by the list of products.

The EC legislation specifies that the inquiry has to cover all activities classified as within the mining and quarrying, manufacturing, and electricity, gas and water supply industries (codes 10 to 41 inclusive of the ISIC equivalent to Sections C, D and E of the EC NACE (rev 1) industrial classification). Information must be collected for each of the roughly 4,800 products in the PRODCOM list and information is required on the physical volume and value of production sold during the survey period.

The main requirement is for annual inquiries but monthly or quarterly inquiries can be required - in 1993 quarterly inquiries are being conducted in 20% of the industries involved in manufacturing activity. A key feature of the PRODCOM list is that each heading is completely consistent with 8-digit combined nomenclature headings used for international trade statistics. Every PRODCOM heading corresponds to one or a single group of CN headings.

PRODCOM will imply a large increase in data collection for the UK. The number of contributors will rise from about 10,000 to nearer 30,000 and the number of identified product headings from 3,200 to about 4,800.

THE SYSTEMS COMPARED

There are many common features of the systems and considerable advantages to be had from basing them both on the same product nomenclature. However there are also a number of differences. Most of these were unavoidable in the circumstances with large differences in the previous systems and, in the case of INTRASTAT, a rush to meet the unavoidable deadline imposed by the start of the European Single Market at the beginning of 1993. When the systems settle down these differences may impede the use of the data and I suggest that Eurostat should be looking for opportunities to reduce them.

Legal Base

Both INTRASTAT AND PRODCOM are specified in European Council Regulations adopted by the Council of Ministers at the end of 1991. There are also a number of European Commission Regulations which specify some of the more detailed implementing measures.

The two sets of Regulations are however different. The INTRASTAT legislation goes into great detail and imposes obligations directly on traders throughout the Community. The PRODCOM legislation mainly imposes obligations on Member States to collect the statistics rather than on individual businesses to supply them. In the light of more recent debates in Europe on the subsidiarity issue in relation to the Maastricht Treaty, the INTRASTAT legislation would probably not specify so much detail if it were being proposed today.

These European Regulations have the direct force of law in each Member State. However some minor national legislation has been necessary to implement INTRASTAT mainly to establish a penalty regime for non- compliance.

Eurostat's current emphasis on creating an overall legislative framework will probably ensure a more uniform approach in future legislation.

Quality Requirements

Both sets of legislation impose quality requirements that the results of the inquiries must meet. This is a major step forward since, in the past, legislation has tended to specify inquiry methodology but not mention quality. It should mean that wherever a trader is established in the EC it faces similar costs in meeting requests for statistics and has similar quality information available to inform its business decisions. Without this it is possible for there to be competitive advantages for some Member States with businesses moving to minimise their net overhead costs.

In both cases the requirements were the subject of a long debate between Member States before they were adopted. Compromises had to be struck and the outcomes were the best available at the time to meet conflicting priorities. Neither was completely ideal for the UK and I feel looking back at them now that improvement and rationalisation would be desirable. Both systems specify minimum quality requirements. INTRASTAT makes an (not very effective) attempt to specify a maximum quality requirement to prevent excessive respondent loads: PRODCOM does not do this.

There is a significant difference between the requirements even though both inquiries are intended to measure data about products. The INTRASTAT requirements refer to the accuracy of estimated trade in individual products. The PRODCOM requirement specifies coverage of production for each industry.

These differences in approach arose mainly because the historic data needed to design the samples were not the same in both cases. Nevertheless I believe that Eurostat should be aiming for more uniformity in future legislation.

It is difficult to make precise comparisons between the effects of the two different approaches to specifying quality because of the use of different registers and reporting units: most INTRASTAT reporting units are enterprises and most PRODCOM units are establishments. In the UK in INTRASTAT about 24% of intra-EC traders are above the threshold accounting for nearly 97½% of intra-EC trade. On the other hand in PRODCOM only 19% of undertakings will be reporting accounting for 90% of employment. Since the two inquiries are both collecting data about products this difference in effect seems unjustified. The UK believes that the INTRASTAT requirements should be relaxed as the high quality trade data will be of limited use to users without matching quality in the production data.

Sample Design

Neither quality requirement encourages best statistical practice. The INTRASTAT legislation unambiguously imposes a cut-off survey. That is that data is collected from all traders above a size threshold but from no traders below that threshold. This is not the most efficient form of sampling.

Similarly the PRODCOM regulations encourage the use of a cut-off survey by saying that "account shall be taken of all undertakings employing at least 20 people". The UK, however, believes that the wording does not rule out sampling providing that the results are grossed up. We therefore intend to develop an efficient stratified random sample design for PRODCOM to meet the quality requirements - this will include some light sampling of those employing less than 20 so that the results can be accurately grossed up to the full population of manufacturers.

Users

Both systems have been designed to meet the perceived needs of users of the statistics. Broadly these fall into three categories.

One category is companies and businesses throughout Europe. Markets work best on the basis of good information. Businesses would like information on trade and output to assist them in making appropriate investment and marketing decisions. To pursue this they would like accurate information at the most detailed level.

A second category is the European Commission and other Community institutions. Their direct interest in intra-EC trade statistics is relatively limited in some product areas in the Single Market although information continues to be needed for the management of the Common Agricultural Policy, Fisheries Policy and some trade policies. They do however have a need for comprehensive aggregate estimates to measure the success of the Single Market and to ensure that Member States are able to continue to make accurate estimates of GNP - levels of GNP are one of the factors which determine contributions by Member States to the EC budget.

The third category is national governments. Both sets of statistics contribute to the National Accounts which need to be measured thoroughly and accurately to guide economic management and fiscal decisions. More detailed information is needed in some product areas for managing other national policies.

Providers

Statistics can only be produced if businesses spend time and effort providing the raw data. In today's very competitive business climate many businesses resent what they see as non-productive time on completing statistical returns. Considerable effort has gone into reducing and simplifying the burdens imposed on industry.

INTRASTAT will cut costs to industry dramatically compared with the previous system based on Customs documents - the cut should be particularly large in the UK which has set its thresholds higher than in most other Member States. However in the short-term businesses are faced with large one off computer and other costs to establish systems to meet a range of changed requirements in the Single Market. Many have chosen to provide the INTRASTAT data themselves where previously Customs Agents completed the documents on their behalf as part of an overall service.

These transitional costs and problems have resulted in a fairly high level of complaints and representations in most Member States. At least one Member State has raised its thresholds as a result.

The effect of PRODCOM will vary depending on the previous scale of production inquiries. In the UK as already mentioned there will be a large increase in costs. However industry will not be faced with the

increase in most cases until the forms for the annual 1993 inquiry go out early in 1994.

Both the EC and national governments have pursued a number of initiatives to simplify the provision of data and cut costs to industry. In INTRASTAT the provision of data by electronic media is strongly encouraged and a range of software has been developed to assist in data preparation. For PRODCOM the UK has developed personalised user-friendly forms. The UK has kept its contributors well-informed throughout. All INTRASTAT contributors received personal visits. All PRODCOM contributors have had personal advance notice letters and some have had visits. A number of other publicity and information campaigns have been pursued.

IMPLEMENTATION

INTRASTAT came into force from 1 January 1993 without a full overlap with the previous system. A partial bridge between the systems had been provided in the UK by including the new boxes on VAT returns from the beginning of 1992. This had assisted the preparation of a register of INTRASTAT traders and advance identification of traders above the threshold. UK traders had received final detailed instructions early in December. The final preparations were a little rushed because of the late agreement in Europe of some of the details.

Some Member States published estimates of their intra-EC trade in January during March. These early estimates were not really INTRASTAT data: they depended on a variety of sources and estimates and only very indirectly on the INTRASTAT returns themselves. The UK made it clear in advance that we would not publish estimates before they were known to be of reasonable quality and that we hoped that this would be in June. In fact we published an estimate of total trade in the first quarter on 11 June.

The UK's implementation problems have mainly concerned the speed of response but also to a lesser extent the quality of response. Returns are due from traders 10 working days after the end of each calendar month. This has not been achieved but initial response was mainly up to expectations with response around 80% for quarter one eight weeks after the quarter end. This was sufficient to support publication of an aggregate quarter one estimate.

However response then plateaued at about this level despite normal postal and telephone reminder action. Estimation for this level of non-response over a period of months was more than our estimation routines could cope with reliably. It also made it impossible to prepare reliable estimates of trade by commodity or by country.

A major staged compliance exercise is therefore now underway. Non-compliant traders are receiving visits at their premises at the rate of 1,000 a month with an initial emphasis on the largest. In addition all of the 3,500 largest traders will receive routine visits by the end of the year to check on the quality of their returns. Initial targets are, first, to raise compliance to 90% within 12 weeks of the month end by the end of August 1993 and then to speed up this level of compliance so that it is achieved within 4 weeks of the month end by Summer 1994. Ultimately we hope to have 90% response by soon after the 10 working day deadline and final response over 95% A hard core of large persistent non-responders are being considered for penalty action during September and October 1993.

For most traders the quality of data appears to be good. Information agrees with data on their VAT returns and is comparable with data on Customs documents in 1992. However some traders are showing trade in quite different commodities in 1993 and 1992 and a few appear to have covered only part of their trade. These doubtful returns are being checked and, where necessary, corrected urgently.

The effect of these compliance difficulties was magnified by some difficulties with the register of traders. Two features were involved. First we had some register errors especially for exports because trader identifiers were not always present or accurate on the trade documents which were the main source of register information. Second there have been more changes that usual in the register at the beginning of 1993 as businesses re-organised their tax reporting because of the start of the Single Market.

These response problems needed to be resolved before we could be certain that further estimates would be of publishable quality. Publication plans were therefore adjusted with monthly estimates of total trade for January to June expected in early September and with the first commodity and country estimates in early October.

