Available online at www.sciencedirect.com
ScienceDirect
International Journal of Forecasting 27 (2011) 923–941
international journal of forecasting
www.elsevier.com/locate/ijforecast

Forecasting (aggregate) demand for US commercial air travel

Richard T. Carson[a], Tolga Cenesizoglu[b,c,*], Roger Parker[d]

[a] Department of Economics, University of California, San Diego, United States
[b] Department of Finance, HEC Montreal, Canada
[c] CIRPÉE, 3000, Chemin de la Côte-Sainte-Catherine, Bureau 4.348, Montréal (Québec), H3T 2A7, Canada
[d] Virtual Minds, SA, Switzerland

Available online 3 June 2010

Abstract

We analyze whether it is better to forecast air travel demand using aggregate data at (say) a national level, or to aggregate the forecasts derived for individual airports using airport-specific data. We compare the US Federal Aviation Administration's (FAA) practice of predicting the total number of passengers using macroeconomic variables with an equivalently specified AIM (aggregating individual markets) approach. The AIM approach outperforms the aggregate forecasting approach in terms of its out-of-sample air travel demand predictions for different forecast horizons. Variants of AIM, where we restrict the coefficient estimates of some explanatory variables to be the same across individual airports, generally dominate both the aggregate and AIM approaches. The superior out-of-sample performances of these so-called quasi-AIM approaches depend on the trade-off between heterogeneity and estimation uncertainty. We argue that the quasi-AIM approaches exploit the heterogeneity across individual airports efficiently, without suffering from as much estimation uncertainty as the AIM approach.

Keywords: Demand forecasting; Disaggregation; Transportation forecasting; Panel data

1. Introduction

Forecasts of air travel demand are important inputs for a wide variety of economic decisions, including, but not limited to, research and development, airplane design and production planning. For a relatively mature product like air travel, where the interest lies in the aggregate demand, the typical empirical practice is to obtain a national level forecast using aggregate level data when individual market data are not easily accessible (Lehmann & Winer, 2001). However, recent empirical evidence (Bronnenberg, Dhar, & Dube, 2007) has shown that individual markets are much more heterogenous than was thought previously, even for well-known national products which are sold in spatially distinct markets, as is the case for air travel.

An aggregate approach is preferable to a disaggregate approach when the computational/analytical burden of producing forecasts for separate markets is

* Corresponding author at: Department of Finance, HEC Montreal, H3T 2A7, Canada. Tel.: +1 514 340 5668; fax: +1 514 340 5632.
E-mail addresses: cenesizo@yahoo.com, tolga.cenesizoglu@hec.ca (T. Cenesizoglu).

0169-2070/$ - see front matter
doi:10.1016/j.ijforecast.2010.02.010

substantial. Furthermore, there are also concerns that in a disaggregate approach, the number of parameters estimated by modeling each market individually quickly becomes large relative to the length of the available time series. On the other hand, the econometric arguments in favor of a disaggregate approach are also fairly strong when disaggregate data are available. Allen and Fildes (2001) review the literature on the advantages of using disaggregate data, one of which is the additional information available due to heterogeneity across individual markets. However, they also argue that the relative performances of aggregate and disaggregate approaches might depend on the specifics of the forecasting exercise.

This paper focuses on exploring the effect of disaggregate information on the accuracy of aggregate air travel demand forecasts. This paper considers the case of US commercial air travel demand, with the objective of predicting the total number of commercial passengers. The data available to us from the US Department of Transportation are the monthly numbers of passengers departing from major US airports between 1990 and 2004. We use individual airport level data for 179 major airports, together with aggregate level data.

We initially consider two extreme approaches. The first follows the FAA's practice of predicting the total number of passengers using a combination of exogenous macroeconomic variables such as population, income, and energy prices in a time series model. The second approach, which we term an AIM (aggregating individual markets) forecast, models the travel demand at the individual market level using exogenous variables specific to that region, then sums the forecasts to produce a forecast of the total number of passengers. Specifically, consider forecasting a variable that is a contemporaneous aggregation of the individual subcomponents at time t:

$$y_t = \sum_{i=1}^{N} y_{it} \quad \text{for } t = 1, 2, \ldots, \tag{1}$$

where y_{it} $(i = 1, 2, \ldots, N)$ are the subcomponents of the aggregate variable y_t. Forecasts of the aggregate variable can be obtained using two different approaches: (1) estimating a reduced-form model for the aggregate variable using aggregate level data, then forming forecasts of the aggregate variable from the estimated model; or (2) estimating a reduced-form model for the subcomponents using individual level data, forming forecasts of the subcomponents from the estimated individual models, then aggregating the subcomponent forecasts to obtain the forecast of the aggregate variable. Not only does the second approach use more information, but also the forecasts at the disaggregate level are readily available, whereas one needs to allocate the aggregate forecast to the individual markets to obtain forecasts at the disaggregate level in the first approach. In this paper, we analyze the out-of-sample forecast performances of these two extreme approaches in forecasting the aggregate variable, along with the performances of other approaches that are between these two extremes.

There has been a revival of interest in the contemporaneous aggregation of disaggregate forecasts to form forecasts of an aggregate variable in both the theoretical and empirical econometrics literature. The empirical literature has focused mainly on forecasting aggregate macroeconomic indices such as Euro-area variables or aggregate US variables by using the information available in the disaggregated subcomponents.[1] In this paper, we not only demonstrate the usefulness of disaggregate information in forecasting an aggregate variable of interest using a new and comprehensive data set, but also discuss different ways of exploiting the heterogeneity across disaggregate level data.

The main contributions of our paper to the existing literature can be summarized as follows. First, applying the two approaches to monthly US air passenger data over the period 1990 to 2002 and then forecasting out-of-sample for the next two years yields a striking contrast: the AIM forecast is far more accurate than the aggregate level forecast. We argue that the performance of the AIM approach depends on the tradeoff between the heterogeneity across markets and the estimation uncertainty due to the number of coefficients estimated. The AIM approach outperforms the aggregate approach, since the forecasting power gained by exploiting heterogenous information across markets dominates the forecasting power lost due to

[1] See Benalal, del Hoyo, Landau, Roma, and Skudelny (2004), Espasa, Senra, and Albacete (2002), Fair and Shiller (1990), Marcellino, Stock, and Watson (2003) and Zellner and Tobias (2000).