Greek hotel industry: customer satisfaction and hotel financial performance

Sklavos Fabien

A thesis presented for the degree of Doctor of Philosophy

Department of Economics
University of Piraeus
Greece

May 2024

ABSTRACT

Customer satisfaction stands out as a paramount objective within the hotel industry. Despite numerous factors associated with service quality influencing customer satisfaction, a notable gap persists in the literature concerning the nexus between hotels' financial performance and customer satisfaction. This thesis endeavors to investigate the correlation between various financial indicators of hotels and customer satisfaction, as manifested through ratings on the TripAdvisor platform. To achieve this objective, data encompassing review scores were extracted from the TripAdvisor platform for a sample of Greek 3- to 5-star hotels (N=554) spanning the years 2011 to 2016 (5,540 observations). Subsequently, the TripAdvisor review scores were aggregated and aligned with the corresponding financial data and ratios derived from the hotels' balance sheets and profit & loss accounts. The analytical framework employed for this investigation included random effects panel models and stochastic frontier analysis. Stratified analyses were undertaken based on hotel size, star classification, and efficiency scores. The results indicate that financial metrics wield significant predictive power over online ratings. Notably, total assets exhibit a non-linear relationship with ratings, and investments in guest amenities emerge as pivotal factors influencing customer satisfaction. Furthermore, a preference for long-term investments in fixed assets and amenities becomes apparent, suggesting their efficacy in enhancing visitor satisfaction. Contrastingly, having the hotel a pool is negatively related to customer satisfaction. Additionally, both the size of the hotel and its star rating seem to impact its efficiency, with larger hotels and those with a 5-star rating being the most significantly influenced. Furthermore, various regions in Greece display distinct technical efficiency scores related to their hotels. It is imperative to note that the generalizability of these findings to diverse hotel populations necessitates further research. In essence, this study sheds light on the intricate interplay between financial performance indicators and customer satisfaction within the context of the hotel industry, providing valuable insights for both academic discourse and practical implications within the hospitality sector.

ΠΕΡΙΛΗΨΗ ΔΙΑΤΡΙΒΗΣ

Η ικανοποίηση των πελατών αναδεικνύεται ως πρωταρχικός στόχος στον ξενοδοχειακό κλάδο καθότι συνδέεται με την παραγωγή πωλήσεων. Παρόλο που ένας σημαντικός αριθμός των παραγόντων που επηρεάζουν την ικανοποίηση των πελατών έχει μελετηθεί, υπάρχει ένα αξιοσημείωτο κενό στη βιβλιογραφία αναφορικά με τη σχέση μεταξύ των οικονομικών δεδομένων και απόδοσης των ξενοδοχείων, και της ικανοποίησης των πελατών τους. Η διατριβή επιχειρεί να διερευνήσει τη συσχέτιση μεταξύ διαφόρων οικονομικών δεικτών των ξενοδοχείων και της ικανοποίησης των πελατών τους, όπως αυτή αποτυπώνεται στις αξιολογήσεις στην πλατφόρμας του TripAdvisor. Προκειμένου να επιτευχθεί ο στόχος της διατριβής, αντλήθηκαν δεδομένα που περιλάμβαναν τις βαθμολογίες στο TripAdvisor για ένα δείγμα ελληνικών ξενοδοχείων τριών έως πέντε αστέρων (N=554) για τα έτη 2011 έως 2016 (5,540 παρατηρήσεις). Στη συνέχεια, οι βαθμολογίες του TripAdvisor του κάθε ξενοδοχείου αντιστοιχήθηκαν με οικονομικά δεδομένα (και επιλεγμένους χρηματοοικονομικούς δείκτες) που αντλήθηκαν από τους ισολογισμούς και τις καταστάσεις αποτελεσμάτων χρήσεων, των ξενοδοχείων. Η μεθοδολογική οικονομετρική προσέγγιση που χρησιμοποιήθηκε περιλαμβάνει "random effects panel models" και "stochastic frontier analysis". Τα δεδομένα διαστρωματώθηκαν και αναλύθηκαν με βάση το μέγεθος του ξενοδοχείου, την κατηγορία αστέρων και την επιχειρησιακή τους αποδοτικότητα. Τα αποτελέσματα υποδεικνύουν ότι οι οικονομικοί δείκτες και τα οικονομικά μεγέθη των ξενοδοχείων παρουσιάζουν σημαντική συσχέτιση με τις βαθμολογίες που λαμβάνουν αυτά στην πλατφόρμα TripAdvisor. Ειδικότερα, το συνολικό ενεργητικό παρουσιάζει μη γραμμική σχέση με τις βαθμολογίες, ενώ οι επενδύσεις σε κεφαλαιουχικές και λειτουργικές δαπάνες, είναι καθοριστικοί παράγοντες επηρεασμού της ικανοποίησης των πελατών. Αντίθετα, η ύπαρξη πισίνας στο ξενοδοχείο σχετίζεται αρνητικά με την ικανοποίηση των πελατών. Επιπλέον, τόσο το μέγεθος όσο και η κατάταξη αστέρων του ξενοδοχείου σχετίζονται με την αποδοτικότητα του, με τα μεγαλύτερα ξενοδοχεία και τα ξενοδοχεία 5 αστέρων να παρουσιάζουν την μεγαλύτερη συσχέτιση. Επιπρόσθετα, τα ξενοδοχεία συγκεκριμένων νομών της χώρας αποδεικνύεται ότι είναι πολύ πιο αποδοτικά και λαμβάνουν σημαντικά υψηλότερες βαθμολογίες στο TripAdvisor σε σχέση με άλλους νομούς, λιγότερο τουριστικά αναπτυγμένους. Είναι βέβαιο ότι η εφαρμογή και η γενίκευση των ευρημάτων της διατριβής σε διαφορετικούς πληθυσμούς ξενοδοχείων καθώς και ξενοδοχεία άλλων χωρών απαιτεί περαιτέρω έρευνα. Εν κατακλείδι, η διατριβή διερευνά την περίπλοκη αλληλεπίδραση μεταξύ των οικονομικών μεγεθών και δεικτών των ξενοδοχείων και της

ικανοποίησης των πελατών τους, παρέχοντας σημαντικά συμπεράσματα τα οποία δύναται να χρησιμοποιηθούν για την αποτελεσματικότερη διοίκηση και διαχείριση των ξενοδοχειακών μονάδων, για την λήψη επενδυτικών αποφάσεων δημιουργίας νέων ξενοδοχειακών υποδομών αλλά και βάση για περαιτέρω ακαδημαϊκή έρευνα.

STATEMENT OF ORIGINALITY

This is to certify that to the best of my knowledge; the content of this thesis is my work. This thesis has not been submitted previously for a higher degree or qualification at any other university or institute of higher learning. I certify that the intellectual content of this thesis is the product of my work and that all the assistance received in preparing this thesis, and the sources used have been acknowledged.

ACKNOWLEDGEMENTS

First and foremost, I thank my academic supervisor, Dr. Konstantinos Eleftheriou, Associate Professor in the Department of Economics at the University of Piraeus, for his consistently strong support throughout the preparation of this thesis. This research has also benefited greatly from the thoughtful insights and feedback provided by Christos Alexakis Associate Professor, Finance & Accounting Academic Area, Rennes School of Business, and V. Pappas Senior Lecturer in Finance and Director at the University of Kent.

PREFACE

A shortened, adjusted version of the study, titled “Do hotel financial factors influence satisfaction?” was included in a paper co-authored with Assoc. Prof. C. Alexakis, et al. The paper has been published in **Annals of Tourism Research**. It is available online from January 5th, 2021, from:

https://doi.org/10.1016/j.annals.2020.103128

Contents

Chapter 1. Introduction

This introductory chapter provides insights into the motivation of the topic, presents the research hypothesis, and presents the contribution to existing research as well as the thesis structure.

1.1 Motivation

Greece is a major tourist destination with a developed hotel infrastructure. The tourism industry is of very high importance to the country's economy since it is the predominant component of Greece's GDP. Customer satisfaction is the driving force for succeeding in a sustainable and qualitative critical mass of visitors flowing to the country as a touristic destination. The central question to be addressed is whether there is a correlation between customer satisfaction and hotel financial metrics.

1.2 Research hypothesis

This thesis investigates how hotel customers' online numerical ratings, extracted from the online travel community, are related to hotels' financial performance indicators, qualitative hotel characteristics, and tourism industry variables. To achieve the study's research goals, a sample of Greek hotels (N=554) was selected. Data on TripAdvisor ratings, financial statements, and hotels' qualitative parameters, as well as the tourism sector's data, were collected and analyzed. The dataset covers the 2007–2016 period and concerns 3-star, 4-star, and 5-star hotel categories. Panel data models were estimated with random effects, using Huber-White robust standard errors. Efficiency was examined with the method of financial ratios using a stochastic frontier analysis (SFA) approach. Statistical results indicated heterogeneous behavior for each category of hotels (3-star, 4-star, and 5-star) for different hotel sizes, in terms of beds and fixed assets as well as for different Technical Efficiency levels.

1.3 Statement of Contribution

To the best of my knowledge, this is the first study examining the relation between eWoM (electronic Word of Mouth) rating—via TripAdvisor data—and financial data derived from financial statements, for different categories of hotels i.e., 3,4 and 5 stars, for different hotel sizes i.e., by total assets and by the number of beds they operate, as well as for different technical efficiency stratifications.

Moreover, this is the first time Stochastic Frontier Analysis is used instead of Data Envelopment Analysis to calculate Technical Efficiency, using stratification by different metrics of size and star category. Finally, an extended sample of 3,4,5-star categories of hotels operating in Greece, representing all regions of Greece were for the first time used.

