香港統計學會有限公司
Hong Kong Statistical Society Limited
c/o Department of Statistics & Actuarial Science,
The University of Hong Kong, Pokfulam Road,
Hong Kong

Bulletin
Volume 23 Number 1
June 2000

Editor's Foreword

First, I thank the associate editors, secretary and those who have offered to contribute to the Bulletin. However, more contributions are still required and if you are interested in writing for the Bulletin please contact me.

This issue features an article by Mr. Mike Cheung on resources for statisticians available on the internet. It is in everyone's interest to explore the use of internet resources in this new age, and the article provides valuable guidance about how internet can be used in both statistical teaching and research.

In the president's forum, Professor W.K. Li discusses work plan of the Society. Also, Mr. Peter Ip from the Census and Statistics Department discusses the compilation of producer price indices of service industries. Ms Agnes Lo gives a report on the 1999/2000 Statistical Project Competition for Secondary School Students.

As always, we encourage your feedback on the Bulletin. We have been very pleased with the response from some of the members and look forward to more contributions to the Bulletin from you in the future.

Alan Wan

CONTENTS

(Vol. 23/No. 1, June 2000)

President's Forum

Professor W.K. Li

Thanks to my predecessor's hard work, the new Hong Kong Statistical Society (HKSS) Limited is now running healthily and smoothly. In this connection, I am happy to report that progress has been made in establishing a local accreditation and examination system for the statistics profession. Dr. W.K. Chiu has kindly accepted our offer as the Executive Secretary of the Examination Office. He will be working with the project team led by Mr. H.W. Fung with the aim of finally establishing the aforementioned system. As you are aware, the Statistical Project Competition (SPC) for Secondary School Students has been very successful in the past fourteen years since its inception. Again, the success of the SPC is impossible without the enthusiastic support of members. This year the SPC will be led by Miss Cecilia Chan of the Census and Statistics Department and I am sure that members will give the team the same strong support as in the past. Regarding membership matters, we now have a newly and revised membership application form that takes into account comments from members at the last Annual General Meeting. This can be downloaded from the website of the Society.

I would like to end this note by wishing you a nice and enjoyable summer.

Producer Price Indices of Service Industries

Peter Ip
Census and Statistics Department

Introduction

Producer price indices (PPI) are currently compiled by the Census and Statistics Department for Hong Kong's manufacturing industries and some selected service industries.

PPI of manufacturing industries has been published quarterly since July 1997. PPI for service industries has been developed phase by phase. In view of the conceptual and technical complexities involved, intensive studies are required before appropriate methods can be established for compiling the relevant statistics. Indices on hotels and boarding houses, telecommunications and miscellaneous communication services for the third quarter of 1998 to the third quarter of 1999 have been released for the first time in April 2000.

What is the producer price index?

Before going into the details of PPI compilation, let's look at the general concept of a **price index**. A price index is calculated as a form of the Laspeyres index:

$$I_t = \frac{\sum_k Q_{k0} P_{kt}}{\sum_k Q_{k0} P_{k0}} \times 100 \qquad \textbf{(1)}$$

where:

P_{k0}, P_{kt} = prices in period 0 (base period) and t (current period) respectively

Q_{k0} = quantity in period 0 (base period)

k = kth commodity or service

The price index can thus be seen in the light of the price of a set of commodities/services in a current period relative to that in the base period. A most widely used price index is the Consumer Price Index (CPI). CPI shows the change over time in the price level of consumer commodities and services generally purchased by households. It reflects the impact of price changes on consumers.

Other types of price indices can be compiled for different purposes, e.g. retail price index measures price changes of commodities sold in retail outlets (retail outlets may also sell to tourists and other visitors as well as to business firms; this index differs from the consumer price index); stock market price index; labour and material price index (for the constructions sector); and so on.

PPI is a measure of the average change in prices received by producers of goods and services during a period over those during the **base period**. Producer prices are transacted prices, net of any discounts, premiums, rebates or allowances given to buyers but including surcharges, received by local producers for their outputs.

The PPI of a **service industry** in respect of a reference quarter measures the output prices received in the reference quarter of the service products delivered by the industry concerned compared with the base period.

PPI is also useful as an economic indicator for monitoring the price movements of local outputs and evaluating their price competitiveness vis-à-vis those delivered in other countries/territories.

