# 18th MEETING OF THE VOORBURG GROUP

**Tokyo, Japan**
**October 2003**

## SESSION 2: PRODUCER PRICE INDEX FOR SERVICES

# PRODUCER PRICE INDEXES FOR COMPUTER SYSTEMS DESIGN, DATA PROCESSING, AND SOFTWARE PUBLISHING

**Fred Barzyk**
**Statistics Canada**

## Introduction

In 1993, Statistics Canada (STC) launched the Informatics Professional Services Price Index (IPSPI) which collects financial, wage and contractor fee information that is used to produce price indexes measuring changes in prices for informatics professional services such as hardware and software consultancy, computer facilities management, and system maintenance.

Upon its launch and subsequently for the next several years, the IPSPI only collected information and produced price indexes for one particular informatics category— NAICS 541510: *Computer Systems and Related Services* (where NAICS stands for the North American Industry Classification System). Under this classification, this industry consists of establishments primarily engaged in computer information technology consultancy, development of custom software, systems and network design, systems development and analysis, computer programming to meet a customer's specifications

Services Division at STC, through its *Annual Survey of Software Development and Computer Services*, has been collecting and producing economic activity statistics on three NAICS categories (since 1997 on the NAICS format)[1],

NAICS 54151: *Computer Systems Design and Related Services*
NAICS 51121: *Software Publishers*
NAICS 51821: *Data Processing, Hosting and Related Services*.

In 2002, Prices Division added the two NAICS categories (51121 and 51821) to their survey in order to produce a complete set of complementary price indexes which could be used as deflators for this group of activity statistics.

## 1. Industry Output

According to the NAICS 2002 definition (see Appendix A for more details) the primary activities for the *Computer Systems Design and Related Services* are to provide expertise in the field of information technologies through activities, such as writing, modifying, testing and supporting software to meet the needs of a particular customer. *Software Publishers* are generally those firms engaged in publishing computer software, usually for several clients and often referred to as 'packaged software', while establishments classified as belonging to *Data Processing, Hosting and Related Services* provide hosting or data processing services to their clients. Of the three types services, the work carried out by firms in the *Data Processing, Hosting and Related Services* group could be considered as more routine and production-oriented than the other two services, where

[1] Actually, the Services Division series are, as of this date, still based on the 1997 NAICS. The price index series discussed in this paper are based on the 2002 NAICS. Other than the change in the NAICS 51821: Data Processing, Hosting and Related Services (2002) from NAICS 51421: Data Processing (1997), there is little difference.

development programming is the main product and therefore the average skill level of the employee is much higher.

The output of these three industries combined rose from 1.5% of nominal GDP in 1997 to 2.5% in 1999. Individually, *Computer Systems and Related Services* has held the largest share (1.7% of nominal GDP in 1999) when compared *Software Publishers* (0.6%) and *Data Processing* (0.2%).

From a price index perspective, the trend in the series for *Computer Systems Design and Related Services* has been upward, with an average annual growth rate of 4.1% over a ten year period . Preliminary results for 2002 show an increase of 1.4% (see Figure1) while preliminary results for the two additional NAICS result in *Software Publishers* rising by 2.8% and *Data Processing* by 2.3%.

The indexes for the two sub-components, labour cost and the realized net multiplier (or change in profitability), are presented in Figure 3 and Figure 4. For *Computer Systems Design and Related Services* the change in the index for labour cost is positive throughout the entire period, while the change in the realized net multiplier has been negative for the most part. For *Software Publishers*, preliminary results for 2002 show an increased labour cost and declining profits, while for *Data Processing*, profitability improved.

## 2. Index Methodology

### *2.1 Prices*

The IPSPI relies on the input approach to price index construction. Prices collected for the IPSPI represent the input costs of labour and the realized profit for the firm. The labour cost is calculated as the geometric mean of the firm's contract fees and wage rates for the year, while the profit portion reported is used to derive the realized net multiplier. Both of these inputs are combined using a geometric average to arrive at a total price index. The most recent two years of published indexes are subject to revision.

The labour cost index is calculated as a weighted mean of changes in the wages and salaries of employees and in the daily fees paid to their contract workers.[2] To produce the labour cost index, information is collected on the annual average of changes in wages and salaries for employees and in the fees paid to contract workers.

