一橋大学 21世紀COEプログラム
社会科学の統計分析拠点構築
Research Unit for Statistical Analysis in Social Sciences

Hi-Stat Newsletter
No.5 (May 2006)

Contents

Introduction to Databases, No.6

Standing at the End of the Tunnel
—Publication of the first volume of the Asian Historical Statistics Project—

Tokihiko Settsu (COE Project, Hitotsubashi University)

We are happy to be able to announce the imminent publication of the first volume of the Asian Historical Statistics Project (AHSTAT Project), the *Long-Term Economic Statistics of Taiwan*, edited by Toshiyuki Mizoguchi and published by Toyo Keizai Shimposha. The volume has been a long time in the making. Originally conceived in 1995, when the AHSTAT Project supported by a government grant-in-aid commenced, work on the volume has been ongoing since 2003.

Working as a research assistant at the Institute of Economic Research at Hitotsubashi University, I have been closely involved in the preparation of the Taiwan volume, and the purpose of this brief article is to provide a brief introduction of the volume and an outline of the work ahead.

It is useful to begin with a brief overview of the Taiwan volume. The book consists of two parts, a descriptive part and a statistical one. The following is a table of contents of the descriptive part (names of authors are given in parentheses).

Introduction (Purpose of this book and examples of analyses of the Taiwanese economy using long-term statistics) [Toshiyuki Mizoguchi]

Chapter 1 Characteristics and development of Taiwan's statistical system [Toshiyuki Mizoguchi, Masahiro Sato and Hiroshi Ikegami]

Chapter 2 Population [Osamu Saito and Tadayoshi Taniguchi]

Chapter 3 Labor force [Konosuke Odaka, Tangjun Yuan and Tadayoshi Taniguchi]

Chapter 4 Production activity: primary industry [Toshiyuki Mizoguchi]

Chapter 5 Production activity: secondary industry [Yasuhiro Hara and Toshiyuki Mizoguchi]

Chapter 6 Production activity: tertiary industry [Toshiyuki Mizoguchi]

Chapter 7 Financial statistics and commodity prices [Toshiyuki Mizoguchi]

Chapter 8 Private consumption and capital formation [Toshiyuki Mizoguchi]

Chapter 9 Foreign trade [Noriyuki Nojima]

Chapter 10 National accounts [Toshiyuki Mizoguchi]

The statistical part contains the estimates underlying each of the descriptive chapters, and a notable feature of the volume is that it comes with a supplemental CD-ROM containing the full data set on which the estimates are based, allowing readers to examine the estimation procedure.

Estimates for the pre-war period are a revision of earlier results for Taiwan published in the *Basic Economic Statistics of Former Japanese Colonies, 1895-1938: Estimates and Findings* (Toyo Keizai Shimposha, 1988) edited by Mataji Umemura and Toshiyuki Mizoguchi. Figures for the post-war period were newly estimated, making sure that they were consistent with official statistics published by the government of Taiwan. In this context, we would like to take this opportunity to express our thanks to Professor Tsong-min Wu of the National Taiwan University for his generous help.

As noted above, the Taiwan volume has been a long time in the making. Looking back, we come across an essay written by the AHSTAT project leader, Professor Konosuke Odaka, in 2000, in which he mentions that the Taiwan volume is the most advanced part of the project. He optimistically noted: "The light at the end of the tunnel has become visible" (*Newsletter of the Asian Historical Statistics Project*, No.16, March 2000, http://www.ier.hit-u.ac.jp/COE/Japanese/ Newsletter/No.16.english/Odaka.html). In other words, the Taiwan volume has been the "front-runner" of the project since 2000 – yet, it has been more than 6 years for it to "cross the finish line." A cynic might say that "the light Professor Odaka saw in 2000 must have been so powerful that it could be seen over a long distance." However, based on my involvement in the project over the past three years, I think there are two very powerful reasons why it took so long to complete this volume.