A number of common implementation issues have arisen in several EC countries including the UK. Good response is easier to achieve for exports than for imports. This is because most companies have been able to take export INTRASTAT data from their invoicing systems but have had to set up new systems to capture import data quickly and comprehensively. The seasonality of data from INTRASTAT is different to that from Custom's documents. The UK has successfully estimated some new seasonal factors from a transformation of the back series but these are less certain than usual.

Quite a large number of traders are providing data

electronically. The data is a lot less aggregated than had been expected - many traders are reporting separately for each consignment rather than aggregating similar consignments over the month. Accurate classification is proving difficult for traders particularly for imports - many traders previously used Customs or other agents to supply their declarations and this is therefore the first time that they have had to use the nearly 10,000 heading CN. (It may be a transitional problem with most traders using only 2-300 headings at most). Reporting of net mass in kilograms is proving difficult for some products. Doubts about the commodity codes and quantities have prevented us from calculating price and volume indices from the first INTRASTAT data.

PRODCOM on the other hand remains at a relatively early stage of implementation. The first quarterly forms have gone out to about 16% of the contributors selected. The first quarterly results will be available later this year and the first annual results will be published in the summer of 1994. All contributors except for the small sample of the smallest businesses have had advance notice letters.

An important result of the two new systems stems from their use of classifications based on the CN. For the first time ever it will be possible to look at detailed product statistics for each country in the EC on a completely harmonised basis for imports, exports and home production for sale. The potential for analysis of market penetration, and analysis and monitoring of company relative performance is very powerful. The UK will present summary INTRASTAT data alongside the PRODCOM results. It will also enable us to improve further the consistency of deflation across the national accounts.

Despite some similarities between the INTRASTAT and PRODCOM systems and the use of linked product nomenclatures the systems have been developed and implemented separately. In the UK they are being run by different Government Departments - INTRASTAT is run by Customs and Excise (who are responsible for the administration of VAT and for the measurement of trade with non-EC countries) and PRODCOM by the Central Statistical Office. There are a number of valid reasons for this difference in approach - different periodicities (monthly v annual), different reporting units (enterprises v establishments) and so on. When the systems were agreed by the Council of Ministers in 1991 this seemed to be the right way to go. However once the systems have been established it would be sensible to reconsider whether it would be more efficient to move towards a more consistent approach viewing the two sets of data as different aspects of a common system.

CONCLUSIONS

Two new systems are being implemented by EC Member States to measure intra-EC trade (INTRASTAT) and production for sale (PRODCOM).

INTRASTAT will reduce respondent load and costs while PRODCOM will generally increase them.

INTRASTAT was introduced without a full overlap with the previous system with acceptance by the Commission and Member States that the smooth flow of monthly trade statistics would be interrupted.

Both systems specify minimum quality requirements for their results but those for INTRASTAT are more demanding than those for PRODCOM.

There are other differences between the systems which need to be reviewed and reduced where possible by amending EC Regulations.

The systems use closely linked product nomenclatures.

INTRASTAT respondents have had early difficulty using the CN which suggests that, at nearly 10,000 headings, it may be too detailed and complex.

In other respects the accuracy of INTRASTAT data should be guaranteed through its links with an administrative revenue collecting system.

These systems are early steps in a move by Eurostat towards a European Statistical System.

BUILDING AND USING AN EXPORTER DATA BASE

Michael P. Risha, Michael G. Farrell, Bureau of the Census
Michael P. Risha, Room 2180-A Bldg 3, Bureau of the Census, Washington, DC 21122

Introduction

KEY WORDS: Firms, establishments, SEDs, Economic Censuses, SSEL

Until recently, the Commerce Department had no comprehensive data on firms or establishments engaged in exporting merchandise. While there was an enormous amount of information on the flow of merchandise exports, little was known about the domestic firms that were engaged in either producing or distributing those exports in the international market place, which industries were exporting to which countries, the relative size of the exporters or their geographic distribution. There were only indirect measures of the underreporting of exports. This paper will discuss the dramatic changes brought about by the 1987 Exporter Data Base (EDB).

The authors developed the EDB, a comprehensive data base of all United States exporters and all shipments that they exported during the latest Economic Census benchmark year. This very large data base resides on a microcomputer in randomly accessible SAS data sets. The EDB provides the Census Bureau with many benefits. It allows the Census Bureau to:

- improve the quality and coverage of the filing documentation requirements
- identify firms that file erroneous data or consistently omit items
- generate a mailing list with measures of size for future samples or surveys of exporters
- provide direct information on the kinds of commodities exported by various industries rather than estimates obtained by using classification concordances
- provide a benchmark for an annual program; an opportunity that will not come around again until 1997
- provide marketers with location of exporter information by ZIP, and information about the enterprises that own them
- give export promotion programs a competitive edge by providing information about domestic producers and their role in the international marketplace
- provide states with reliable measures of the origin of grain and other shipments of fungible commodities
- answer questions about the size, location, industrial composition, method of shipping goods, and trading partners of exporting firms and establishments
- contribute to the effort to estimate export-related employment more expeditiously
- incorporate other data, including a potential linkage of the data to the Bureau of Economic Analysis (BEA) Foreign Direct Investment Survey on trade between related parties

History of Available Data

Establishment Data

The first information about industries engaged in exporting appeared in the report "Survey of the Origin of Exports of Manufactured Products: 1960." This report was done under contract with the International Trade Administration (ITA), but funding stopped in 1981. Census began publishing it on an annual basis in 1983. To obtain the estimates for this series, Census provides BEA with the merchandise trade data and concordances to convert the commodity information to an establishment basis. BEA uses their input/output model to estimate direct and supporting exports and export-related employment for the report. Census allocates the results of these estimates on a state-by-state basis based on export data collected in the Annual Survey of Manufactures and geographic information from the most recent Census of Manufactures. This report series, renamed to "Exports from Manufacturing Establishments," provides state-by-state estimates of the value of exports produced by manufacturing establishments in those states. It is released approximately three years after the reference year.

In May 1986, the ITA prepared a decision Memorandum for then Deputy Secretary of Commerce Clarence J. Brown that stated,

In response to frequent questions about the number of manufacturing firms that export, Administration officials currently rely on data from a 1978 ITA report, "A Survey of Business Needs in Export Marketing." The report estimated that 30,000 U.S. manufacturing firms, or about 10 percent of the total

300,000 manufacturing firms, export...The report also included distributions of manufactures exporters by 2-digit SIC (Standard Industrial Classification) code and firm size (defined as employment size). It is important to note that the data in the 1978 report actually cover the year 1972. We are, therefore, still publicly quoting information which is now nearly 15 years old...Information updates at regular intervals have clearly been needed for some time.

In May 1988, the ITA published a special supplemental report of the Annual Survey of Manufactures (ASM), entitled "U.S. Manufactured Exports And Export-Related Employment." The report stated,

A landmark survey taken by the Bureau of the Census for all U.S. manufacturing enterprises has determined that some 58,760 manufacturing establishments nationwide had export sales in 1984. These establishments constituted roughly 17 percent of the estimated total of 350,000 domestic manufacturing plants (organized into some 300,000 companies) in existence that year, indicating that significantly more than one out of every six manufacturing establishments in the United States engaged in export activity during 1984.

The study provided estimates, by state, of the number of exporters by the employment size of the firm that owned the plant. It also provided 2-digit SIC detail, total shipments and the free on board (f.o.b.) value of manufactured exports at the point of production, and measures of export-related employment, but no commodity or destination detail.

In December 1988, the firm of Jack Faucett and Associates published the report "Small Business Exports Of Manufactured Products, 1985" under a Small Business Administration (SBA) contract. The rationale for the study was,

The participation of small businesses in U.S. exports is of considerable current interest and efforts are underway to promote a larger participation. It is believed that small businesses have highly marketable specialized products, and high technology products, and are missing opportunities for export sales due to lack of information on foreign markets and the inability to realize economics of scale in foreign marketing and distribution... Data are relatively scarce on the extent of small business exports in industry detail, by U.S. regional location and by foreign areas of destination. Such data are needed to evaluate existing shares of small businesses in U.S. exports and to target efforts to increase these shares...Unfortunately data available permit the development of this detail only for exports by vessel, excluding exports by air cargo and overland exports to Mexico and Canada (the leading country for U.S. exports)...The Bureau of the Census collects data on direct exports by manufacturing firms (including sales offices and branches) by establishment and publishes data by 2-digit SIC industries by state, covering 70-75 percent of exports of manufactured products. These data cover exports by all modes, i.e. by vessel, air and overland. Data on foreign destinations are not collected. In order to identify shares for small business firms Census could match the establishment data with records for the parent firms as collected and published in Enterprise Statistics and publish the export data by firm size.

Finally, in August 1991, the Dun and Bradstreet company published its annual export outlook survey for 1991 and 1992. The survey used a 5,000-firm sample, selected and weighted so it was statistically representative of all U.S. businesses. Their sample estimates that there are between 650,000 and 750,000 exporters (?).

Merchandise Trade Data

Measurement of exports is one of the government's oldest statistical programs, starting back in 1790. From 1790 to 1820, the compilation of the value of exports was essentially an accumulation of "adjusted estimates" by the customs collectors. In 1820, the process was formalized and individual export shipments began to be recorded on a standard document. This document was filed with the U. S. Customs Service, along with the ship's manifest and other shipping documents needed by the carrier when shipments crossed the international border. The information collected formed the basis for computing the merchandise trade balance with U.S. trading partners and the U.S. current account balance. This method of collecting data on exports is essentially the same around the world. Currently, the Census Bureau processes a million export documents a month and publishes, in electronic form, 600 megabytes of data monthly.