Readers may note that CPI also covers services. However, CPI for services is different from PPI for service industries. CPI for services measures the changes in the total cost of a basket of services representative of the average expenditure pattern of households. It is useful for analyzing price movements of services generally consumed by households. On the other hand, the output of service industries can be either an input to other industries or consumed by households. The outputs of some service industries are not normally consumed by households (e.g. freight transport services) while some services predominantly cater for needs of visitors (e.g. hotel services) and are therefore not covered by the CPI system.

PPI is a particularly useful reference in evaluating the movement in input cost of business, but one has to ensure the relevance in each case of study as to what the PPI actually covers.

Development of PPI in respect of service industries

So far, there has been less statistical information on output prices for service industries than for manufacturing industries. PPI for manufacturing industries has been compiled and published since 1997, based on data on the producer prices of industrial goods/services collected from the Quarterly Survey of Industrial Production.

With a view to enhancing the statistical infrastructure in support of Government's initiative of services

promotion, a study was conducted in late 1997 to develop price indices in respect of service industries.

Compared with manufactured products, output of the service industries is more difficult to measure. Notable examples include management consultancy, accountancy and legal services. These industries usually produce "one-off" outputs and may not regularly repeat exactly the same assignment or service. Changes in quality are also less discernible. Thus it is more difficult to obtain comparable price data for different periods.

In view of the conceptual and technical complexities involved in compiling PPI in respect of service industries, a few major industry groups were first selected for PPI compilation, including hotels and boarding houses, telecommunications and miscellaneous communication services. They were selected because the respective product specifications were more straightforward and the price data over a long period of time could be available, which in turn would facilitate the compilation.

Prominent establishments in these selected service industries were consulted to identify the availability of the service products and ascertain the practicability of collecting the relevant price information.

Subsequently, price and product data on these service industries were collected as from the third quarter of 1998 through the Quarterly Survey of Services Industries (QSSI) launched by the Census and Statistics Department.

How is PPI compiled?

The PPI in respect of selected service industries of Hong Kong are compiled mainly from data on producer prices of service products collected through the QSSI. The survey also collects business receipts data for compiling the quarterly business receipts indices of service industries, which measure changes in the value of local services output.

For some service products where price data are currently collected through the CPI system, relevant data are extracted from the CPI system and there is no need to collect them through the QSSI.

The PPI of a particular industry is compiled on the basis of a selected "bundle" of the more important service products of that industry. Formula (1) can be re-written as:

$$I_t = \frac{\sum_k Q_{k0} P_{k0} \cdot \left(\frac{P_{kt}}{P_{k0}} \right)}{\sum_k Q_{k0} P_{k0}}$$

$$= \frac{\sum_k w_k \cdot \left(\frac{P_{kt}}{P_{k0}} \right)}{\sum_k w_k}$$

where:

(P_{kt}/P_{k0}) = **price relative** in period t compared with period 0

w_k = $Q_{k0}P_{k0}$ or business receipts derived from the kth commodity or service

In actual compilation, the index is obtained by first computing the price relative of each of the component service products pertaining to that industry. The "price relative" refers to the ratio of producer price for the reference quarter to the average producer price of the same product in the preceding year.

The price relative of a service product is obtained by aggregating the price relatives of each of the establishments producing that service product, using the business receipts of the service produced by the establishments in the preceding year as weights:

$$PI_{t-1,i}^{q,t} = \frac{\sum_j (V_{ij}^{t-1} \times PI_{t-1,ij}^{q,t})}{\sum_j V_{ij}^{t-1}}$$

where:

$PI_{t-1,i}^{q,t}$ = price relative for product (i) in current quarter q of year (t) compared with year (t-1)

$PI_{t-1,ij}^{q,t}$ = price relative for product (i) in establishment (j) in current quarter q of year (t) compared with year (t-1)

V_{ij}^{t-1} = gross-up business receipts of establishment (j) of product (i) in year (t-1)

The price relatives of the component service products are then aggregated, using the business receipts of the products as weights, to form the price relative of the respective industry. The business receipts that are used as weights for aggregation are obtained from the relevant annual surveys, which include the Annual Survey of Storage, Communication, Financing, Insurance and Business Services and the Annual Survey of Wholesale, Retail and Import and Export Trades, Restaurants and Hotels.

$$PI_{t-1,ind}^{q,t} = \frac{\sum_i (W_i \times PI_{t-1,i}^{q,t})}{\sum_i W_i}$$

where:

$PI_{t-1,ind}^{q,t}$ = PPI for industry *ind* in current quarter (q) of year (t) with year (t-1) as comparison year

$PI_{t-1,i}^{q,t}$ = PPI for product (i) in current quarter (q) of year (t) with year (t-1) as comparison year

W_i = sales value/business receipts of product (i)

The result, being an index with the preceding year as the comparison base, is converted to the PPI of the industry at base period by the method of chaining:

$$PI_o^{q,t} = \frac{PI_{t-1}^{q,t} \times PI_o^{t-1}}{100}.$$

The chained index formula is considered most appropriate for compiling PPI given the rapid changes of service products over time.

Quality change

In general, PPI should not be affected by changes in quality or in sales conditions. Quality differences have to be identified and eliminated in calculating price change for inclusion in the index. Service product changes that are regarded purely as style changes are not generally considered to be quality changes. Minor changes in specification that do not affect the contents of the service are also regarded as having no effect on quality.

To separate pure price movements from other changes when an item is replaced by a substitute item of different quality, the method of splicing is used. Prices for both the old and the new variety are collected in an overlapping quarter. The difference between the prices of the two varieties in the overlapping quarter is assumed to represent the value of their quality differences. For the purpose of calculating the price index, price relative for the new variety in the quarter following the overlapping quarter over the overlapping quarter will be linked up with the price index of the old variety in the overlapping quarter.

For an analysis of the PPI in respect of the hotels and boarding houses, telecommunications and miscellaneous communication services, interested readers may refer to the feature article published in the April 2000 issue of the Hong Kong Monthly Digest of Statistics of the Census and Statistics Department.

Future development

In addition to PPI in respect of hotels and boarding houses, telecommunications and miscellaneous communication services, plans are in hand to publish PPI in respect of the transport sector later this year.

The feasibility of further expanding the coverage of PPI to other service

industries is now being studied. However, compiling PPI for other service industries is conceptually more challenging. More detailed study of the relevant overseas experience especially in respect of data collection and compilation methodology, as well as extensive consultation with potential data suppliers, will be made to explore the best way of collecting meaningful price data.

Internet Resources and Statisticians

Mike W.L. Cheung
The Chinese University of Hong Kong

Introduction

Information Technology, especially Internet, is a hot topic in Hong Kong. The number of people using Internet is increasing and increasing. There is no doubt that Internet is changing how people think, do and communicate with each others. With the advancement of the technology and Internet, the work of statisticians is expected to be different from that in the past. Knowing the statistical theories or data analysis techniques may not be sufficient to be a competent statistician. In this article, I would like to discuss how statisticians could be benefited from the resources in Internet.

Teaching and Learning

Traditionally, we learn statistics mainly from teachers and textbooks. Teachers prepare lecture notes and teach the concepts in the class. Students may not grasp the concepts easily and lose interest in Statistics. Nowadays, teachers can explain the statistical concepts easier with the help of resources in Internet. Teaching and learning Statistics is more interesting than before with the help of multimedia. Another advantage of using Internet is that students can visit the sites after lectures to understand the materials deeply.

Using correlation as an example. By running Java applets with an Internet browser like Netscape and Internet Explorer, it is easy to show the correlation graphically. For instance, in the site http://noppa5.pc.helsinki.fi/koe/corr/cor.html, scatter plots of different correlation coefficients can be shown easily by clicking the desired values. Another example is the concept of Set Theory. Venn diagrams are always used to illustrate the concepts of "Subset," "Intersection," "Union" and "Complement." In the site http://noppa5.pc.helsinki.fi/koe/boole/boolea.html, it is easy to show the relationships of different elements and operations by using a simple click. Learners can experience the effects of the operations themselves for better understanding of these abstract concepts.

Teachers can also choose suitable electronic materials for students. There

are lots of "online textbooks" for introductory statistical concepts. For instance, http://www2.chass.ncsu.edu/garson/pa765/statnote.htm, http://davidmlane.com/hyperstat/index.html and http://www.statsoft.com/textbook/stathome.html are some sites providing general statistical concepts.

Data Analysis

Data analysis and result interpretations are typical work of statisticians. SAS and SPSS are the two commonly used software programs. By visiting their sites http://www.sas.com/ and http://www.spss.com/, we can obtain the updates and macro from them. For example, SPSS does not provide post-hoc tests for the adjusted means in ANCOVA. We can get the macro from http://www.spss.com/tech/macros/index.html to help us to do it in syntax.