The first reason is that the objective of the AHSTAT Project is not to simply compile existing statistics, but to newly estimate and assemble long-term historical statistics based on the System of National Account (SNA). Actually, with the exception of Chapter 10, drafts of all of the chapters had been completed by the end of 2003. However, the statistical tables required myriads of minor revisions, painstakingly undertaken by Professor Toshiyuki Mizoguchi, and as anyone working in this field will know, such revisions are extremely time-consuming. In fact, Professor Mizoguchi's passion for precise and consistent estimation reminded me of the enthusiasm of the "masters of estimation" which would have been passing through our Institute since the publication of the *Estimates of Long-term Economic Statistics of Japan since 1868* (in 14 volumes, compiled by Kazushi Okawa, Miyohei Shinohara and Mataji Umemura and published by Toyo Keizai Shimposha, between 1965 and 1988). I am certain that readers taking the trouble to re-examine the process of estimation based on the dataset included in the supplemental CD-ROM would have a similar impression.

The second reason that the publication of the Taiwan volume has taken such a long time is that it provides the template for future publications. Following the Taiwan volume, similar volumes on China, Korea and Vietnam are due to be published and eventually the *Long-term Asian Macro-Economic Historical Statistics Series*, will comprise 12 volumes in total (see *Newsletter of the Asian Historical Statistics Project*, No.17, April 2002). Thus, the Taiwan volume assumes the role of determining the basic format of these subsequent volumes: not only does it provide a model for the methodology of estimation, but also the templates for the texts and statistical tables. In fact, deciding on these templates required much time because of the difficulties involved in finding a common style for table headings, since these contained both English and Japanese. Although this work may seem rather mundane, we believe that it has made it much easier to publish the other volumes.

In this manner, the front-runner deliberately spent a long time "walking in the darkness" but arrived at the end of the tunnel at long last. Although the "front-runner" has left useful guideposts for those who follow, more time and effort will be required to goad all runners to the goal. Nevertheless, the publication of the Taiwan volume will mark a huge step toward the successful completion of our project.

Introduction to Databases, No.7

Use of Government Microdata: Comments from Users

<div align="right">Editors</div>

One reason why the number of empirical studies on the Japanese economy based on microdata has been small is that statistical regulations have limited the use of government microdata. In order to improve the situation and provide better access to government microdata for academic researchers throughout Japan, the Statistical Survey Department of the Statistics Bureau of the Japanese Ministry of Internal Affairs and Communications has asked the Research Centre for Information and Statistics of Social Science (RCISSS) at the Institute of Economic Research, Hitotsubashi University, to construct a trial data access system.[1] This system is part of the Hi-Stat Project and the trial access to the microdata will maintain the confidentiality of the data and hence safeguard the privacy of survey respondents.

The past three rounds of the following three surveys were available during the trial: (1) *The Employment Status Survey*, (2) *The Survey on Time Use and Leisure Activities*, and (3) *The National Survey of Family Income and Expenditure*. First applications for use of the microdata were received by November 2004 and six research proposals were chosen. These were followed by four further research projects approved in April 2005 and another twelve in October 2005. Currently, the RCISSS is processing applications that were due in April 2006.

The number of applications so far has been slightly below the expectations of the RCISSS. One of the reasons for this could be the lack of publicity. With more publicity for the trial system, probably more researchers would have applied. Another reason could be the inconvenience of the procedure in obtaining access to the data. Therefore, the editors of this newsletter asked actual users of the trial data access system to send their comments on the system. In this note, we report their comments in the hope that such information will help the RCISSS to improve the system.

The Employment Status Survey

<div align="center">Toshio Kosaki (Tokai University)</div>

Significant aspects of the trial access to government microdata

The first significant aspect of the trial system is that it provides equal opportunity to researchers. Previously, only a limited number of researchers (mainly researchers under contract by the government) were able to conduct research based on government microdata. This situation has led to disparities in the quality of research that have nothing to do with researchers' ability but are the result of unequal access to data. The trial access system has partially resolved this problem.

A second point, related to the first, is that the verifiability of empirical research has been improved. Without access to government microdata, ordinary researchers had no means to re-examine the empirical findings claimed by researchers with personal access to government microdata. Thus, with more equal access, more research can be scientifically verified.

Third, microdata are frequently used in policy-oriented documents, such as the *OECD Employment Outlook* and the *White Paper on the Labor Economy* by the Japanese Ministry of Health, Labour, and Welfare. I think that, in designing government policies, it would be more efficient to base decisions on a greater diversity of research by allowing researchers more liberal access to microdata.

Fourth, allowing researchers more liberal access to microdata helps them to improve their analytical skills. Therefore, the trial access to government microdata is significant in terms of scientific progress.