For nearly two centuries, merchandise trade data has dealt only with measures of country-to-country commodity flow and traffic through the U.S. Customs districts and ports. Trade documents did not request data about the firms or industries that exported these commodities. The Shipper's Export Declaration (SED), the official government export document, is the source of statistics on the export of merchandise. The data that

are keyed from the SED concern the value, shipping weight, quantity, mode of transportation, and other information on commodity flow but nothing about the exporter. While the name and address of the exporter are reported on the SED, they are not keyed; thus to physically sort the millions of SEDs received during the processing year into firm order to determine the dollar volume of merchandise accounted for by each firm is not practical.

Commodity transactions by firm was, however, exactly the information that the Bureau of Labor Statistics (BLS) needed to develop an export price index. The Census Bureau, at the same time, needed the export price index produced by BLS to develop a price deflated series for merchandise exports adjusted to a constant dollar basis.

In order to provide this needed information, the Census Bureau added new questions to the SED in 1985. Those questions included a requirement for the exporter to report the state where the export journey began ("state of origin of movement"), his Employer Identification Number (EIN), the three-digit zip from which he is located, and whether the transaction was between related parties. Commodity records could then be summarized by firm on the computer, sampled and the EIN supplied, under contract, to BLS for the export price survey.

By 1987, the question on the domestic origin of movement of exports before they got to the port was of enough interest to the transportation sector that they provided funds to key and tabulate the information. They wanted to know "to which state they should send the truck to move the goods to the port." As a result, effective with the first quarter, 1987 statistics, the Census Bureau released the "State of Origin of Movement" data series of U. S. exports; which provides the transportation industry state-by-state estimates of the value of exports from the state where the commodity began its' export journey. Aggregated state value totals are published on a monthly basis and industry detail are made available quarterly in machine-readable form. The data, reflecting the state where the merchandise begins its export journey, provides value and shipping weight statistics on the domestic transportation of exports. It was never intended to provide the information about the state where the exporter is located as required by the Omnibus Trade and Competitiveness Act of 1988 nor was it intended to supply information about where the commodities were grown, produced or assembled. [The data, adjusted for the nonreporting (20-25 percent) of the state field via mathematical algorithms developed by the Massachusetts Instituite for Social and Economic Research (MISER), is available from MISER or, beginning with the first quarter, 1990 on the Commerce Department's National Trade Data Bank (NTDB) CD-ROM (Compact Disk-Read Only Memory).]

Initial investigation addressed the question of "What state is the exporter reporting as the 'origin of movement' or pick-up point?" It might be the state where the manufacturer is located, where the wholesaler or other intermediary is located, a warehouse location, the farm where the goods were grown, the quarry where the minerals were mined, the port of export or some place else. A direct consequence of developing the EDB answered these questions.

Status of Available Data

As stated before, the Census Bureau publishes two data series that provide information on the role of states in U. S. export trade: the Foreign Trade Division "State of Origin of Movement" (OM) series and the Industry Division "Exports from Manufacturing Establishments" (PO) series. Both series provide state-level information on U. S. exports, but from two different points of view regarding the definition of origin: movement and production. The OM series is more timely, provides more detailed commodity information, provides country of destination information, includes both exports of nonmanufacturing commodities, and exports by wholesaling and other intermediate establishments, while the PO series provides characteristics/number of manufacturing establishments and firms exporting merchandise, indirect exports, and employment and export-related employment. If possible, the two data series should be linked into one, i.e., the EDB.

The basic conceptual difference between the two series, that of export movement of the goods verses production of the goods, has brought about much confusion among the data user community. Data users, for reasons of timeliness, tend to use the OM series as though it were PO data. Exports from the PO series are directly reported by manufacturing establishments on their 1987 Census of Manufacturers (CM) questionnaire. The OM manufactured exports value, which includes total exports of manufactured commodities by all types of establishments; not only manufacturing establishments, is 25 percent higher than the directly reported PO value. To inflate the PO series estimates to reflect the OM manufactured export total, the value difference (OM value - PO value) was allocated to geographic areas at the 3-digit SIC level, and added to the total PO value of exports by state. It is assumed that this "value difference" of exports includes intrastate manufacturer to wholesaler/intermediary trade of commodities destined for export.

A recent study (Risha, 1991) investigated the possibility that the OM and PO series are comparable

substitutes for each other. Overall findings revealed that from a nonparametric standpoint, the two data series reflect the same population of establishments and the principle products they produce and export -- directly or through a distribution channel(s), but are comparable only on an aggregate level. Large value deviations prevent the proportions from being used to estimate the total value of exports for a particular state or for a particular two-digit SIC. It can be suggested that the geographic reporting of manufacturer to wholesaler/intermediary trade of commodities destined for export, as reflected by the OM series, indicates interstate trade. That is, contrary to the assumption made in the PO series as to allocation of the "value difference," wholesalers and other intermediaries may be located in different states than the manufacturers from which they purchase their commodities.

A direct result of the creation of the EDB revealed the true relationship between production origin and origin of movement for both manufactured and non-manufactured exports and the validity of the Schedule B commodity classification to industry SIC concordance. Linking the SED records to the 1987 Census records was made and found to be successful for two thirds of the SED records. Results indicate that approximately 68 percent of the establishments identified, accounting for about 76 percent of the total U. S. export value of manufactured commodities, are manufacturers who do their own exporting. The commodities exported by these manufacturers reflect the principle commodities they produce from which their industry SIC code is based. About one quarter of the exports are attributable to wholesalers and other intermediaries.

It was also found that the ZIP code reflects the physical location of the exporter approximately 91.1 percent of the time. The ZIP code reported by manufacturers, wholesalers and other intermediaries reflects the location of the plant or establishment that exported directly to a foreign country, not the headquarters of the firm. ZIP codes reported by wholesalers and other intermediaries reflect their own location and not the location of the manufacturing plant, farm or quarry where the merchandise that they exported was made, grown or mined. While the intent under the OM series was for exporters to report shipments of merchandise based on the point where the export journey to the port of exportation began, most exporters reported the state where they were physically located. Other exporters reported the state where the port was located. Owing to this, starting with release of the first quarter 1993, data on U. S. exports by state based on the 5-digit ZIP code reported in item [1a.] of the SED, the name and address of the exporter will replace the "state of origin of movement" series.

The 1987 EDB: A New Methodology to Identify Exporters

It is evident that the information available from both private sector and public sector sources on the size, composition and geographic dispersion of the exporter community is fragmented and episodic. Clearly efforts by the ITA, the SBA, and by state development agencies to increase the participation of small and medium sized businesses in the field of exporting and to monitor the effects of their programs are limited by the lack of current, continuing, and comprehensive data about the exporter community. There is no ongoing program to chart significant changes and trends in the size and industrial composition of exporters. As the United States becomes more involved in international trade, the need to know the "demographics" of the businesses engaged in international trade from a comprehensive data base of known exporters is crucial.

In 1987, the ITA provided funds for a special tabulation of data using the EIN reported on the SEDs. They hoped to develop a "Small Business Data Base" that would supply current information about the size characteristics of exporters and their foreign market participation. [Census keys the EIN, but no edit for accuracy or attempt to follow up for bad or blank reporting was made. The response rate to the EIN question was 75 percent on a value of exports basis. This meant that a minimum of 25 percent of the value of exports could not be linked to exporters using the EIN.] A small sample (n = 5000) of nonresponse records was developed and the name and address reported on these SEDs was further researched to provide universe estimates as to the number of establishments that exported manufactured commodities. Census published the summary results of this study in a July, 1989 press release entitled, "More Than 100,000 Exporters Ship U. S. Manufactured Goods, Census Bureau Reports."

The ITA then requested that another study be conducted to cover all export shipments, not just manufactured goods, to determine whether the EIN's reported were valid, to determine the number and employment size of firms and establishments represented by the EINs, and to develop data by geographic area. The authors suggested that this could be accomplished by matching the SED file to the Standard Statistical Establishment List (SSEL), the Census Bureau's directory of all firms and establishments with paid employment. The Bureau uses the SSEL as a mailing and control list for the Economic

Censuses, as a universe for the County Business Patterns program and as a source for current economic surveys. The authors further suggested the information about exports that would soon be available from the various 1987 Economic Censuses Questionnaires be incorporated into the data, as a means of filling in the missing information caused by the EIN not being reported on all SEDs. ITA provided funds for the research, which was called the EDB.

Spinoffs

Invaluable "spinoffs" for Foreign Trade Division's (FTD) quality assurance program were discovered when all of the various foreign trade and economic census files were linked. The techniques used in matching the data about commodity shipments to the exporting firms also provided new insights into the activities of the exporting community. In an internal memorandum (April, 1992), Bruce Walter, Assistant Chief of FTD, summarized the benefits of the EDB to the division's internal quality assurance programs. He wrote,

In recent years, Census has instituted efforts to measure the most serious export quality issue, the nonreporting of exports. Our efforts to estimate nonreporting have involved several indirect measures. The principal measures of nonreporting have come from the reconciliation studies that we have conducted with major trading partners....As a second measure of undercount we have audited export transactions at a number of major ports of exit. We have developed, as a result of the audits, some further understanding of export underreporting. Although we can draw inferences about overall underreporting from these efforts, they have proven to be expensive, time consuming and inexact.

Walter goes on to suggest that,

The EDB provides the opportunity for a much more direct way to measure export undercount. By using the database, we can link the export trade statistics with data on exports compiled from manufacturers and wholesalers during the quinquennial Economic Census. This linkage will allow us to make direct comparisons of export data from two independent sources. .. Although this comparison currently can be made only for data from the Economic Census, we also can build linkages with data collected as part of the ASM...This would provide comparisons between Census years for a substantial segment of export trade...In addition to measuring export undercoverage the database provides a way of measuring the completeness of reported data...Finally, as we begin to discuss alternate methods for collecting trade data, the database will become a valuable tool. Defining and analyzing the exporter population seems a very reasonable first step in beginning research on alternate data sources.