Apart from using statistical packages, statisticians often refer to many different statistical tables in their work. For instance, in an attitude survey, a researcher wants to know if there are differences in the attitudes among different groups. After running an ANOVA test, he/she finds an nonsignificant result. He/she wants to estimate the number of respondents needed in order to obtain a significant result given the effect sizes and Types I and II errors. Traditionally, he/she refers to the formulas and tables for power analysis, e.g., Cohen (1988). However, with the help of Internet, e.g., http://www.math.yorku.ca/SCS/Demos/power/index.html, we can estimate the required sample without painful calculations and table references. With an browser and Internet, statisticians can refer to sites, e.g, http://www.stat.ucla.edu/calculators/cdf/, for critical values.

A recent trend of data analysis is the on-line statistical software. By using a Java-enabled browser, statisticians can submit and analyze data without the statistical package installed on their own computer. For example, Rweb, http://www.math.montana.edu/Rweb/ uses R, a statistical analysis package almost compatible with S or Splus in source code level, as its programming language. Another one is the WebStat http://www.ms.uky.edu/~mai/java/webstat/version1.0/. You can analyze or plot data with it. In some specific packages, e.g., Mx, a combination of a matrix algebra interpreter and a numerical optimizer useful in Structural Equation Modeling or Covariance Structure Analysis, http://views.vcu.edu/mx/cgi-bin/mxi.cgi, provides an environment for users to input their codes and run in the server's computer. Then the output can be displayed in the users' browsers.

For the time being, it is not a good idea to use Web-based statistical software

to analyze large data set or to do intensive computing as running Java or CGI is quite resource demanding. However, it is expected that these short-coming can be overcome very soon as the computer is becoming more and more powerful. The main advantage of using Web-based statistical software is that we do not need to update or purchase the software ourselves. The software provided by the vendors is the most updated version and bugs are minimized. For the vendors, they need to develop software for one platform only. This minimizes the cost for development. Moreover, the problems of transferring data in different versions are also reduced.

Online Data Set

Besides teaching and data analysis, statisticians and social researchers sometimes need to provide suggestions for the policy makers or the government. It is not easy to collect large and good quality data. However, with the help of Internet, it is easy to obtain useful information from different areas or countries. For example, http://www.un.org/databases/ from the United Nations, http://www.worldbank.org/data/ from the World Bank, and http://www.census.gov/ from the U.S. Census Bureau are some good data banks. There are valuable demographic and economic data collected from different countries. Berinstein (1998) classified statistics which could be accessed from the Internet. Interested readers can refer to it in finding suitable statistics online.

Conclusion

In sum, Internet can make the life of statisticians easier in teaching, learning and data analysis. The sites mentioned above are just a few in the whole Internet. There are new sites open or old ones closed daily, therefore, it is wise to use some tools to help us to find the right ones.

Listserv, newsgroups and search engines are some useful tools for us to keep updated information. Listserv uses email to communicate or discuss about some common topics. For example, Semnet, http://www.gsu.edu/~mkteer/semnet.html and Multilevel, http://www.ioe.ac.uk/multilevel/email.html are two e-mail based discussion groups for Structural Equation Modeling and Multilevel Data Analysis, respectively (see Chan, 1999 for the introduction of these two statistical methods). You need to join the Listserv before you can read or send messages while Newsgroup allows anyone to read and send messages to it without joining it. For example, sci.stat.consult, sci.stat.edu and sci.stat.math are the common places for Statisticians. If you still cannot find information in a particular topic, search engines, for instances, http://www.google.com/ and

http://www.yahoo.com/ are some good starting points.

It seems that Internet can help us do everything but we should keep in mind about the truth: Everyone can set up his/her own sites; therefore, it is hard to verify the validity of information from Internet. Information from Internet should only be used with cautions. Lastly, an updated links of the above sites can be found in http://www.psy.cuhk.edu.hk/user/wlcheung/statistics.html.

References

Berinstein, P. (1998). Finding Statistics Online: How to Locate the Elusive Numbers You Need. Medford, N.J.: Information Today, Inc.

Chan, W. (1999). The current role of statistical methods in the social sciences. *Bulletin of the Hong Kong Statistical Society*, *21(4)*, 10-16.

