Tourism Management 69 (2018) 69–87

Contents lists available at ScienceDirect

Tourism Management

journal homepage: www.elsevier.com/locate/tourman

Measuring hotel performance: Toward more rigorous evidence in both scope and methods

A. George Assaf[a,*,1], Mike Tsionas[b,c,1]

[a] Isenberg School of Management, University of Massachusetts-Amherst, 90 Campus Center Way, 209A Flint Lab, Amherst, MA, 01003, United States
[b] Lancaster University Management School, United Kingdom
[c] Athens University of Economics and Business, Greece

ARTICLE INFO

Keywords:
Hotel performance
Efficiency
Productivity
Heterogeneity
Bayesian stochastic frontier

ABSTRACT

This paper extends the literature on hotel performance in both scope and methods. We introduce a model that accounts for heterogeneity in a flexible way and allows for the measurement of both efficiency and productivity. The model also accounts for the endogeneity problem in inputs and the issue of unobserved prices. We use a large sample of hotel companies that spreads across multiple geographical regions and locations, and accounts for some interesting and key determinants of hotel performance. We provide more validation to some contradictory findings in the literature. We show that large hotels do not necessarily outperform small hotels, and that hotel efficiency differs based on location, geographical region and type of service. The results further indicate that productivity growth is not a driving force in the industry.

1. Introduction

Over the last decade, there has been a remarkable growth in the use of frontier methods to measure tourism and hotel performance (Sainaghi, Phillips, & Zavarrone, 2017). In contrast to simple performance methods, frontier methods measure performance relative to a frontier of best practices, and allow the inclusion of multiple inputs and outputs in the measurement of hotel performance. While these methods have their advantages, they can be sensitive to sample characteristics and the selection of appropriate inputs and outputs. Assaf and Josiassen (2016) emphasized that most studies in the literature seem to ignore these limitations, focusing only on one destination or one specific region within a destination, making it hence difficult to generalize the findings to hotels from other destinations. There is also the problem of small sample size and data limitations. Since it is always challenging to collect reliable data on hotels, most studies seem to rely on small samples and a limited number of inputs and outputs.

The aim of this paper is to address these limitations. We focus on providing a more comprehensive representation of the operational characteristics of the hotel industry while addressing several contradicting hypotheses regarding the determinants of hotel performance (e.g. size, location, type of service, etc.). For the first time, we use a unique sample that covers more than one destination, spreading across the US, Europe, the Asia Pacific and the Middle East. The sample is unique in that it does not only cover different destinations but also various locational characteristics (e.g. urban, resort, airport, etc.), hotel classifications (e.g. luxury, economy, independent, etc.), and a large list of input and output variables.

Methodologically, we also present several important contributions. Given the unique characteristics of our sample, which includes heterogeneous hotel groups that vary in terms of size/classification and location, we develop a new stochastic frontier model that accounts for such heterogeneity. Most studies in the literature have so far measured the frontier technology without accounting for firm heterogeneity. Here, we introduce a new SF, developed in a Bayesian framework, to account for such heterogeneity. We provide measures of both efficiency and productivity growth and assess how they vary with various hotel characteristics (e.g. size, location, etc.). These two performance metrics are different. The aim of measuring efficiency "is to separate production units that perform well from those that perform poorly. Whereas efficiency measures firm performance relative to the existing production, cost, or revenue frontier, productivity measures shifts in the frontier over time" (Cummins & Xie, 2013, p. 143). Hence, each of these measures provides an important source of information and has different policy implications.

In terms of methodological contributions, our point of departure is to model heterogeneity in a flexible way when measuring productivity and efficiency. Existing alternatives are the finite mixture model (as

* Corresponding author.
E-mail addresses: assaf@isenberg.umass.edu (A.G. Assaf), m.tsionas@lancaster.ac.uk (M. Tsionas).
[1] Both authors have contributed equally to the paper.

https://doi.org/10.1016/j.tourman.2018.05.008
Received 29 January 2018; Received in revised form 8 May 2018; Accepted 22 May 2018
0261-5177/

refined in Geweke & Keane, 2007) and the random coefficient approach (Tsionas, 2002). Here, we opt for a more flexible approach, which allows environmental variables to directly influence heterogeneity. The model is an artificial neural network (ANN) with G nodes and it is known that as G increases it can approximate any functional form. Inefficiency and productivity are related through a vector autoregressive (VAR) scheme, so that we can examine impulse responses from one variable to the others for different groups and also for different hotels. We also account for the potential endogeneity problem of inputs using the first order conditions from an input distance function and cost minimization (Atkinson & Tsionas, 2016). In this context, a commonly encountered problem is that most if not all input prices are unobserved. We handle the problem by assuming that relative prices are latent and can be related to input-specific and time-specific effects. The resulting model is highly non-linear and has a non-trivial Jacobian of transformation, which has to be taken into account when we develop likelihood-based inference. We develop efficient Markov chain Mote Carlo (MCMC) procedures for Bayesian inference in the model. MCMC is needed because the likelihood function depends on multivariate integrals that cannot be expressed in closed form.

The rest of this paper proceeds as follows: Next, we discuss the current gaps in the literature. We then present the model and the sample characteristics, followed by the results, discussion and implications of the findings.

