

Improving Sampling Efficiency for the Annual Survey of Services

By

Andy Chiang, Cui Hui Min, Neo Siok Hoong and Tay Boon Shang
Business Statistics Division
Singapore Department of Statistics

Introduction

The Annual Survey of Services (AS) collects vital business data used for the study of financial structure and performance of the services industries and the compilation of Singapore's economic statistics. The survey sample size for the Annual Survey of Services has grown exponentially since 1998. This article discusses the recent improvement in sampling design of the AS that will be implemented for reference year 2007.

Optimal Sampling for Business Surveys

In business surveys, populations have skewed distributions where a small number of units account for a large

share of the total of the study variable (y). Following sampling theory, it is more efficient to employ a stratified sampling design where y is used as a size stratification variable by constructing boundaries in the range of y . In business surveys conducted by the Singapore Department of Statistics, operating receipts is frequently used for such size stratification as it is positively correlated to several key variables of interest such as operating expenditure, employment, value added and profit.

Survey statisticians have developed various approaches for optimal size stratification. Horgan (2006) and Kozak & Verma (2006) reviewed the best methods, which include the Hidiroglou (1986) and Lavallée-Hidiroglou (1988) methods. These two methods adopt a common strategy of stratifying

the population into two or more size strata. The firms in the large size stratum are selected with certainty whereas the firms in the smaller size strata are sampled using simple random sampling without replacement. Both methods produce optimal sample size and allocation for a given level of precision¹. The difference is that the Hidiroglou (1986) method allows for two size strata (large and small) while the Lavallée-Hidiroglou (1988) method provides two or more² size strata.

Sampling Design of Annual Survey of Services

Before 1997 : Stratified Systematic Sampling

Before 1997, the AS used a stratified systematic sampling design. All establishments were stratified by SSIC³ industry and operating receipts as follows. Within each SSIC industry, establishments belonging to enterprises with multiple branches or divisions were selected with

certainty. For the remaining establishments, they were selected based on their operating receipts relative to a pre-determined cut-off value. Those above the cut-off value were selected with certainty while those below were selected with a probability of one in ten systematically. Based on this method, the sample size grew proportionately with the population in the survey frame from 22,000 for 1986 to 27,000 establishments for 1996.

1997 to 2006 : Stratified Random Sampling Using Hidiroglou Method

The steady increase in sample size has resulted in increased respondents' burden, longer processing time before survey results become available as well as the need for increasingly scarce manpower resources. To manage the situation, a sampling design review⁴ was conducted and resulted in the adoption of the Hidiroglou (1986) method for stratified sampling of enterprises with coefficient of variation of 10 per cent at SSIC 5-digit industry level. This brought about a reduction of sample size to 16,400

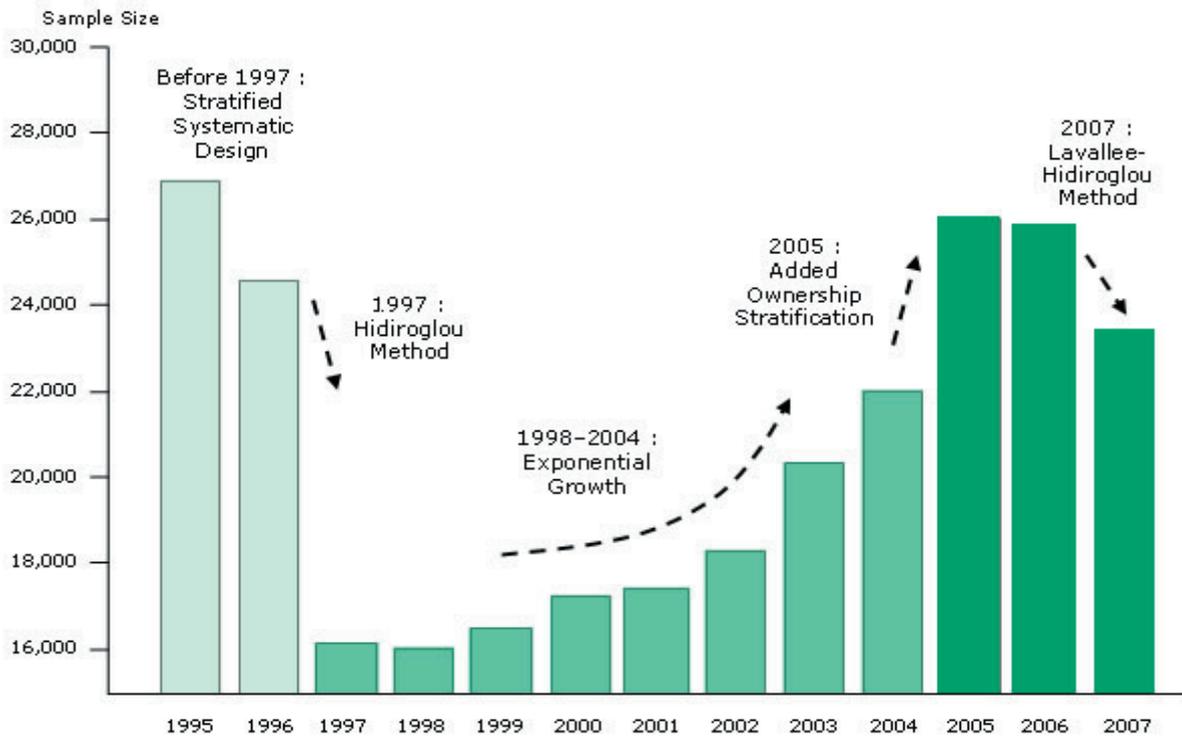
¹ Precision is measured by a coefficient of variation which is expressed in per cent format. A higher (lower) coefficient of variation corresponds to lower (higher) precision for the survey estimate.