Results provide insights regarding the efficiency of the tourism industry, which can be particularly helpful for investors, entrepreneurs, hotel managers, as well as policymakers in countries where tourism is a leading industry and an important source of national income. Moreover, efficiency in the supply of touristic services can lead to increased guest satisfaction, more targeted investment choices concerning hotel infrastructure, and better management decisions in hotel operations.

1.4. Thesis structure

The present thesis has the following structure: Chapter 2. Introduces preparatory issues helping to conceive holistically the hypothesis tested. A brief presentation of tourism in Greece, and issues concerning how customer satisfaction is perceived and measured are included. Moreover, it addresses the use of big data in tourism and how hotels are classified. Chapter 3. entails a detailed literature review of studies related to the topic of the thesis. Chapter 4. presents details regarding data collection, variables configuration, and relevant descriptive statistics. In Chapter 5, the methodology of the analysis, namely, panel data methods and Stochastic Frontier Analysis, are presented. Chapter 6 discusses the results and Chapter 7 concludes the findings of the study, presenting hints for further research and the limitations of the thesis. Lastly, an Appendix is provided, featuring an overview of the demand and supply dynamics of Greece's tourism and hotel industry, to clearly illustrate the context in which the hotels included into the sample operate.

Chapter 2. The Greek tourism industry, Hotel classification and customer satisfaction

This chapter provides an overview of tourism in Greece and the star categorization of hotels. It explores the utilization of big data in the tourism sector and examines the perception and measurement of customer satisfaction through the TripAdvisor platform.

2.1 International tourism and tourism in Greece

The tourism industry includes different sub-sectors, for example, hospitality, transportation, and tour operators/travel agencies (Corne, 2015). Tourism is one of the most dynamic industries in the global economy and has become an important source of income for many countries and regions globally. According to World Travel & Tourism Council (WTTC) data, in 2017, tourism represented 10% of the world's GDP, 7% of the world's exports, and 30% of services exports. According to the previous source, the tourism industry ranks third after chemicals and fuels and ahead of automotive products as a worldwide export category. During 2017 tourism generated globally 1/10th of jobs.

Over the last five decades (1970s to 2020), global tourism grew significantly as a direct effect of the increase in the social and economic status of the general population, the improvements and investment in transport means, as well as due to the amelioration of living standards in the developed countries. International tourist arrivals, defined as overnight visitors, saw a remarkable growth from 166 million in the 1970s to 1.4 billion by 2018.

According to United Nations World Tourism Organization (UNWTO) data, tourism surpassed the threshold of 1 billion arrivals in 2012 while, in 2017, global international arrivals stood at 1.322 million. Moreover, global tourism receipts, in 2017, amounted to $1.332 billion. In 2017, Europe accounted for 50.8% of total international arrivals, welcoming 674 million international tourists, followed by Asia and the Pacific at 24.3%, the Americas at 15.9%, Africa at 4.7%, and the Middle East at 4.3%. Europe, the world's top tourist destination, in 2017, experienced its eighth consecutive year of sustained growth. That year, international tourist arrivals in Europe reached 672 million, with international receipts amounting to $519 billion. This growth was fuelled by increased travel demand from both within and outside Europe.

Focusing on Southern Europe, total arrivals were 267 million in 2017. Spain led the region with 30.6% of Southern Europe's arrivals, followed by Italy (21.8%), Turkey (14.1%), Greece (10.2%), Portugal (7.9%), Croatia (5.8%), and other Southern European destinations (9.6%).

Tourism is the most important economic sector of the Greek economy and a major source of income during the years of economic crisis (2010–2019). During that period, the tourism industry significantly contributed to the country’s GDP and helped mitigate the impact and implications of the severe economic recession. On the course of this turbulent time, tourism also aided in the reduction of unemployment. Moreover, due to the dispersion and multitude of tourist destinations across Greece, tourism, in 2017, was the main source of regional income for more than half of the Greek territory.

From 2007 to 2012, Greece's tourism sector suffered due to the global financial crisis and the country's debt crisis, leading to a decrease in tourist arrivals impacted by economic instability and negative media image. Concurrently, there was a global uptick in budget travel, positioning Greece as an attractive destination due to its lower costs during the crisis. After 2012, the Greek tourism industry saw a significant revival, reaching record numbers of inbound tourists. This resurgence was fuelled by Greece's cost-effective appeal, its classic attractions of sea, sun, and culture, and marketing efforts by the government and the sector’s business to rebrand Greece as a secure and inviting destination.

According to the Association of Greek Tourism Enterprises (SETE Intelligence), in 2017, arrivals in Greece stood at 27.2 million, producing 88 million overnight stays in accommodation establishments. Foreign tourists were most of the occupancy share, constituting 84% of overnight stays. During the decade from 2007 to 2017, Germany and the UK consistently led as the primary sources of tourists to Greece, comprising 25% of all international visitors in 2017. Italy and France followed, accounting for another 10% of arrivals that year. Predominantly, in 2017, 68% of tourists to Greece were from EU countries, with the largest portion of overnight stays and tourism revenues generated by visitors from Germany, the UK, France, and Italy. These tourists not only spent more per visit but also tended to stay longer.

While the influx of tourists to Greece increased, the revenue per tourist did not rise proportionately, suggesting a reduction in per capita spending. In 2017, the average expenditure per stay was €68.00, with Crete and the Southern Aegean registering the highest rates, and Central Macedonia and Central Greece the lowest. The tourism industry in Greece is markedly seasonal, with 85% of visits occurring from May to October. The Southern Aegean, Central Macedonia, Crete, Attica, and the Ionian Islands, known for their well-developed tourism infrastructure, accounted for most overnight stays and arrivals.

In 2017, Greece's tourism revenue amounted to €14.202 million, predominantly generated by the following regions: Southern Aegean, Crete, Attica, Central Macedonia, and the Ionian Islands. Domestic tourism also played a significant role, although exhibiting a decrease in spending from €3.9 billion in 2008 to €1.4 billion in 2017, reflecting the economic crisis's impact on local tourism overnight stays and spending.

The positive track of the incoming tourists in Greece over the last years had been the result of the concurrence of a variety of factors, such as civil security and safety in the Eastern

Mediterranean region, the economic growth that took place in the countries of origin of incoming tourists, the increase of the level of infrastructure and the upgrade in the country's hotel capacity as well as the implementation of solid promotional activities abroad by the Greek state and private businesses. The rapid growth in the tourism sector over the last 10 years has had a positive impact on the improvement and development of the hotel industry infrastructure.

The Greek Hotel Sector is regarded as an integral part of the "Tourism Industry" since it affects a wide spectrum of associative services/businesses such as travel agencies, car rentals, restaurants, and leisure facilities. According to the Association of Greek Tourism Enterprises (SETE Intelligence), as of 2017, Greece boasted 9,783 active hotel units, encompassing 414,127 rooms and 806,045 beds. Of these, five-star hotels accounted for 19% of the bed capacity, four-star hotels for 26.2%, and three-star hotels for 18.6%. The majority were two-star hotels, making up 39.9% of the total. Over the decade from 2007 to 2017, the number of Greek hotels increased by 6.26%, and the total bed capacity rose by 15%, indicating a shift towards larger hotel establishments. Notably, the count of five-star hotels in Greece surged by 149% between 2007 and 2017, in contrast to the decline in one-star and two-star hotels by 11% and 16%, respectively. During this period, the capacity for beds and rooms in five-star hotels expanded considerably. Most of this capacity in 2017 was found in the Southern Aegean, Crete, the Ionian Islands, Central Macedonia, and Attica, regions predominantly housing the bulk of five-star and four-star hotels.

According to SETE Intelligence, seasonality is a key characteristic of the Greek tourism industry since most arrivals (69% for 2017) take place during the summer period (June to September). The highest seasonality observed in the Southern Aegean, Crete, and the Ionian Islands. The average occupancy rate of Greek hotels in 2017 stood at 52.80%, which was lower compared to neighbouring countries like Turkey, Italy, and France.

According to the “Developments in Key Figures for the Greek Hotel Industry 2018 Survey” conducted by the Hellenic Chamber of Hotels, investment by the hotel sector for the construction of new and for renovation/repair of existing hotel rooms, in 2017, was estimated at approximately € 1,541 million.

Travel receipts in 2017 amounted to €14.6 billion, accounting for 74.0% of Greece's trade balance and 52.0% of the nation's export income. A study by KEPE underscored a 2.5 multiplier effect for accommodation services specifically. According to SETE Intelligence, the direct contribution of the tourism industry, in 2017, to the Greek GDP was estimated to be 10.3%. The tourism industry's impact was particularly prominent in three island regions of Greece - Crete, the Ionian Islands, and the South Aegean - where it formed a significant portion of the regional GDP (47.2%, 71.2%, and 97.1%, respectively).

The tourism sector also played a crucial role in employment, with its employment rate growing annually by 0.9% between 2008 and 2017, in contrast to a 2.5% decline in other sectors. By the third quarter of 2017, tourism employment in Greece reached approximately 398.7 thousand individuals according to data from INSETE.

It is worthwhile to mention that following the period under study (2007 – 2016) and leading up to the end of 2023, several significant events took place, greatly impacting the tourism and hotel industry. Key among these were the COVID-19 pandemic beginning in 2019, the conflicts in Ukraine in 2021 and Gaza in 2023, escalating fuel costs and inflation rates, environmental changes, and the growing popularity of Airbnb. All these factors are recognized as major disruptions to the sector. On the contrary, the period from 2007 to 2016 was a more "business-as-usual with challenges" timeframe for Greek tourism. In that context, it is assumed that the current analysis reflects in a more comprehensive and "sterilized" manner the relation between customer satisfaction as exhibited by TripAdvisor's ratings and financial metrics of the companies operating the hotels.

