

2012/39



Total tourist arrival forecast:
Aggregation vs. disaggregation

Shui-Ki Wan, Shin-Huei Wang
and Chi-Keung Woo



CORE

DISCUSSION PAPER

Center for Operations Research
and Econometrics

Voie du Roman Pays, 34
B-1348 Louvain-la-Neuve
Belgium

<http://www.uclouvain.be/core>

CORE DISCUSSION PAPER
2012/39

**Total tourist arrival forecast:
aggregation vs. disaggregation**

Shui-Ki WAN¹, Shin-Huei WANG²
and Chi-Keung WOO³

October 2012

Abstract

Total tourist arrivals are the sum of disaggregate subcomponent arrivals by country of origin. We use seven time-series models to assess whether the aggregate approach that directly forecasts the total tourist arrivals outperforms the disaggregate approach that produces the total arrival forecast as an unweighted sum of its subcomponent forecasts. Based on Hong Kong's monthly tourist arrival data, we find (a) the seasonal autoregressive integrated moving average model outperforms the other non-seasonal and seasonal models under the aggregate approach, and (b) forecast performance can be improved by the disaggregate approach.

Keywords: tourism demand, aggregate and disaggregate approaches, forecast combination, seasonal ARIMA, Holt-Winters.

JEL Classification: C01, C18

¹ Department of Economics, Hong Kong Baptist University.

² Université catholique de Louvain, CORE, B-1348 Louvain-la-Neuve, Belgium; CEREFIM, Facultés Universitaires Notre-Dame de la Paix, Namur, Belgium.

³ Department of Economics, Hong Kong Baptist University.

The first author acknowledges the financial support from the Hong Kong University Grant Council (HKBU255212). We thank Ira Horowitz of the University of Florida and Luc Bauwens of the Academie Louvain for their constructive comments. We also thank Alisha Woo for her editorial assistance. All remaining errors are ours.

1. INTRODUCTION

An accurate forecast of tourist arrivals is critical for the tourism industry, chiefly because of the perishable nature of tourism-related services, as exemplified by empty hotel rooms that cannot be used for future sales (Archer, 1987). Witt and Witt (1995) report that some commercial organizations produce international tourism demand forecasts, focusing on the total number of tourists traveling from a country of origin to a country of destination (e.g., Brooke, Buckley & Witt, 1985; Means & Avila, 1988). Large companies (e.g., major airlines and hotel chains) often generate forecasts internally, with varying time horizons (e.g., short- vs. long-term) and levels (e.g., local vs. regional) to suit their decision making. For example, short-term forecasts are required for scheduling and staffing, medium-term forecasts for planning tours and vacation packages, and long-term forecasts for planning capital expenditures. For economic growth and planning reasons, governments are also interested in total tourist-arrival forecasts, preferably differentiated by season and local destination.

Total tourist arrivals to a destination are the unweighted sum of disaggregate subcomponent arrivals by origin. Using seven time-series models and Hong Kong tourist arrival data, this paper investigates whether the aggregate approach that directly forecasts total tourist arrivals outperforms the disaggregate approach that derives the total tourist-arrival forecast as an unweighted sum of its subcomponent forecasts. The policy implication is clear: if the disaggregate approach outperforms the aggregate approach, it should be used in forecasting Hong Kong's total tourist arrivals.

The paper contributes to the tourism forecasting literature as follows. First, it uses Hong Kong's monthly tourist-arrival data to document the performance of alternative forecasts under the aggregate and disaggregate approaches. To the best of our knowledge, this is the first analysis of its kind for Hong Kong, one of the most popular tourist destinations in the world. Second, it finds that the seasonal autoregressive integrated moving average model (SARIMA) outperforms the other models under the aggregate approach. Third, it illustrates that the disaggregate approach outperforms the aggregate approach, corroborating the empirical evidence in other applications (e.g., Zellner and Tobias, 2000). Finally, it proposes a variant of the disaggregate approach that forms the total tourist forecast as a weighted sum of the subcomponent forecasts. While the variant generally improves the forecast performance of six of the seven models considered herein, it is not the case for the SARIMA model. Hence, we recommend applying SARIMA and the disaggregate unweighted-sum approach for forecasting Hong Kong's total tourist arrivals.

The rest of the paper proceeds as follows. To provide a contextual background, section 2 reviews the literature on the debate of the aggregate vis-à-vis the disaggregate approach in forecasting. Section 3 presents our empirical results, and Section 4 concludes.

2. AGGREGATE VS DISAGGREGATE APPROACH

The debate on aggregate versus disaggregate forecast modeling is not new (Theil, 1954; Grunfeld & Griliches, 1960). Many observed macroeconomic variables are simply an aggregate over a large number of micro units. An important issue addressed in this literature is the problem of aggregation bias, defined by the deviation of the macro parameters from the average of the corresponding micro parameters. One area of particular interest is the conditions under which macro models can accurately provide interpretable information on the underlying behavior of the micro units such as households and firms. For example, Hsiao, Shen, & Fujiki (2005) use Japanese aggregate and disaggregate money-demand data to show that conflicting inferences can arise. The aggregate data appear to support the contention of an unstable money-demand function. The disaggregate data, however, shows a stable money-demand function.

A closely related issue is the choice of a forecasting approach first discussed by Grunfeld & Griliches (1960), whose focus is whether to forecast aggregate variables using macro (the aggregate time series) or micro equations (the subcomponent time series of the aggregate variable). Since then, a series of empirical work comparing these two approaches emerge. For example, Fair & Shiller (1990) compare these approaches for United States GNP; Zellner & Tobias (2000) for industrialized countries' GDP growth; Marcellino, Stock, & Watson (2003) for GDP growth using disaggregated data from the Eurozone countries; Espasa, Senra, & Albacete (2002), Hubrich (2005) and Bruneau, De Bandt, Flageollet, & Michaux (2007) for inflation; and Kim & Moosa (2005) for tourist flows to Australia.

When the data--generating process (DGP) is known, aggregating subcomponent forecasts has a better performance and lower mean square error (MSE) than directly forecasting the aggregate, thanks to the larger information set hidden in the subcomponents (Kohn, 1982; Lütkepohl, 1984b, 1987, 2006; Granger, 1987; Pesaran, Pierse, & Kumar, 1989; Garderen, Lee, & Pesaran, 2000). The DGP, however, is usually unknown, and therefore has to be estimated. Suitably specifying the unknown DGP is necessary for the disaggregate approach to outperform the aggregate approach. Thus, deciding which approach has a better forecast performance becomes an empirical issue.

Motivated by the renewed interest in the aggregate vs. disaggregate issue, this paper uses Hong Kong's total tourist-arrival data to investigate whether contemporaneous aggregation of the subcomponent forecasts is preferable to directly forecasting the total tourist arrivals. Contemporaneous aggregation of forecasts can be divided into: (a) the aggregation of alternative forecasts generated by different forecasting models; or (b) the aggregation of subcomponents forecasts produced by a single forecasting model.

