Bootstrapping Australian inbound tourism

Y.H. Cheung[a] and G. Yap[a]

[a]*School of Finance, Accounting, and Economics, Faculty of Business and Law, Edith Cowan University, 270 Joondalup Drive, Joondalup, Western Australia, WA 6027, Australia.*

Email: y.cheung@ecu.edu.au

Abstract: This paper adopts a model-free approach to forecasting monthly international tourist arrivals to Australia from four major origin countries: New Zealand, UK, the USA, and China. While most researchers use parametric methodologies to model tourism demand, this study proposes a non-parametric approach that employs a 'Partitive Simulation Process' or PASIP by partitioning the original monthly time series into 12 sub-series according to the month. Both ordinary and time-weighted non-parametric bootstraps are used, to resample the observed samples 2,000 times, to estimate the underlying population statistics. We then conduct PASIP's forecasts and compare them with ARIMA forecasts. There are four significant observations. First, the partitive process is a viable way of handling seasonality. Second, PASIP performs well in predicting the turning points and data trends. Third, weighted PASIP generally outperforms non-weighted PASIP in terms of forecast errors. Fourth, PASIP produces smaller forecast errors than ARIMA for UK, USA, and China data.

1. INTRODUCTION

Surveys of tourism demand (Li et al., 2005; Song and Li, 2008) suggested that no single model could generate the best forecasts for all cases. In fact, a study by Witt and Witt (1995) discovered that naïve and autoregressive models outperformed most of the analytical approaches. Neither type of model employs any variable other than the tourism demand data itself. Accordingly, this paper forecasts Australian inbound tourism, without specifying an analytical or time-series model. The approach is particularly useful in situations where little information is available.

This study focuses on the top four countries contributing to Australian inbound tourism in 2009: New Zealand, the UK, the USA, and China. We partition the data from January 1991 to December 2009 into 12 monthly sub-series and resample the data 2,000 times using ordinarily and time-weighted non-parametric bootstrap processes for forecasting the inbound tourist numbers during January 2010 to December 2010. We call our approach the 'Partitive Simulation Process' or PASIP. The forecasts are thereafter compared with the outcomes of ARIMA models.

2. NON-PARAMETRIC BOOTSTRAP

The method of non-parametric bootstrap was originally conceived by Efron (1979, 1981, and 1982) as an alternative to jackknife and further developed by Efron and Tibshirani (1993). Bootstrap uses the bootstrap samples (or resamples) to efficiently approximate the underlying population distribution given the following conditions. First, the observed sample is a reasonable representation of the underlying population. Second, the observed sample size is reasonably large, no less than 10 and ideally greater than 30 (see Schenker, 1985; Mooney and Duval, 1993, p.44).[1] Third, the number of bootstrap samples is large but the improvement in the estimated statistic increases at a decreasing rate when the number of bootstrap samples exceeds 1,000, that being the minimum number of bootstrap samples required (Babu and Singh, 1983).[2]