Requests regarding government microdata and the trial access system

First, in compiling microdata and constructing

statistics based on such data, it is critically important to ensure that time-series analysis is possible. I would say that the value of purely cross-sectional analyses as scientific evidence can be very limited. In the microdata of *The Employment Status Survey*, however, several regional variables are available only in the 1992 dataset but not in the 2002 dataset. Similarly, in the statistical tables published in the reports of *The Employment Status Survey*, many variables are not comparable across the 1992, 1997, and 2002 reports. The comparability across time should be maintained whenever new variables or new questions are added.

Second, when researchers use government microdata for their research, trial-and-error is an important part of the process. The current procedure does not allow researchers to try the data in advance. I feel that a more flexible procedure would be helpful.

Finally, as new variables or new questions to be added in future surveys, I would recommend to include variables and questions that enable us to scientifically evaluate policies. In searching for such variables and questions, research findings of the research community should be incorporated.

The Survey on Time Use and Leisure Activities

Masaki Katsuura (Meijo University)

Using the trial data access system from March to September 2005, I used the three rounds of *The Survey on Time Use and Leisure Activities* conducted in 1991, 1996, 2001. I am not sure why I was not aware of the trial access system. I may not have heard about it because I work in Nagoya, far away from Tokyo, or I may have heard about it but then forgotten again. In any case, I was not actively looking whether there may have been any changes in the use of government microdata. Fortunately, however, I met the person in charge of the trial access system just before the application deadline and rushed to apply.

Since I already had several research topics in mind that can be analyzed using the microdata of *The*

Survey on Time Use and Leisure Activities, and I also had gained experience in using the dataset under the "non-statistical" purposes access system previously available,[2] I did not have difficulty in preparing the application in time. In addition, the procedural requirements in the application form under the new system were much simpler than under the previous system, as far as I remember. Especially as a researcher who usually depends on trial-and-error, I felt relieved to see that under the new system, I did not need to completely specify the list of explanatory variables that I would use in the regression analyses. On the other hand, even under the new procedure, I needed to specify the form of statistical tables that I would prepare. This requirement, however, was beneficial to me because it helped me to clarify the research objective, which had been rather vague before the preparation of the application.

Overall, I was satisfied with the data access service for the following reasons. First, the microdata set I obtained was based on re-sampled observations from the original survey. The re-sampling ratio was as high as 80% so that the statistical reliability of the re-sampled data was sufficiently high. I was able to confirm the reliability myself before conducting my own analysis, which I greatly appreciated. Second, the documentation on the dataset (including the information on the web) was easy to understand and rich in content. The names of variables were almost the same across the three rounds of the survey. Third, the support system was excellent. Not only during the time of the application and orientation, but also during the time of data analysis, support staff responded quickly by e-mail whenever I had questions.

On the other hand, I found several aspects that could be improved in the near future. The first regards the way the confidentiality of the data and hence the privacy of survey respondents are safeguarded. Geographical information is available only for the "three metropolitan regions" (Tokyo, Osaka, and Nagoya) and "other regions"; information by prefecture is not available at all. Although I knew this in advance, it severely limited the scope of my analysis.

Second, the information on time use and the information on leisure activities cannot be combined at the micro level due to the design of the re-sampling method. Since time use and leisure activities are intimately related, this is another drawback.

The third problem is the inflexibility regarding the time period the microdata can be used. Although I understand that imposing a time limit is inevitable, greater flexibility regarding the starting date (and hence the end date) of the project would be desirable. Since most users will probably be most devoted to the project toward the end of the period, it would be helpful to be able to schedule the project in such a way that the end falls into a relatively free period. The current procedure does not allow this.

Finally, there are many other types of government microdata I would like to use, and I therefore hope that access to statistics other than the current three surveys will be improved.

The National Survey of Family Income and Expenditure

Reishi Maruya (University of Kobe)

Around ten years ago, or more, I visited my professor in Germany with my research article on the distribution structure of the German economy. I expected some complimentary remark, but instead the professor asked me in his first comment, "Don't you know the SOEP (*Das Sozio-oekonomische Panel*) dataset?" The research article I had brought was not based on microdata. I thus decided to re-analyze the distribution structure of the German economy using the SOEP dataset. I also thought it was a good idea to conduct a similar analysis for the Japanese economy for comparison, so I applied to use the government dataset under the "non-statistical" purposes access system. My application was immediately rejected – a memorable experience that has stayed with me until today.