² A three size strata may be labeled as large, medium and small respectively.

³ Singapore Standard Industrial Classification (SSIC) is the official classification of economic activities undertaken by economic units and used for censuses of population, household and establishment surveys and administrative records. SSIC was established in 1958 and last revised in 2005.

⁴ Singapore Department of Statistics (1999). New sample design for the annual surveys of commerce and services. *Occasional Paper on Business Statistics*.

CHART 1 1995-2007 ANNUAL SURVEY OF SERVICES DESIGN SAMPLE SIZE



establishments for 1997 (see Chart 1). Subsequently, due to increases in the size and variability in the business population, AS sample size grew exponentially from 16,400 for 1997 to 22,000 establishments for 2004. Since 2005, in order to produce more precise estimation at ownership level, AS population was stratified by industry and ownership (local and foreign). Each industry-ownership stratum was size-stratified by operating receipts using the Hidiroglou method. The addition of ownership stratification raised sample size to 26,000 for 2005 as the number of strata before Hidiroglou size stratification was doubled.

Improving Sampling Efficiency Using Lavallée-Hidiroglou Method

We will use two examples to illustrate how the new sampling design using Lavallée-Hidiroglou method for size stratification improves over the old sampling design using Hidiroglou method. We will examine two populations (A and B) that are representative of large and small populations in services industries. Chart 2 shows the skewness in the distribution of operating receipts, which is typical in services industries.

Size Stratification in Population A

First, consider population A from a transport industry in Table 1. Assuming 10 per cent coefficient of variation for estimating total industry operating receipts, the Hidiroglou method will stratify the population by operating receipts into two size strata : large (over S\$5.3m) and small (under S\$5.3m). Lavallée-Hidiroglou method stratifies the population more efficiently into three operating receipts size strata : large (over S\$11.2m), medium (between S\$1.3m to S\$11.2m) and small (less than S\$1.3m). By having medium size strata, fewer firms are sampled : the medium and small size strata had a combined sample of 29 units, compared with 65 units for the small size stratum in the old design. Together with the reduction in large

(take-all) units due to higher cut-off for operating receipts, the overall sample size is reduced from 110 in the old design to only 44 in the new design. This represented a 60 per cent reduction in sample size for the same level of 10 per cent coefficient of variation.

Size Stratification in Population B

Next, we examine population B from a wholesale industry in Table 2. Hidiroglou method has two size strata : large (over S\$3.1m) and small (under S\$3.1m) while Lavallée-Hidiroglou method gives three : (large = over S\$13.1m, medium = S\$1.8m to S\$13.1m and small = under S\$1.8m). As the industry distribution of operating receipts is highly skewed (see Chart 2), the Hidiroglou cut-off of S\$3.1m is too low to be efficient. Consequently,

CHART 2 DISTRIBUTION OF OPERATING RECEIPTS IN POPULATION A AND B

Table 1

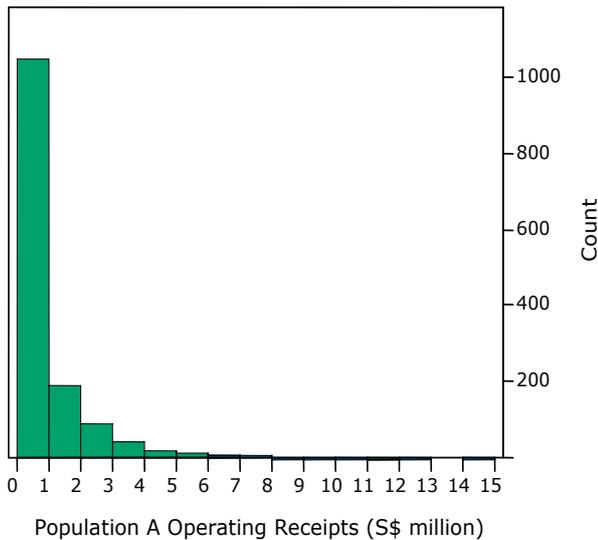


Table 2

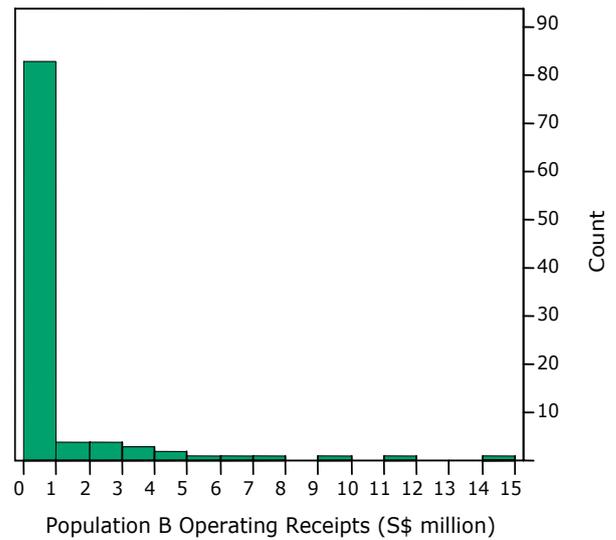


TABLE 1 POPULATION A CUT-OFF VALUES AND SAMPLE SIZES

	Size Strata	Cuf-Off Value (S\$) [1]	Population Size [2]	Sample Size [3]	Sampling Fraction (%) [4]
	Total		1,517	110	7
Hidiroglou Method	Large	5,296,186	45	45	100
	Small	0	1,472	65	4
	Total		1,517	44	3
Lavellée-Hidiroglou Method	Large	11,188,698	15	15	100
	Medium	1,307,368	273	17	6
	Small	0	1,229	12	1

TABLE 2 POPULATION B CUT-OFF VALUES AND SAMPLE SIZES

	Size Strata	Cuf-Off Value (S\$)	Population Size	Sample Size	Sampling Fraction (%)
	Total		103	16	16
Hidiroglou Method	Large	3,081,631	13	13	100
	Small	0	90	3	3
	Total		103	13	13
Lavellée-Hidiroglou Method	Large	13,057,417	5	5	100
	Medium	1,832,494	11	5	45
	Small	0	87	3	3

[1] Cut-off value refers to the lower boundary value of operating receipts for the size strata.