The essence of the bootstrap method is as follows. Suppose we observe a sample $x = \{x_1, x_2, \cdots, x_n\}$ from an unknown population distribution function $F(X)$ with mean μ and standard deviation σ in the real world. To obtain an unbiased maximum likelihood (ML) estimation of a population statistic θ (which may be a mean, a variance, a 95th percentile), we can resample the observed sample x B times, with replacement, to obtain the bootstrap samples $x_b^* = \{x_1^*, x_2^*, \cdots, x_n^*\}$, $b = 1, 2, \cdots, B$. Resampling mimics the random process of the underlying system that generated $F(X)$. If x_i^*'s are assumed to be uniformed distributed, that is, $P(x_i^* = x_j | x) = 1/n \ \forall \ i, j \in [1, n]$, we have ordinary non-parametric bootstrap. If x_i^*'s are assumed to be non-uniformed distributed, that is, $P(x_i^* = x_j | x) = p_j \ \forall \ i, j \in [1, n]$ and $p_i \geq 0 \ \forall i$ and $\sum_i p_i = 1$, we have weighted non-parametric bootstrap. With the B bootstrap samples, we can calculate a bootstrap sample statistic $\hat{\theta}^*$ to be our bootstrap estimator of the underlying population statistic θ. To evaluate bootstrap estimator, we examine the standard error of the bootstrap sample statistics $\sigma_{\theta^*} = \sqrt{\sum_{b=1}^{B} (\theta_b^* - \hat{\theta}^*)/(B-1)}$. The smaller the standard error, the better is the estimate. For example, if we wish to estimate the mean μ of $F(X)$, we calculate the bootstrap sample means μ_b^*. The mean of the B bootstrap samples, given by $\hat{\mu}^* = \left(\sum_{b=1}^{B} \mu_b^*\right)/B$, is the bootstrap estimate of μ. If $F(X)$ is skewed, then the average of the B bootstrap sample median will be used as the bootstrap estimate of the median of $F(X)$. To estimate the $(1 - \alpha)$ per cent confidence interval for the estimated mean, we calculate the $\left[\mu_b^{*\left(\frac{\alpha}{2}\right)}, \mu_b^{*\left(1-\frac{\alpha}{2}\right)}\right]$ percentile confidence interval by ranking the resample means ordinally.

3. THE PARTITIVE SIMULATION PROCESS

Retail sales, for example, exhibit a seasonal pattern. And approaching Christmas, public media ask retailers about their sales expectations for Christmas. Repeatedly, anecdotic evidence shows that retailers base their

expectations largely on two variables: last year Christmas sales and current consumer sentiment (sales momentum).

Similar to retail sales, inbound tourism from each origin country follows its own seasonal pattern and momentum (see Figure 1). Instead of using time-series analysis, this study adopts a different strategy to capture similar forces at work in the tourism data. The time series of each origin country is partitioned into 12 sub-series each containing the year-on-year data for a particular month, from which the year-on-year rates of change are calculated. Two motives for partitioning are present. Firstly, it enables us to circumvent seasonality yet capture the momentum of each month. Secondly, since there is no general concept about the distribution of the year-on-year rates of change, partitioning allows for the possibility that the 12 monthly probability density functions can differ.
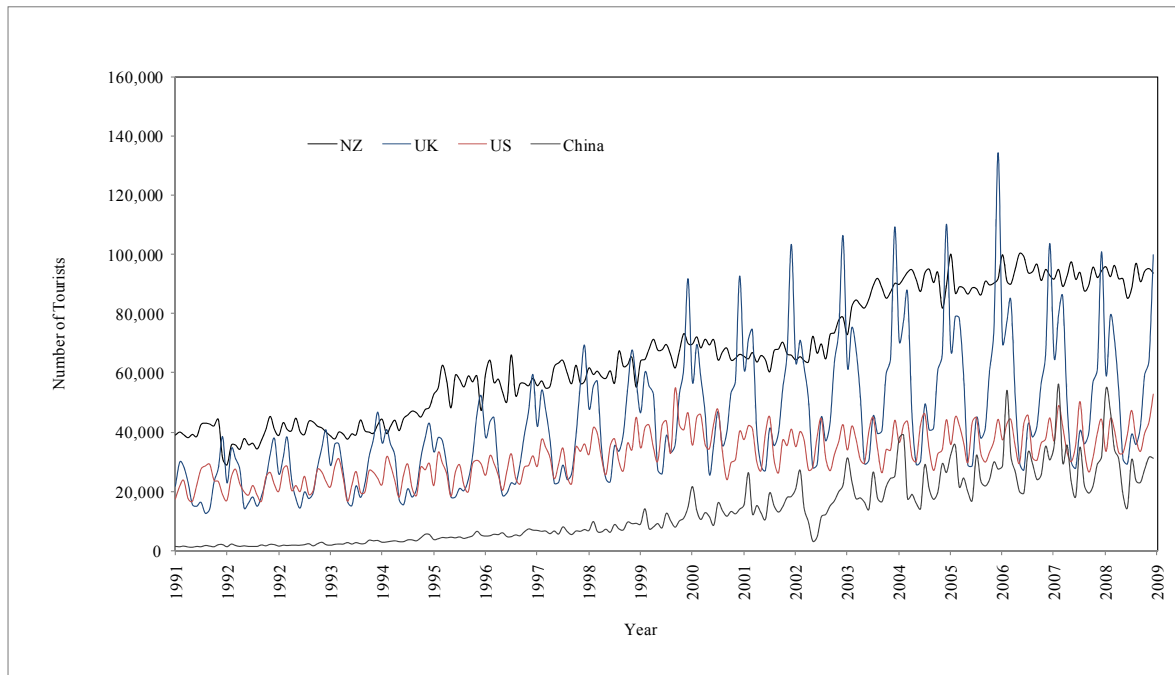