The realized net multiplier index measures changes in the mark-ups to cover profit and overhead charged on labour and other items. This index is calculated from ratios of the annual revenue from contracts for informatics professional services to the expenses incurred to complete these contracts. This index is intended to measure the impact of market conditions on prices for informatics professional services.

[2] Research revealed that most firms use some combination of employee and contract labour to work on contracts, so their (contract labour) inclusion is warranted.

*2.2 Collection*

The collection of price information is carried out annually, through a questionnaire mail-out with telephone follow-ups. The timing of the mail-out and the subsequent collection of data are determined by the fiscal year of the respondent, so that in actual fact there can be three to four mail-outs for a given reference year. This differs from other price surveys at Statistics Canada (e.g. Consumer Price Index and Industrial Product Price Index) in that they have different sources of price information (i.e. product catalogues, pricing schedules or observed prices) more readily available on a current basis. Relying on the fiscal year end information leads to a trade-off between timeliness and the size of revisions.

## 3. Weights and sampling

*3.1 Aggregation Weights*

The current weighting scheme is based on the number of establishments (or units) within a company, a province/territory or at the national level depending on the level of aggregation (see Figure 5). At the lowest level, the establishment level, relatives for the labour cost and realized net multiplier are calculated based on the data from respondents. The labour cost index is actually an average of two components—the average annual change in salaries and wage rates and the average annual change in fees paid to contract workers—weighted together by the expenses reported for both.

The geometric mean is then applied across all units to arrive at a company level estimate. The number of units within a company are then applied to aggregate up to the provincial/territorial level. Finally, the number of units within a province or territory is used to produce a national index.[3]

The decision to switch from revenue-based weights to unit-based weights arose for two reasons. First, there is a significant degree of birth/death activity in this industry in a given year, with many mergers/acquisitions, making longitudinal sampling difficult. Compounded by this is the fact that the revenue reported and residing on the STC Central Frame Database is often incorrect or outdated due to the high company turnover. Relying on the number of units within a company represents the more stable and certain of the options for a weighting strategy over time.

*3.2 Sampling*

For the IPSPI, the target population consists of all establishments primarily engaged in informatics professional services, as identified on Statistics Canada's Central Frame Database. The respondents are selected through a cut-off sample survey based on the significance of their operating income and based on the region they are located in. Prior

[3] While weighting occurs at the provincial level, the indexes are published at the Canada level only.

to the addition of NAICS 51121 and 51421, the sample size was approximately 400 establishments, whereas after the addition, the sample size grew to just over 1,000 establishments. The sample for each of the three NAICS is drawn every year.

Cut-off sampling is a form of judgement sampling where all units falling above a predetermined revenue cut-off range are chosen to be in the sample. Each unit represents itself with a "sampling" weight of one. Unlike probability sampling, with cut-off sampling there is a higher sampling bias since not all of the population has representation. On the positive side, however, there is no sampling variance to contend with.[4] In general, cut-off sampling can be used when:

1. the index is strongly influenced by the changes in the items or establishments which contribute the largest revenues;
2. the allowable sample size is small due to cost factors;
3. there is "reliable" information on the revenue of all businesses in the population, and
4. the revenue of the population is skewed,

all of which is the case of the IPSPI. With such a volatile industry frame, cut-off sampling provides an easier method for accommodating mergers and acquisitions. Estimation becomes much simpler than with a probability sample, mainly because there are no sampling weights to calculate.

One drawback with cut-off sampling is that there is no easy way to estimate the underlying sampling bias incurred by a specific cut-off sample. If the distribution is *not* skewed, the estimates can be of very low quality.

Cut-off sampling can also increase response burden. The cut-off sample is not designed to rotate respondents, and the only way for a rotation to occur is to have a dramatic shift in revenue distribution that results in a current respondent falling below the cut-off line, or in having the cut-off line shift due to substantial improvements of new and existing companies, or both. Efforts are made to control response burden and one initiative involving large companies is to request information from only one establishment per enterprise per province and impute the data from this response to all establishments of the same enterprise in the province.