[2] Population size refers to the number of units in the respective strata in the population.

[3] Sample size refers to the number of units in the sample for the respective strata.

[4] Sampling fraction refers to the percentage of units in the population selected into the sample, i.e. [4] = [3] / [2] x 100%.

13 firms are treated as large and to be selected with certainty. In contrast, the Lavallée-Hidiroglou method requires only the top 5 large firms (with over S\$13.1m in operating receipts) to be selected with certainty and samples another 5 out of 11 possible medium firms. Both methods sample only 3 out of 87-90 small firms. The new design using Lavallée-Hidiroglou method achieves a 19 per cent reduction in sample size while maintaining the same level of precision of 10 per cent coefficient of variation.

Final Sample Design for Annual Survey of Services 2007

The Lavallée-Hidiroglou size-stratification for operating receipts with three size strata will be implemented in the AS for reference year 2007. While there

is potential to reduce sample size dramatically (as high as 60 per cent in the example for Population A) for fixed coefficient of variation at 10 per cent, part of the reduction in sample size will be used to trade-off for improvement in estimation precision at broader groupings such as SSIC 2-digit industry divisions⁵ via a reduction in coefficient of variation from 10 per cent to 8 per cent (typically) and as low as 5 per cent in some cases.

Overall, the new sampling design will reduce total sample size by about 10 per cent (23,500 in 2007 compared to 26,000 in 2006) while improving estimation precision: SSIC 5-digit industry coefficient of variation by at least 2 per cent (5-8 per cent for 2007 compared to 10 per cent for 2006).

References

Hidiroglou, M.A. (1986). The construction of a self-representing stratum of large units in survey design. *The American Statistician* 40(1), 27-31.

Horgan, J.M. (2006). Stratification of skewed populations: a review. *International Statistical Review* 74(1), 67-76.

Kozak, M., Verma, M.R. (2006). Geometric versus optimization approach to stratifica-

tion: comparison of efficiency. *Survey Methodology* 32(2), 157-163.

Lavallée, P. & Hidiroglou, M. (1988). On the stratification of skewed populations. *Survey Methodology* 14, 33-43.

Singapore Department of Statistics (1999). New sample design for the annual surveys of commerce & services. *Occasional Paper on Business Statistics*.

⁵ SSIC 2-digit industry division refers to a broad grouping of 5-digit industries with the same first 2-digit code. For instance, "51" retail trade industry division contains the industries ranging from "51101" (Non-specialised retail trade in department stores) to "51609" (Other non-store retail sale). For AS 2007, coefficient of variation is set at 5 per cent for estimation at SSIC 2-digit industry division level.

Purchasing Power Parity and the 2005 International Comparison Program

By

Suzanne Wong Sook Han
Economic Accounts Division
Singapore Department of Statistics

Introduction

The International Comparison Program (ICP) is a global statistical project to collect price and expenditure data for the estimation of purchasing power parity (PPP). The main objective of the ICP is to enable PPP-based international comparisons of macro-economic aggregates such as income and output (Gross Domestic Product, GDP), productivity and standards of living, which take into account relative price levels between economies. This article presents the concept of PPP and an overview of the benchmark 2005 ICP project. It reviews the final global PPP comparisons from the 2005 ICP (released in February 2008) and highlights, in particular, Singapore's relative price level and per capita measures of GDP, household consumption and gross fixed capital formation.

Purchasing Power Parities

PPP refers to the number of currency units required to purchase an amount of goods and services equivalent to what can be bought with one unit of currency of the base country, for example, the US dollar (a commonly used base currency).

Unlike market exchange rates, PPPs adjust for differences in price levels between countries/economies and enable more robust cross-country comparisons of economic output, productivity and standards of living, based on a common set of average international prices.

In contrast, comparisons based on market exchange rates which are determined by the demand and supply of currencies in international transactions do not necessarily reflect the real value of an economy's output and the standard of living of its residents. As market exchange rates do not take into account relative prices between economies which are typically lower in developing economies due to cheaper non-traded (labor-intensive) goods and services, they tend to exaggerate income and productivity differentials between economies at different levels of development.

The use of PPPs for international comparisons of macro-economic data is generally considered more appropriate vis-à-vis market exchange rates. Hence, PPP-based macro-economic aggregates are featured prominently as key fundamental macro-economic

indicators in many international studies, programs and targets. For example, the World Bank's *World Development Indicators*, the International Monetary Fund's (IMF) *World Economic Outlook* and the United Nation's Millennium Development Goals. However, PPPs should not be used for all types of economic analysis. Market exchange rates which reflect demand for currencies as a medium of exchange, store of value (investments) or official reserves should be used, for example, for comparisons of international trade, capital flows and foreign debt.

2005 ICP Programme

The construction of PPPs is a complex and challenging exercise, requiring detailed high quality price and expenditure data (on products that are both representative of an economy and comparable with other economies) from countries/economies whose PPPs are being calculated. Launched in 1968, the ICP spearheads the development of PPPs and has, over the years, carried out several rounds of international comparison.