Figure 1. Top Four Origin Countries' Contribution to Australian Inbound Tourism, 1991-2009

Source: Australian Bureau of Statistics

We bootstrap the year-on-year rates of change to arrive at the ordinary bootstrap estimates of the means and the medians of the growth rates of inbound tourism, upon which the estimate with the smallest standard error is used to perform one-step ahead forecast. For example, the coming January inbound tourist number from USA is forecasted by multiplying last January number with one plus the estimated growth rate for January, and this pattern is repeated for each of the twelve months. We call this procedure the Partitive Simulation Process or PASIP to reflect that the twelve months are simulated separately on a year-on-year basis.

Ordinary bootstrap implicitly assumes that tourism preferences remain the same over time and that recent events are no more important than ancient events. To recognise that preferences may change and recent events may have larger impacts on future inbound tourism, the year-on-year rates of change are weighted by age. The *ad hoc* time-varying weights are generated by a truncated geometric series: $G(n) = p(1-p)^{n-1}$ with $p = 0.5$. It is truncated at the point where $f(n) = 0$ corrected to four decimal places. To differentiate the time-weighted and the non-time-weighted procedure, we call the former Weighted PASIP and the latter Non-weighted PASIP. Figure 2 summaries the simulation procedure using PASIP.
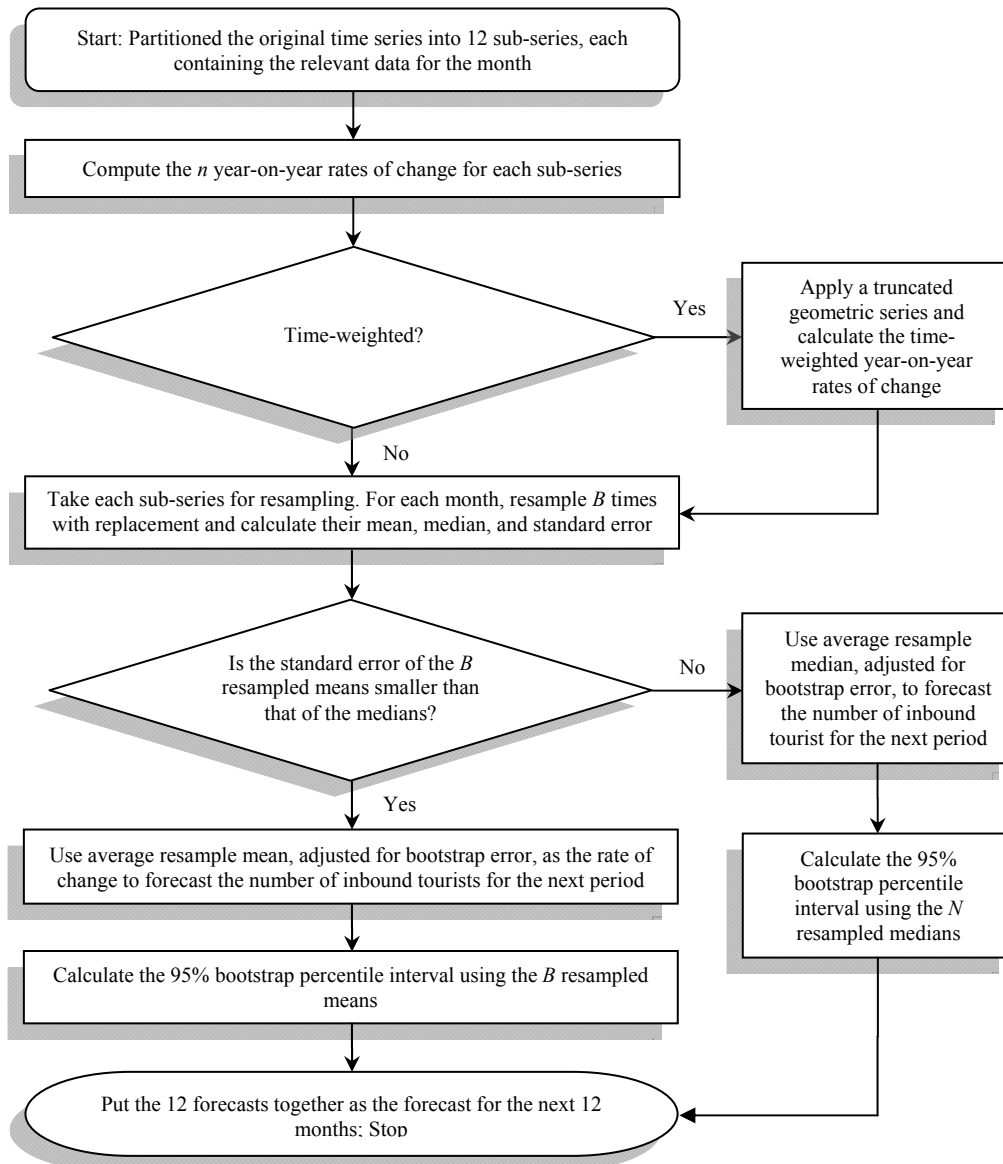
```
┌─────────────────────────────────────────────────┐
│ Start: Partitioned the original time series into │
│ 12 sub-series, each containing the relevant data │
│ for the month                                    │
└─────────────────────────────────────────────────┘
                        │
                        ▼
┌─────────────────────────────────────────────────┐
│ Compute the n year-on-year rates of change for   │
│ each sub-series                                  │
└─────────────────────────────────────────────────┘
                        │
                        ▼
             ◇ Time-weighted? ◇ ──Yes──▶ ┌──────────────────────┐