## 4. Issues in maintaining constant quality

Given the pricing methodology of the IPSPI, there *is no direct method or effort* to control for quality change. With an indirect approach, the price is constructed based on the sum of labour inputs and profit inputs across all contracts for a respondents' fiscal year. Rolling all the activities into one has the unfortunate result of tangling quality change with price change. As a result, the index will be biased upward when quality is increasing, and biased downward when the opposite occurs.

---

[4] This discussion on cut-off sampling is really a synopsis of a review of the subject written by Susana Rubin-Bleuer and Stephen Heney of the Business Survey Methods Division at Statistics Canada.

Early on in the development of the IPSPI, the issued of controlling for quality change was addressed from two perspectives, *contract pricing* and *model pricing.* These pricing methods were discussed with various industry members and associations, in order to assess their viability from a respondent's perspective.

The results led to the realisation that *contract* pricing—tracking the price change in a unchanging contract or bundle of contracts over time—would be very difficult to pursue in this type of service industry since nearly each contract is unique to the job required. The value of a contract typically comprises the days of a professional's time, their per diem rates, as well as the value (cost plus mark-up) of other charges to the client necessary to fulfill the terms of the contract (e.g. travel, software, hardware). The per diem rates cover daily salary/wage and benefit or fee (contract worker) charges and a mark-up for profit and overhead.

Under the contract pricing approach, firms are asked to select contracts that are representative of the type of work that they do and to price these contracts over time. However, the general indication from the informatics professional services firms contacted was that indexes derived from re-priced contracts would not provide a good source for measuring price change, mainly because such contracts don't exist to the degree the statistician would like to see. In fact, any contracts which could be tracked over time would be very small as a proportion of the overall business, or if significant, would not remain so for long since rapid changes in hardware and software technology continually impact on what and how professional services are provided.

The other option considered was the use of *model* pricing, where the hypothetical 'bid' price for reproducing a contract is collected through time (i.e. what a firm would bid as a price for carrying out the same contract in another time period). In reality, actual prices can differ significantly from bid prices due to such factors as competitive conditions and the importance of the client, and this difference can be very large for the relatively smaller companies that make up the majority of firms offering these types of services. As a result, the firms interviewed did not see the usefulness of an index based on either approach and felt that they would impose a significant response burden.

## 5. Price measurement challenges.

The most difficult challenge to this survey is to establish and maintain a 'clean' and up-to-date survey frame from which to draw a robust sample. Experience to date has shown that a high rate of company turnover makes this difficult. Past samples have suffered from a high percentage of out-of-scope, dead and merged establishments selected. For instance, these three categories together represented 48% of the sample selected in 1997, 47% of the sample selected in 1998 and 44% of the sample selected in 1999. This has improved somewhat in the current sample (now 33%) with better coding and more resources dedicated to clean the frame. To combat this problem, a sample will be drawn every year.

Respondent relations also poses difficulties. As with any survey, contacting and convincing respondents to participate can be demanding. Recent changes to the questionnaire have made reporting easier, and increased telephone contact with respondents has been effective, resulting in the response rate increasing from 59% for the 2001 reference year to 72% for 2002.

Trying capture and adjust for quality change represents a long term challenge. Given the complexity of this service, one option being considered would be to conduct a special study of the industry periodically and then use the results to benchmark quality change.