Creating the EDB

Computer Requirements

To develop the EDB within a microcomputer environment, the minimum computer requirements needed include a 386 personal computer (PC) with 400 million bytes of storage, 33 megahertz processing speed, 4 megabytes random access memory (RAM), the OS/2 operating system, and SAS-OS/2 software package (version 6.06). [Recommended computer upgrades: 486DX PC, EISA 32-bit with 1 gigabyte of storage and 16 megabytes of RAM.]

The mainframe computer (UNISYS) was used to aggregate the 9.7 million record SED file into a workable, 2.65 million record, microcomputer file. Aggregation of all applicable numeric fields was based on 13 critical character fields.

Firms and Establishments Selected

The EDB includes all establishments and firms that reported merchandise exports in 1987 from two data sources: The Economic Censuses Questionnaires or the SEDs. The database is made up of:

1. All multiunit (MU) establishments and single unit (SU) establishments or firms who reported export activity on the 1987 Census of Manufactures or Census of Wholesale Trade;

2. The SU establishments who filed one or more SEDs for commodities they exported during 1987 as determined by linking SEDs to the SSEL SU records based on the EIN;

3. The MU establishments who filed one or more SEDs for commodities they exported during 1987 as determined by linking SEDs to the SSEL MU records based on the CFN6 (enterprise identifier), and variations of the SIC, ZIP and state location of the exporter.

The data fields shown in Appendix C -- extracted from three databases: the 1987 SEDs, the 1987 Economic Censuses, and the SSEL -- make up the EDB.

1987 SED EIN File. FTD created a file of "potentially valid" EINs [Record Count (RC) = 194,173] from the SED records. "Potentially valid" is defined as EINs with a 9-digit numeric code where the first two digits reflect a valid Internal Revenue District (IRD). [At this point, it is not known from the EIN if the establishment is a SU or a MU. It was later found that a large number of the EINs were found to be Social Security Numbers, reflecting single individuals who are "one time exporters," or invalid EINs that met the validity criteria.] FTD was provided (in ASCII format) the Economic Censuses and SSEL records. State location of the exporter was derived from the three-digit ZIP codes that concord to one state (accounting for 82.7 % of the total export value and 84.2% of the total number of shipments). For three-digit ZIP codes that concord to multiple states, the state location of the exporter was derived from the three-digit ZIP code only if the state of origin of movement for those SEDs matched a state from the zip-to-state concordance file (accounting for 3.8 % of the total export value and 5.8% of the total number of shipments).]

Manufacturing File. From the final 1987 Census of Manufacturers, Industry Division (IND) provided FTD all MU and SU establishment records with an imputed or reported "Value of Products Exported" greater than "0" or for the MC-9610s, the sample of small manufacturing establishments (SUs only), an imputed or reported "Percent of Products Exported" greater than 0. [RC = (SUs + MC9610s) + MUs -- 33,875 = (12,490 + 611) + 20,774].

Wholesale File. From the Final 1987 Census of Wholesalers, Business Division (BUS) provided all MU and SU establishment records with an imputed or reported value for "Export Sales" greater than "0." [RC = SUs + MUs -- 38,884 = 27,051 + 11,833].

Matching to the SSEL. From the above, FTD created two input files for matching to the SSEL:

1. An unduplicated EIN list of SU establishment records from the manufacturing and wholesaling files, and from the SED records [RC = (SED EINs - Dup EINs) + (Man SUs + MC9610s + Whole SUs) - 218,684 = (194,173 - 15,641) + (12,490 + 611 + 27,051)]; and

2. A list of CFNs corresponding to all MU establishments from the combined manufacturing and wholesaling files. [RC = Man MUs + Whole MUs - 32,607 = 20,774 + 11,833].

Both files were sent to the Directory and Census Mailing List Branch, Economic Processing Division (EPD) and matched to the SSEL. The SSEL, which resides on the Digital Equipment Corporation (DEC) mini-computer network, is a relational database of SU and MU establishments and enterprises (firms) that are broken down into modules of data. EPD provided FTD with all SU and MU records that matched the EIN list and all MU records that matched the CFN list. [The mailing list of establishments for the Economic Censuses is created from the SSEL. Thus, the Economic Censuses files are subsets of the SSEL and a 100 percent linkage for these establishments is realized.] Appendix A shows the flow of establishment records from selection -- matching to the SSEL -- Output Files A, B, and C. The establishment and firm output files were then linked to the SEDs.

Linking SEDs to Establishments and Firms

The basic means of linking the SEDs to the firm and establishment records was through the EIN. The EIN is a 9-digit number used by firms and establishments to report payroll withholding to the Treasury Department. The EIN is also the exporter's number on the SED, and part of the SSEL and Economic Censuses records. Also included in the SSEL and Economic Censuses records is the CFN, an internal 10-digit number that identifies each establishment and firm. As stated before, establishments can be single units (SUs) or multiunits (MUs). SU establishments are firms that operate at ONLY ONE location and possess a unique EIN. Their CFN is a 10-digit number consisting of the EIN preceded by a "0." Thus, linking SEDs to SUs based solely on the EIN is valid.

However, MU establishments are seperate locations within an enterprise (firm). The enterprise uses one or more EINs for payroll withholding, i. e. several MUs within an enterprise may possess the same EIN or there may be separate EINs for each, depending on the structure of the firm. The enterprise operates at TWO OR MORE locations. Each location's CFN is a 10-digit number consisting of a 6-digit number (CFN6) that identifies the enterprise to which it belongs, followed by a 4-digit number that identifies itself. In fact, for a number of MUs, the EIN reported on the SED reflected a sales arm or headquarters location of the enterprise and not really the actual exporting establishment. Thus, linking SEDs to the enterprises of MU establishments based solely on the EIN is plausible, but linking SEDs to specific MU establishments within an enterprise is dubious at best.

Based on the above, the EIN was used to link (1) SEDs to SU establishments or firms and (2) SEDs to

enterprises, followed by a **RELINK** of the enterprise-grouped SEDs to corresponding MU establishments using the CFN6, variations of the SIC, the three-digit ZIP and the state location of exporter to complete the linking process. Appendix B illustrates the linking of SEDs to exporting establishments and firms from three output files: A, B and C.

The summary results in Appendix B show that 60.3 percent of the total 1987 export value from the SEDs were linked to individual establishments. Linkages were achieved as the matching criteria were relaxed. The degree of relaxation is reflected by assigning reliability codes, with "1E" and "1C" being the most stringent. Reliability codes, definitions, f.a.s. "free alongside ship" export values which include inland freight, insurance, and other charges in placing the goods along side the exporting carrier; number of shipments and percents are given in Appendix D.

After the linking process was completed, FTD provided a list of CFN6s taken from the MU records to the Enterprise Statistics Branch, Statistical Research Division (SRD) and to the Directory and Census Mailing List Branch, EPD. Each branch provided all enterprise records that matched the CFN6 list. FTD combined these two files into an enterprise file called File D. The summary results in Appendix B show that additional SEDs, consisting of another 5.1 percent of the total 1987 export value were linked only to enterprises. These additional SEDs were included in all enterprise level statistics.

Results

It was found that 104,564 firms operated 129,883 establishments who exported $241.3 billion in f.o.b. "free on board" or plant value of merchandise during 1987. Eighty-six percent of the f.o.b. value was as reported by manufacturers and wholesalers on their Economic Census forms, and fourteen percent of the value was estimated using the f.a.s. value from the SEDs deflated (where applicable) to an f.o.b. basis for those exporters who did not receive or who did not complete a Census form. Manufacturers accounted for $135.5 billion (56.2%)[reported + estimate = $125.8 + ($11.2 x 0.8738)], wholesalers for $88.4 billion (36.6%)[reported + estimate = $81.6 + $6.8], and retailers and other intermediaries for $17.4 billion (7.2%)[100% estimated] of the total f.o.b. export value.

Of the 129,883 establishments that exported merchandise in 1987, 84,995 (65.4%) filed one or more SEDs; 44,888 (34.6%) did not -- reporting export activity **ONLY** on their Census questionnaires; and 34,724 (26.7%) establishments did both. A breakdown by establishment type and filing status is given in Table 1. Table 1 includes number of exporting establishments, f.o.b. reported and actual values of exports, f.a.s. value of total, manufactured, agricultural, and other exports as reflected by the linked SEDs, and percent of total f.a.s. exports represented by the linked value.

The 1987 U.S. total f.a.s. export value and number of shipments (before revisions) as tabulated from the SEDS was $244.4 billion and 8,790,934 shipments. From this total, 66.3 percent (N=69,354) of the number of enterprises that reported export activity during 1987 were linked to SEDs. The total linkage accounts for 65.4 percent ($160 billion) of the total value and 59.2 percent(5,202,652) of the total number of shipments.

A comparison between the values derived from the establishment records to the values derived from the commodity records shows that one is within 0.4 percent of the other and almost exactly equal to the commodity-based total of $244 billion for exports published in February, 1988. The Industry Division had independently estimated the value of exports by manufacturing plants totaling $136 billion using the Census of Manufactures records supplemented by a sample survey. The estimated value of $135.5 billion by manufactures' exports from the establishment data set in the EDB was within 1.0 percent of the Industry Division estimate. The Commerce Department made public its' findings in a press release issued on June 24, 1992 by Census and ITA entitled, "Commerce Announces Statistical Data Available on Exporters," which provided accurate statistics profiling the U. S. exporter population. An indepth, 40 page ITA report which supplements the press release was made available in August, 1993. In addition, the ITA will package and disseminate state-by-state and consolidated metropolitan area profiles of exporters in September.

Funding

Census has attempted to obtain funds for an ongoing EDB program that would provide current and comprehensive data about exporters without success. Attempts to obtain funds from other sources have had disappointing results. These funds are needed to launch an exporter education program on the necessity of reporting a valid EIN, for the cost of editing the ZIP Code, and for analyzing the data on a state-by-state basis. Most of these costs will disappear if the U.S. Customs Service has as much success with the Automated Export System (AES) as they have had with the Automated Broker Interface (ABI), the electronic data transmission program for 92 percent of all imports, which requires and edits the reporting of the importer's EIN.