                        │                │ Apply a truncated    │
                        No               │ geometric series and │
                        │                │ calculate the time-  │
                        ▼                │ weighted year-on-year│
                                         │ rates of change      │
                                         └──────────────────────┘
┌─────────────────────────────────────────────────┐
│ Take each sub-series for resampling. For each    │
│ month, resample B times with replacement and     │
│ calculate their mean, median, and standard error │
└─────────────────────────────────────────────────┘
                        │
                        ▼
      ◇ Is the standard error of the B            ┌──────────────────────┐
        resampled means smaller than ◇ ──No──▶    │ Use average resample │
        that of the medians?                      │ median, adjusted for │
                        │                         │ bootstrap error, to  │
                        Yes                        │ forecast the number  │
                        │                         │ of inbound tourist   │
                        ▼                         │ for the next period  │
┌─────────────────────────────────────────────┐  └──────────────────────┘
│ Use average resample mean, adjusted for      │              │
│ bootstrap error, as the rate of change to    │              ▼
│ forecast the number of inbound tourists for  │  ┌──────────────────────┐
│ the next period                              │  │ Calculate the 95%    │
└─────────────────────────────────────────────┘  │ bootstrap percentile │
                        │                         │ interval using the N │
                        ▼                         │ resampled medians    │
┌─────────────────────────────────────────────┐  └──────────────────────┘
│ Calculate the 95% bootstrap percentile       │
│ interval using the B resampled means         │
└─────────────────────────────────────────────┘
                        │
                        ▼
        ( Put the 12 forecasts together as the forecast
          for the next 12 months; Stop )
```