Because of this experience, I was thrilled to hear the news in April 2005 about the initiation of the trial data access system for government microdata. I was really impressed with the effort by the RCISSS staff led by Prof. Hiroshi Matsui. In my research this time, I estimated equivalence scales in household consumption behavior, using the microdata of *The National Survey of Family Income and Expenditure*. Next time, I am planning to apply for permission to use the data to analyze changes in the distribution structure of the Japanese economy, using the estimated equivalence scale parameters. A researcher with a strong background in statistics or econometrics may have been able to complete this analysis in one six-month period. However, since I am not specialized in statistics or econometrics, I was not able to reach the second stage of my research. I therefore really appreciate that a researcher can apply to use microdata repeatedly.

Following the very strict restrictions applied to users of the SOEP microdata, I always disconnect from the internet when processing the microdata of *The National Survey of Family Income and Expenditure*. Some may wonder if this is really necessary, considering the compact space of our research laboratory. I believe that such precaution is better than running the risk of having a confidentiality problem.

Notes

1. See Hiroshi Matsui, "Use of Government Microdata and the Statistics Law in Japan," *Hi-Stat Newsletter*, No.2, September 2005, pp.4-7.
2. The Japanese word is "Mokuteki-gai Shinsei." See the reference in footnote 1 for details of the "non-statistical" purposes access system.

Hi-Stat Discussion Papers (November 2005 - May 2006)

2005/06 (those issued after Newsletter No.4)

No.126 (November 2005, revised January 2006) Hiroaki Chigira, "A Test of Cointegration Rank Based on Principal Component Analysis"

No.127 (November 2005) Satoru Kanoh, "A Further Extension of Duration Dependent Models"

No.128 (November 2005) Takeshi Miyazaki, "Shichoson Gappei niha Saishutsu Sakugen koka ga arunoka" [in Japanese]

No.129 (January 2006) Kazuhiko Hayakawa, "The Asymptotic Properties of the System GMM Estimator in Dynamic Panel Data Models When Both N and T are Large"

No.130 (January 2006) Kazuhiko Hayakawa, "Efficient GMM Estimation of Dynamic Panel Data Models Where Large Heterogeneity May Be Present"

No.131 (January 2006) Eiji Ogawa and Junko Shimizu, "AMU Deviation Indicator for Coordinated Exchange Rate Policies in East Asia and its Relation with Effective Exchange Rates"

No.132 (January 2006) Kyoji Fukao, Debin Ma, and Tangjun Yuan, "Real GDP in Pre-War East Asia: A 1934-36 Benchmark Purchasing Power Parity Comparison with the U.S."

No.133 (January 2006) Kyoji Fukao, Hyeog Ug Kwon and Miho Takizawa, "Tainichi Kokunai Kigyokan M&A to Hibaisyukigyo no performance" [in Japanese]

No.134 (January 2006) Takashi Kurosaki, "Labor Contracts, Incentives, and Food Security in Rural Myanmar"

No.135 (January 2006) Osamu Saito, "Land, labour and market forces in Tokugawa Japan"

No.136 (January 2006) Shunsuke Sakamoto, "Parental Attitudes toward Children and Child Labor: Evidence from Rural India"

No.137 (February 2006) Satoru Kanoh and Chakkrit Pumpaisanchai, "Listening to the Market: Estimating Credit Demand and Supply from Survey Data"

No.138 (February 2006) Yukako Murakami, "Are Multinational Enterprises More Productive? A Test of the Selection Hypothesis"

No.139 (February 2006) Osamu Saito and Tokihiko Settsu, "Money, credit and Smithian growth in Tokugawa Japan"

No.140 (February 2006) Rene Belderbos, Kyoji Fukao, and Tomoko Iwasa, "Foreign and Domestic R&D Investment"

No.141 (February 2006) Ralph Paprzycki, "The Impact of Foreign Direct Investment in Japan: Case Studies of the Automobile, Finance, and Health Care Industries"

No.142 (February 2006) Tangjun Yuan, "Seisanyososhijo no bundan to bukkasuijun: Chugoku ni okeru hiboekizai boekizai no sotaikakaku ni kansuru bunseki" [in Japanese]