A more detailed presentation of demand and supply characteristics of the Greek tourism sector as well as data concerning the justification of the importance of tourism for the Greek sector is presented in Appendix A of the study.

2.2 Hotel classification

Hotel classification systems are used in the hotel industry to provide both consumers and sector stakeholders with a comparable indicator of the infrastructure and services standards to be found at individual establishments. There is a wide variety of rating systems used by different countries and/or organizations around the world. Many have a system involving stars, with a greater number of stars indicating greater luxury. Hotels in Greece are ranked on a scale of 1 to 5 stars. The Greek star rating scheme differs from that of other star rating systems in the EU and worldwide. Eventually, there is not a horizontal, homogenized system applied worldwide or uniformly in the EU territory.

In Greece, 5-star hotels offer luxurious infrastructure and service keeping almost uniformly the standards of their category with negligible exceptions. 4-star hotels offer upper-class accommodation presenting some within-variation in the overall experience offered. 3-star hotels offer above-average service and facilities, but the quality of the final product cannot be considered homogeneous throughout the category. A considerable percentage of 3-star category hotels outperform their category standards, offering better overall experience and some underperform. During the 2007–2016 decade, a significant shift from 1- and 2-star hotels to 3,4- and 5-star hotels took place in the Greek hotel sector. 1- and 2-star hotels decreased in units by 41%, while in the same period, the number of 4- and 5-star hotels increased by 58% in numbers according to data provided by SETE. Considering this thesis, 5,4-, and 3-star hotels category is investigated, leaving 1- and 2-star categories off the sample for reasons explained in Chapter 4.

Regarding focused studies that concern 5,4-, and 3-star hotels, despite heterogeneity in the criteria of the hotel star-rating classification system throughout the world, a relationship does exist with guest satisfaction, based on the scores awarded by former customers both on TripAdvisor and on Booking (Martin-Fuentes, 2016). According to Rhee and Yang (2015), hotel guests' expectations and actual experiences of hotel service quality often fail to coincide, and this could be augmented depending on the hotel star-classification and the overall rating from previous guests. Furthermore, Lopez and Bedia (2016) state that "quality is associated with the delivery of a service according to client expectations more than it is with establishment's category." Finally, Blomberg-Nygard and Anderson (2016) accord with the fact that traditionally, classification systems have been about amenities, whereas guest reviews are about meeting expectations; thus, guest reviews may be able to provide a quality check on the amenities that are required as part of the classification system.

2.3 Guest experience and customer satisfaction

With a direct impact on reputation, loyalty, and revenue, guest satisfaction is the ultimate measure of success for lodging providers, and other types of hospitality businesses. In the travel industry, "customer's lifecycle", is a measure of accountability when one assesses the overall guest experience. For hotels, the customer lifecycle can be divided into five stages: research, shopping, booking, in-stay, and post-stay. Throughout the vacation lifecycle, the traveler contacts the hotel in person, on the phone, or digitally. Each contact with the hotel creates impressions. If all goes well, on the next visit, the customer may skip the shopping and research phase and rebook the same hotel. Moreover, the guest will share his experience with other travelers via eWoW.

The traveler's expectations of the hotel are formed in the shopping, booking, and pre-stay stages. Upon arrival at the hotel, the moment of truth comes, where the lodge will either meet the expectations or will disappoint the traveler. The outcome is determined by a variety of factors, including value-for-money perception, location, infrastructure, service provided, comfort, amenities, and food & beverage experience. Expectations are confirmed or broken, usually (but not always), upon the guest’s arrival at the hotel. Essentially, the hotel the traveler visits will either:

- Fail to meet expectations. The guest is disappointed and may become a detractor. He will not come back. He will not recommend the hotel. Even worse, he may even warn others to stay away.

or

- Meet expectations. In that case, the guest’s basic needs are met, but he doesn’t feel passionate about the hotel. He is passive or neutral. He might come back, but he also may not. He may recommend the hotel, but the feedback will probably be lukewarm or mixed.

or

- Exceed expectations. The guest is delighted with his stay and may become a promoter. Not only will he come back, but he may also recommend the hotel to others. In that case, the satisfied guest becomes an asset for the hotel.

Consequently, a happy guest is valuable to hotels as he tends to be a repeater, loyal customer, and likely to recommend. Moving from a pleasant guest experience to a delightful guest experience can lead to an increase in guest likelihood of a guest returning and recommending.

When it comes to measuring guest satisfaction, hotels have two different options. The first is to collect feedback delivered by visitors on social media and through digital channels. Online reviews are posted publicly to review platforms like TripAdvisor and Booking. com—often unsolicited by the hotel. The second option is guest surveys. These are performed by hotels via email, telephone interview, or text. Results are kept private. Both types of feedback provide content that is easily measured and quantified. Whether it is sleeping quality, value, or service, if a gap exists between what guests expect and what the hotel delivers, the hotel needs to figure it out. Considering a gap between expectations and deliverables, two different cases may be held: marketing overpromising or operation underdelivering. In that context, the way a hotel or brand sets expectations falls to marketing, sales, revenue management, and the brand. The way staff deliver on expectations primarily falls to operations. In some cases, both setting and delivering on expectations underperform. Management needs to take measures to ensure that the marketing and sales departments set expectations to a realistic level as well as that staff receive the resources and support needed to upgrade the services provided. Both actions have as a final goal to bridge the gap. The concept of knowing the guests' preferences is fundamental for a hotel to deliver exceptional guest experience. Fortunately, collecting the relevant information has become much easier, since today's guests are more willing than ever to provide it via social media and review platforms.

The basic need for a quiet, well-maintained, clean room in a safe and interesting location remains critical to a satisfactory guest experience. Parallel, new standard comforts such as sustainability practices, modern fitness facilities, destination interactive experiences, etc., play an augmented role in customer satisfaction. Engaging guests meaningfully and creating lasting, positive impressions is the path to obtaining and retaining guest loyalty and advocacy. Moreover, the challenge of differentiation lies beyond the expected comforts.

Hotels can elevate guest satisfaction by knowing their preferences and adjusting the experience accordingly. In that sense, loyalty programs can be redefined to attain a more personalized approach. Empowering guests to customize their experience is of equal importance as it offers them the flexibility and the freedom to satisfy their needs when and whenever they want to. Finally, “surprise and delight” the guest is the “over and above” that makes a memorable experience. Simple gestures or actions that are pleasantly unexpected and custom-made, can certainly exceed guests’ expectations, and promote their loyalty.

Overall, knowing the guest is the foundation for delivering on all guest needs. Knowing the customer involves getting familiar with the guest’s personal preferences. Consequently, guests are more accustomed to hotels that interact, understand, remember, and predict their preferences. Moreover, the way the hotel personnel engages with guests is essential in establishing a personalized, authentic, and attentive relationship. It is a fact that despite all the technological advancements, human engagement remains critical for the final perception of satisfaction a visitor forms. Guests are more likely to recommend when hotel teams are friendly and attentive, proactively communicate with them, and offer personalized experiences. Of equal importance is the way hotels listen to guests' needs, handle difficult and unpredictable cases, empathize with their situations, and react accordingly. The hotel employees must establish a unique, tailormade and authentic, interaction with the hotel guest. The hotel personnel must be communicative, proactive, and friendly.

Realizing the impact of "guest knowledge", "guest experience", and "guest satisfaction" hotels must develop their cognitive capabilities to enhance the guest experience and consequently boost turnover and profit.

2.4 The TripAdvisor rating system

Textual reviews give a sentiment on the perceived products or services. It can be either positive, negative, or neutral. According to Öğüt and Taş (2012), online reviews posted on well-known online travel platforms are perceived as more credible and more useful compared to less unknown sources. Moreover, online reviews of hotels listed in "best hotels rankings" have more chances to receive bookings when compared to hotels in lists with less high-ranked hotels (Casaló, Flavián, Guinalíu, & Ekinci, 2015). Geetha et al. (2017) showed that a strong interrelation exists between the customer's online textual reviews and star ratings. In general, most of the research proves that customer textual reviews using star ratings are more consistent and useful compared with textual reviews alone. Star ratings represent various levels of sentiments—for example, one star represents the extremely poor quality of the service, whereas a five-star rating represents relatively good service. For average services anticipated, consumers tend to use the middle value (three stars) to represent their mixed feelings (Mudambi and Schuff, 2010).

Nowadays, consumers making purchasing decisions in the tourism universe are heavily based on the Electronic Word of Mouth (eWOM) shaped through online textual reviews shared by previous customers (Baka, 2016). eWOM seems to be more reported if the reviews are positive compared with negative ones (Lee & Youn, 2009). In the online world, eWOM is found in different social media such as Facebook, travel blogs, online travel communities, and specially designed platforms such as TripAdvisor. Moreover, they exist in different presentational forms such as reviews, comments, rating videos, and photos. Customer reviews and ratings are posted on online related platforms such as TripAdvisor and have a significant effect on future customers' decisions (Xie, Zhang, & Zhang, 2014). According to Xiang, Schwartz, Gerdes, & Uysal, 2015 researching ratings is the most straightforward method to understand the customer's overall satisfaction.

TripAdvisor is one of the most popular online Travel websites used throughout the world. People will refrain from using such platforms when planning their trips, and they will continue to use them even when at their destination to get relevant and credible information (Yoo, Sigala, and Gretzel, 2016). Travelers around the world check on TripAdvisor for the availability of accommodation, explore different destinations, and cruise on information. Moreover, they share useful information, photos, and videos. TripAdvisor, Inc. operates a website and mobile app with user-generated content and a comparison-shopping website. It also offers online hotel reservations and bookings for transportation, lodging, travel experiences, and restaurants. TripAdvisor.com reached 463 million average monthly unique visitors in 2019. The website has versions in 48 markets and 28 languages worldwide. As of September 2019, it featured approximately 859 million reviews and opinions on approximately 8.6 million establishments—including 1.4 million hotels, inns, bed and breakfasts, and specialty lodging, 842,000 rental properties, 5.2 million restaurants, and 1.2 million travel experiences worldwide (source: Wikipedia, TripAdvisor, 2019).