Figure 2. Flow Chart of PASIP

4.   EMPIRICAL FINDINGS

Preliminary data analysis shows that very few of the probability distributions of the year-on-year rates of change are symmetrically distributed. This view is further confirmed by the substantial differences between the bootstrap means and the bootstrap medians in the non-weighted PASIP. The finding also justifies the use of non-parametric approach in forecasting Australian inbound tourism.

We perform a simulation according to the flow chart presented in Figure 2. We use both non-weighted and weighted PASIP to conduct out-of-sample forecasts for January to December 2010. We resample the year-on-year rates of change 2000 times to arrive at the bootstrap estimates of the mean and the median of the growth rate of inbound tourism for each month and for each origin country.

For all four origin countries, the standard errors of the bootstrap means of the year-on-year rates of change of the non-weighted PASIP tend to be relatively smaller than those for the bootstrap medians, whereas the opposite pattern is evident for the weighted PASIP. As a result, the bootstrap means are largely used for forecasting out-of-sample numbers in non-weighted PASIP, while the bootstrap medians are employed to forecast out-of-sample numbers in weighted PASIP. This dichotomous phenomenon is due to the *ad hoc* truncated geometric weighting scheme applied which skews the distribution to the left in weighted PASIP. A

different weighting scheme would have produced very different results.[3] Since the probability distributions are skewed to the left in the weighted PASIP, the growth rates used to forecast the out-of-sample numbers would be smaller than that of those in the non-weighted PASIP. As a result, inbound tourist numbers tend to be smaller under the weighted PASIP (see Table 1A and 1B for comparison).

We also performed a time-series analysis of the original data using ARIMA models as a benchmark for examining the performance of PASIP. A variety of models were fitted to the original data and the best performing ARIMA model, in terms of R-square and selection criteria, is ARIMA (0,1,1) for New Zealand, ARIMA (6,1,1) for UK, Seasonal ARIMA (2,1,2) for the USA, and Multiplicative ARIMA $(0,1,0)$ $(1,1,12)_{12}$ for China. These ARIMA models have also undergone rigorous diagnostic testing to ensure reliable estimations. More information can be obtained upon request.

Table 1A and 1B present the 12-month forecasts with the different methods. Under each origin country, three forecasts are presented in the order of non-weighted PASIP, weighted PASIP, and ARIMA. Note that since the Australian authority reports the tourism data in the hundreds, we round digits less than 50 down and rounds digits larger than or equal to 50 up to produce numbers that are consistent with format of the reported numbers.

It is also interesting to observe that many of the forecasts in the non-weighted PASIP in UK, USA, and China come from using the bootstrap medians. Furthermore, few of the 96 bootstrap estimates of the underlying population means and underlying population medians are biased, and therefore, require adjusting for the bootstrap errors (which equal average of resampled statistic minus the observed statistic).

Table 1A and 1B reveal that PASIP is capable of picking up the ebb and flow of inbound tourism and performs well in predicting the turning points and trends of the data. For instance, the weighted PASIP for UK is able to predict lowest arrivals in June 2010 and a subsequent gradual increase. Similarly, the weighted PASIP traces the upward and downward movements well in the USA data. As for the Chinese tourist numbers, although the forecasting errors are much larger, the non-weighted PASIP predictions also follow the observed data well.