In this latest 2005 round which began in 2003, the worldwide ICP program covered over 140 economies and includes for the first time, the participation of China and India – the two most populous economies in the world. Due to the scale and complexity of the project, the 2005 ICP in effect comprises two separate PPP programs, namely :

- (a) the World Bank coordinated global ICP program, which covered over 100 (mostly developing) economies from five geographical regions (Africa, Asia Pacific, Commonwealth of Independent States, Latin America and Western Asia); and
- (b) the Statistical Office of the European Communities (Eurostat) and the Organization for Economic Co-operation and Development (OECD) PPP program, which included 46 (mostly member) economies.

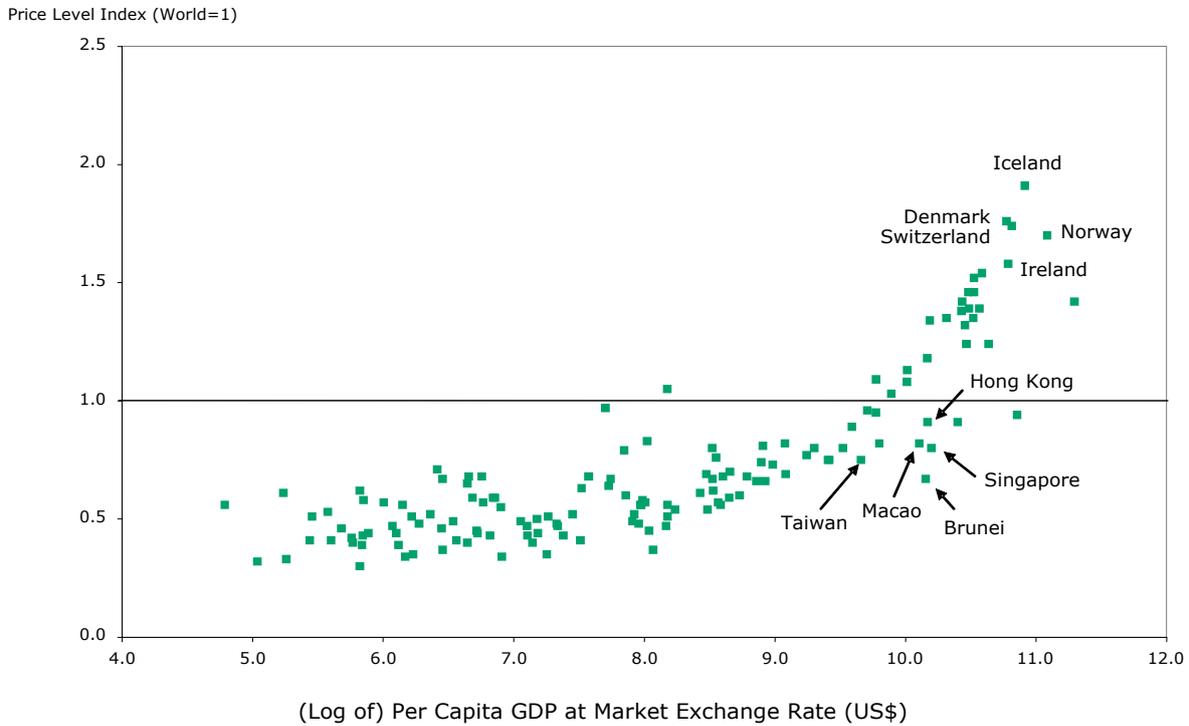
With the ICP conducted on a regional basis, overall global comparisons are obtained by linking regional results through "ring comparisons" – where a few economies from each region participate in a separate parallel international comparison organized specifically to provide a link between regions.

2005 ICP Global Results

Relative Price Levels

Price level indices (ratios of PPPs to market exchange rates) which are used to express the relative price levels between economies generally correlate positively with income i.e. per capita GDP (Chart 1). For example, high-income economies such as Iceland, Denmark, Switzerland, Norway and Ireland, were the most expensive places to live and work in 2005 with price levels as high as two times the world average.

CHART 1 RELATIVE PRICE LEVELS AND PER CAPITA GDP, 2005



Source : 2005 International Comparison Programme : Tables of Final Results, 2008

In comparison, price levels in Singapore were about one-fifth lower than the world average. They were also relatively lower than price levels in Hong Kong and Macao but higher than price levels in Taiwan and Brunei – the other high-income Asia Pacific economies.

GDP at PPP

Cross-country comparisons of GDP at PPP show that the United States and China were the world’s largest economies in 2005, with world GDP shares of 23 per cent and 10 per cent respectively (Table 1). This was followed by Japan with a world share of 7 per cent, Germany with 5 per

cent and India with 4 per cent. These five economies collectively accounted for nearly half of global GDP valued at approximately US\$55 trillion at PPP in 2005.

In contrast, the standard (poorer) currency conversion method using market exchange rates results in comparisons that under or overstate the relative sizes of some economies. For example, the 2005 GDP levels of China, India and the Russian Federation at market exchange rates grossly understated the sizes of these economies (with low domestic price levels) relative to other high-income economies.