TripAdvisor relies heavily on user-generated content. Travelers can create accounts and contribute reviews, ratings, photos, and videos of their experiences with hotels, restaurants, attractions, and other travel-related services. The main operation of TripAdvisor's platform is the collection of reviews and ratings. Users share their opinions and experiences, giving a rating out of five stars and writing detailed reviews about various aspects of their stay or visit. Users can search for specific destinations, hotels, restaurants, or attractions on the TripAdvisor website or app. They can also browse different categories and locations to find relevant information. TripAdvisor uses an algorithm to determine the popularity and ranking of businesses based on factors like the quantity and quality of reviews, ratings, and user engagement. This information helps users identify popular and well-reviewed options. Moreover, TripAdvisor allows users to book hotels, flights, and other travel services directly through its platform. It partners with various online travel agencies and booking platforms to facilitate these bookings. Businesses, including hotels, restaurants, and attractions, can claim their listings on TripAdvisor and access management tools. They can respond to reviews, update their business information, and engage with customers. Users can interact with reviews by upvoting helpful ones and marking reviews as "helpful" or "not helpful." This helps filter out potentially biased or unhelpful content. Users often utilize TripAdvisor as a trip-planning tool. They can create itineraries, save places they want to visit and read travel forums for advice and tips from other travelers. TripAdvisor employs content moderation to ensure that the reviews and content posted on its platform meet certain standards and guidelines.

TripAdvisor generates income through several revenue streams. Initially, it offers advertising opportunities to businesses in the travel and hospitality industry. These businesses can promote their services through display ads, sponsored listings, and other advertising formats on the TripAdvisor website and app. These ads are often displayed alongside search results and business listings. Furthermore, TripAdvisor partners with online travel agencies (OTAs) and booking platforms to facilitate hotel, flight, and other travel service bookings directly through their platform. When a user makes a booking through TripAdvisor, the company receives a commission from the partner for referring the customer. Moreover, TripAdvisor offers a feature called "Sponsored Placements" that allows businesses to have their listings prominently displayed in search results and on specific pages. This provides greater visibility and increases the chances of attracting users' attention. The concrete service is provided by TripAdvisor with a fee.

In 2019, TripAdvisor revealed the results of a study that uncovers how significantly reviews influence booking decisions. The global study polled more than 23,000 TripAdvisor users across 12 markets on the use of online reviews and their role when booking hotels, restaurants, and experiences. The results show that traveller reviews remain a go-to source of information, with 72% of respondents always or frequently reading reviews before deciding on places to stay and eat, or things to do. The figure is even higher when it comes to accommodation bookings, with four out of five participating travellers (81%) always or frequently reading reviews before booking a place to stay. Nearly 8 out of 10 TripAdvisor users (79%) are more likely to book a hotel with a higher rating when choosing between two otherwise identical properties, and more than half (52%) agree that they would never book a hotel with no reviews. When researching their travel, respondents indicate that the most important thing they are looking for when reading reviews is recent content.

When submitting a review on TripAdvisor, travelers submit ratings on a scale from 5 (Excellent) to 1 (Terrible). The average review rating in 2020 was 4.30 out of 5.0, up from 4.22 out of 5.0 in 2018. In 2020 26 million reviews were submitted to TripAdvisor, 12 million for restaurants, 8 million for hotels, and 4 million for experiences, attractions, and activities. Most reviews are submitted for user experiences in Europe (54.1%), North America (23.5%), and Asia and the South Pacific (13.7%).

TripAdvisor has been the subject of controversy for allowing unsubstantiated anonymous reviews to be posted about any hotel, bed and breakfast, inn, or restaurant. In October 2021, TripAdvisor published for the first time a transparency report, according to which 9% of reviews are fraudulent or biased, requiring them to be removed. TripAdvisor reported that this is a global problem and they found fake reviews in 131 countries.

2.5 Big data and tourism

Today, information on guest experience and satisfaction is massively found on social media and tourism platforms such as TripAdvisor and Booking.com. Moreover, performance results can be found in hotels' financial statements. For hotels to extrapolate valuable information and patterns on guest experience from the above-mentioned sources, they must combine big data technologies in an efficient and meaningful way. Cognitive insights gained via massive platforms data interpretation means understanding and exploiting hidden patterns in datasets (extracted by large-scale bundles of datasets found on rating platforms, hotels financial data, big data provided by sectoral bodies, statistical services, and government intelligence agencies). Using pattern detection and cross-referencing, hotels may draw valuable conclusions.

The new world of online data-driven guest experience can also enable the development of a partnership ecosystem powered by digital engagement. Guests with digital portable devices can live self-created experiences that involve the hotel more than the hotel-offered amenities. So, the hotel can act as the middle agent between guests and a "partnership ecosystem" that extends beyond its services and facilities. Guests seek the benefit of local knowledge and experience. The hotel's role is to identify how to make this process easier. In a partnership ecosystem, hotels can use new business relationships and mobile technology, for example, to offer their guests access to a boutique fitness experience operating individually outside the hotel business. Moreover, a partnership ecosystem helps the hotels to compete more effectively in the sense that hotels through the concrete build-up procedure create a portfolio of assets owned by third parties, helping them offer a diverse set of experiences without an extensive capital outlay.

Nowadays, there is a great potential for improvement in attaining the ultimate guest experience. That kind of information can be extracted from big data sources such as the ratings and reviews from tourism platforms, from the elaboration of statistics collected from renowned bodies such as the state's statistical authorities, and hotel associations, and even from the financial statements of the hotels' operating companies. The identification of relations between them and consequently the efficient combination of conclusions shaped by artificial and human intelligence can deliver new management methods and implications for designing high-quality experiences and services for hotel guests.

A significant recent body of literature has investigated 'electronic word of mouth' for hotels, particularly TripAdvisor ratings of experiences. Mainly, two branches of research on electronic word of mouth have so far been examined. The first focuses on the impact of ratings, e.g., how online ratings influence decisions of customers who observe these ratings (Filieri, 2016; Liang et al., 2020). The second branch seeks to explain the process by which hotel guests generate their ratings. This thesis is situated in this latter area. Prior studies have primarily collected hotel features available through an online platform such as TripAdvisor and used these factors to explain ratings on that platform (Banerjee and Chua, 2016; Mellinas and Nicolau, 2020; Radojevic et al., 2017).

Missing from prior studies is a link between 'offline' information and online ratings. This thesis contributes to the literature by concentrating on the offline information and especially on hotel financial measures, which have not previously been investigated as a driver of guest ratings. The motivation is that the financial accounts of a hotel are a snapshot of the long-term (assets) and short-term (expenses) investment of a hotel in infrastructure and service quality. Hotel service quality has been linked to guest satisfaction (Choi and Chu, 2000; and Choi and Chu, 2001) but the survey approaches utilized in both these studies are time-consuming to implement. It is also not clear how guests relatively value short- and long-term investment, as well as different aspects of service investment, in terms of judging service quality. The choice of financial variables in this thesis allow to explore the latter, including the quality of amenities (in terms of infrastructure) and experiences (fixed assets and sales expenses) and service efficiency (cost of goods sold and administration expenses). Thus, while prior research by Chang and Sokol (2020) shows that financial investment is used by hotels to improve service quality, the study can show which form of service investment is particularly important.

Chapter 3. Literature review

This chapter presents a selection of the empirical literature regarding tourism's effects on the economy, tourism demand, hotel performance and efficiency, hotel classification, as well as eWOM and hotel performance.

3.1 Tourism and the Economy

Tourism is viewed as an important tool in adding economic value to a country, through improving the balance of payments and creating income, taxes, hard currency, and employment. A connection between tourism and economic growth has long been established. A substantial part of the literature discovered a positive relationship between tourism demand and economic growth for developed and developing economies (Alam and Paramati 2017; Jambor and Leitão 2017; Mitra 2019; Shaheen et al. 2019). This association was true not only for Western countries but also for countries in Eastern Europe (Jambor and Leitão, 2017) and South Asian countries (Amin et al. 2019). The results were also held for studies in Turkey and Taiwan (Gunduz and Hatemi, 2005; Kim and Chen, 2006). Moreover, the connection is particularly strong for island economies, where economic performance is mainly driven by tourism (Santana-Gallego et al., 2011; Seetanah, 2010). Additionally, smaller countries that specialize in tourism appear to grow faster (Brau et al., 2007). Eugenio-Martin et al. (2004), also claim that the tourism development effect is more pronounced in countries with lower GDP levels.

Seetanah (2010), presented the possibility of a dynamic linkage between tourism and economic growth through the usage of a dynamic panel data framework of island countries over the period between 1990-2007, namely the Generalized Method of Moments (GMM) method. The results of the analysis revealed that tourism development is a major factor in explaining economic performance in island economies and the results are consistent with earlier work on developing countries by Gunduz and Hatemi (2005) and Tosun (1999) for the case of Turkey, Kim et al. (2006) for Taiwan, and Eugenio-Martin et al. (2004) for a sample of Latin American countries. Furthermore, the number of visitors was able to predict other macroeconomic variables such as exports and money supply, and to a lesser extent, the exchange rate and GDP of a country (Sharma and Bannigidadmath, 2013).