No.143 (February 2006) Yukako Murakami and Kyoji Fukao, "Inward Foreign Direct Investments and Productivity Growth in Japan"

No.144 (February 2006) Tangjun Yuan, "Nokokan koekijoken no naigaikakakusa to shigen iten: Naze chugoku no nomin ha mazushiinoka" [in Japanese]

No.145 (March 2006) Satoru Kanoh and Asuka Takeuchi, "An Analysis of Option Pricing in the Japanese Market"

No.146 (March 2006) Ho-il Moon, "Shokuminchiki chosen no nanboku jinkouhi: Chosen sotokuhu kokuseichosasiryo no bunkatsu fomatto" [in Japanese]

No.147 (March 2006) Kyoji Fukao, Kozo Kiyota and Ximing Yue, "China's Long-Term International Trade Statistics: By Commodity, 1952-1964 and 1981-2000"

No.148 (March 2006) Hiroaki Chigira and Taku Yamamoto, "Cointegration, Integration, and Long-Term Forcasting"

No.149 (March 2006) Bongchan Ha and Hak K. Pyo, "Data Structure of Korea for Estimating Productivity in KLEMS Model"

No.150 (March 2006) Yukihiko Kiyokawa and Jian Wang, "Senzennihon no chihotokeisoshiki no seiritsu to tokeichosain seido: Noshomutokei wo chushin ni" [in Japanese]

No.151 (March 2006) Rui Takahashi, "Betonamu ni okeru nogyosensasu no jisshi to sono hyoka: Nihon tono hikaku ni miru mondaiten" [in Japanese]

No.152 (March 2006) Jean-Pierre Dormois, "Tracking

the elusive French productivity lag in industry 1840-1973"

No.153 (March 2006) Masuyo Takahashi, "Nicchiki taiwan no tokeichosaseidoshi (ko): Taiwan sotokufu no tokeichosajigyo, tokuni 'Hokokurei' ni tsuite" [in Japanese]

No.154 (March 2006) Tetsuro Narita, "Remittances to Latin America: Moving towards More Precise Data Collection"

No.155 (March 2006) Kuan-Pin Lin, Zhi-He Long and Mei Wu, "A Spatial Investigation of σ-Convergence in China"

No.156 (March 2006) John M. de Figueiredo and Charles M. Cameron, "Endogenous Cost Lobbying: Theory and Evidence"

No.157 (March 2006) Zekeriya Eser and Joe Peek, "Reciprocity and Network Coordination: Evidence from Japanese Banks"

No.158 (March 2006) Siddhartha G. Dastidar, Raymound Fisman and Tarun Khanna, "Limits to Policy

Reversal: Privatization in India"

No.159 (March 2006) Maria Guadalupe and Francisco Perez-Gonzalez, "The Impact of Product Market Competition on Private Benefits of Control"

No.160 (March 2006) Erik Feijen and Enrico Perotti, "The Political Economy of Financial Fragility"

2006/07

No.161 (April 2006) Takahiro Ito and Takashi Kurosaki, "Weather Risk and the Off-Farm Labor Supply of Agricultural Households in India"

No.162 (April 2006) Kyoji Fukao, Hyeog Ug Kwon and Miho Takizawa, "M&A to hibaishukigyo no pafomansu: Tainichi M&A to kokunaikigyokan M&A no hikaku" [in Japanese]

No.163 (May 2006) Boris Lokshin, Rene Belderbos, and Martin Carree, "Internal and external R&D: complements or substitutes? Evidence from a dynamic panel data model"

Seminars and Meetings (November 2005 - May 2006)

Hi-Stat Lecture Series

5th (Jan. 10-11, 2006) H. Tsurumi (Rutgers University) "Recent Developments in Bayesian Econometrics"

6th (Feb. 13, 16, 2006) A. Ueda (Waseda University) "Dynamic Models and Numerical Analysis"

7th (March 14-15, 2006) In Choi (Hong Kong University of Science and Technology), "Non-Sationary Panel Data Models"

8th (May 9, 16, 23, 2006) T. Amemiya (Stanford University) "Microecnometrics: Qualitative Response Models, Tobit Models, Duration Models"

Hi-Stat Research Seminars

(Only those seminars with English papers/handouts are listed here. There were fifteen regular research seminars and two junior research seminars held during November 2005 – May 2006.)