TABLE 1 RELATIVE SIZE OF ECONOMIES, 2005

Economies	GDP at PPP			GDP at Market Exchange Rates		
	US\$ Billion	Share (World=100)	Rank	US\$ Billion	Share (World=100)	Rank
United States	12,376.1	22.51	1	12,376.1	27.93	1
China	5,333.2	9.70	2	2,243.8	5.06	5
Japan	3,870.3	7.04	3	4,549.2	10.27	2
Germany	2,514.8	4.57	4	2,791.3	6.30	3
India	2,341.0	4.26	5	778.7	1.76	12
United Kingdom	1,901.7	3.46	6	2,244.1	5.06	4
France	1,862.2	3.39	7	2,136.3	4.82	6
Russian Federation	1,697.5	3.09	8	764.4	1.73	14
Italy	1,626.3	2.96	9	1,769.6	3.99	7
Brazil	1,583.2	2.88	10	882.5	1.99	10

Source : 2005 International Comparison Programme : Tables of Final Results, 2008

Per Capita GDP at PPP

Per capita measures of GDP at PPP are useful for comparing the average standard of living in different economies. In 2005, the top five economies with the highest per capita GDP at PPP were Luxembourg, Qatar, Norway, Brunei and Kuwait (Chart 2). These economies, all very small and mostly oil-rich, accounted for less than 1 per cent of the world economy.

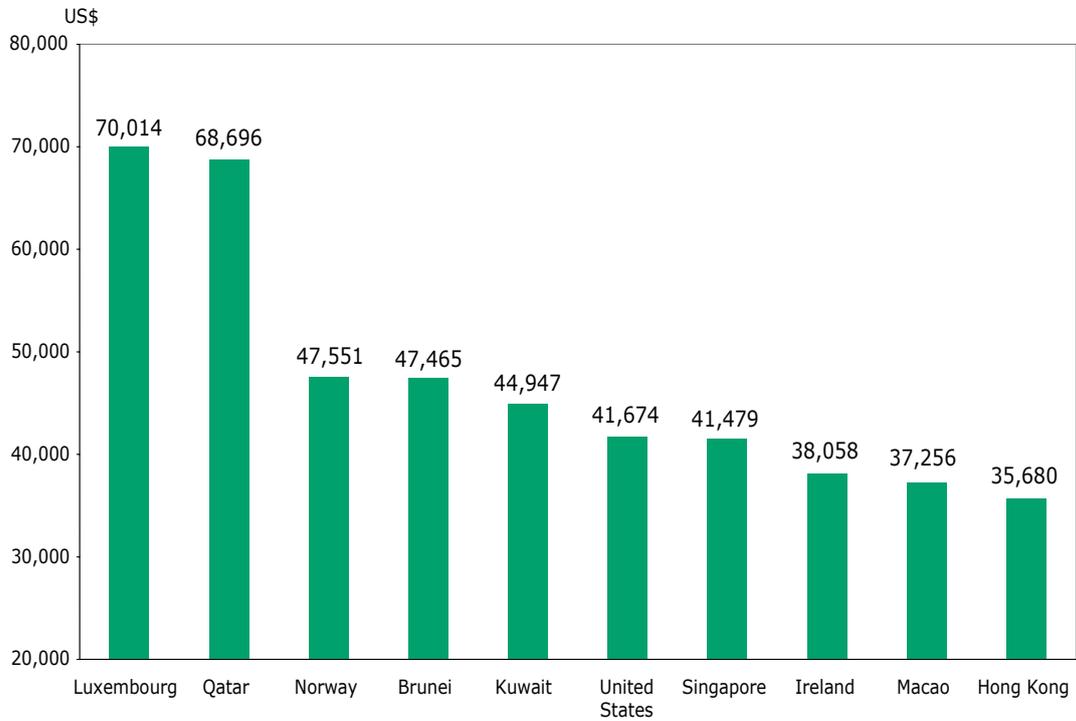
Singapore's PPP-based per capita GDP at US\$41,479 was also amongst the highest in the world and comparable to the United States. Within the Asia Pacific region, Singapore's per capita GDP at PPP was ranked second, behind Brunei (US\$47,465)

and above Macao (US\$37,256) and Hong Kong (US\$35,680).

Per Capita Actual Final Consumption of Households (AFCH)

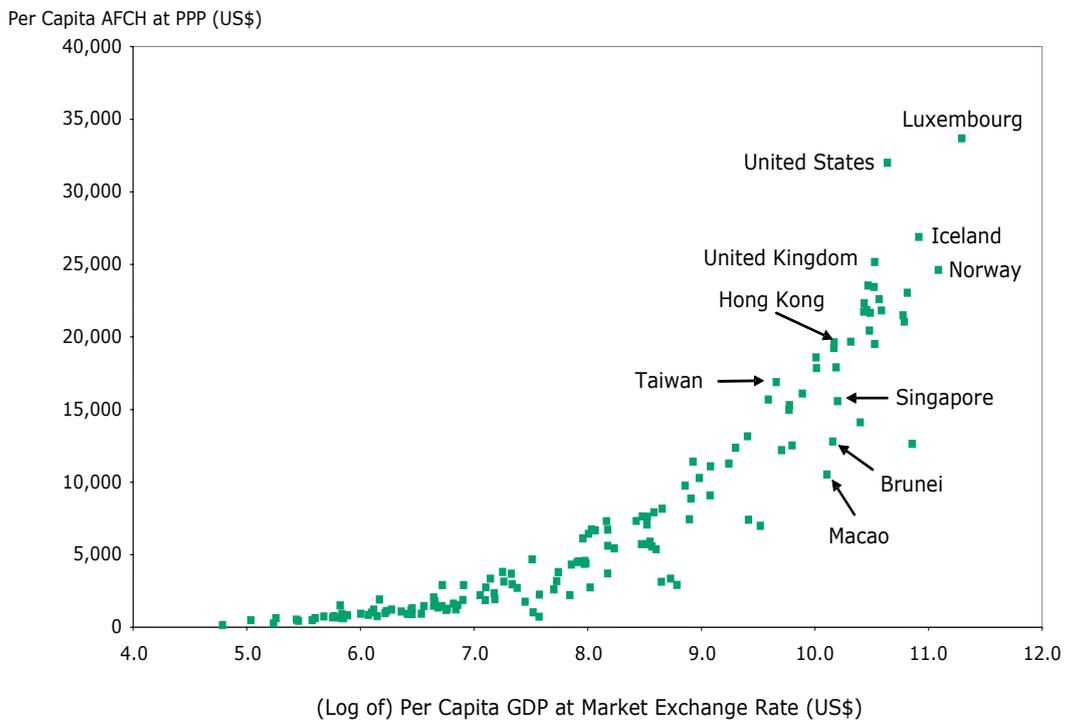
AFCH measures the amount of goods and services consumed by households. High-income economies generally have higher per capita AFCH. Based on the 2005 ICP, Luxembourg, United States, Iceland, United Kingdom and Norway had the highest per capita AFCH at PPP while Singapore's per capita AFCH at PPP (US\$15,564) was about two and a half times the world average (US\$6,095) and the third largest among high-income Asia Pacific economies (i.e. Brunei, Macao, Hong Kong and Taiwan) (Chart 3).