The relationship between trade openness and tourism has been explored in recent years, with researchers finding that international tourism promotes international trade between countries. Santana-Gallego et al. (2015) present empirical and theoretical evidence that tourism matters for trade openness. Moreover, the empirical evidence suggests that tourism increases both the probability of two countries trading with each other and the volume of international trade between them. A Korean study also confirms the link between the trade of goods and tourism (Keum, 2011).

On a regional level, panel data evidence suggests that tourism development also affects the neighbors. Yang (2012) identified agglomeration effects from tourism development in Chinese provinces. Moreover, Eleftheriou and Sambracos (2019), also found that tourist arrivals can create spillover effects to nearby (NUTS-3) regions in terms of economic growth. This section is by no means exhaustive but serves as a background on the importance of tourism for the economy.

3.2 Tourism demand

Customer satisfaction is one of the important determinants of tourism demand. The estimation of tourism demand has several challenges since it involves the consumption of multiple and heterogeneous services. Additionally, the link with transportation and the autocorrelation between sequential periods make modeling even more difficult. The choice of tourist flows as the dependent variable creates un-separability between leisure consumption and labor supply and/or between tourism demand and the demand for other goods and services. All these issues are compounded by the existence of unquantifiable factors influencing demand, as well as by the inaccuracy or unavailability of data for those that are, in principle, measurable. Moreover, it is very difficult to find a precise indicator for measuring external demand, but prices and competitor countries perform well prediction-wise (Gonzalez and Moral, 1995). Due to all these methodological complexities, the use of panel data methods has prevailed (more on these methods in Chapter 4).

For the time, studies choose either a predictive time series approach, i.e., demand forecasting between one or several pairs of countries (Dritsakis, 2004; Kulenderan and Witt, 2001; Seddighi and Theocharous, 2002; Song et al., 2003), or explore the determinants of tourism demand (Couch,1994; Witt and Witt, 1995; Lim, 1997). Tourist arrivals/departures and tourism receipts/expenditures by country of origin usually serve as the dependent variable. The number of tourists is the most prevalent dependent variable (for example see Crouch and Shaw, 1992). Other approaches use, as the depended variable, foreign visitors crossing the borders, the number of nights spent by visitors from abroad, the receipts originating from the visitors' spending, or the length of stay of tourists visiting a country. Ledesma-Rodriguez et al. (2011) tried an alternative approach where the number of visitors lodged in the destination country was used as the dependent variable. According to them, this variable is preferable since it considers the length of stay and excludes tourists who are hosted at family or friend's houses. The number of overnight tourists staying in their destination countries and the average length of stay has also been studied, but much less frequently.

Nevertheless, none of these measures is fully satisfactory in encompassing all the aspects that characterize the demand for tourism in a specific location. The comprehensive reviews of the empirical literature on tourism demand by Crouch (1994a, 1994b, 1995), Witt and Witt (1995), and Lim (1997, 1999) substantially agree on both the tourism demand measures and the variables that are important in explaining international tourism flows. A more recent review on tourism demand forecasting by Song and Li (2008), indicates that new methodologies have been applied and highlights the importance of modeling unexpected events such as crises and disasters.

Tourism demand has been studied in several countries, or groups of countries, through different perspectives. For the explanatory variables, empirical models of tourism demand borrow heavily from the consumer theory, which predicts that the level of consumption depends on the consumer's income, the price of the good/service in question, the prices of related goods, and other demand shifters. As a result, income and prices are the most used variables in terms of the major factors influencing tourism demand. These two factors (GDP *per capita* and relative prices) were found to be the two most relevant variables in explaining tourism arrivals (Martins et al., 2017).

Aside from a positive association with income and GDP, which was found in many other studies (Ibrahim, 2011; Khoshnevis-Yazdi and Khanalizadeh, 2017; Leitão, 2010; Surugiu et al., 2011), other economic conditions can also affect tourism demand (Seetaram, 2012). As expected, prices had a negative effect on tourism demand (Ibrahim, 2011; Khoshnevis-Yazdi and Khanalizadeh, 2017). Khoshnevis-Yazdi and Khanalizadeh (2017), also found a negative association with the real exchange rate in the US. A connection with trade flows is also well established (Ibrahim, 2011; Leitão, 2010; Santana-Gallego et al., 2011; Surugiu et al., 2011).

A country's risk was also related to tourism demand (Sequeira and Nunes, 2008). In this complex index exist several non-economic factors such as political stability and personal safety (Naudé and Saayman, 2004). Demand is also affected by events that might have taken place in the past such as terrorist attacks (Khoshnevis-Yazdi and Khanalizadeh, 2017; Teresa and Martin, 2007).

The role of transport and tourism infrastructure has also been highlighted (Khadaroo and Seetanah,2007; Naudé and Saayman, 2004). Kaul (1985) was among the first to recognize the importance of transport infrastructure as an essential component of successful development in that it induces new attraction creation and the growth of existing ones. Tourism and transport infrastructure can help reduce tourism congestion which has been positively linked to the number of tourist arrivals (Albaladejo et al., 2016).

Studies have found that population size is also positively related to tourism demand (Leitão, 2010; Surugiu et al., 2011). Finally, information and communications technology (ICT) as well as the country's marketing efforts may also play a role (Naudé and Saayman, 2004; Ramos and Rodrigues, 2013).

3.3 Hotel performance and technical efficiency

The hospitality management literature has attracted increasing research interest in financial performance and efficiency during the past two decades (Jang and Park, 2011), with emphasis on quantitative analyses for risk management, capital structure, financing, and bankruptcy (Manasakis et al., 2013). Studies benchmarking tourism destinations have initially focused on issues such as customer satisfaction (Milman and Pizam, 1995), competitiveness (Kozak and Rimmington, 1999; Ritchie and Crouch, 2003), or some simple indicators such as tourism arrivals, tourism receipts, foreign exchange earnings, changes in market share and tourism satellite accounts (Dwyer and Kim, 2003).

Furthermore, performance evaluation is a necessary part of management control since it can be used as a reference in decision-making, but also serves as the basis for improvement. The literature on tourism performance is well-established (Tsionas and Assaf, 2014). Conventionally, research uses average occupancy rates and average room/rates as indicators of performance. Performance-evaluation factors relevant to the hospitality industry are multidimensional and cannot simply be aggregated using price or cost figures (Chen, 2009). In a review of the literature regarding hotel performance, Sainaghi (2010), identified four main areas of interest, namely, strategy, production, marketing, and organization.

Hotel performance has been associated with several factors such as the organizational type (Barros et al., 2011; Botti et al., 2009; Chen, 2007; Fuentes, 2011; Hwang and Chang, 2003; Koksal and Aksu, 2007; Min et al., 2009; Perrigot et al., 2009; Shang et al., 2010), hotel location (Barros, 2005; Chen, 2007; Fuentes, 2011; Hwang and Chang, 2003; Shang et al., 2010; Wang et al., 2006), hotel size (Chen, 2007; Hwang & Chang, 2003; Pulina et al., 2010; Sardo et al., 2018; Serrasqueiro and Nunes, 2016), e-commerce (Shang and Wang, 2008a;

Shang et al., 2010), age and experience (Fuentes, 2011; Shang et al., 2010), stock returns, capital structure (Sardo et al., 2018; Serrasqueiro and Nunes, 2014), and gross domestic product (Baik et al., 2012; Chen, 2010; Chen and Kim, 2010).

Additionally, hotel performance has also been positively related to occupancy rates (Chen, 2010), growth rates of total foreign tourist arrivals (Chen, 2010; Hwang and Chang, 2003), and socio-economic factors such as education, county's international attractiveness, and the payment levels of employees (Huang et al., 2012). Moreover, factors such as being publicly traded, adopting mergers and acquisitions strategies, and endowments in touristic regions (Barros et al., 2011), are all related to superior hotel performance.

Measuring efficiency is an important and broad-scope subject. Managers, economists, and other researchers have attempted to accurately measure the efficiency of the hotel industry for many years (Hwang et al., 2003). Since the late 90s, many papers have been published concerning tourism efficiency and several studies have used production frontier models (Assaf, 2012; Guetat et al., 2015; Pérez-Granja and Inchausti-Sintes, 2023). Effectiveness scores have been positively related to overall hotel performance (Singh et al., 2020). Wassenaar and Stafford (1991) advocate the use of a lodging index indicator for the hotel/motel industry. The lodging index is defined as the average revenue realized from each room, vacant or occupied, within a region or city during a given period.

When it comes to quantitative analysis of efficiency two econometric methods are most used, the stochastic frontier approach (SFA) and the data envelopment analysis (DEA). The DEA approach is far more popular since, unlike the SFA, it does not impose any functional form on the data, nor does it make distributional assumptions for the inefficiency term (Charnes et al., 1978, Charnes and Cooper 1985). However, the DEA is sensitive to outliers. Moreover, it is essential to recognize that the specifications DEA can be notably influenced by how the data is structured in a panel format, encompassing considerations such as data attrition, among other factors. It also deserves mentioning that some of the DEA deficits can be addressed by non-traditional DEA models, such as the cone-ratio DEA model of Charnes et al. (1990) and the assurance region DEA model of Thompson et al. (1986, 1990), which however are not particularly used due to computational difficulties.[1]

The DEA method has several more limitations. Many studies have used the deterministic DEA method which is biased by design and is sensitive to the sampling variations of the obtained frontier (Simar and Wilson, 1998, 2000). Moreover, Simar and Wilson (1998, 2000) identified strong dependence in statistical terms of efficiency scores generated by the DEA, and therefore their use in a seconds-stage regression would violate the assumption of the basic model

[1] Both models, cone-ratio DEA and assurance region DEA contribute to the broader field of efficiency analysis by offering methods to handle uncertainty in the input-output data, thereby enhancing the robustness and applicability of DEA in practical settings. The cone-ratio DEA model is an extension of the classic Data Envelopment Analysis (DEA) introduced by Charnes, Cooper, and Rhodes in 1978. The cone-ratio model is designed to handle situations with uncertainty in the data, such as measurement errors or variations. It employs a mathematical framework to model this uncertainty by considering a cone of potential variations around the observed data points. The model aims to find the efficient frontier within this cone of uncertainty, providing a more robust analysis of efficiency. The assurance region DEA model is another variant of the traditional DEA methodology. Proposed by Thompson and colleagues, this model introduces the concept of an "assurance region" to account for imprecision or noise in the data. In essence, the assurance region defines a permissible range within which the decision-making units (DMUs) are allowed to deviate from their observed inputs and outputs. The model seeks to identify the set of efficient DMUs within this defined assurance region, providing a more flexible and realistic approach to dealing with uncertainties inherent in real-world data.

required by the regression. Several studies have also used censored models to investigate the determinants of the efficiency estimates. However, according to Simar and Wilson (2007), DEA estimates used in a second stage are biased and serially correlated, which implies that standard methods for inference in the second-stage regression are invalid. The motivation for the present study comes from the fact that most DEA studies in the tourism industry are static, based mainly on cross-sectional data at the individual level of hotel entities.