Additional matching problems may be encountered with exports to Canada. Effective, January 1990 with the Canadian Data Exchange, the EIN field is not included in the Canadian Import Tape (Canada's import tape is used in lieu of collecting paper documents on U.S. Exports). Instead, a thirty character field reflecting the exporter name, and a seven character field reflecting the exporter ZIP/state location is used. Linking algorithms must be developed to match these fields to the corresponding fields in the SSEL.

Conclusion

The 1987 EDB proved to be a valuable source of information on the structure and composition of the exporting community. A great deal was learned about how exporters were answering the questions on the SED and which firms accounted for significant numbers of exports. The ITA strongly support the technique the authors developed of matching the information about the origin and destination of commodity shipments that exporters reported on their SEDs; the information about operating characteristics that these firms provided on their Economic Censuses questionnaires and the classification, geographic location and corporate structure information Census maintains on the SSEL. When combined, these three sources provided a comprehensive list of nearly all of the firms and establishments engaged in exporting.

Plans are underway to make the EDB an integral part of the 1992 Economic Census. The Foreign Trade Division, the Industry Division, the Business Division, and the Economic Census and Surveys Division will develop a 1992 cross-divisional EDB from which various analyses and publication of reports and press releases can be made. The title "Characteristics of Exporters: 1992" was selected for the primary joint publication. It will take the place of the publication "Selected Characteristics of Manufacturing Establishments That Export: 1987."

References

Farrell, M. & Risha, M. (1992, May). The Exporter Data Base (EDB). The Exporter, Vol 13 No. 01 4-5, 25.

U.S. Department of Commerce, Bureau of the Census (1962), "Survey of the Origin of Exports of Manufactured Products: 1960." [Current Industrial Reports Series M-161(60)-1], Industry Division, Washington, DC.

U. S. Department of Commerce, Bureau of the Census (1992), "Exports from Manufacturing Establishments: 1987." [Manufacturing, Analytical Report Series (AR87-1)], Industry Division, Washington, D. C.

Excerpts from a Decision Memorandum to Deputy Secretary Clarence J. Brown from the Under Secretary For International Trade entitled,"Critical Need to Update U.S. Information Base on Numbers of Manufacturing Firms That Export, By Firm Size, Industry, and State," May 19, 1986.

U.S. Department of Commerce, International Trade Administration, "U.S. Manufactured Exports and Export-Related Employment: Profiles of the 50 States and 35 Selected Metropolitan Areas for 1984." Staff report of the Office of Trade and Investment Analysis, May 1988.

Jack Faucett Associates, "Small Business Exports Of Manufactured Products, 1985." Prepared for the Office of Advocacy, U.S. Small Business Administration, Washington D.C., December 1988.

U.S. Department of Commerce, Bureau of the Census (1988), "Highlights of U.S. Export and Import Trade," Report FT990 (Monthly), Washington, D.C.

U.S. Department of Commerce, Bureau of the Census (1988), "Highlights of U.S. Export and Import Trade," Report FT900 & Supplement (Monthly), Washington, D.C.

Risha, M. P. (1991). A comparison of the "Origin of Movement" Series and the "Exports from Manufacturing Establishments" Series. [Working Papers Series. U.S. Department of Commerce, Bureau of the Census, Industry Division, Washington, D.C.]

U.S. Department of Commerce (1989, July). More than 100,000 exporters ship U. S. manufactured goods, Census Bureau reports. Commerce News, CB89-92. Washington, D.C.: Bureau of the Census, Foreign Trade Division.

Excerpts from an internal memorandum to Foreign Trade Division Chief Don L. Adams from Assistant Chief Bruce C. Walter entitled, "A New Way of Looking at the Quality of the Merchandise Export Statistics," April 8, 1992.

U.S. Department of Commerce (1992, June). Commerce announces statistical data available on Exporters. Commerce News, ITA 92-40. Washington, D.C.: International Trade Administration.

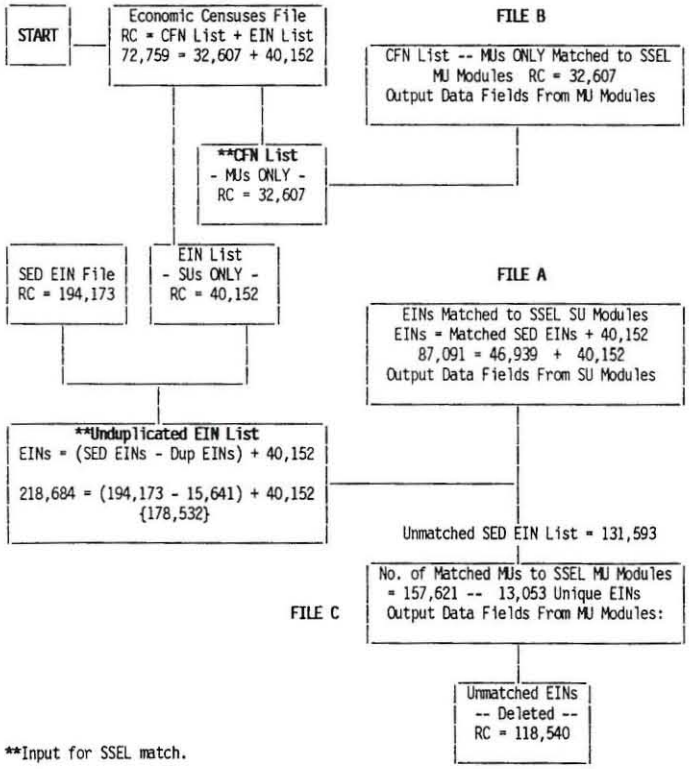
Table 1

Filing Status of Establishments that Exported Merchandise in 1987: (1) Shipper's Export Declarations (SEDs) (2) Economic Census Questionnaire or (3) Both [values in billions of dollars]

Industrial Division of Establishment / Filing Status	Number of Exporters	Pct	Category	F.o.b. Export Value	Pct	Linked F.a.s. Export Value: Man Comm	Pct	Agr Comm	Pct	Oth Comm	Pct	Total	Pct	**Pct Total Exports
Universe of Exporters:	129,883	100	-	$241,272	100	$130,877	100	$11,048	100	$5,465	100	$147,391	100	60.3
(1) Filed Census/No SED	44,888	34.6	-	$26,034	10.8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
(2) Filed SED/No Census	50,271	38.7	-	*$33,859	14.0	$26,481	20.2	$2,480	22.5	$1,867	34.2	$30,829	20.9	12.6
(3) Filed Both	34,724	26.7	-	$181,379	75.2	$104,396	79.8	$8,568	77.5	$3,598	65.8	$116,562	79.1	47.7
Manufacturers:	50,847	39.2	100	$135,486	56.2	$93,050	71.1	$790	7.2	$763	14.0	$94,603	64.2	38.7
(1) Filed Census/No SED	11,294	8.7	22.2	$13,386	5.6	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
(2) Filed SED/No Census	20,014	15.4	39.4	*$9,735	4.0	$10,962	8.4	$80	0.7	$99	1.8	$11,141	7.6	4.6
(3) Filed Both	19,539	15.1	38.4	$112,365	46.6	$82,088	62.7	$711	6.5	$663	12.2	$83,462	56.6	34.1
Wholesalers:	52,630	40.5	100	$88,426	36.6	$24,434	18.7	$8,199	74.2	$2,794	51.1	$35,427	24.0	14.5
(1) Filed Census/No SED	25,476	19.6	48.4	$12,645	5.2	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
(2) Filed SED/No Census	13,746	10.6	26.1	*$6,767	2.8	$5,519	4.2	$830	7.5	$418	7.6	$6,767	4.6	2.8
(3) Filed Both	13,408	10.3	25.5	$69,014	28.6	$18,916	14.5	$7,369	66.7	$2,376	43.5	$28,660	19.4	11.7
Other Intermediaries:	26,406	20.3	100	*$17,360	7.2	$13,393	10.2	$2,058	18.6	$1,909	34.9	$17,360	11.8	7.1

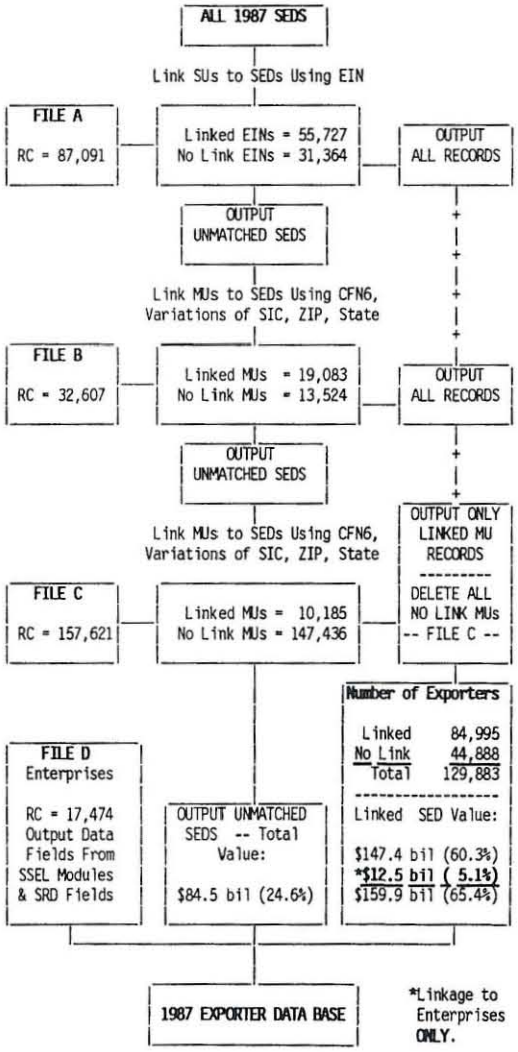
*Estimated f.o.b. value from SEDs. **Linked value as a percent of U.S. exports totaling $244.4 billion.