CHART 2 ECONOMIES WITH HIGHEST PER CAPITA GDP AT PPP, 2005



Source : 2005 International Comparison Programme : Tables of Final Results, 2008

CHART 3 PER CAPITA AFCH AT PPP, 2005



Source : 2005 International Comparison Programme : Tables of Final Results, 2008

Per Capita Gross Fixed Capital Formation (GFCF)

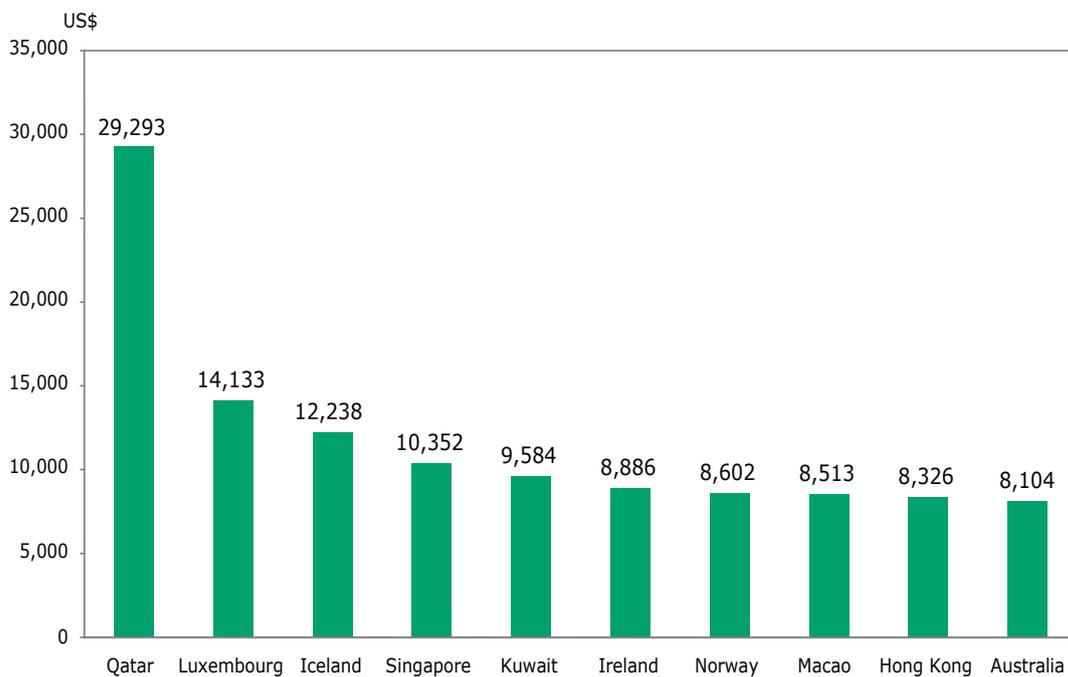
In terms of GFCF – that is, investments in buildings & infrastructure, machinery, equipment and software, which are important drivers of economic growth – economies with the highest per capita spending were Qatar, Luxembourg, Iceland, Singapore and Kuwait (Chart 4). Singapore’s per capita GFCF at PPP, for instance, was roughly five and half times the world average.

Conclusion

Unlike market exchange rates, PPPs adjust for differences in price levels between countries/economies and enable more robust and meaningful cross-country comparisons of macro-economic aggregates.

The updated PPPs with 2005 as the reference year will serve as important inputs to the World Bank, IMF and other international agencies in developing and monitoring policies, programs and targets.

CHART 4 ECONOMIES WITH HIGHEST PER CAPITA GFCF AT PPP, 2005



Source : 2005 International Comparison Programme : Tables of Final Results, 2008

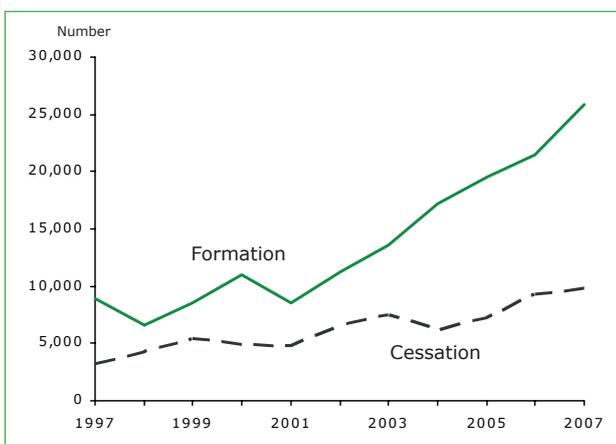
Formation and Cessation of Companies and Businesses, 2007

Companies

The number of companies incorporated in 2007 was 25,900, an increase of 21 per cent from 2006. All major industries recorded higher number of company formations. In particular, double-digit growths were observed in industries such as real estate, rental & leasing (76 per cent), financial & insurance (44 per cent), construction (32 per cent), hotels & restaurants (32 per cent) and education, health & social work (29 per cent).