The forecast-combination literature addresses (a) by finding the optimal weights for model-specific forecasts, so as to minimize the MSE of the combined forecast. While combining model-specific forecasts tends to reduce MSE (Timmermann, 2006), there is no general consensus on how to derive the optimal weights (Bates & Granger, 1969; Granger & Ramanathan, 1984; Shen, Li, & Song 2011; Hsiao & Wan, in press). Moreover, there is empirical evidence that a simple unweighted average of model-specific forecasts may yield a lower MSE than a presumably "optimal" estimator (Huang & Lee, 2010).

The approach under (b) is not about finding optimal weights for the subcomponent forecasts. Rather, it postulates that exploiting the heterogeneity of the underlying subcomponent time series can improve the forecast of the aggregate variable (Theil, 1954). To be sure, Grunfeld & Griliches (1960) find that if the subcomponent forecast models are misspecified, the aggregate approach may outperform the disaggregate approach.

To crystalize the two approaches, consider y_t , the aggregate variable of interest, which in this paper is the total tourist arrivals in Hong Kong. The contemporaneous aggregate can be written as:

$$(1)$$

where $y_{i,t}$, $i=1,\dots,n$ are the n subcomponents of y_t . As will be seen below, each subcomponent in our empirical analysis is the total arrivals by region of origin.

All $y_{i,t}$ and y_t are observable. The aggregate variable can be forecasted in two ways. First, it can be a direct forecast \hat{y}_t as, for example, produced by a univariate time-series model. Alternatively, we can form the disaggregate forecasts, $\hat{y}_{i,t}$, $i = 1, \dots, n$ and then use the unweighted sum to form the final forecast $\hat{y}_t = \sum_{i=1}^n \hat{y}_{i,t}$. To gauge forecasting performance, we use the root-mean-square-error (RMSE) defined below:

$$(2)$$

where P is the number of out-of-sample forecast values.

There are two arguments that favor a disaggregate approach (Hubrich, 2005; Clements & Hendry, 2002). First, disaggregation can better capture the individual

dynamic properties, which likely differ across disaggregate variables. Second, disaggregate forecast errors may offset, leading to more accurate predictions of the aggregate.

When the forecast model is misspecified, especially in the presence of common shocks to the correlated disaggregate variables, the disaggregate approach can have a worse forecast performance than the aggregate approach. Thus, which approach should be used is an empirical issue, to be determined by the trade-off between: (a) the aggregate approach's error in ignoring the heterogeneity in the subcomponent time series; and (b) the disaggregate approach's error in specifying the disaggregate DGP. A case in point is Lütkepohl (1984a, 1987), whose Monte Carlo simulations show that the small-sample rankings of the two approaches are mixed and DGP-dependent.

Potential misspecification of the forecast model affects the accuracy of the resulting forecast, thus explaining that theoretical results on predictability are not always supported by empirical applications (Hendry, 2004; Clements & Hendry, 2006). Even when the disaggregate processes are appropriately specified, it does not necessarily mean that the disaggregate approach is always preferable to the aggregate approach. When there are unexpected shocks that commonly affect the disaggregate forecast errors in the same direction, adding up all the subcomponent forecasts does not lead to offsetting forecast errors.

To further explore the aggregation versus disaggregation debate, we propose a variant of the disaggregate unweighted-sum approach. Instead of simply adding the n individual forecasts, we apply ordinary least square (OLS) to the within-sample predictions to find the minimum variance weights w_i through the following ordinary least squares (OLS) regression:

$$(3)$$

The final forecast is $y_t = \sum_{i=1}^n w_i y_{it}$. The disaggregate unweighted-sum approach and its weighted-sum variant would yield the same total forecast under the null hypothesis below:

$$w_i = 1/n, \text{ for } i=1, \dots, n. \quad (4)$$

Rejecting this null hypothesis suggests that the weighted-sum variant may be preferable to the unweighted-sum approach.

In summary, the literature does not offer a clear answer regarding the relative performance of the two approaches in their application to forecasting tourist arrivals. Hence, the next section evaluates their relative performance using Hong Kong's monthly tourist-arrival data.

3. EMPIRICAL ANALYSIS

3.1 Tourism demand forecast methods

Studies on tourism demand can be divided into: (1) non-causal time-series modeling based on historic trends; and (2) causal modeling of the underlying factors that drive tourism demand (Song, Wong, & Chon, 2003).

With regard to (1) non-causal time series modeling, Dharmaratne (1995) uses the autoregressive integrated moving average (ARIMA) model to analyze the aggregate tourism arrivals in Barbados. Martin & Witt (1989) evaluate the performance of seven forecasting models for outbound tourism from France, Germany, U.K. and U.S.. Chu (1998) investigates six time-series forecasting techniques for tourist arrivals in ten different destinations. Using data on Australia's tourist arrivals from three regions, Lim & McAleer (2001) examine the forecasting performance of several exponential-smoothing models, including the Holt-Winters additive and multiplicative seasonal models. Further, Goh & Law (2002) use data on Hong Kong tourist arrivals from 10 regions to evaluate the forecasting performance of several time-series models. Chan, Lim, & McAleer (2005) use three multivariate static specifications to model the conditional mean and conditional variance of the logarithm of the monthly tourist arrivals from four countries to Australia. All these studies suggest the presence of heterogeneity in the disaggregate data series.

With regard to (2) causal modeling, Song & Witt (2000) and Song et al. (2003) use an autoregressive distributed lag model (ADLM) to analyze Hong Kong tourist arrivals by region of origin, so as to capture the dynamic effects of regional economic activities. Furthermore, Kulendran & Wilson (2000) apply an error correction model (ECM) to analyze business trips to Australia from four of Australia's most important travel and trade partners. Using data on inbound tourism to South Korea from four major origin countries, Song & Witt (2003) find that a single model specification across origin countries is not appropriate. Song & Witt (2006) apply a vector autoregressive (VAR) model to study the relationship between Macau's tourist arrivals by region of origin and a set of macroeconomic variables. Li, Wong, Song & Witt (2006) advance the causal models by integrating the ECM with a time-varying parameter specification. All these causal studies also indicate the presence of heterogeneity in the disaggregate data series.

While there is a wide range of tourism demand forecasting models, there is no single model that can outperform all others in all situations (Witt & Witt, 1995; Li, Song & Witt, 2005; Song & Li, 2008). Hence, a useful line of inquiry is to study the forecast performance of a combination of forecasts produced by alternative forecast models. Shen, Li, & Song (2008), Song, Witt, Wong & Wu (2009), and Wong,

Song, Witt & Wu (2007) document that forecast combination tends to improve forecast performance.

3.2 *Disaggregate data on Hong Kong tourist arrivals*

This paper extends the inquiry of forecast combination in a different direction. In particular, it recognizes that an aggregate time series is the unweighted sum of its disaggregate subcomponents. An important case in point is Hong Kong's total tourist arrivals, which is the unweighted sum of tourist arrivals by region of origin. To place our analysis in context, a brief background is given below.