APPENDIX A
Selection and Matching to SSEL -- Establishments with Export Activity: 1987



**Input for SSEL match.

APPENDIX B -- Linking the SEDs to Exporting Establishments & Enterprises (Firms)



*Linkage to Enterprises ONLY.

SOURCE: Department of Commerce, Bureau of the Census, Foreign Trade Division, Data Analysis and Planning Staff. Contact Name: Michael P. Risha, Phone: (301) 763-2724.

APPENDIX C -- Key Data Fields That Make Up the 1987 Exporter Data Base

SED FILE STRUCTURE

Field	Width	Type	Description
ORSTATE	2	C	State of Origin of Movement
*EIN	9	C	Employer Identification Number
D/F	1	C	Domestic/foreign Shipment Code
SCHB	7	C	Seven-Digit Schedule B Commodity
CTRY	4	C	Country of Destination
DIST	2	C	U.S. Customs District of Export
PORT	2	C	U.S. Port of Export
FTZ	4	C	Foreign Trade Zone
MOT	1	C	Method of Transportation
CONT	1	C	Containerized/Noncontainerized Vessel Shipment Code
R/N	1	C	Related/Nonrelated Party Trade
*SIC	4	C	1972 Four-Digit Standard Industrial Commodity Classification
*ZIP	3	C	Three-Digit ZIP Code -- Location of Exporter
*STATE	2	C	State Location of Exporter
RELIAB	3	C	Code Assigned to SEDs Linked to Establishments and Enterprises
VALUE	8	N	F.a.s. Dollar Value
QTY1	8	N	First Unit of Quantity
SWT	8	N	Shipping Weight -- Vessel & Air
NOFSHIP	8	N	Number of Shipments
CFN	10	C	Census File Number -- **ONLY** for SEDs that were **LINKED** to Establishments and Enterprises
*CFN6	6	C	Enterprise Identifier -- **ONLY** for SEDs that were **LINKED** to Establishments and Enterprises

ESTABLISHMENT FILE STRUCTURE
-- SSEL AND ECONOMIC CENSUS RECORDS --

Field	Width	Type	Description
CFN	10	C	Census File Number
*CFN6	6	C	Enterprise Identifier (Alpha)
*EIN	9	C	Employer Identification Number
NAME1	36	C	Primary Name of Business
NAME2	36	C	Additional Name of Business
STREET	36	C	Number and Street of Mailing Address
PLCE	20	C	City, Town, etc. of Mailing Address
STATE	2	C	State of Mailing Address
ZIP	9	C	ZIP of Mailing Address
PSTRT	36	C	Physical Street Address
*PSTATE	2	C	Physical State Address
PPLCE	20	C	Physical City, Town, etc Address
*PZIP	9	C	Physical ZIP Address
CYTE	2	C	Current Year Employment Size Code
CTOC	2	C	1987 Census Type of Operation Code
CSIC_87	6	C	1987 Census Standard Industrical Classification Code
*CSIC_72	6	C	1972 Census Standard Industrical Classification Code
CLFO	2	C	1987 Census Legal Form of Organization
CSTGEO	2	C	1987 Census State Geographic Code
CCTYGEO	3	C	1987 Census County Geographic Code
CPLCEGEO	4	C	1987 Census Place Geographic Code
CEMP	8	N	1987 Census Total Employment
MEXP_WEXP	8	N	Value of Commodities Exported, 1987 for Manufacturers & Wholesalers Only
MSHP_WSLS	8	N	Value of Commodities Produced or Shipped, 1987 for Manufacturers & Wholesalers Only

ENTERPRISE FILE STRUCTURE
— SSEL AND ENTERPRISE STATISTICS RECORDS —

Field	Width	Type	Description
*CFN6	6	C	Enterprise Identifier (Alpha)
NAME1	36	C	Primary Name of Enterprise
NAME2	36	C	Additional Name of Enterprise
STREET	36	C	No & Street of Mailing Address
PLCE	20	C	City & Town of Mailing Address
STATE	2	C	State of Mailing Address
ZIP	9	C	ZIP of Mailing Address
LFO	2	C	Legal form of Organization
FA	1	C	Foreign Affiliates
FO	1	C	Foreign Ownership
FCC	4	C	Foreign Country Code
EIC	4	C	Enterprise Classification Code
TOTEMP	8	N	Total Number of Employees
TOTEST	8	N	Total Number of Establishments

*Variables and/or variations of used in the linking process.

THE 1987 EXPORTER DATA BASE

APPENDIX D
Results of the SED/Establishment Linking Process for Merchandise Exports, 1987

Reliability Code	F.a.s. Value (mil $)	Pct	Number of Shipments	Pct	Definition
1E	$25,492.2	17.3	1,287,970	26.6	Single unit establishments with unique EINs.
1C	$9,302.9	6.3	418,267	8.7	Multiunit establishments with unique CFN6 (enterprise identifier). Findings indicate that the commodity and zip reported on the SEDs reflect each multiunit establishment's industrial classification and location 95 percent of the time.
2C	$24,415.9	16.6	683,043	14.1	Multiunit establishments with a unique CFN6, 4-digit SIC, and 3-digit zip. Linkage based on enterprise identifier, principal commodity classification and geographic location.
2C*	$18,227.9	12.4	398,295	8.2	Multiunit establishments with the same CFN6, 4-digit SIC, and 3-digit zip. The number of of establishments is accurate for principal commodity exported and geographic location--export value could not be allocated to specific establishments.
3C	$3,002.7	2.0	97,599	2.0	Multiunit establishments with a unique CFN6, 3-digit SIC, and 3-digit zip. Linkage based on enterprise identifier, principal commodity classification and geographic location.
3C*	$1,129.3	0.8	39,289	0.8	Multiunit establishments with the same CFN6, 3-digit SIC, and 3-digit zip. The number of establishments is accurate for principal commodity exported and geographic location--export value could not be allocated to specific establishments.
4C	$3,997.6	2.7	179,638	3.7	Multiunit establishments with a unique CFN6, 2-digit SIC, and 3-digit zip. Linkage based on enterprise identifier, principal commodity classification and geographic location.
4C*	$4,135.9	2.8	137,658	2.8	Multiunit establishments with the same CFN6, 2-digit SIC, and 3-digit zip. The number of establishments is accurate for principal commodity exported and geographic location--export value could not be allocated to specific establishments.
5C	$19,179.4	13.0	560,889	11.6	Multiunit establishments with a unique CFN6, 3-digit zip. Linkage based on enterprise identifier and geographic location.
5C*	$16,656.6	11.3	536,902	11.1	Multiunit establishments with the same CFN6, 3-digit zip. The number of establishments is accurate for geographic location--export value could not be allocated to specific establishments.
6C	$4,725.8	3.2	77,938	1.6	Multiunit establishments with a unique CFN6, 4-digit SIC, and state code. Linkage based on enterprise identifier, principal commodity classification and state location.
6C*	$6,536.0	4.4	52,197	1.1	Multiunit establishments with the same CFN6, 4-digit SIC, and state code. The number of establishments is accurate for principal commodity exported and state location--export value could not be allocated to specific establishments.
7C	$373.4	0.3	14,064	0.3	Multiunit establishments with a unique CFN6, 3-digit SIC, and state code. Linkage based on enterprise identifier, principal commodity classification and state location.
7C*	$604.9	0.4	24,536	0.5	Multiunit establishments with the same CFN6, 3-digit SIC, and state code. The number of establishments is accurate for principal commodity exported and state location--export value could not be allocated to specific establishments.
8C	$902.3	0.6	31,853	0.7	Multiunit establishments with a unique CFN6, 2-digit SIC, and state code. Linkage based on enterprise identifier, principal commodity classification and state location.
8C*	$1,735.3	1.2	50,883	1.1	Multiunit establishments with the same CFN6, 2-digit SIC, and state code. The number of establishments is accurate for principal commodity exported and state location--export value could not be allocated to specific establishments.
9C	$2,076.3	1.4	73,124	1.5	Multiunit establishments with a unique CFN6 and state code. Linkage based on enterprise identifier and state location.
9C*	$4,896.2	3.3	169,085	3.5	Multiunit establishments with the same CFN6, state code. The number of establishments is accurate for state location--export value could not be allocated to specific establishments.
TTl	$147,390.5	100	4,833,230	100	Total linkage with establishment integrity is 60.3% of total export value and 55.0% of total shipments.
Match					
ENT	$12,522.7	5.1	369,422	4.2	Linkage to enterprises **ONLY** using enterprise identifier (CFN6). Percentages based on total exports of $244 billion.

ALTERNATE WAYS OF MEASURING INTERNATIONAL MERCHANDISE TRADE A BLUEPRINT FOR THE FUTURE

V. Crompton, C. Graziadei, D. Kotsovos, S. Mozes, Statistics Canada
S. Mozes, Statistics Canada, 9th Floor, Jean Talon Bldg.,Ottawa, Ontario, K1A 0T6

I BACKGROUND

CUSTOMS BASED TRADE DATA

International trade data represent an integral element of every country's statistical program and are almost universally derived from Customs administrative documents. Because of this, trade data embody a number of unique features.

Foremost, they are among the very few current economic statistics that are drawn from a census of all relevant transactions. Because of this, they lend themselves to a multiplicity of very detailed breakdowns which makes them one of the few economic indicators that has immediate relevance for small as well as large business.

Secondly, the statistical data derived from Customs documents are generally very reliable because there is an immediate operational benefit to business in providing accurate and complete information to facilitate the movement of their goods across international borders and because the provision of this information is enforced by Customs.