56th (Dec. 3, 2005) *Workshop on regional inequality and growth in France and in Japan (1870-2005)* (Chorus French-Japanese Program supported by the

JSPS, the French Ministry of Education, and the COE Hi-Stat program)

1. Kyoji Fukao (Institute of Economic Research, Hitotsubashi University) and Yue Ximing (China Academy of Social Sciences, Beijing), "Income convergence in Japan, 1950-2000"

2. Jean-Pascal Bassino (Maison Franco-Japonaise, and Institute of Economic Research, Hitotsubashi University) and Noriko Kato (National Institute of Public Health), "Regional inequality in post-WWII Japan; income, life style, and stature"

3. Jean-Pierre Dormois (Marc Bloch University and Institut Universitaire de France), "Wage dispersion and the standard of living in France, 1850-1913"

4. Laurent Heyberger (Belfort-Montbelliard University), "Stature and real wages in 19th century France: a regional approach"

5. General discussion and round table

57th (Dec. 15, 2005) K. Kiyota (Yokohama National University) "Effects of Offshore Production and Foreign-ownership on Employment in Japan" (with Toshiyuki Matsuura)

59th (Jan. 7-8, 2006) *Hitotsubashi Conference on Econometrics*

1. Yukitoshi Matsushita (University of Tokyo), "Comparing tests of coefficients in microeconometric models"
2. Naoto Kunitomo (University of Tokyo), "The asymptotic expansion of distribution of empirical likelihood estimator and its consequences in econometrics"
3. Yoshihiko Nishiyma (Kyoto University), "Statistical properties of rank size rule regression under Pareto distribution"
4. Chang-Jin Kim (Korea University) with Charles R. Nelson, "Estimation of a forward-looking monetary policy rule: A time-varying parameter model using ex-post data"
5. Hiroki Tsurumi (Rutgers University), "Bayesian analysis of TARMA and FARMA nonlinear time series models"
6. Myoung-Jae Lee (Chinese University of Hong Kong), "Difference in quasi-differences with panel data: Reversed effects of private school"
7. Keiko Yamaguchi (Hitotsubashi University), "Testing for change of the long memory parameter for nonstationary processes"
8. Shinya Tanaka (Hitotsubashi University) and Takayuki Shiohama (Hitotsubashi University), "Effects of the age distribution on the long run relationship between consumption and income in Japan"
9. Teruo Nakatsuma　Keio University), "A Bayesian model averaging approach for portfolio selection"
10. Naohiko Baba (Bank of Japan) and Hiromichi Goko (Bank of Japan), "Survival analysis of hedge funds"
11. Hidehiko Ichimura (University of Tokyo) with R. Blundell, A. Gosling, and C. Meghir, "Changes in the distribution of male and female wages: Accounting for employment composition using bounds"
12. Ryo Okui (Hong Kong University of Science and Technology), "Shrinkage GMM estimation in conditional moment restriction models"
13. Kazuhiko Hayakawa (Hitotsubashi University), "Efficient GMM estimation of dynamic panel data models where large heterogeneity may be present"
14. Koichiro Kamada (Bank of Japan), Wataru Hirata (Bank of Japan), and Hajime Wago (Nagoya University), "The land price correlation across prefectures in Japan: An analysis of spatial econometrics"

62nd (Feb. 10, 2006) Jean-Pascal Bassino (Australian National University), "Market integration and Famines in Tokugawa Japan"

64th (March 4, 2006) *Institutions, Politics and Corporate Governance* (conference jointly held by COE Hi-Stat Project and Center for Economic Institutions, IER, Hitotsubashi University)

65th (March 17, 2006) In Choi (Hong Kong University of Science and Technology), "Inference for Predictive Regression with Nearly $I(1)$ variables via Subsampling"

Hi-Stat Junior Research Seminars

6th (Dec. 6, 2005) H. Chigira (Hitotsubashi University) "A New Approach to Estimation of Dynamic Panel Models in Small Samples"

7th (May 9, 2006)

A. Tsukada (Hitotsubashi University) "Understanding Voting Behavior in Japan"

K. Sakamoto (The Institute for Research on Household Economics) "Familial Support for Unemployed Youth"