Hong Kong, a special administrative region (SAR) of China that houses seven million people within its 1,000 square kilometers, boasted a per capita income of approximately US\$34,386 (\approx HK\$268,213 at HK\$7.8 per US\$1) in 2011.¹ Tourism is a major component of this vibrant city's GDP. As shown in Figure 1, there were over four million tourist arrivals in December 2011, representing a 16.4% year-to-year increase that affirms Hong Kong as one of the world's favorite travel destinations.² For the entire 2011, Hong Kong saw 41.9 million arrivals, with 22 million (52.6%) staying overnight and the remaining 19.9 million (47.4%) leaving on the same day of arrival.³ Total tourism expenditure associated with the 41.9 million arrivals in 2011 was about US\$33 billion (\approx HK\$263 billion).⁴

Tourism is the fourth key sector in Hong Kong, behind trading and logistics, professional services and other producer services, and financial services. In 1990-1996, it accounted for 3.2% to 3.6% of Hong Kong's GDP. Its GDP share fell to 2.8% in 1997 due to the Asian financial crisis and remained low at 2.2 to 2.5% between 1998 and 2001. In 2002, it rebounded to 3.0%, thanks to the regional economic growth and rising visits by tourists from Mainland China. This rebound, however, was sharply disrupted in 2003 by the SARS epidemic, causing monthly total tourist arrivals to plummet from 1.5 million to 0.5 million. In 2004, Hong Kong's tourist arrivals rose, due chiefly to the Individual Visit Scheme (IVS) under the Closer Economic Partnership Agreement (CEPA) signed between China and Hong Kong. Despite the 2007 subprime crisis in the U.S. and Europe, the tourism industry remained strong, contributing about 3.3% to Hong Kong's GDP. In 2010, tourism accounted for 4.4% of the GDP.

¹ Hong Kong Census and Statistics Department

² 2011 Hong Kong Economic Background and 2012 Prospects

³ Tourists staying overnight directly drive the demand for hotel rooms (Qu, Xu & Tan, 2002).

⁴ <http://www.legco.gov.hk/yr11-12/english/panels/fa/papers/fac1-1849-1-e.pdf>

The rising trend of total tourist arrivals in Figure 1 is accompanied by a changing mix of tourists by region of origin. Figure 2 shows that the last 10 years saw relatively more tourist arrivals from China and fewer from Europe, Africa and the Middle East, North Asia and the Americas, mirroring the 2004 IVS and China's unsurpassed economic growth of almost 10% per year. In 2011, the Mainland region's visitor arrivals surged by 23.9% to 28.1 million, accounting for 67% of the total arrivals.

3.3 Sample description

Our evaluation uses Hong Kong's total monthly arrival data by origin region, which is available from the Hong Kong Monthly Digest of Statistics. The aggregate variable of interest, denoted by "TOTAL", is the sum of the tourist arrivals from eight origin regions: (1) China; (2) South and Southeast Asia ("SASIA"); (3) North Asia ("NASIA"); (4) The Americas; (5) Europe, Africa & the Middle East ("EAM"); (6) Taiwan (7) Australia, New Zealand and the South Pacific ("ANS"); and (8) Macao.

Figure 3 shows the monthly data series for tourist arrivals for the period from August 2002 to December 2011, yielding a total sample of 113 monthly observations. China, SASIA and Macao display a relatively smooth upward linear trend. Taiwan fluctuates around the average. A preliminary analysis suggests that the remaining regions exhibit a quadratic trend. All series exhibit strong seasonality, with more arrivals in July through August and October through January than the remaining months, reflecting the relatively heavy summer and winter travel demand. All series have a large dip, reflecting the SARS epidemic in 2003.

3.4 Models

The disaggregate series portrayed by Figures 1 and 2 are likely heterogeneous, chiefly due to (a) Hong Kong's tourism promotion policies that vary by region and (b) each region's own domestic economic health, exchange rates against the Hong Kong dollars, and diverse preferences. A substantive question thus arises: can a forecast of Hong Kong's total tourist arrivals be improved by first modeling its subcomponents and then summing the individual subcomponent forecasts?

To answer this question, we use seven forecast models to evaluate the total tourist arrival forecast performance under the aggregate and disaggregate approaches: (1) a random-walk model; (2) a 12-month moving average model; (3) a historical mean model; (4) a seasonal random-walk model; (5) the Holt-Winters additive model; (6) the Holt-Winters multiplicative model; and (7) the seasonal ARIMA model.

The first three models do not adjust for seasonality and are regarded as benchmark (Goh & Law, 2002). The last four models adjust for seasonality. As shown

in Lim & McAleer (2001), the Holt-Winters additive seasonal model and the Holt-Winters multiplicative seasonal model outperform single and double exponential-smoothing models. The last model is the seasonal ARIMA(p, d, q).

We use one model specification for forecasting the aggregate series and its subcomponents. Thus, the aggregate approach generates seven model-specific forecasts for Hong Kong's total tourist arrivals. The disaggregate approach generates seven model-specific sets of eight subcomponent forecasts. Summing the eight subcomponent forecasts within a model-specific set yields the model-specific total-arrival forecast under the disaggregate approach. Our goal here is to compare their relative forecast performance under the two approaches.

3.4.1 Random walk model

A random-walk model says that the best prediction of the immediate future is the most recent past observation. This model is the Naïve 1 model in Goh & Law (2002). The m -step ahead forecast is:

3.4.2 Historical mean model

The historical average model gives equal weight to every past observation, from the beginning to the most current historical observation. The m -step ahead forecast is:

3.4.3 Twelve-month moving average model

The 12-month moving average model gives equal weight to the 12 most recent past observations. The m -step-ahead forecast is:

3.4.4 Seasonal random walk model

Called the Naïve 2 model by Goh & Law (2002), this model improves on the simple random-walk model by adjusting seasonal effects through a multiplicative factor. The m -step-ahead forecast is:

3.4.5 Holt-Winters additive seasonal model

Since the Hong Kong tourist-arrival data exhibit seasonality and trends, we use the Holt-Winters seasonal models that generalize the exponential-smoothing recursions.

The Holt-Winters model with additive seasonality is:

where the level of the series at , the trend bt and seasonal element st have the following dynamics:

To capture seasonality, st is a function of $st-12$. Since the disaggregate time series and aggregate time series evolve differently, the parameters α , β and γ likely have different values in each of the time series. The m -step-ahead forecast is obtained through:

3.4.6 Holt-Winters multiplicative seasonal model

When the seasonal effect has a time trend, the Holt-Winters multiplicative seasonal model below is more appropriate than the additive model:

Similar to the additive model, there are three dynamics for the level of the series at , the trend bt , and the seasonal element st :

The m -step-ahead forecast is obtained through:

3.4.7 Seasonal ARIMA model

The ARIMA model encompasses a wide range of non-stationary models and is useful for representing data with a trend. Estimated with Box-Jenkins procedures, it is commonly denoted as $ARIMA(p, d, q)$, where p , d and q are non-negative integers that respectively refer to the order of the autoregressive, integrated, and moving average parts of the model. In particular, a random-walk model can be represented by $ARIMA(0, 1, 0)$. Standard exponential-smoothing models are special cases of ARIMA. For example, single exponential smoothing corresponds to $ARIMA(0,1,1)$.