Innovative tools, derived from DEA, have also been applied to measure efficiency in the tourism sector (Peypoch, 2007; Peypoch and Solonandrasana, 2008). However, no studies of the tourism hospitality sector have used the hierarchical category DEA model, except for a paper by Mansourirad (2013) which applied a fuzzy DEA method. The hierarchical category DEA model is useful for analyzing a sample where the units are grouped. Previous studies on hospitality efficiency failed to consider the heterogeneity of the hotels based on their star rating (Barros and Dieke, 2008; Hathroubi et al., 2014). For the hospitality sector, some of the main contributions that have used DEA in the last decade are, among others, Barros and Mascarenhas (2005), Barros (2005), Assaf et al. (2010), Assaf and Agbola (2011), Assaf and Barros (2011), Barros et al. (2011) and Hathroubi et al. (2014). See Assaf et al. (2012) for a complete overview of this topic.

Other research papers that studied efficiency using DEA in the tourism sector include Bell and Morey (1995) for corporate travel departments, Anderson et al. (2000) for US hotels, Sigala (2004) for the 3-star hotel sector in the UK, Hokey et al. (2008) for hotel chains in Korea, Botti et al. (2009) for French hotels, and Wu et al. (2011) for 4- and 5-star hotels in Taipei. Several studies also exist for Taiwan (Tsaur and Tsai, 1999; Lee et al., 2011; Fu et al., 2011; Wang et al., 2007) and China (Zhou et al., 2008; Tsai, 2009; Huang et al., 2012; Li et al., 2014; Luo et al., 2014).

More recent works that utilize the SFA framework include Guetta et al. (2015), who studied how corporate governance affected hotel industry performance in Tunisia, and Pérez-Granja and Inchausti-Sintes (2023), who studied the importance of specialization for hotel efficiency in Spanish provinces. Assaf et al. (2017) also make a case for the benefits of using a Bayesian approach in SFA, which was applied for the benchmarking of the Asia Pacific tourism industry in combination with the DEA approach (Assaf, 2012). The combination of the SFA with tourism forecasting has also been advocated recently (Wu et al., 2019; Wu et al., 2020). Assaf and Tsionas (2019) offer the latest review to date of performance modeling in tourism with a special focus on frontier models.

Regardless, both methods for estimating technical efficiency (DEA and SFA) assume that the production function of the fully efficient decision unit is known. In practice, this is not the case, and the efficient isoquant must be estimated from the sample data. Under these conditions, the frontier is relative to the sample considered in the analysis (Barros et al., 2009). The overall research using frontier models applied to the tourism industry is enlarging the traditional models used within this methodological approach. There are several papers proposing the frontier model to evaluate hotel performance and examine its determinants, considering different inputs and outputs. For example, among others, Morey and Dittman (1995), Johns et al. (1997), Anderson et al. (1999; 2000), Brown and Ragsdale (2002), Wöber (2002), who discussed the overall operational efficiency of hotels industry. For more information see among others Hwang and Chang (2003), Reynolds (2003), Barros and Alves (2004), Barros (2005), Chiang-Ming et al. (2004), Wang et al., (2006a), Barros and Santos (2006), Cracolici et al. (2006), Reynolds and Thompson (2007), Gunjan (2007), Chen (2007), and Botti et al. (2009).

The main factors that affect hotel efficiency according to several researchers are market competition (Aissa and Goaied, 2016; Chaabouni, 2019; Huang et al., 2012), geographic location (Anderson et al., 1999; Lado-Sestayo and Fernandez-Castro, 2019; Oliveira et al., 2013; Parte-Esteban and Alberca-Oliver, 2015), the existence of international attractions (Aissa and Goaied, 2016; Huang et al., 2012), the ownership type (independent vs. chains, public vs. private) (Barros, 2004, 2005a; Barros and Dieke, 2008; Chang, 2003; Chen, 2002, 2007; Chen and Huang, 2001; Chiang-Ming et al., 2004; He, 2003; Luo et al., 2014; Pien, 2001; Wang, 2002), and not surprisingly hotel management (Baker and Riley, 1994; Brotherton and Mooney, 1992; Donaghy et al., 1995; Hwang and Chang, 2003). Other factors include the source of customers (Hwang and Chang, 2003), hotel size (Anderson et al., 1999; Sanjeev, 2007; Neves and Lourenco, 2009; Parte-Esteban and Alberca-Oliver, 2015), the education and payment levels of employees (Huang et al., 2012), hotel class (Corne, 2015), agglomeration and urbanization (Lado-Sestayo and Fernandez-Castro, 2019), the political hierarchy and tourism dependence (Luo et al., 2014), past profitability (Baik et al., 2012), resource allocation (Chiang, 2006), as well as trade openness (Huang et al., 2012) and climate change (Chaabouni, 2019). Furthermore, Yan (1997), along with Chen (2002) and Yang and Lu (2006) conclude that to increase hotel efficiency, hotels must achieve more optimal uses of inputs, such as rooms, and staff numbers in the accommodation and catering departments.

Lastly, acknowledgment and understanding of the underlying structures of tourism are necessary for sustainable tourism according to Bramwell and Lane, 1993 In a relevant study, Miller (2001) focused on the development of indicators to measure the movement of the tourism product at a company/resort level towards a position of greater or lesser sustainability. The results show considerable disagreement over sustainability and where the borders of the concept exist. In addition, the research identified contrasting views on the use of qualitative versus quantitative indicators and the role that consumer pressure can play. Moreover, using a modified Delphi technique, Choi and Sirakaya (2006), developed indicators to measure community tourism development (CTD) within a sustainable framework.

Finally, a thorough and deep analysis of the classification and the different performance variables and measures in the hotel industry has been performed by Pnevmatikoudi and Stavrinoudis (2016). They address the lack of hotel performance indicators that capture the multidimensional nature of the subject by introducing a novel classification after analyzing the content of seventy-nine scientific papers. This study benefits from Pnevmatikoudi and Stavrinoudis's (2016) study by using several hotel performance indicators identified by the two researchers.

3.4 Hotel star classification

This section of the literature review aims to present the texture and differences between hotel star categories (3, 4, or 5-star hotels), how these signs are anticipated and interpreted by hotel guests and whether an internationally common understanding of these signs holds.

According to specific criteria, hotels can be classified according to their overall services and quality. The classification system is a good indicator of price and, it is assumed to be also one of quality as it equates more quality with luxury and higher price (Israeli, 2002). The hotel classification system is generally producer-driven rather than customer-driven (Briggs et al., 2007). It is employed to serve customers who could conveniently raise or lower their expectations of hotel attributes based on the rating results. The most common rating system is Star and Diamond, which range from one to five, with higher values indicating better quality. Internationally, the star system is usually more employed. Under this system of classification, hotels receive up to five stars according to the minimum technical requirements relating to their hotel facilities, communication areas, guest areas, staff areas, and general services offered. Terms such as ¨hotel rating¨ ¨hotel grading¨ ¨hotel classification¨ and ¨hotel segment¨ are used interchangeably to distinguish hotels for their price, service, and facility levels (Cser and Ohuchi, 2008).

The origin of this rating system was developed gradually during the 1950s and 1960s, by the International Union of Official Tourist Organizations (IUOTO, the predecessor of the World Tourism Organization and the UNWTO today), as one of the mechanisms to face the growing tourism activity around the world after World War II and to advance the contribution of tourism to world economy. The result of this historical debate is what is known today as the formal hotel classification based on the "five stars" ranking (Talias, 2018).

A multitude of organizations are involved in the rating of hotels. These organizations can be divided into two categories, national and independent bodies. Consequently, hotel star rating is used to classify hotels according to their quality in accordance to laws approved by national or local governments, or by applying criteria established by independent organizations, hotel associations, national consumer travel organizations, guidebooks, travel websites, and volunteer organizations (Denizci et al., 2010). The reason behind governments' involvement in hotel rating lies in the fact that tourist lodging has a major impact on the tourist experience and thus its quality must be managed. The rise of online feedback on social media, however, shifts the focus to the consumer and, therefore the focus of hoteliers to these platforms.

Quality certifications based on the adoption of the ISO 9000 standards were introduced by the International Organization for Standardization (ISO) in 1987. A model of quality assurance is proposed to rationalize quality issues in contractual business-to-business relations and establish a quality system (Barnes, 1998; Conti, 1999; Zhu and Scheuermann, 1999; Tsekouras et al., 2002; van der Wiele, 2002; van der Wiele et al., 2006).