And, not least in importance, trade data are virtually free to statistical agencies because they are a by-product of regular Customs operations.

Detailed, accurate and timely trade data are critical to decision making in both the public and the private sector. Governments require trade data to effectively manage taxation and foreign exchange policies and to negotiate multinational trade agreements and tariff negotiations. In the business sector, trade data are used continually to measure market shares, monitor price trends, identify import and export markets and to guard against unfair trade practices.

WINDS OF CHANGE

Recent developments in the international and domestic trading environments, however, such as Europe 1992, the Canada/U.S. Free Trade Agreement, NAFTA, Customs 2000, electronic data interchange and pressure from the business community to reduce paper burden are going to influence both the way in which Customs will collect information in the future and the way in which statistical agencies can derive and produce international trade statistics.

ALTERNATE DATA SOURCES PROJECT

To ensure that it can continue to produce international trade data at the level of detail required by its user community, Statistics Canada has established its *Alternate Data Sources Project* to develop a long term contingency proposal for measuring trade statistics in a situation of significantly reduced or non-existent Customs documentation.

To accomplish this task, the Project has focused its activity on four integral sub-projects:

i maintaining ongoing liaison with Customs to ensure that the requirements of the trade statistical program are considered in the development of all "Customs 2000" initiatives;

ii investigating the documentation and record keeping practices of exporters and importers to determine the availability and accessibility of trade data elements for a survey driven trade statistical program;

iii measuring user demand for trade data and assessing user response to the possibility of reduced detail and timeliness of trade data in the future in order to evaluate the impact of a survey driven trade program on the utility of the data; and,

iv developing and analyzing well defined and properly stratified frames of exporters and importers as a basis for survey methodologies and cross-sectional analysis.

The analysis and integration of these sub-projects provide the necessary framework for the development and evaluation of survey-based strategies for the trade program in terms of both collection methodology and user requirements.

II ANALYTICAL FINDINGS

USER CONSULTATION

The user consultation program substantiates that the demand for trade data is widespread at all levels of aggregation. It is in the business sector, however, that the requirement for detailed and timely trade statistics is especially crucial.

Users who depend heavily on trade data are willing to provide whatever support they can to preserve the trade statistics program as it presently exists. Their strong conviction is that any conversion to a survey-based trade program would seriously impair the utility of the data, particularly in terms of the current quality and detail. Frequency of release seems to be the only field for compromise and even here a move from monthly to quarterly would have a critical impact on those users who require trade data to monitor immediate market conditions.

RESPONDENT CONSULTATION

In tracking the linkages between commercial and Customs documentation, it is clear that the commercial invoice is the key document for deriving the basic data elements required by Statistics Canada - commodity, value, quantity, origin and destination. The invoice, however, will not consistently provide the HS code, mode of transport, freight detail or Customs clearance information.

The commodity detail provided by the 10-digit HS code is one of the key attributes that make trade data so valuable to the user community. Because the 10-digit HS code incorporates tariff identifiers, it is a required part of the commodity description on all Customs documents. The majority of companies, however, use the services of a broker in assigning HS codes and will have limited in-house familiarity with or need for the classification system.

Mode of transport and freight detail are not always included on the invoice but are usually available on related transportation documents. Customs clearance information (port of clearance, date of clearance) would also require reference to transportation documents.

While it is apparent that both importers and exporters can provide the necessary data elements for maintaining the trade statistics program, the ease with which respondents can compile these data elements for survey purpose varies considerably. For firms with integrated sales, purchasing and traffic systems, the burden of supplying this information to Statistics Canada on a regular survey basis would be minimal, provided sufficient lead time is given to adjust their output systems. Firms that are decentralized in structure or operating at a low level of automation (and these represent the majority) anticipate a substantial "response burden" being imposed by a Statistics Canada survey questionnaire or any other collection method.

IMPORTER AND EXPORTER DATABASES

The principal objectives in developing importer and exporter databases were to establish an information base on the trading universes and a mechanism for the storage and maintenance of frame information in a survey environment. The utility of these databases, however, has expanded into the analytical field and they have proven to be powerful tools for longitudinal and cross-sectional analyses of trading activities and, via linkage with other Statistics Canada files, offer the potential to analyze trade, production and ownership data in a fully integrated environment.

To date, the importer and exporter databases have been used to carry out several analytical studies: the size and structure of the trading communities; the impact of selected survey thresholds on commodity and value coverage; trade by industry sector; trade among business affiliates; and, a longitudinal analysis of importers.

This paper, however, focuses on the size and structure of the trading communities and the impact of selected survey thresholds on commodity and value coverage, because these are the two studies which will most influence the evolution of a strategy for the future within the context of the *Alternate Data Sources Project*. Although no distinction is made in this paper between trade among NAFTA partners and those outside NAFTA, the final blueprint should be equally applicable to the measurement of trade within and outside the trading block. For measurement of trade with off-shore countries, a Customs based solution may still be applicable.

Size and Structure of the Canadian Trading Community

The Canadian trading community comprises a small number of trading units[1] which generate a large share of the total import/export value and transaction volume coupled with a large number of trading units who generate only a small share of value and volume.

Table 1

Canadian Imports and Exports for 1990:
Number of Importers and Exporters, Their Size and Associated Trade Values. Preliminary

	Imports				Exports			
Size of Trade Value Range($)	No. of Trading Units	%	Value $mill.	%	No. of Trading Units	%	Value $mill.	%
Over $50 M	277	0.2	77,411	54.3	352	0.5	92,510	68.1
$5 M to $50M	2,809	1.9	37,088	26.0	1,995	2.7	28,524	21.0
$0.5M to $4.9M	13,871	9.5	21,617	15.2	7,010	9.4	11,453	8.4
Less than $0.5M	128,877	88.4	6,518	4.5	64,823	87.4	3,463	2.5
Total	**145,832**	**100.0**	**142,634**	**100.0**	**74,180**	**100.0**	**135,950**	**100.0**

In 1990, there were 145,832 trading units who imported $142.6 billion and 74,180 trading units who exported $135.9 billion. The concentration of these trading units by value range is illustrated in Table 1. Large importers, those importing more than $50 million per year, represent 54.3% of total imports but only 0.2% of the total number of importers; and, a small number of exporters, representing 0.5% of the total, account for 68.0% of total export values. Conversely, 88.4% of the total importer community generate only 4.5% of import values and 87.4% of the exporters account for only 2.5% of export values.

This dichotomy in the distribution of trading units, which is evident in both value ranges and in volume of shipments, indicates that a survey of a small proportion of the trading population is likely to yield high coverage in terms of value, shipments, and transactions.

Survey Scenarios: Potential Value and Commodity Coverage at Different Thresholds

To assess the value and commodity coverage that might be achieved under different survey scenarios, three arbitrary thresholds - $1 million, $5 million and $10 million - were examined. The population in each scenario includes all trading units with annual imports or exports greater than or equal to the survey threshold.

Table 2, which measures the impact of the three thresholds on sample size, value and commodity coverage[2], shows that a $1 million threshold for imports would yield 92.5% value and 99.1% commodity coverage with a sample of 11,044 trading units (7.6% of the total universe) and that for exports, this $1 million threshold would provide 96% value but only 83.1% commodity coverage with a sample representing 8.4% of the universe (6,463 trading units). A $10 million threshold, with a sample of 1,571 importers (1.1% of the universe), could achieve 72.9% value and 94.3% commodity coverage while for exports, the sample size of 1,394 (1.9%), could yield value coverage of 84.1% but only 57.6% commodity coverage.

Table 2

Survey Scenarios: Imports 1990. Preliminary

	No. of Trading Units	**%**	**Value $bill.**	**%**	**No. of HS6 Commodities**	**%**
Total	145,832	100.0	142.6	100.0	5,012	100.0
M>1Mill.[3]	11,044	7.6	131.9	92.5	4,965	99.1
M>5Mill.	3,084	2.1	114.5	80.3	4,856	96.9
M>10Mill.	1,571	1.1	103.9	72.9	4,727	94.3

Table 2 - cont'd

Survey Scenarios: Exports 1990. Preliminary

	No. of Trading Units	%	Value $bill.	%	No. of HS6 Commodities	%
Total	74,180	100.0	135.9	100.0	4,135	100.0
X>1Mill.[4]	6,463	8.7	130.4	96.0	3,435	83.1
X>5Mill.	2,347	3.2	121.0	89.0	2,781	67.3
X>10Mill.	1,394	1.9	114.3	84.1	2,383	57.6

These thresholds exclude many exporters and their commodities. Some commodities disappear completely, while of others, only a small percentage remains. Commodity disappearance for both imports and exports rises as the threshold is increased and the selection of a static threshold denotes an inverse relationship between commodity coverage and sample size.

A closer look at commodity coverage is provided in Table 3, which shows how many HS-6 commodities would be lost completely under the different survey scenarios.

Table 3

Lost Commodities: Imports 1990. Preliminary

	Value $mill.	%	No. of Comm.	No. of Lost Comm.	%	Value of Lost Commodities $mill.	%	Total Imports
Total	142,643	100.0	5,012		100.0			
M>1Mill.	131,942	92.5	4,965	47	0.9	2.5		0.002
M>5Mill.	114,498	80.3	4,856	156	3.1	53.6		0.038
M>10Mil.	103,943	72.9	4,727	285	5.7	133.5		0.094

Lost commodities: Exports 1990. Preliminary

	Value $mill.	%	No. of Comm.	No. of Lost Comm.	%	Value of Lost Commodities $mill.	%	Total Exports
Total	135,951	100.0	4,135		100.0			
X>1Mill.	130,436	96.0	3,435	700	16.9	73.2		0.05
X>5Mill.	121,035	89.0	2,781	1,354	32.7	547.5		0.40
X>10Mill.	114,317	84.1	2,383	1,751	42.3	1,284.9		0.91

For all thresholds, exports show a higher proportion of lost commodities. The range for lost commodities in exports extends from 17% to 42%, implying that large exporters specialize in a limited number of product lines while smaller traders export more diversified commodities. In comparison, the 1% to 6% range in commodity loss in imports implies that large importers not only dominate but are also importing diversified product lines.