Since the data exhibit seasonality, we use seasonal ARIMA to model the tourist-arrival time series, which are found to be non-stationary. We denote the seasonal ARIMA as $SARIMA(p, d, q)(P, D, Q)s$, where capitalized letters respectively represent the seasonal components of the model and s indicates the order of periodicity or seasonality. As we have monthly data, s is equal to 12. Our SARIMA

is a simple extension of ARIMA, where seasonality in the data is taken into consideration by first differencing. We apply the Box and Jenkins (1976) technique to the 113-month sample to select each time series' own parameters, so as to allow for the presence of heterogeneity.

3.5 Results

After estimating each model, we form 1-, 3- and 12-step-ahead forecasts of total tourist arrivals by approach (aggregate vs. disaggregate) for the 12-month period of January to December 2011. Each approach's forecast performance is evaluated based on the RMSE in equation (2).

Tables 1a to 1c report the RMSE of total-arrival forecasts by model type based on the level data. The first column of each table identifies the model type, the second column shows the RMSE under the aggregate approach, and the third column shows the RMSE under the disaggregate unweighted-sum approach.

The last column reports the RMSE under a variant of the disaggregate weighted-sum approach. The weights are found by estimating the OLS regression described by equation (3). The disaggregate, unweighted-sum approach and its weighted-sum variant yield the same total arrival forecast under the null hypothesis given by equation (4). Hence, the last column reports the F -statistic (in bracket) and its p -value (in square bracket) for testing the null hypothesis.

Now, consider the second to fourth rows of Table 1a. For each of these three non-seasonal models, the aggregate and disaggregate unweighted-sum approaches yield the same RMSE. The 12-month moving average model has the lowest RMSE among the three non-seasonal models because it partially captures the time trend in the tourist-arrival data series.

The last column of Table 1a shows that the weighted-sum variant yields the lowest RMSE for the three non-seasonal models. Moreover, the p -values in the last column indicate statistically-significant ($\alpha < 0.01$) rejection of the null hypothesis given by equation (4), thus lending support to the performance superiority of the weighted-sum variant.

The fifth row of Table 1a shows that including the seasonal effect in the random-walk model worsens the forecasting performance for all approaches. As a result, this model is not useful for forecasting Hong Kong's total tourist arrivals.

The sixth row of Table 1a shows that when the Holt-Winters additive model is used to form the 1- step-ahead forecast, the RMSE of 275,627 is the same for the aggregate and disaggregate unweighted-sum approaches. This is because the reduction in aggregation error due to directly forecasting total arrivals is exactly offset by the increase in the estimation error in the disaggregate models. The F -test statistic

for the null in equation (4) is 4.59 with a p -value less than 0.0001. Using the regression-based weights to form the total arrival forecast substantially reduces the RMSE from 275,626 to 182,847.

The seventh row of Table 1a shows that for the Holt-Winters multiplicative model, the disaggregate unweighted-sum approach's RMSE is 268,535, marginally less than the aggregate approach's RMSE of 275,942. These RMSEs are similar to those for the corresponding Holt-Winters additive model, indicating that the seasonal effect varies little over time. The F -test statistic for the null in equation (4) is 7.01 with a p -value less than 0.0001. The last column of Table 1a indicates that using the regression-based weights to form the TOTAL forecast can substantially reduce the RMSE from 268,535 to 176,624.

The last row of Table 1a shows that the SARIMA model has a RMSE of 212,197 under the aggregate approach and 186,322 under the disaggregate unweighted-sum approach. The SARIMA model's RMSE of 186,322 under the disaggregate unweighted-sum approach for the 1-step-ahead forecast is above the Holt-Winters multiplicative model's RMSE of 176,624 under the weighted-sum variant. This, however, is not the case for the 3-step-ahead and 12-step-ahead forecasts described in Tables 1b and 1c below.

The F -test statistic for the null in equation (4) is 3.00 with a p -value 0.004. Using the regression-based weights to form the total-arrival forecast, however, actually increases the RMSE from 186,322 to 268,150. This result is counter-intuitive, since the F -test statistic rejects the null and the weights are supposedly "optimally chosen" via the OLS regression. We attribute this finding to the well-known forecast combination puzzle in Stock and Watson (2004): a simple average of model-specific forecasts outperforms an optimally-weighted combination of forecasts.

The results in Table 1b for the 3-step-ahead forecast and Table 1c for the 12-step-ahead forecast are similar to those in Table 1a. There is, however, an important exception. The SARIMA model's RMSE under the disaggregate unweighted-sum approach are substantially lower than the Holt-Winters models' RMSE under weighted-sum approach.

Taken together, Tables 1a – 1c show that the non-seasonal models have the worst forecast performance, failing to capture the tourist-arrival series' strong seasonality. Moreover, the RMSE of the Holt-Winters models can be much higher than those of the SARIMA models, suggesting that the latter should be used to characterize the tourist-arrival data. Further, the disaggregate unweighted-sum approach implemented via the SARIMA model generally has the lowest RMSE among all of the forecast models and approaches considered in this paper.

As a final check, we recalculate the RMSE values in Tables 1a-1c when the last three forecast models are applied to the logarithm of the tourist-arrival data. The first four models are ignored in this comparison due to their poor performance, as reported in Tables 1a to 1c. We construct the forecast levels based on the expected-value formula for a lognormal random variable (Mood, Graybill, & Boes, 1974). The results are given in Tables 2a to 2c, showing that logarithmic data transformation does not improve the 1- and 3-step-ahead forecast performance of all three models under all three approaches. It may, however, marginally improve the 12-step-ahead forecast performance.

4. CONCLUSION

Recognizing the heterogeneity in Hong Kong's disaggregate time series of total tourist arrivals, we postulate that first modeling the disaggregate series, and then adding up the forecasts may outperform directly forecasting the aggregate variable. Using seven model specifications, we evaluate the forecasting performance of aggregate and disaggregate approaches for Hong Kong's total tourist arrivals.

Our results show that seasonal models outperform non-seasonal models. Moreover the disaggregate unweighted sum approach improves the forecast performance of the Holt-Winters multiplicative model, though not the Holt-Winters additive model. For these two Holt-Winters model, however, the disaggregate weighted-sum approach yield large RMSE reductions.

Our results also show that under the aggregate approach, the SARIMA model should be used because it generally has a much lower RMSE than do the Holt-Winters additive and multiplicative models. The SARIMA model's performance can be further improved by adopting the disaggregate unweighted-sum approach that recognizes the heterogeneity in subcomponent arrival data by region. The weighted-sum variant, however, does not further reduce the RMSE, echoing the forecast combination puzzle where a presumably optimally-weighted combination of forecasts may underperform a simple combination of forecasts.

Both government and tourism-related businesses require accurate tourism demand forecasts for planning and pricing purposes. Our empirical results suggest that accurately forecasting Hong Kong's total tourist arrivals requires an SARIMA model specification that can suitably characterize the DGP of the data series. Moreover, a simple additive combination of properly specified disaggregate forecasts can outperform an aggregate forecast because of the heterogeneity that exists in the regional time series. Finally, there is no clear evidence to support the view that a logarithmic data transformation can improve the performance of a total-arrival

forecast. Taken together, these findings lead us to recommend applying SARIMA and the disaggregate unweighted-sum approach to the level data for forecasting Hong Kong's total tourist arrivals.