**Figure 1 - Informatics Professional Services Price Indexes**

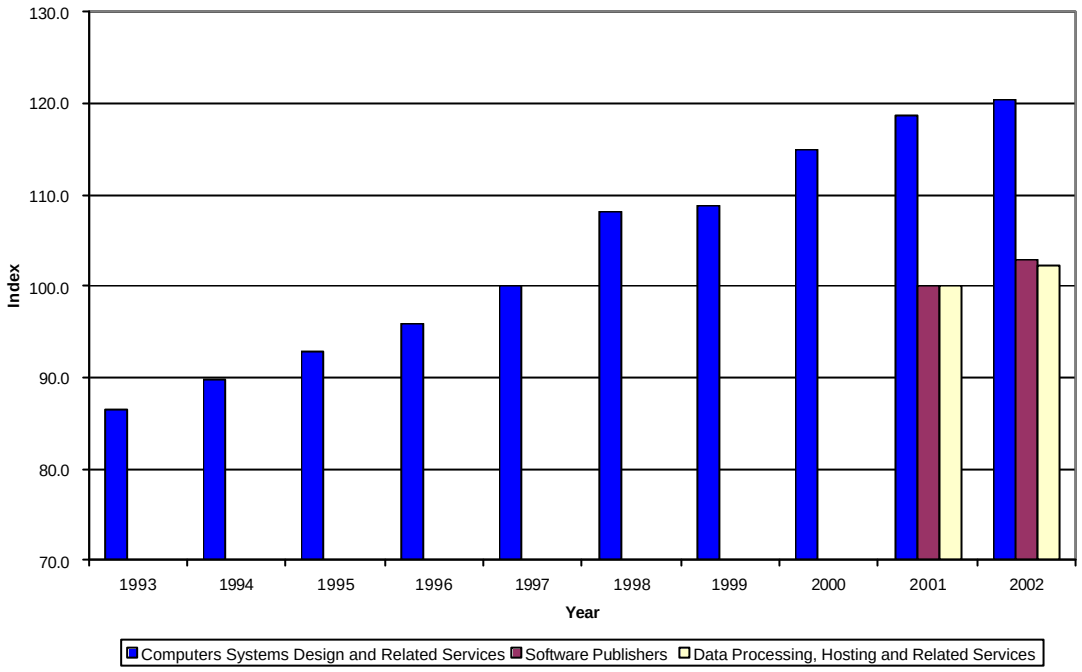


**Figure 2 - Percent Change in IPSPI for Labour Cost**

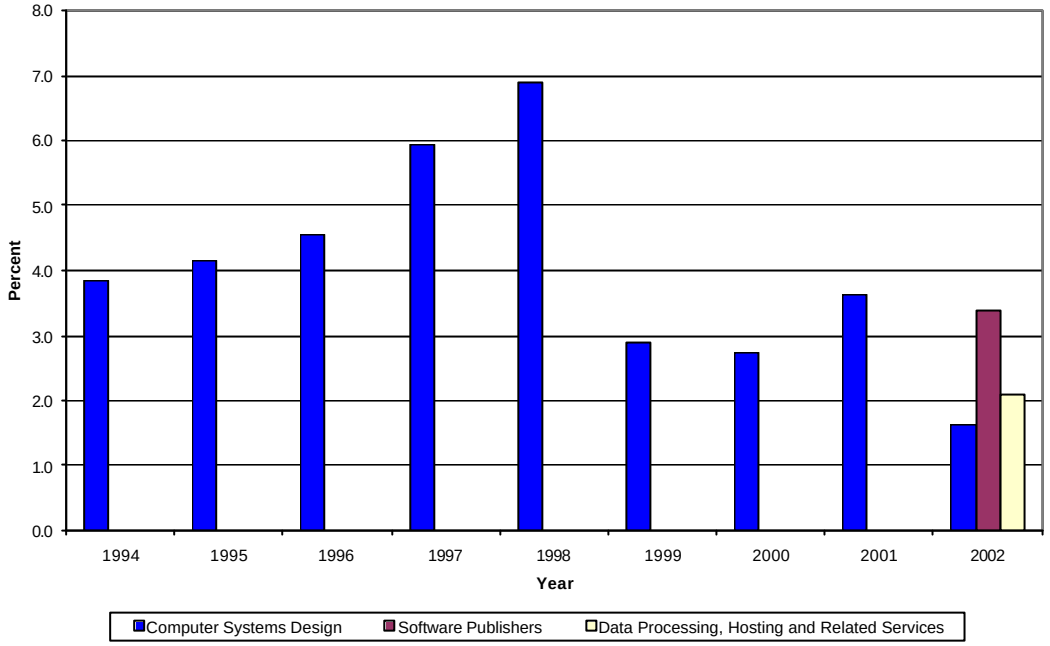
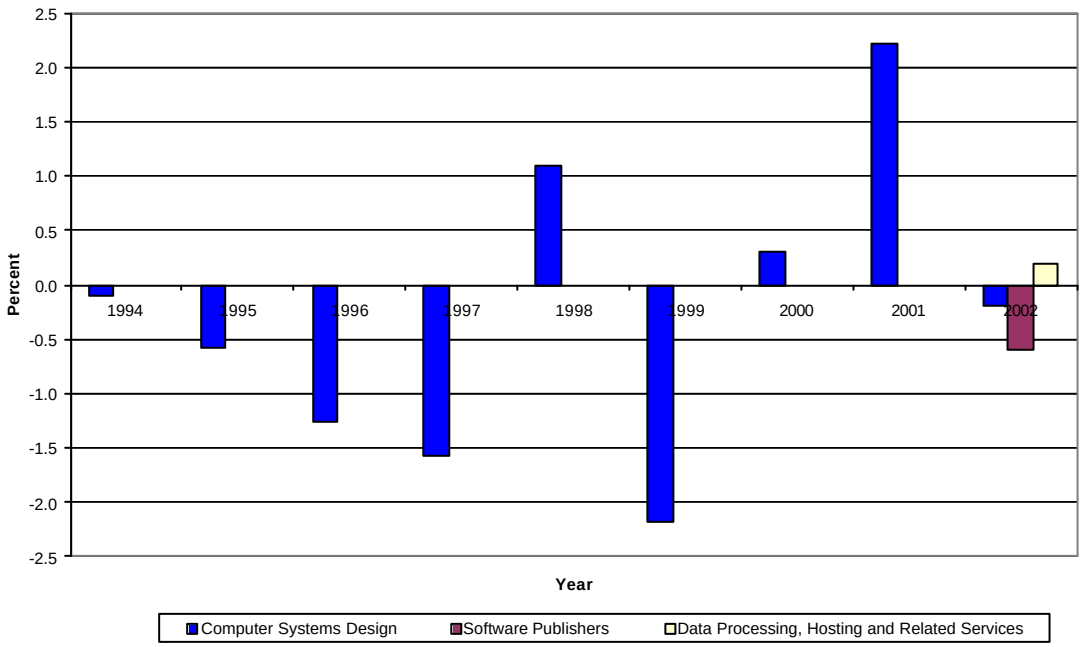