Cohen, J. (1988). *Statistical Power Analysis For the Behavioral Sciences* (2nd Ed.). Hillsdale, N.J.: L. Erlbaum Associates.

A Report on the 1999/2000 Statistical Project Competition for Secondary School Students

Agnes K.M. Lo
Chairperson, Organizing Committee 1999/2000

The 1999/2000 Statistical Project Competition has come to a successful conclusion with its prize presentation ceremony held on 22 April, 2000 at the Wei Hing Theatre of the City University of Hong Kong.

Like those of the previous rounds, the prize presentation ceremony of the 1999/2000 Statistical Project Competition was well attended. Not only students of the winning teams but also many other contestants, teachers, academics from local universities and statistics professionals had showed their keen support to the event by attending the ceremony. We were much honoured to have the presence of our distinguished guests, Mr. Frederick W.H. Ho, the Commissioner of the Census and Statistics Department of the HKSAR Government, the Honourable Eric K.C. Li, Member of the Legislative Council and Mr. Moses M.C. Cheng, the representative of the Hang Seng Bank, to present the prizes to the award-winners.

Since its inception in 1986, the Statistical Project Competition for Secondary School Students has been one of the main annual programmes in the Society's activity calendar. The Competition, together with its related activities such as exhibition of winning projects in its previous rounds and seminars for potential entrants, is aimed to promote students' social awareness and civic-mindedness. The Competition provides a good opportunity for students to apply those statistical concepts and skills that they have been taught in the classroom to real life situations so that they can learn how to understand the community in a scientific manner through the proper use of statistical data and analytical tools. This is indeed very important in this era of information explosion where information processing plays an essential part in our daily lives.

This round's Competition is particularly memorable, as a record high of 171 entries was received. With 97 entries in the Senior Section and 74 in the Junior Section, the themes of the projects covered a wide spectrum of areas including environment, health, transport, crime, population, entertainment, tourism and economic issues. The large variety of

subjects covered reflected the broad range of interest that students had directed at the various aspects of Hong Kong's community.

The very encouraging number of entries, together with the high standard of work in the project reports, had certainly posed a very challenging and painstaking task to the adjudication panel. The panel, which had a larger membership than ever, comprised 37 statistics academics from local universities and experienced professional practising statisticians. Headed by Dr. Lo Hing-po of the City University of Hong Kong, the adjudicators had worked extremely hard in scrutinizing the project reports before the more distinguished ones could be short-listed from the bunch of high-quality entries. In the second round of adjudication, the adjudicators had interviewed the participating students of each of the short-listed entries.

Taking into account the scores of the short-listed project reports and the performance of the students in the interviews, the results of the Competition were eventually finalized in early April 2000. A total of 6 winning teams emerged from the Senior Section and 5 from the Junior Section. All winning teams received a cash prize together with a Certificate of Merit. In addition, a trophy and statistical publications donated by the Census and Statistics Department were presented to the winning schools.

This year, two new awards were granted to the entry with the best creativity and the school with the highest number of entries submitted respectively for each of the Senior and Junior Sections. Furthermore, as a token of appreciation of their enthusiastic participation and encouragement for their continued support of the Competition, all participating students were awarded a souvenir.

In conclusion, the 1999/2000 Statistical Project Competition had been very successful under the great effort of all parties concerned. I would therefore like to extend, on behalf of the Organizing Committee of the 1999/2000 Competition, my sincere thanks to all the persons who had rendered their assistance to the event. In particular, I would like to express gratitude to Mr. Frederick W.H. Ho, the Honourable Eric K.C. Li and Mr. Moses M.C. Cheng, who had given their valuable support in being the patrons of the Competition. Last but not least, I must thank the Hang Seng Bank, which had generously sponsored all the expenses of the event.

The coming round of the Competition is its 15th round. I am confident that this meaningful annual event will continue to receive support from the community in the new millenium. I wish the 2000/01 round Competition every success.

News

1. Appointment

Department of Statistics, The Chinese University of Hong Kong

Professor Fan Jianqing from the University of California has joined the Department of Statistics as the Chair of Statistics.

Department of Management Sciences, The City University of Hong Kong

Professor H.S. Lau from the Oklahoma State University has joined the department as Chair Professor of Management Sciences.

Dr. Ronald Lau from the South Dakota University has joined the department as temporary visiting fellow.