The present research can be extended to compare the aggregation and disaggregation approaches using subcomponents decomposed by purpose (vacation vs. business), duration of stay (same day vs. over night), and travel mode (land, sea, and air). The extension will likely be fruitful because these subcomponents are likely to be heterogeneous, as in the case of disaggregate arrivals by region of origin that we have documented herein.

Acknowledgement

The first author acknowledges the financial support from the Hong Kong University Grant Council (HKBU255212). This paper has been released as CORE Discussion Paper 2012/xx. We thank Ira Horowitz of the University of Florida and Luc Bauwens of the Academie Louvain for their constructive comments. We also thank Alisha Woo for her editorial assistance. All remaining errors are ours.

References

- Archer, B.H. (1987). Demand forecasting and estimation, In Ritchie, J.R.B. and Goeldner, C.R. (Eds.) *Travel, tourism and hospitality research* (pp. 77 – 85). Wiley: New York.
- Bates, J.M., & Granger, C.M.W. (1969). The combination of forecasts. *Operations Research Quarterly*, 20, 451 – 468.
- Box, G., & Jenkins, G. (1976). *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco.
- Brooke, M.Z., Buckley, P.J. & Witt, S.F. (1985). *The International Travel and Tourism Forecast*. Industry Forecasts, London.
- Bruneau, C., De Bandt, O., Flageollet, A. & Michaux, E. (2007). Forecasting inflation using economic indicators: The case of France. *Journal of Forecasting*, 26, 1 – 22.
- Chan, F., Lim, C., & McAleer, M. (2005). Modelling multivariate international tourism demand and volatility. *Tourism Management*, 26, 459–471.
- Chu, F.L. (1998). Forecasting tourism demand in Asian–Pacific countries. *Annals of Tourism Research*, 25, 597–615.
- Clements, M., & Hendry, D. F. (2002). Modeling methodology and forecast failure. *Econometrics Journal*, 5, 319 – 344.
- Clements, M. P. & Hendry, D. F. (2006). Forecasting with breaks, in Granger, C. W. J., Elliott, G. & Timmermann, A. (Eds), *Handbook of Forecasting* (pp. 605–657). Amsterdam: Elsevier.
- Dharmaratne, G.S. (1995). Forecasting tourist arrivals in Barbados. *Annals of Tourism Research*, 22, 804–818.
- Espasa, A., Senra, E. & Albacete, R. (2002). Forecasting inflation in the European monetary union: A disaggregated approach by countries and by sectors. *European Journal of Finance*, 8, 402 – 421.
- Fair, R. C. & Shiller, J. (1990). Comparing information in forecasts from econometric models. *The American Economic Review*, 80, 375 – 389.
- Garderen, V. K. J., Lee, K. & Pesaran, M. H. (2000). Cross-sectional aggregation of non-linear models. *Journal of Econometrics*, 95, 285 – 331.
- Goh, C., & Law, R. (2002). Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. *Tourism Management*, 23,

499–510.

Granger, C. (1987). Implications of aggregation with common factors. *Econometric Theory*, 3, 208 – 222.

Granger, C.W.J., & Ramanathan, R. (1984). Improved methods of combining forecast accuracy. *Journal of Forecasting*, 19, 197 – 204.

Grunfeld, Y. & Griliches, Z. (1960). Is aggregation necessarily bad? *The Review of Economics and Statistics*, 42, 1 – 13.

Hendry, D. F. (2004). Unpredictability and the foundations of economic forecasting, Working paper, Economics Department, Oxford University.

Hsiao, C., Shen, Y. & Fujiki, H. (2005). Aggregate vs. disaggregate data analysis – A paradox in the estimation of a money demand function of Japan under the low interest rate policy. *Journal of Applied Econometrics*, 20, 579 – 601.

Hsiao, C., & Wan, S. (in press). Is there an optimal forecast combination? *Journal of Econometrics*, forthcoming.

Huang, H., & Lee, T.H. (2010). To combine forecasts or to combine information? *Econometric Reviews*, 29, 539 – 570.

Hubrich, K. (2005). Forecasting Euro area inflation: Does aggregating forecasts by HICP component improve forecast accuracy? *International Journal of Forecasting*, 21, 119 – 136.

Kim, J.H., & Moosa, I.A. (2005). Forecasting international tourist flows to Australia: a comparison between the direct and indirect methods. *Tourism Management*, 26, 69 – 78.

Kohn, R. (1982). When is an aggregate of a time series efficiently forecast by its past? *Journal of Econometrics*, 18, 337 – 349.

Kulendran, N., & Wilson, K. (2000). Modelling business travel. *Tourism Economics*, 6, 47 – 59.

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

Li, G., Wong, K. F., Song, H., & Witt, S. F. (2006). Tourism demand forecasting: A time varying parameter error correction model. *Journal of Travel Research*, 45, 175 – 185.

Lim, C. & McAleer, M. (2001). Forecasting tourist arrivals. *Annals of Tourism Research*, 28, 965 – 977.

- Lütkepohl, H. (1984a). Forecasting contemporaneously aggregated vector ARMA processes. *Journal of Business and Economic Statistics*, 2, 201–214.
- Lütkepohl, H. (1984b). Linear transformations of vector ARMA processes. *Journal of Econometrics*, 26, 283–293.
- Lütkepohl, H. (1987). Forecasting aggregated vector ARMA processes. Berlin: Springer-Verlag.
- Lütkepohl, H. (2006). Forecasting with VARMA processes. In Elliott, G., Granger, C. W. J. & Timmermann, A. (Eds.), *Handbook of Economic Forecasting* (Vol. 1, pp. 287–326). Amsterdam: Elsevier.
- Marcellino, M., Stock, J. H., & Watson, M.W. (2003). Macroeconomic forecasting in the Euro area: Country specific versus area-wide information. *European Economic Review*, 47, 1–18.
- Martin, C.A., & Witt, S.F. (1989). Accuracy of econometric forecasts of tourism. *Annals of Tourism Research*, 16, 407–428.
- Means, G., & Avila, R. (1988). Globalization redirects the travel industry: The 1988 TRAM analysis, *World Travel Overview 1988/89* (pp. 86-100). New York: American Express Publishing.
- Mood, A., Graybill, F.A., & Boes, D.C. (1974). Introduction to the Theory of Statistics. New York : McGraw-Hill.
- Pesaran, M. H., Pierse, R. G. & Kumar, M. S. (1989). Econometric analysis of aggregation in the context of linear prediction models. *Econometrica*, 57, 861–888.
- Qu, H., Xu, P., & Tan, A. (2002). A simultaneous equations model of the hotel room supply and demand in Hong Kong. *Hospitality Management*, 21, 455–462.
- Shen, S., Li, G., & Song, H. (2008). An assessment of combining tourism demand forecasts over different time horizons. *Journal of Travel Research*, 47, 197–207.
- Shen, S., Li, G. & Song, H. (2011). Combination forecasts of international tourism demand. *Annals of Tourism Research*, 38, 72–89.
- Song, H., & Li, G. (2008). Tourism demand modeling and forecasting – a review of recent research. *Tourism Management*, 29, 203–220.
- Song, H., & Witt, S.F. (2000). *Tourism demand modeling and forecasting: Modern econometric approaches*. Oxford: Pergamon.
- Song, H., & Witt, S. F. (2003). Tourism forecasting: The general-to specific approach.