There were 9,880 company cessations in 2007, a rise of 6.3 per cent from the 9,290 recorded in 2006. Most major industries recorded increases in company cessations, ranging from 3.1 per cent (hotels & restaurants) to 30 per cent (arts, entertainment, recreation & other service). Conversely, declines in company cessations were registered in construction (-9.2 per cent), transport & storage (-3.1 per cent) and administrative & support service (-3.0 per cent).

CHART 1 FORMATION AND CESSATION OF COMPANIES

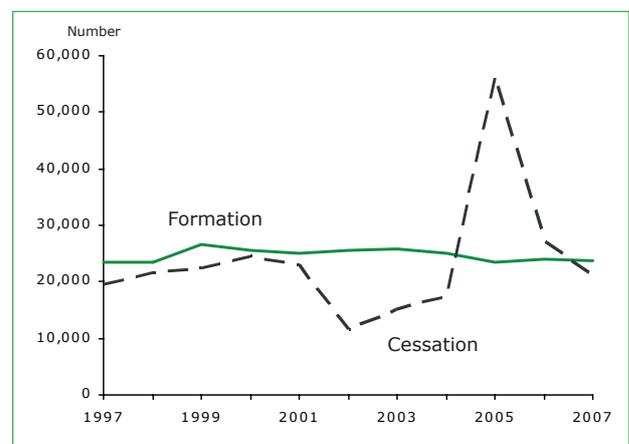


Businesses

The number of businesses formed declined slightly from 24,090 in 2006 to 23,850 in 2007. Industries which recorded declines in business formations included wholesale & retail trade (-6.7 per cent), information & communications (-6.1 per cent), and arts, entertainment, recreation & other service (-5.9 per cent). In contrast, growths in business formations were observed in real estate, rental & leasing (34 per cent), transport & storage (16 per cent) and construction (10 per cent).

The number of business cessations declined by 22 per cent from 27,040 in 2006 to 21,160 in 2007. All major industries recorded decreases in business cessations, with the larger declines recorded in administrative & support service (-30 per cent), construction (-29 per cent) and professional, scientific & technical (-27 per cent).

CHART 2 FORMATION AND CESSATION OF BUSINESSES



2007 in Brief

Singapore's population

... reached a total size of 4.59 million in June 2007.

Mean years of schooling

... was 9.9 years for male resident non-students aged 25 years and over and
8.8 years for female resident non-students aged 25 years and over.

Home ownership rate

... was 91 per cent.

Labour force participation rate

... was 76.5 per cent among males and
54.3 per cent among females.

Resident unemployment rate (seasonally adjusted)

... was 3.1 per cent.

Per capita Gross National Income

... was S\$51,119.

Official foreign reserves

... increased to S\$235 billion.

Visitor arrivals

... was 10 million.

Mobile phone subscribers

... reached 1,225 per 1,000 population.

Residential broadband subscribers

... increased to 252 per 1,000 resident population.

Crime rate

... declined to 695 per 100,000 population.

Change in 2007 (Year-on-Year)

Economic Growth

... was 7.7 per cent.

Annual growth in total trade (at 2006 prices)

... reached 7.0 per cent.

Unit Labour Cost Index

... was 3.8 per cent higher.

Consumer Price Index

... increased by 2.1 per cent.

Domestic Supply Price Index

... rose by 0.1 per cent.

Retail Sales Index

... was 3.8 per cent higher.

Catering Trade Index

... was 5.5 per cent higher.

Domestic Wholesale Trade Index

... grew by 8.6 per cent.

Foreign Wholesale Trade Index

... rose by 9.1 per cent.

Business Receipts Index

... increased by 14.8 per cent.

Industrial Production Index

... rose by 5.9 per cent.

Singaporeans are living longer...

Find out more about life expectancy in Singapore and how it compares with developed countries/areas in our latest Information Paper (IP), *Complete Life Tables 2003–2006 for Singapore Resident Population*.

Softcopy of this IP may be downloaded from the SingStat website at

<http://www.singstat.gov.sg/pubn/papers/people/ip-s13.pdf>

Household income grew amidst strong economic performance...

Highlights on the key trends in household income from work in 2007 among employed households are available from our latest Occasional Paper (OP), *Key Household Income Trends, 2007*.

Softcopy of this OP may be downloaded from the SingStat website at

<http://www.singstat.gov.sg/pubn/papers/people/op-s14.pdf>

Overseas Visitors

The Singapore Department of Statistics (DOS) received the following visitors over the past seven months.

Topics discussed included consumer price index and census activities in Singapore, multi-modal data collection strategy in business surveys, data anonymisation, and scope, coverage and data collection of surveys on the services sector.

Overview of Singapore's statistical system, the conduct of annual services survey and compilation of value added for the services industries, as well as the compilation of Singapore Gross Domestic Product by output and expenditure approaches were presented.