For imports, the lowest threshold would result in a loss of 47 commodities (0.9%) associated with a value of $2.5 million, representing 0.002% of total imports. The same threshold for exports would cause the disappearance of 700 commodities (16.9%) worth $73.2 million, representing 0.05% of total export value.

The $10 million threshold on imports would raise the number of lost commodities to 285 (5.7%). These lost commodities represent $133.5 million, (0.09%) of total import values. The $10 million threshold for exports raises the number of lost commodities to 1.751 (42.3%). While these lost commodities account for $1.2 billion, they represent only 0.9% of total value exports.

The H.S. chapters affected by the selection of these static thresholds involve raw materials, textile and manufactured products. Chapter 84 (Machinery and Mechanical Appliances, parts thereof) shows the highest value of lost commodities for both imports and exports, for almost all thresholds. Other chapters affected include animal, vegetable and mineral products. These chapters, however, comprise many small importers and exporters the fall outside the threshold.

Nonetheless, while the proportion of commodities that disappear appears to be significant (especially for the export threshold of $10 million), the commodities that disappear represent less than 1% of the total value of imports or exports, regardless of the selected threshold.

A survey of a small number of trading units could achieve adequate levels of both value and commodity coverage. The commodity losses resulting from the different thresholds, although higher for exports, are not excessive and account for relatively low values. Moreover, their composition shows that no major commodities in terms of trade value are included. It would be possible, therefore, by refining various thresholds, to achieve an acceptable balance between value coverage, commodity disappearance and the sample size.

III A BLUEPRINT FOR THE FUTURE

As trade statistics will still be needed if the traditional revenue protection role of Customs disappears, an alternative system, capable of providing information similar to that based on Customs sources, will have to be developed.

Analyses of the supply and demand for trade statistics and of the size and structure of the trading communities shows:

- that there is a demonstrated demand for reliable detailed monthly trade statistics;
- that respondents can generally provide the information required for statistical purposes but that in a survey environment the timeliness, quality and detail of this information may deteriorate;
- that the structure and behaviour of the trading communities is such that different survey methods, concepts and methodologies would have to be used for collecting data from the various distinct segments of the trading communities;
- that a survey approach providing detailed, complete and monthly data covering all traders may not be feasible and or justifiable given the size and structure of the trading communities;
- that the statistical information should be collected from respondents (traders) and not from documents accompanying the goods since merchandise may cross borders without accompanying documents; and,
- that new and innovative data collection, such as computerized reporting systems and survey methods, may be required.

At the same time, however, the *Alternate Data Sources Project* has to ensure that the following requirements are considered in developing its blueprint for a "survey driven trade statistical program":

- to minimize and/or reduce response burden;.
- to preserve timely, detailed information for a very high portion of trading activity;
- to maintain international comparability;
- to match information requirements to the internal commercial systems of respondents;
- to implement highly flexible data collection procedures that are automated to the fullest extent; and,
- to safeguard historical continuity.

PRINCIPAL ELEMENTS OF THE BLUEPRINT

- Large traders can provide detailed, timely information to Statistics Canada from their internal commercial systems. These traders show consistent behaviour over time, represent a very large proportion of total trade activity and, as a general rule, are good respondents. There is no reason to believe that they would not be able to provide the same of amount of detail with the same periodicity as present. However, in order to facilitate the collection of data in a constantly evolving informatics environment, Statistics Canada would have to develop highly flexible collection instruments such as custom made software and must also be able to communicate with traders using Electronic Data Interchange (EDI).

- Small traders account for a very small percentage of total trade - 129,000 importers, for example, account for almost 90% of the universe but less than 5% of the total value of imports. They are reluctant respondents and are often not capable of providing the information required. Their behaviour pattern, however, is very similar to that of large traders.

Given their relative importance and stable behaviour the following options can be considered: exempt them from statistical reporting; obtain annual summary data from them either on a census or on a sample basis and estimate monthly movements based on previous behaviour patterns or the behaviour patterns of larger traders; or, use the Goods and Services Tax (GST) system to estimate their activity.

- Medium size traders represent an important contribution to trade flows, they are numerous and their reporting capabilities vary considerably. Any survey strategy would have to take into account: the number of traders and commodities they trade; data availability and ease of obtaining it; response burden; resource availability; and, the demand for data. Based on the present imperfect knowledge of this segment of the population the following options are being considered:

i to collect monthly summary (not transaction based) statistics from all respondents regardless of their size or the commodities they trade;

ii to collect monthly summary statistics from a sample of traders where the sample size would be a function of the size of the trader;

iii to collect monthly summary statistics from a sample of traders where the sample size would be a function of the size of trader and the commodities traded; or,

iv to collect summary data either on a census or a sample basis on a quarterly basis only.

Obviously, the option of accepting transaction level details in machine readable form is an option not to be discarded.

Table 4 provides a brief comparison between the existing Customs based system and the possible future survey based systems.

Table 4

Comparison of Existing (Customs) and Future (Survey) Systems

	Present (Customs)	Future (Survey)
Data Source	Administrative, Customs	Survey form(s), respondent
Timing	Time goods cross border	After goods crossed border
Statistical period	Calendar month	Calendar month, year or quarter
Responsible party	Exporter/Importer (may be delegated to others)	Exporter/Importer
Coverage	Full coverage of commercial imports Almost full coverage of exports	Partial coverage of both exports and imports
Frame	Not known	Well defined
Data elements	Extensive	Reduced
Quality	Imports better and more complete than exports	Coverage of exports and imports may be the same but quality of exports better than imports
Data collection	Tapes and paper documents	Various (eg. E.D.I., Customs software, etc.)

Three principal differences exist between the two systems. These can be summarized as follows.

i Relationship between respondents and Statistics Canada

Presently there is very little interface between traders and Statistics Canada as compliance (for both imports and exports) and the quality of imports data are the responsibility of Customs. With a survey based system, a direct and close relationship will have to be established among respondents and the statistical agency.

ii Establishment of a frame of traders.

Establishing a universe of shippers and consignees involved in trade is essential in developing and implementing a survey based statistical system. The definition and delineation of the respondent (e.g. firms, establishments) both from a methodological as well as from a legal point of view is critical.

Derivation of a frame for previous years, based on existing Customs documentation, presented a difficult but manageable task. What is imperative, however, is the development of an update mechanism independent of the present Customs sources. One option may be to use the GST filing system with modifications to allow the separate capture of exporter identification information whenever an exporter applies for a GST drawback. Identification of importers via the GST system is a viable option.

iii Reduce Response Burden

The present statistical system is a by product of the traditional role of Customs which has been one of "protection of revenue". The same data that Customs requires in its day to day operation lends itself admirably to the compilation of trade data, without putting extra burden on the respondent. It is recognized that with liberalization of trade rules (F.T.A., NAFTA), there will be considerable pressure on the statistical agencies to reduce and minimize the response burden on business. The blueprint envisaged reduces the response burden by exempting certain respondents from reporting requirements, by decreasing the number of data elements demanded, by introducing highly automated reporting systems and adopting sampling methodologies, and by possibly changing to quarterly periodicity.

IV CONCLUSION

The development and implementation of a survey based trade statistical program presents an formidable undertaking which is critically dependent on the successful completion of the following key tasks.

- Develop an update mechanism for the databases (births).
- Profile all units on the frame and define and delineate reporting units
- Promote the concept of a survey driven program to users and respondents
- Negotiate and establish criteria for i) inclusions and exclusions ii) value thresholds for defining large, medium and small traders and associated data elements and periodicity, and iii) thresholds for commodity coverage.Ideally, it would also be of benefit to align statistical and fiscal obligations under the GST in order to establish "benchmarks" and "checks" in a situation where the fiscal population (G.S.T.) would correspond to the statistical population of exporters and importers.The *Alternate Data Sources Project*, as described, addresses the issue of how to derive trade statistics without Customs documentation. It assumes the worst case scenario and provides a contingency plan for such an occurrence. Even though this worst case scenario may never transpire, it is likely that the Customs based statistics will undergo significant changes over the next few years, a few of which include: self assessment combined with audit; periodic accounting; summary reporting; and, greater use of E.D.I., etc. The blueprint for the future, consequently, will also provide a contingency plan for measuring trade statistics in a situation of substantially reduced or modified Customs documentation.

The successful completion of the *Alternate Data Sources Project* in late 1993 will prepare Statistics Canada to start moving towards a survey driven trade statistics program. The importer and exporter databases, which support the critical frame information for a survey driven program and which already provide a means for controlling export undercoverage and estimating for data, will also serve as powerful tools to analyze trade, production and ownership data in a fully integrated environment.

Statistics Canada is not a pioneer in the endeavour to develop a survey based trade program. The European experience provides an important prototype for us to study, and fortunately, we do not currently face the same stringent implementation schedule as our European counterparts. Against these principal assets of experience and time, however, balances our most serious liability. Canada has no parallel for the Value Added Tax (VAT) to serve as the framework for collection activity and as a mechanism for monitoring its accuracy and coverage.

1. A trading unit is defined as an entity engaged in importing or exporting merchandise. It may be an individual, a company, establishment, an agent, or division of a company, government agency, etc.

2. Commodity coverage is estimated on an "all or nothing" basis. Commodities are considered "covered" when there are values greater than zero reported in a selected target population.

3. M>1(5,10)Mil denotes imports of $1(5,10) million or more.

4. X>1(5,10)Mil implies exports of $1(5,10) million or more.