Journal of Travel Research, 42, 65 – 74.

Song, H., & Witt, S. F. (2006). Forecasting international tourist flows to Macau. *Tourism Management*, 27, 214 – 224.

Song, H., Witt, S.F., Wong, K.F., & Wu, C. (2009). An empirical study of tourism demand forecast combination. *Journal of Hospitality and Tourism Research*, 33, 3 – 29.

Song, H., Wong, K. & Chon, K. (2003). Modeling and forecasting the demand for Hong Kong tourism. *International Journal of Hospitality Management*, 22, 435 – 451.

Stock, J.H., Watson, M.W. (2004). Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting*, 23, 405_430.

Theil, H. (1954). Linear aggregation of economic relations. Amsterdam: North Holland.

Timmermann, A. (2006). Forecast combinations. In Elliott, G., Granger, C.W.J., Timmermann, A. (Eds.) *Handbook of Economic Forecasting*, (Vol. 1, pp. 135 – 196). Amsterdam: Elsevier.

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

Wong, K.K.F., Song, H., Witt, S.F., & Wu, D.C. (2007). Tourism forecasting: To combine or not to combine? *Tourism Management*, 28, 1068 – 1078.

Zellner, A. & Tobias, J. (2000). A note on aggregation, disaggregation and forecasting performance. *Journal of Forecasting*, 19, 457 – 469.