Moreover, one can consider hotel branding as an important element that communicates a certain level of quality to the customer and creates value and guest loyalty (O'Neill and Mattila, 2010). Even if today the brand is not yet one of the most considered attributes in the customer purchasing process (Akan, 1995; Callan and Bowman, 2000; Yesawich, Pepperdine, et al., 2004), the situation is changing due to the development of leading brands competition in the same location. This phenomenon may increase the importance and influence of brands on travelers purchasing behavior (O’Cass and Grace, 2004).

The star rating system may reflect the comfort level and range of services (Foris, 2011), but also suffers from discrepancies. Even though the star rating is based on objective criteria such as infrastructure, services, amenities, and the sizes of the rooms, and common spaces, variability exists between and even within countries. Consequently, on an international level, there is no common standard concerning what a hotel from each category should provide. The system for classifying hotels is different in each country and even within a country due to local regulations. Such an example is Spain where the autonomous governments are empowered to legislate and use different criteria to assign stars to the hotels.

Many European nations led by the United Kingdom, Germany, and Switzerland have established a coherent hotel classification system within their countries by appointing government agencies or private organizations to oversee drawing up the regulations (Cser and Ohuchi, 2008). According to Rhee and Yang (2015), Asian countries are following Europe’s footsteps through diversified approaches. US on the other hand, lacks a uniform system and does not maintain a homogenic hotel stratification code, but instead employs the rating of two private rating agencies (Forbes Travel Guide Hotel Ratings, formerly known as Mobil Travel Guide Hotel Ratings) and AAA (American Automobile Association) Diamond ratings. AAA's rating system relies on the 5 diamond-tier system, with the greatest number of diamonds implying the highest classification level (see among others Su and Sun, 2007), but supporting only the top three levels (five and four stars and recommended) plus two extra designations for their potential quality (soon-to-be-rated and editors' pick) (McCarthy et al., 2010).

Despite all these discrepancies, the star rating system remains the prevalent system regarding hotel classification categories, quality assessment, and choice (Dioko et al., 2013; Fang et al., 2016; Núnez-Serrano et al., 2014). Callan (1998), first focused on exploring the relationship between the attributes identified by consumers when they select a hotel, and those provided by the different categories. Consumer preferences are based on the perceived importance of the service quality (staff service quality, room quality, general amenities, business services, value, security, and international direct dial (IDD) facilities) (Choi and Chu, 1999). Except for the attribute categories of security and IDD facilities, a higher classification level leads to a higher attribute of value (Choi and Chu, 1999). For high-tariff hotels, staff service quality and room quality were the top two critical attributes, whereas, for medium-tariff hotels, security and then, room quality were the two most salient features (Choi and Chu, 1999). Finally, customers may express lower satisfaction for lower-ranking hotels but not be able to discern satisfaction between the 5- and 4-star hotels (Qu et al., 2000).

The importance of hotel features varies, depending on the type of amenity judged and the intrinsic characteristics of the traveler. One consistent finding is that the category of clean/comfortable room is the most important attribute regardless (Knutson, 1988). According to Minazzi (2010), a customer's perception of service quality is the result of the comparison between expectations and experiences (Grönroos, 2000; Zeithaml et al., 2006). Research demonstrates that customer satisfaction is not linked to a specific quality category but depends on the hotel's ability to meet customer expectations (Lopez Fernández et al., 2004). The attributes contributing to the significant differences in customer satisfaction between the higher-star hotels (5- and 4-star) and the lower-star hotels (3-star) are staff service, room features, and value for money, in order of importance. Nasution and Mavondo (2008) collated customers' perspectives with hotel managers to evaluate three distinctive hotel attributes (reputation for quality, value for money, and prestige) of Indonesian hotels belonging to three different hotel classifications: prime (equivalent to 5-star), standard (4- and 3-star), and budget (2- and 1-star). Accordingly, the managers' assessment was more optimistic than the customers' for all the attributes across all the hotel segments. Significant differences between expectations, perceptions, and the various hotel categories may exist, therefore the ranking of the groups does not always correspond with the categories (Fernandez and Bedia, 2004). In sum, the definition of quality is related to the satisfaction of client expectations and necessities and therefore, it is not possible to recognize quality only in luxury services but rather in all those that respond to what the consumer seeks (Parasuraman et al., 1985).

The existence of hotel classification systems is well-established throughout the world, however, research projects that focus on analyzing these systems are scarce. (Callan, 1998, 1999; Israeli and Uriely, 2000; Israeli, 2002). Nevertheless, a few studies demonstrate that the hotel classification in the star category is an indicator of price rather than quality (Israeli and Uriely, 2000; Israeli, 2002, Danziger et al., 2004). From the customer's point of view, price, star category, and brand were solid predictors of expectations regarding room price (Israeli, 2002; Danziger et al. 2006) and performance in Israeli hotels (Israeli, 2002). Therefore, when a customer pays a higher price to stay at an upper-class hotel, it is expected to be more demanding and have higher expectations (Fernández and Bedia, 2005; Fernandez Barcalà et al., 2009; Davutyan, 2007).

The star ranking system has faced a lot of criticism and not all research confirms the relation between the star rating classification system and quality. Its reliability as a measure of quality has been questioned by studies such as that of Núnez-Serrano et al. (2014) who analyzed 7783 hotels from the Official Guide to Hotels in Spain (OGHS). Moreover, a qualitative study conducted by Torres et al. (2014) suggests that hotel rating systems might become obsolete in the future, according to the opinion of general managers. Another criticism is that the criteria used to determine hotel quality are sometimes outdated. Martin-Fuentes (2016), studying more than 14.000 hotels in 100 cities around the world, tried to measure whether the classification system of hotel rating could meet the satisfaction levels. Despite the overall differences, a relationship between satisfaction rate and scores awarded by the two most important websites (Booking and TripAdvisor) does exist (Bulchand-Gidumal et al., 2011).

Despite the criticism, the higher star rating can be considered as being an indicator of higher quality (Abrate et al., 2011) and can be useful in reducing the adverse effects of asymmetric information (Martin-Fuentes, 2016; Nicolau and Sellers, 2010; Ogüt and Tas, 2012). Finally, star rating helps distinguish the factors that influence consumer purchasing decisions according to hotel star classification (Tanford et al., 2012).

The future of the star classification remains unknown. Online tourism platforms will continue to proliferate and dominate when it comes to customer decision-making. Consequently, traditional hotel rating systems are expected to diminish in importance for the sector’s stakeholders even further, unless they find a way to integrate themselves into the current internet-dominated reality. Therefore, customers' opinions via online Word of Mouth will continue to be at the center stage, and expert opinion will fall further into the background. This will in turn force hotels to monitor and to respond to consumer trends with more speed and efficiency, since guest opinion becomes the sole determinant of a hotel's quality (Hensens, 2015).

3.5 TripAdvisor and hotel performance

A travel platform is a digital ecosystem in the form of an online marketplace, that provides various services and resources related to travel and tourism. Travel platforms serve as intermediaries connecting travelers, service providers, and businesses within the travel industry. These platforms offer a wide range of tools, information, and services to facilitate travel planning, booking, and experiences.

As of the end of 2016, the following seven travel platforms were widely recognized and considered to be among the most famous and popular on the internet. The list included:

1. Booking.com is a well-known online travel agency that offers bookings for hotels, flights, rental cars, and other travel services. It is popular for its extensive selection of accommodations and its user-friendly interface.
2. Expedia is another major online travel agency that provides a wide range of travel services, including flights, hotels, vacation packages, and activities. It operates various brands, including Hotels.com and Vrbo.
3. TripAdvisor is a platform where travelers can read reviews, compare prices, and book accommodations, flights, and other travel services. It's known for its user-generated reviews and ratings.
4. Airbnb is a popular platform that connects travelers with hosts offering unique accommodations, such as homes, apartments, and other lodging options. It's known for promoting "home-sharing" experiences.
5. Agoda is an online travel booking platform that focuses on accommodations, particularly in Asia. It offers a wide range of hotels, resorts, and other lodging options.
6. Hotels.com is a subsidiary of Expedia Group and specializes in hotel bookings. The platform offers a loyalty program where users can earn rewards for booking nights.
7. Priceline is an online travel agency that offers discounted hotel bookings, flights, rental cars, and vacation packages through its "Name Your Own Price" feature and other deals.

TripAdvisor is more of a review and information platform where travelers can share and gather insights, while Booking.com and Expedia.com are comprehensive booking engines for accommodations and travel services including flights, vacation packages, and activities. As TripAdvisor focuses mainly on reviews and operates on a less extended scope as a hotel booking platform, was assessed as more appropriate to be the provider of ratings and comments on hotels than booking.com or expedia.com for this study. Moreover, Booking.com and Expedia.com use algorithms to rank and present hotels while a user is searching. The latter is either the result of a direct, paid by the hotel promotion effort or an indirect boost by Booking.com and Expedia.com to compensate hotels that offer a higher number of overnight stays to the platform yearly round than others.

Online guest feedback on social media platforms was initially the subject of much discussion and controversy. TripAdvisor, as the largest and fastest-growing platform, was at the epicenter of the debate. Research conducted by Hensens et al. (2010), suggests that the focus of guest reviews and ratings on TripAdvisor is different from the focus of most conventional rating criteria. Whereas most of the comments made in TripAdvisor reviews focus on service quality, most conventional rating systems tend to focus primarily on objective tangible criteria such as the availability and size of facilities and services. Occasionally TripAdvisor's criteria include subjective measures such as cleanliness and state of maintenance, but rarely on service quality (Hensens et al., 2010). Although TripAdvisor's methods were subject to intense criticism (see for instance: Elliott, 2009; Frommer, 2009; O'Neill, 2009), its growth has led to an acceptance by the hotel industry and the travel trade alike.