Bahrain

- *Bahrain Central Informatics Organisation*
- Dr Nabeel Shams
Director of Statistics

China

Officials from statistical agencies in Beijing visited DOS. The agencies are :

- *Beijing Municipal Bureau of Statistics*
- *National Bureau of Statistics*

Switzerland

- *University of Fribourg*
- Prof Hans Wolfgang Brachinger
Department of Quantitative Economics

United States

- *IBM Software Group*
- Mr Jeff Jonas
IBM Distinguished Engineer and
Chief Scientist, Entity Analytic Solutions

Contents

Improving Sampling Efficiency for the Annual Survey of Services	1
Purchasing Power Parity and the 2005 International Comparison Program	7
Formation and Cessation of Companies and Businesses, 2007	13
2007 in Brief	14
Overseas Visitors	15

The *Statistics Singapore Newsletter* is issued half-yearly by the Singapore Department of Statistics. It aims to provide readers with news of recent research and survey findings. It also serves as a vehicle to inform readers of the latest statistical activities in the Singapore statistical service.

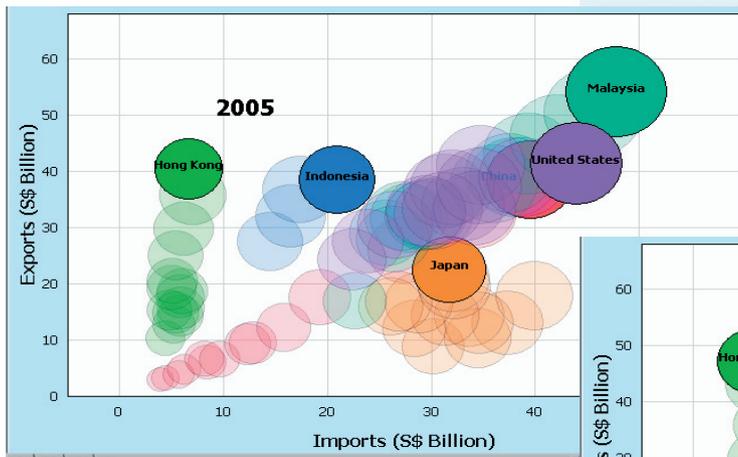
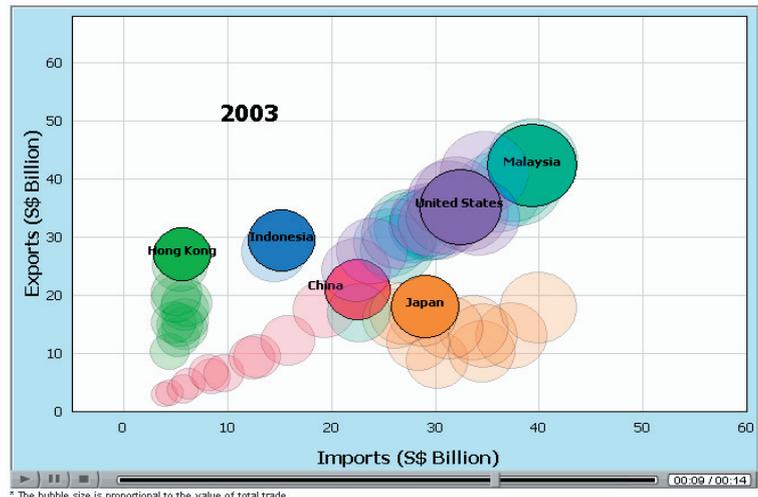
Contributions and comments from readers are welcomed. Please address all correspondence to :

The Editor
Statistics Singapore Newsletter
100 High Street #05-01
The Treasury
Singapore 179434

Fax : 65 6332 7689
Email : info@singstat.gov.sg

Trade with Major Trading Partners

Do you know that total trade between Singapore and the major trading partners have grown over the past decade?

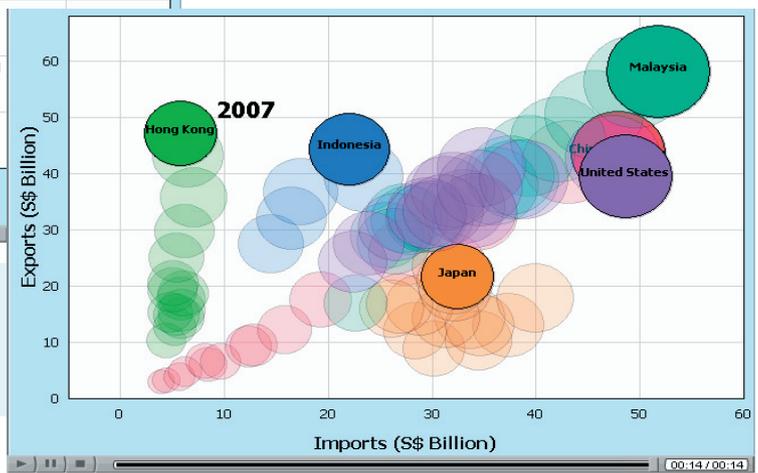


Over the past decade, total trade between Singapore and the major trading partners such as Malaysia, China, US, Indonesia, Japan and Hong Kong have grown. Among them, Singapore's trade with China expanded most rapidly. The values of exports exceeded imports for Hong Kong during 1993 - 2007.



* The bubble size is proportional to the value of total trade.

Over the past decade, total trade between Singapore and the major trading partners such as Malaysia, China, US, Indonesia, Japan and Hong Kong have grown. Among them, Singapore's trade with China expanded most rapidly. The values of exports exceeded imports for Hong Kong during 1993 - 2007.



* The bubble size is proportional to the value of total trade.

Over the past decade, total trade between Singapore and the major trading partners such as Malaysia, China, US, Indonesia, Japan and Hong Kong have grown. Among them, Singapore's trade with China expanded most rapidly. The values of exports exceeded imports for Hong Kong during 1993 - 2007.



See these and more data trends move through time with dynamic charts posted on the "Statistics Visualiser" section of the SingStat website at

<http://www.singstat.gov.sg/stats/stats.html#explorer>