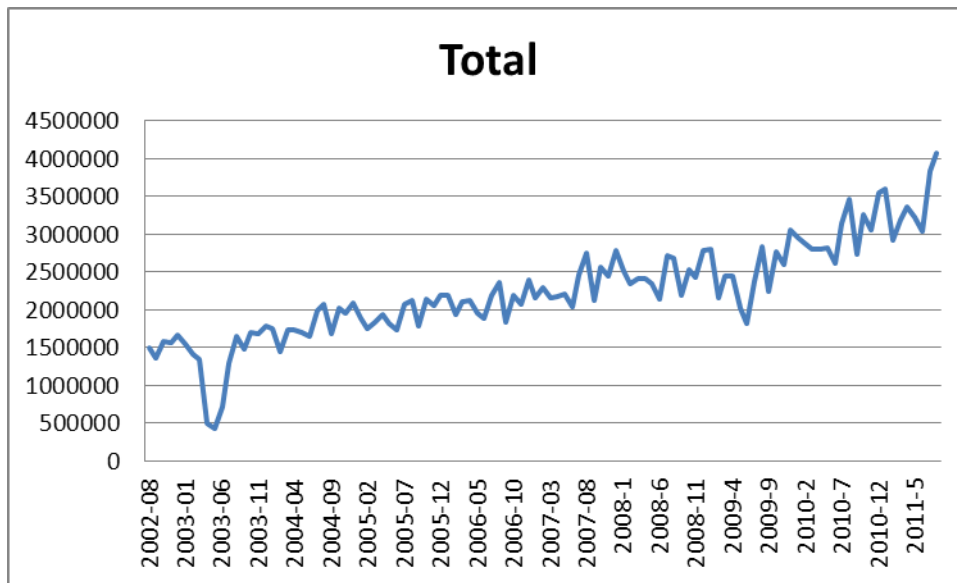


Figure 1. Monthly tourist arrivals to Hong Kong

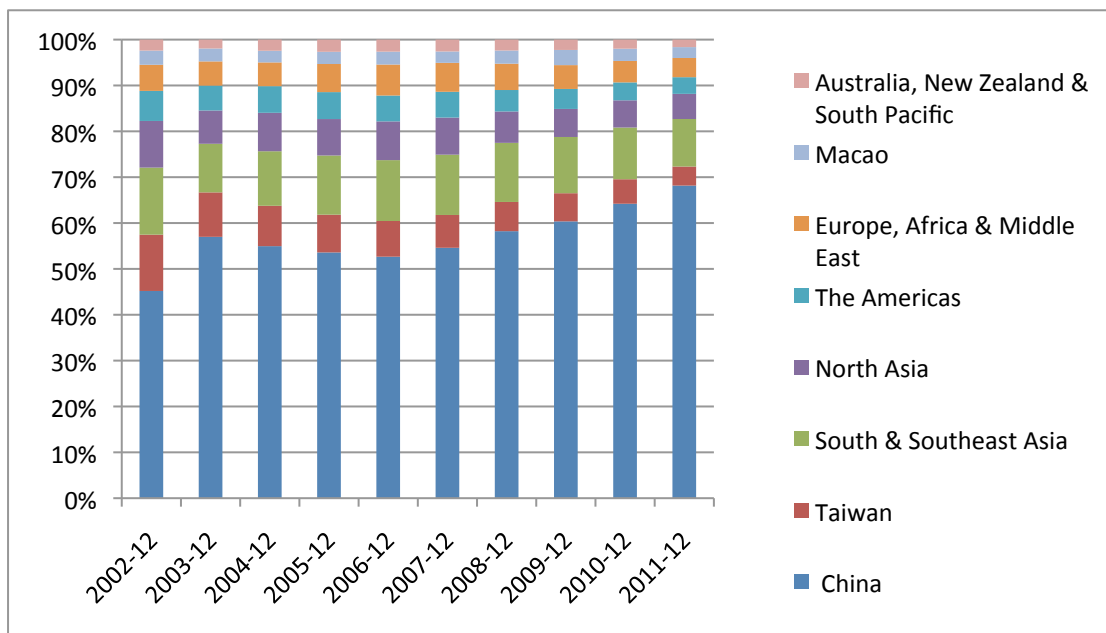


Figure 2. Distributions of annual total tourist arrivals to Hong Kong

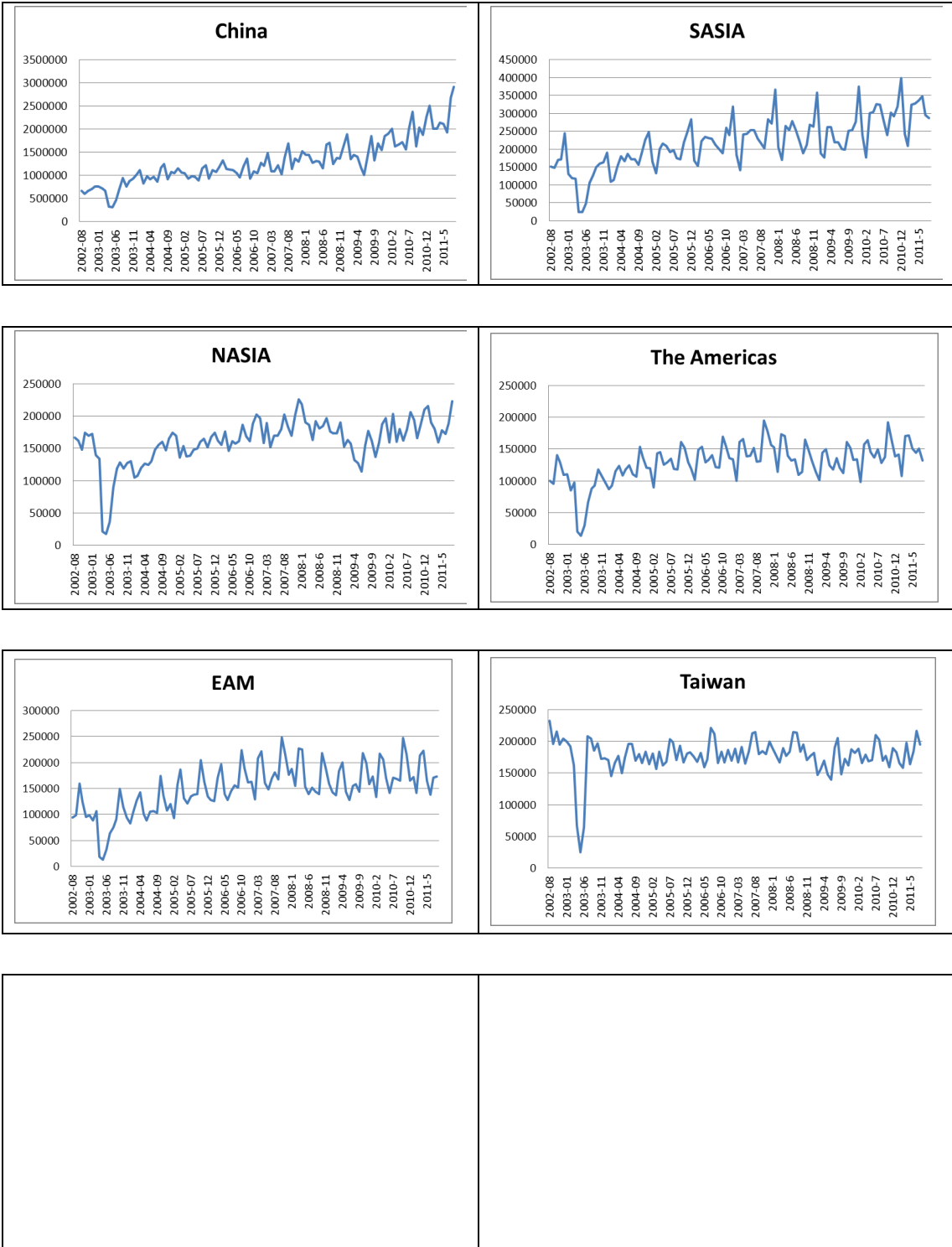


Figure 3. Monthly tourist arrivals by region

Table 1a. RMSE for 1-step ahead forecast accuracy: Level

1. Model type	Aggregate	Disaggregate: Unweighted Sum	Disaggregate: Weighted Sum
2. Random Walk	488,741	488,741	351,571 (3.41) [0.0012]
3. Historical Mean	1,335,676	1,335,676	362,696 (49.88) [<0.0001]
4. 12-month Moving Average	433,440	433,440	329,618 (7.88) [<0.0001]
5. Seasonal Random Walk	801,614	819,618	651,977 (15.01) [<0.0001]
6. Holt-Winters Additive	275,627	275,627	182,847 (4.59) [<0.0001]
7. Holt-Winters Multiplicative	275,942	268,535	176,624 (7.01) [<0.0001]
8. SARIMA	212,197	186,322	268,150 (3.00) [0.0040]

Table 1b. RMSE for 3-step ahead forecast accuracy: Level

1. Model type	Aggregate	Disaggregate: Unweighted Sum	Disaggregate: Weighted Sum
2. Random Walk	463,842	463,842	397,756 (8.04) [<0.0001]
3. Historical Mean	1,357,797	1,357,797	323,356 (58.62) [<0.0001]
4. 12-month Moving Average	488239	488239	332,273 (11.32) [<0.0001]
5. Seasonal Random Walk	791,119	808,979	615,663 (17.63) [<0.0001]
6. Holt-Winters Additive	334,649	334,649	183,000 (4.47) [<0.0001]
7. Holt-Winters Multiplicative	355,672	347,419	213,368 (6.99) [<0.0001]
8. SARIMA	172,355	158,838	231,808 (2.88) [0.0055]

Table 1c. RMSE for 12-step ahead forecast accuracy: Level

1. Model type	Aggregate	Disaggregate: Unweighted Sum	Disaggregate: Weighted Sum
2. Random Walk	516,772	516,772	446,044 (25.68) [<0.0001]
3. Historical Mean	1,449,192	1,449,192	383,502 (87.90) [<0.0001]
4. 12-month Moving Average	810,311	810,311	329,775 (23.24) [<0.0001]
5. Seasonal Random Walk	519,306	468,456	552,962 (21.4) [<0.0001]
6. Holt-Winters Additive	632,101	632,101	457,847 (2.79) [0.0068]
7. Holt-Winters Multiplicative	620,728	620,627	484,055 (5.11) [<0.0001]
8. SARIMA	291,148	210,518	254,046 (2.97) [0.0049]

Notes: The values in () in the last column are F -statistics for testing the null hypothesis given by equation (4) and the values in [] are their p -values.

Table 2a. RMSE for 1-step ahead forecast accuracy: Logarithm

Model type	Aggregate	Disaggregate: Unweighted Sum	Disaggregate: Weighted Sum
Holt-Winters Additive	277,915	265,099	193,888 (12.81) [<0.0001]
Holt-Winters Multiplicative	285,734	274,606	202,830 (14.35) [<0.0001]
SARIMA	233,495	186,073	244,812 (5.09) [<0.0001]

Table 2b. RMSE for 3-step ahead forecast accuracy: Logarithm

Model type	Aggregate	Disaggregate: Unweighted Sum	Disaggregate: Weighted Sum
Holt-Winters Additive	351,662	335,342	227,129 (12.65) [<0.0001]
Holt-Winters Multiplicative	361,228	346,164	237,880 (14.16) [<0.0001]
SARIMA	285,107	162,734	247,416 (5.22) [<0.0001]

Table 2c. RMSE for 12-step ahead forecast accuracy: Logarithm

Model type	Aggregate	Disaggregate: Unweighted Sum	Disaggregate: Weighted Sum
Holt-Winters Additive	547,891	520,581	465,659 (10.77) [<0.0001]
Holt-Winters Multiplicative	541,120	512,503	466,916 (12.27) [<0.0001]
SARIMA	220,538	193,156	307,305 (5.92) [<0.0001]

Notes: The values in () in the last column are F -statistics for testing the null hypothesis given by equation (4) and the values in [] are their p -values.