2. Current gaps in the literature

There is now an extensive literature on frontier methods in the hospitality and tourism literature. As recent studies have presented an extensive review on this topic, we do not intend to reiterate everything here.[2] We focus instead on some of the main gaps in the literature. Table 1 summarizes and groups some of the key studies based on several criteria, including the methodology used, the country covered, the sample size, as well as the assumptions made on the model. Table 2 provides some key findings about the determinants of hotel performance (e.g. location, class, type of service, region and size).

Several important gaps can be observed from Tables 1 and 2:

1 First, it is clear that most studies have used the DEA approach to estimate hotel efficiency. As noted above, while DEA has several advantages, it does not allow for some advanced assumptions (e.g. heterogeneity; endogeneity in inputs) to be made on the frontier model. As indicated by several studies (Tsionas & Kumbhakar, 2014), ignoring such key assumptions can result in significant bias, particularly in contexts like ours where factors such as size, location, classification or star rating can affect the shape and estimation of the frontier model.

2 It is clear that even studies that used the stochastic frontier approach have adopted simplistic assumptions, and largely ignored heterogeneity. Barros, Dieke, and Santos (2010) have estimated a random frontier model to account for heterogeneity in the context of Luanda hotels, but their approach does not account for heterogeneity in a flexible manner as we do here. In this paper, we opt for a more advanced approach, which allows environmental variables to directly influence heterogeneity and addresses the issue of unobserved prices and endogeneity in inputs.

3 Only a few studies have adopted the Bayesian approach despite its ability to handle more complicated stochastic frontier models such as the one we introduce in this study. For instance, our model is highly non-linear and has a non-trivial Jacobian of transformation, which makes the use of frequentist-based estimation methods such as Maximum Likelihood (ML) highly challenging in implementation.

4 From Table 1, it is clear that with the exception of a few studies, most studies have focused only on one destination or used a limited number of hotels. This is probably due to data limitation and may justify why most studies in the literature have used the DEA approach (Coelli, Rao, O'Donnell, & Battese, 2005).

5 It is also important to note that existing studies have focused mainly on the estimation of efficiency. Here we derive measures of both efficiency and productivity from the same model and in a parametric (albeit highly flexible) fashion. We believe that providing these two measures is important for policy implications. Efficiency is "only one component of productivity-productivity growth is not driven by efficiency alone, but also by other factors such as innovation and output growth" (Assaf & Tsionas, 2018, p. 132). In our model, we relate efficiency and productivity through a vector autoregressive (VAR) scheme so that we can examine impulse responses from one variable to the others for different groups and also for different hotels.

6 Overall, it is clear from Table 2 that the literature has so far provided contradictory evidence about how some commonly used "determinants" correlate with hotel performance (e.g. size, hotel classification, type of service and location). Using a much richer sample that covers multiple locations and geographical regions, our aim is to provide a more comprehensive assessment of how these determinants affect hotel performance. Importantly, this paper does not only assess how these determinants influence efficiency, but also test their effect on productivity growth. In this way, we achieve two objectives: 1-follow the logical assumption in the literature that these determinants are actual sources of heterogeneity, and 2- more accurately reflect their impact on efficiency and productivity.

[2] See for example Assaf and Josiassen (2016).

3. The model

Our point of departure is to model heterogeneity in a flexible way and then measure productivity and efficiency. The classical approach, without heterogeneity, rests upon the following specification:

$$y_{it} = x_{it}'\beta + v_{it},\ i = 1,\ \ldots, n,\ t = 1,\ \ldots, T, \tag{1}$$

which is the classical linear model, where x_{it} is a $k \times 1$ vector of covariates, β is a $k \times 1$ vector of parameters, and v_{it} is an error term.

Here, we propose a model to account for heterogeneity. Known alternatives are the finite mixture model (which has been made more flexible in Geweke & Keane, 2007) and the random coefficient approach (Tsionas, 2002). Here, we opt for a more flexible approach, which allows environmental variables to directly influence the heterogeneity (called, for this reason, observed heterogeneity).

Our model is:

$$y_{it} = x_{it}'\beta(z_{it}) + v_{it},\ i = 1,\ \ldots, n,\ t = 1,\ \ldots, T, \tag{2}$$

where z_{it} is $p \times 1$ vector of environmental variables (size, classification, etc.). Here, $\theta \in \Theta \subseteq \mathbb{R}^d$ will denote, thereafter, the parameter vector. Hence in our formulation we make β depend on z_{it}.[3] Here, β is also random itself and has also random error term that it different by hotel. We elaborate further on this formulation in more detail below.

Moreover, we modify the model in (2) as follows:

$$y_{it} = x_{it}'\beta(z_{it}) + v_{it} + \omega_{it} \pm u_{it},\ i = 1,\ \ldots, n,\ t = 1,\ \ldots, T, \tag{3}$$

where a common specification is:

$$u_{it} \sim N_{+}(0, \sigma_u^2), \tag{4}$$

and $u_{it} \geq 0$ is an error component that stands for technical inefficiency.[4] The specification in (4) is highly restrictive, mainly for two reasons.

[3] This is known as a smooth coefficients model in the literature. See for example Li, Huang, and Fu (2002).

[4] We have + in the case of cost functions and $-u_{it}$ in the case of production functions. This corresponds also, respectively, to output distance functions and input distance functions.