Table 1A. Twelve-month Forecasts for New Zealand and UK, January to December 2010

Year 2010	NZ				UK			
	Observed	Non-weighted-PASIP	Weighted-PASIP	ARIMA (0,1,1)	Observed	Non-weighted-PASIP	Weighted-PASIP	ARIMA (6,1,1)
Jan	93,800	101,700	*96,000*	94,300	62,800	62,900	*59,200*	41,300
Feb	97,100	97,900	*92,700*	98,200	73,900	84,000	*79,300*	84,700
Mar	93,900	102,100	*96,500*	93,400	76,900	76,900	*71,600*	67,900
Apr	94,200	97,200	91,900	96,400	43,100	57,600	*54,000*	24,700
May	90,900	96,800	*91,800*	92,700	30,100	*31,700*	*30,800*	29,700
Jun	94,900	89,900	*85,300*	98,600	26,300	*30,400*	*29,100*	20,300
Jul	97,000	92,900	*88,500*	97,500	35,700	41,200	39,000	46,500
Aug	99,100	102,500	*97,200*	99,000	33,600	*37,200*	*35,700*	28,800
Sep	99,500	95,200	*91,000*	99,100	38,900	44,800	*41,700*	37,200
Oct	97,900	99,000	*94,500*	98,200	55,300	*60,600*	*59,200*	71,400
Nov	98,400	99,900	*95,400*	100,300	64,200	67,600	*64,200*	78,400
Dec	100,000	100,000	*93,900*	101,600	105,900	*106,500*	*99,800*	127,400

Note: Italicised numbers in the PASIP columns represent projection from median. The observed data is retrieved from the Australian Bureau of Statistics.

Table 1B. Twelve-month Forecasts for USA and China, January to December 2010

Year 2010	USA				China			
	Observed	Non-weighted-PASIP	Weighted-PASIP	Seasonal ARIMA (2,1,2)	Observed	Non-weighted-PASIP	Weighted-PASIP	Multiplicative ARIMA (0,1,0) $(1,1,12)_{12}$
Jan	38,800	34,900	*33,600*	30,200	37,200	*69,200*	*54,700*	45,400
Feb	45,900	46,500	*44,600*	30,200	73,200	59,500	*47,000*	62,200
Mar	42,900	42,300	*40,900*	51,000	35,300	41,200	*34,300*	21,100
Apr	31,900	34,600	*33,200*	23,900	32,500	*38,000*	*30,900*	23,400
May	32,900	*35,000*	*33,100*	23,000	26,200	*22,000*	*18,800*	26,400
Jun	41,200	40,900	*39,300*	43,100	22,800	*16,800*	*14,400*	24,500
Jul	42,500	49,000	*47,400*	53,300	47,600	38,100	*30,700*	78,800
Aug	36,800	*39,300*	*37,900*	35,000	36,800	*27,700*	*23,600*	38,300
Sep	32,300	*35,000*	*33,600*	29,400	29,000	27,100	*23,000*	21,600
Oct	34,500	*41,300*	*39,700*	23,600	29,300	32,900	*27,400*	33,400
Nov	39,900	45,300	*43,700*	33,800	42,400	*36,700*	*31,500*	48,300
Dec	52,600	*55,000*	*52,900*	43,900	41,500	36,900	*31,200*	40,400

Note: Italicised numbers in the PASIP columns represent projection from median. The observed data is retrieved from the Australian Bureau of Statistics.

To ascertain the out-of-sample forecasting power of PASIP, we compare the out-sample forecast performance of non-weighted and weighted PASIP as well as ARIMA models for the four origin countries from January to December 2010. To measure their forecasting accuracy, we adopt the Root Mean Squares Errors (RMSE) as the criterion. Table 2 presents the RMSE's for each model/process given the origin countries. We find that weighted PASIP outperforms ARIMA and non-weighted PASIP for the UK and the USA data, whereas non-weighted PASIP generates the best forecasts for China. Nevertheless, ARIMA (0,1,1) still produce the best predictions for tourist arrivals from New Zealand. In general, PASIP successfully generates better forecasts than ARIMA models for three out of four origin countries during the nominated period.