Recent titles

CORE Discussion Papers

- 2011/69. Per J. AGRELL and Adel HATAMI-MARBINI. Frontier-based performance analysis models for supply chain management; state of the art and research directions.
- 2011/70. Olivier DEVOLDER. Stochastic first order methods in smooth convex optimization.
- 2011/71. Jens L. HOUGAARD, Juan D. MORENO-TERNERO and Lars P. ØSTERDAL. A unifying framework for the problem of adjudicating conflicting claims.
- 2011/72. Per J. AGRELL and Peter BOGETOFT. Smart-grid investments, regulation and organization.
- 2012/1. Per J. AGRELL and Axel GAUTIER. Rethinking regulatory capture.
- 2012/2. Yu. NESTEROV. Subgradient methods for huge-scale optimization problems.
- 2012/3. Jeroen K. ROMBOUTS, Lars STENTOFT and Francesco VIOLANTE. The value of multivariate model sophistication: An application to pricing Dow Jones Industrial Average options.
- 2012/4. Aitor CALO-BLANCO. Responsibility, freedom, and forgiveness in health care.
- 2012/5. Pierre PESTIEAU and Grégory PONTIERE. The public economics of increasing longevity.
- 2012/6. Thierry BRECHET and Guy MEUNIER. Are clean technology and environmental quality conflicting policy goals?
- 2012/7. Jens L. HOUGAARD, Juan D. MORENO-TERNERO and Lars P. ØSTERDAL. A new axiomatic approach to the evaluation of population health.
- 2012/8. Kirill BORISSOV, Thierry BRECHET and Stéphane LAMBRECHT. Environmental maintenance in a dynamic model with heterogenous agents.
- 2012/9. Ken-Ichi SHIMOMURA and Jacques-François THISSE. Competition among the big and the small.
- 2012/10. Pierre PESTIEAU and Grégory PONTIERE. Optimal lifecycle fertility in a Barro-Becker economy.
- 2012/11. Catherine KRIER, Michel MOUCHART and Abderrahim OULHAJ. Neural modelling of ranking data with an application to stated preference data.
- 2012/12. Matthew O. JACKSON and Dunia LOPEZ-PINTADO. Diffusion and contagion in networks with heterogeneous agents and homophily.
- 2012/13. Claude D'ASPREMONT, Rodolphe DOS SANTOS FERREIRA and Jacques THEPOT. Hawks and doves in segmented markets: A formal approach to competitive aggressiveness.
- 2012/14. Claude D'ASPREMONT and Rodolphe DOS SANTOS FERREIRA. Household behavior and individual autonomy: An extended Lindahl mechanism.
- 2012/15. Dirk VAN DE GAER, Joost VANDENBOSSCHE and José Luis FIGUEROA. Children's health opportunities and project evaluation: Mexico's *Oportunidades* program.
- 2012/16. Giacomo VALLETTA. Health, fairness and taxation.
- 2012/17. Chiara CANTA and Pierre PESTIEAU. Long term care insurance and family norms.
- 2012/18. David DE LA CROIX and Fabio MARIANI. From polygyny to serial monogamy: a unified theory of marriage institutions.
- 2012/19. Carl GAIGNE, Stéphane RIOU and Jacques-François THISSE. Are compact cities environmentally friendly?
- 2012/20. Jean-François CARPANTIER and Besik SAMKHARADZE. The asymmetric commodity inventory effect on the optimal hedge ratio.
- 2012/21. Concetta MENDOLICCHIO, Dimitri PAOLINI and Tito PIETRA. Asymmetric information and overeducation.
- 2012/22. Tom TRUYTS. Stochastic signaling: Information substitutes and complements.
- 2012/23. Pierre DEHEZ and Samuel FERREY. How to share joint liability: A cooperative game approach.
- 2012/24. Pilar GARCIA-GOMEZ, Erik SCHOKKAERT, Tom VAN OURTI and Teresa BAGO D'UVA. Inequity in the face of death.
- 2012/25. Christian HAEDO and Michel MOUCHART. A stochastic independence approach for different measures of concentration and specialization.
- 2012/26. Xavier RAMOS and Dirk VAN DE GAER. Empirical approaches to inequality of opportunity: principles, measures, and evidence.

Recent titles

CORE Discussion Papers - continued

- 2012/27. Jacques H. DRÈZE and Erik SCHOKKAERT. Arrow's theorem of the deductible : moral hazard and stop-loss in health insurance.
- 2012/28 Luc BAUWENS and Giuseppe STORTI. Computationally efficient inference procedures for vast dimensional realized covariance models.
- 2012/29. Pierre DEHEZ. Incomplete-markets economies: The seminar work of Diamond, Drèze and Radner.
- 2012/30 Helmuth CREMER, Pierre PESTIEAU and Grégory PONTIÈRE. The economics of long-term care: a survey.
- 2012/31 Mathieu VAN VYVE, Laurence A. WOLSEY and Hande YAMAN. Relaxations for two-level multi-item lot-sizing problem.
- 2012/32 Jean-François MERTENS and Anna RUBINCHIK. Equilibria in an overlapping generations model with transfer policies and exogenous growth.
- 2012/33 Jean-François MERTENS and Anna RUBINCHIK. Pareto optimality of the golden rule equilibrium in an overlapping generations model with production and transfers.
- 2012/34 Hélène LATZER and Florian MAYNERIS. Income distribution and vertical comparative advantage. Theory and evidence.
- 2012/35 Pascal MOSSAY and Takatoshi TABUCHI. Preferential trade agreements harm third countries.
- 2012/36 Aitor CALO-BLANCO. The compensation problem with fresh starts.
- 2012/37 Jean-François CARPANTIER and Arnaud DUFAYS. Commodities volatility and the theory of storage.
- 2012/38 Jean-François CARPANTIER and Christelle SAPATA. Unfair inequalities in France: A regional comparison.
- 2012/39 Shui-Ki WAN, Shin-Huei WANG and Chi-Keung WOO. Total tourist arrival forecast: aggregation vs. disaggregation.

Books

- G. DURANTON, Ph. MARTIN, Th. MAYER and F. MAYNERIS (2010), *The economics of clusters – Lessons from the French experience*. Oxford University Press.
- J. HINDRIKS and I. VAN DE CLOOT (2011), *Notre pension en héritage*. Itinera Institute.
- M. FLEURBAEY and F. MANIQUET (2011), *A theory of fairness and social welfare*. Cambridge University Press.
- V. GINSBURGH and S. WEBER (2011), *How many languages make sense? The economics of linguistic diversity*. Princeton University Press.
- I. THOMAS, D. VANNESTE and X. QUERRIAU (2011), *Atlas de Belgique – Tome 4 Habitat*. Academia Press.
- W. GAERTNER and E. SCHOKKAERT (2012), *Empirical social choice*. Cambridge University Press.
- L. BAUWENS, Ch. HAFNER and S. LAURENT (2012), *Handbook of volatility models and their applications*. Wiley.
- J-C. PRAGER and J. THISSE (2012), *Economic geography and the unequal development of regions*. Routledge.
- M. FLEURBAEY and F. MANIQUET (2012), *Equality of opportunity: the economics of responsibility*. World Scientific.
- J. HINDRIKS (2012), *Gestion publique*. De Boeck.

CORE Lecture Series

- R. AMIR (2002), Supermodularity and complementarity in economics.
- R. WEISMANTEL (2006), Lectures on mixed nonlinear programming.
- A. SHAPIRO (2010), Stochastic programming: modeling and theory.