**Figure 3 - Percent Change for ISPI for Realized Net Multiplier**



## APPENDIX: North American Industry Classification (NAICS) 2002

### Classification Definitions

### 54151 Computer Systems Design and Related Services

This industry comprises establishments primarily engaged in providing expertise in the field of information technologies through one or more activities, such as writing, modifying, testing and supporting software to meet the needs of a particular customer, including the creation of Internet home pages; planning and designing computer systems that integrate hardware, software and communication technologies; on-site management and operation of clients' computer and data processing facilities; providing advice in the field of information technologies; and other professional and technical computer-related services.

Exclusion(s): Establishments primarily engaged in:

- retailing computer hardware and software and providing support services (44312, Computer and Software Stores)
- publishing packaged software (51121, Software Publishers)
- providing data processing services (51821, Data Processing, Hosting, and Related Services).

### 51121 Software Publishers

This industry comprises establishments primarily engaged in publishing computer software, usually for multiple clients and generally referred to as packaged software. Establishments in this industry carry out operations necessary for producing and distributing computer software, such as designing, providing documentation, assisting in installation and providing support services to software purchasers. These establishments may design and publish, or publish only.

Exclusion(s): Establishments primarily engaged in:

- mass duplication of software (33461, Manufacturing and Reproducing Magnetic and Optical Media)
- reselling packaged software (41731, 44312, )
- publishing software exclusively on the Internet (51611, Internet Publishing and Broadcasting)
- providing access to software for clients from a central host site (51821, Data Processing, Hosting, and Related Services)
- custom designing software to meet the needs of specific users (54151, Computer Systems Design and Related Services).

### 51821 Data Processing, Hosting, and Related Services

This industry comprises establishments primarily engaged in providing hosting or data processing services. Hosting establishments may provide specialized hosting activities, such as web hosting, streaming services or application hosting, or may provide general time-share mainframe facilities to clients. Data processing establishments may provide complete processing and preparation of reports from data supplied by the customer; specialized services, such as automated data entry; or they may make data processing resources available to clients on an hourly or time-sharing basis.

Exclusion(s): Establishments primarily engaged in:

- processing financial transactions (52232, Financial Transactions Processing, Reserve and Clearing House Activities)
- computer facilities management (54151, Computer Systems Design and Related Services)
- providing data keying or keypunch services, text processing or desktop publishing (56141, Document Preparation Services)
- providing access to microcomputers and office equipment from a retail location (56143, Business Service Centres)