Table 2. RMSE of various models

Country	Non-weighted PASIP	Weighted PASIP	ARIMA	Verdict
NZ	4,560	5,284	1,831	ARIMA (0,1,1)
UK	6,132	4,769	18,710	Weighted PASIP
USA	3,704	2,980	15,097	Weighted PASIP
China	11,452	12,403	14,923	Non-weighted PASIP

Note: The figures are the square root of forecast errors accumulated from January to December 2010.

For the Chinese data, none of the models presented perform well. One of the possible reasons is that the tourist numbers are not market determined. The Chinese authority has tight administrative control on outbound groups and individual tourists by imposing an annual quota for each type of tourists in every province and municipality. In fact, such quotas may be subject to change overnight. Therefore, such administrative control is beyond the capability of any quantitative modelling.

Even if we exclude the weighted PASIP, three of the four RMSE's of the non-weighted PASIP are still lower than that of the ARIMA models, rendering the PASIP an attractive alternative in forecasting Australian inbound tourism.

5.   SUMMARY AND CONCLUDING REMARKS

This paper explores an alternative way of forecasting Australian inbound tourism. It proposes the use of non-parametric bootstrap in conducting the PASIP. In the process, either the bootstrap estimates of the mean or the median of the underlying probability distribution functions are employed to perform one-step ahead forecasting.

We find four significant observations about the proposed PASIP. First, the partitive process is appropriate for dealing with seasonality. Second, it predicts the ebb and flow of the time series and performs well in predicting the turning points and data trends. Third, weighted PASIP generally outperforms non- weighted PASIP in terms of forecast errors. Fourth, it produces relatively smaller forecast errors than ARIMA for the UK, the USA, and the China data, which confirms the observation made by Witt and Witt (1995).

This exercise may be the first step into a new direction for modelling tourism demand. Some potential improvements are evident. First, future research can investigate the possible reasons of why non-weighted PASIP tends to outperform weighted-PASIP for some countries, and vice versa. Second, it can investigate the derivation of an *a priori* mechanism for combining the non-weighted and weighted PASIP forecasts. Third, developing computer codes for PASIP using Microsoft Excel VBA could be useful for tourism stakeholders to generate forecasts.

NOTES

1.   For cross-sectional data with sample size greater than 30 and time-series data with sample size greater than 120, we can justify using the asymptotic results.

2.   When the number of bootstrap samples exceeds 1,000, the improvement in the estimated statistic increases at a decreasing rate.

3.   Note that the issue of optimal weights is considered to be beyond the scope of this paper.

REFERENCES

Babu, G.J. and K. Singh. (1983). Inference on means using the bootstrap. *Annals of Statistics*, 11(3), 993-1003.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, 7(1), 1-26.

Efron, B. (1981). Non-parametric estimates of standard error: The jackknife, the bootstrap and other methods. *Biometrika*, 68 (3), 589-599.

Efron, B. (1982). *Bootstrap the jackknife, the bootstrap, and other resampling plans*. Society of Industrial and Applied Mathematics CBMS-NSF Monographs, Vol 38, SIAM, Philadelphia.

Efron, B, and R.J. Tibshirani. (1993). *An Introduction to the bootstrap*. Chapman and Hall, London.

Li, G., H. Song, and S.F. Witt. (2005). Recent developments in econometric modelling and forecasting. *Journal of Travel Research*, 44 (1), 82-99.

Mooney, C.Z., and R.D. Duval. (1993). *Bootstrapping*: *A non-parametric approach to statistical inference*. Quantitative Applications in the Social Sciences, Vol 95. Sage Publications, Newsbury Park, California.

Schenker, N. (1985). Qualms about bootstrap confidence intervals. *Journal of the American Statistical Association*, 80 (380), 360-361.

Song, H., and G. Li. (2008). Tourism demand modelling and focasting: A review of recent research. *Tourism Management*, 29 (2), 203-20.

Witt, S.F., and C.A. Witt. (1995). Forecasting tourism demand: A review of empirical research. *International Journal of Forecasting,* 11 (3), 447-475.