Since several hotel managers are willing to manipulate their ratings, the problem becomes cumbersome (Gössling et al., 2018). Schuckert et al. (2016), suggest that one in five reviews in TripAdvisor is suspicious. Hotels' manipulation tactics also differ under the threat of a common enemy such as AirBnB (Nie et al., 2022). Nevertheless, platforms where only verified customers can post reviews seem to curb the problem (Mayzlin et al., 2014; Xu et al., 2020).

Consequently, several researchers have studied the associations between eWOM and hotel performance with the use of big data from online platforms (Aakash and Gupta, 2022). In one of these studies, the TripAdvisor rating was found to be the best predictor of hotel performance (Kim and Park, 2017).

Tuominen (2011) first explored the correlation between TripAdvisor ratings and hotel performance in Scandinavia. He found that the number of reviews, the average review rating, and customer recommendations were positively correlated with the average daily rate, the revenue per available room, and hotel occupancy. TripAdvisor ranking, on the other hand, was not correlated with hotel performance. Yang et al. (2018) offers a meta-analysis of eWOM and hotel performance.

Other researchers focused on the effects of the managerial response on customers' comments and how these responses can affect hotel performance through the increased ratings and volume of consumer eWoM (Xie et al., 2016), as well as how these are moderated by hotel class (Xie et al., 2017). The time, volume and length of the response also appear to play a role (Xie et al., 2017b). Overall, digital marketing strategies may influence both the volume and ratings of online reviews and, therefore, hotel performance (De Pelsmacker et al., 2018).

Contrary to previous research which used TripAdvisor data mainly as predictors of hotel performance, the concrete study adds a new dimension by examining the impact of hotel financial measures on the generation of positive word-of-mouth and vice versa. The study investigates whether long-term investment by hotels as well as day-to-day expenditure influences the extent to which guests rate their experiences highly on online reviewing platforms.

Chapter 4. Data and variables description

The primary objective of this thesis is to identify the connections between customer satisfaction ratings of Greek hotels and their financial performance. Greece has been chosen as the primary research ground for this study due to its prominent status as a major international tourist destination. The country boasts a significant number of hotels operating within its borders. Moreover, as previously established, tourism plays a pivotal role in contributing to the overall economy of Greece (De Vita and Kyaw, 2016). In our analysis of customer satisfaction, we rely on data extracted from the TripAdvisor platform, encompassing a comprehensive range of customer evaluations, which we subsequently employ to construct our rating metrics. Concurrently, financial data are sourced from the year-end financial statements of these hotels.

The requisite data of Greek hotels along with their characteristics and attributes were readily accessible through various means, including online sources and prominent institutional bodies. The sources used encompass entities such as the Association of Greek Tourism Enterprises (SETE), its research division INSETE, the Hellenic Chamber of Hotels, the Hellenic Statistical Authority, the Bank of Greece, as well as reputable private consultancy firms specializing in the Greek tourism sector.

Initially, a comprehensive dataset was assessed, including all operational hotels within Greece, provided by the Hellenic Chambers of Hotels. The dataset underwent categorization based on the hotels' star ratings and geographical locations. Additionally, the Hellenic Chambers of Hotels furnished website URLs for hotels possessing an online presence.

To ensure the integrity of the dataset, hotels falling within the 1 and 2-star categories were intentionally excluded. The concrete decision was imposed by the observation that a significant percentage of these hotels either did not report any financial data or reported data in a non-standardized manner. Furthermore, a substantial portion of the 1 and 2-star hotels operated with fewer than 20 rooms, and a considerable proportion of them lacked websites or a presence on TripAdvisor.

The resultant dataset, representing "3, 4, and 5-star hotels operating in Greece," underwent a meticulous quantitative refinement process. This entailed scrutinizing the availability of financial data for each hotel throughout each year of the entire 2007 to 2016 period, as well as confirming the presence of both a website and a listing on TripAdvisor. Several hotels were excluded from this refined "3, 4, 5-star hotels sample" due to various other factors rendering them unsuitable for research.

Ultimately, after careful consideration of the criteria, a substantial and statistically robust sample emerged. This sample encompassed 3,147 hotel units, constituting 72.7% of all 3, 4, and 5-star hotels in Greece as of 2016-2017 figures (see Appendix A.). Importantly, this sample was geographically distributed in close alignment with regional patterns of tourist activity across the diverse expanse of the Greek territory.

4.1 Dependent variable: Hotels' ratings on TripAdvisor

Customer satisfaction is arguably an important performance measure in the hotel industry. Researchers who have dealt with the measurement of hotel performance categorize the satisfaction element into two groups: "Customer satisfaction" and "Employee satisfaction".

Garrigos-Simon, Palacios-Marques, and Narangajavana (2005) incorporate the above groups into the generalized sense of "Stakeholders satisfaction". On the other hand, Chen, Hsu, and Tzeng (2011); Bolat and Yilmaz (2009); and Espino-Rodriguez and Padron Robaina (2005) distinguish between "Customer satisfaction indicators" and "Employee satisfaction indicators". In the concrete thesis, customer satisfaction is the dependent variable of the regression model. Customer satisfaction is expressed as a continuous variable which ranges from 1 to 5. This corresponds to the customers' score on the TripAdvisor platform, after leaving the hotel. More precisely, this variable takes the form of the yearly average rating of each hotel on the TripAdvisor platform, for the period stretching from 2007 to 2016.

To extract the unstructured rating data from TripAdvisor's website, Python version 3 and particularly the 'Scrapy' framework were used. Python is an interpreted, high-level, and general-purpose programming language. Python supports multiple programming paradigms such as structured (particularly procedural), object-oriented, and functional. '*Scrapy*' is an open-source application, generally used for fast web crawling and scraping. Originally designed for web scraping, it can also be used to extract data using an application programming interface (API) or as a general-purpose web crawler '*Scrapy'* project architecture is built around "spiders", which are self-contained crawlers operating under a given set of instructions. Following the spirit of other "don't repeat yourself frameworks" '*Scrapy'* builds and scales large crawling projects. '*Scrapy'* crawls websites using instructions defined by the programmer. It crawls websites by matching domain names (in this study the hotels' names on TripAdvisor and/or the hotels' URLs) and extracting structured series (ratings) from unstructured data found on the preselected hotels' webpages. '*Scrapy'* can work asynchronously and can simultaneously extract data from multiple pages. The '*Scrapy'* framework defines a spider class that incorporates all the user's information (such as a list of URL names) as well as structured data in "words" to be extracted, either specified as CSS selector or Xpath variables. Selectors or Xpath variables are incorporated in the Python script. Spiders generally return the extracted data in a Python dictionary format, that can be stored in different file formats like CSV and, XLS, The following figure illustrates exactly how the '*scrapy'* framework works.

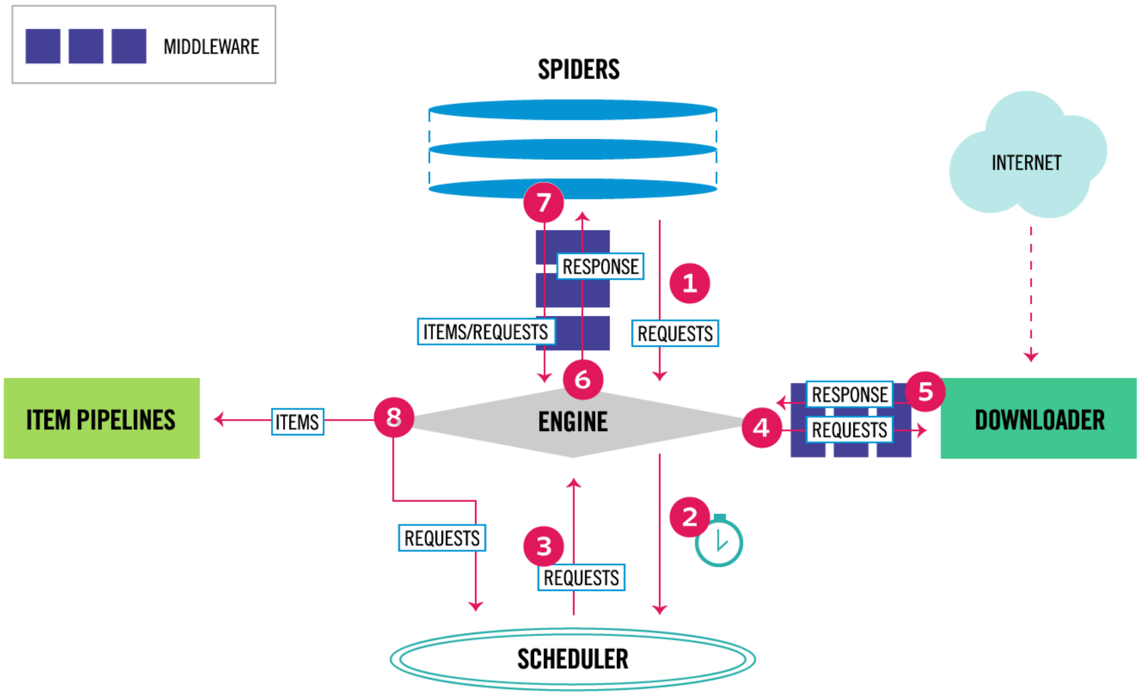


Figure 4.1. Architecture of the Scrapy framework.

Source: https://doc.scrapy.org/en/latest/topics/architecture.html

Considering the present thesis, '*Scrapy'* was programmed to extract from TripAdvisor data, only from the English language reviews. '*Scrapy'* spider, collected information for the pre-defined hotel names, referring to TripAdvisor's username, review date, overall rating, user location, review title, review full text, date of stay, and individual star rating for sleep quality, cleanliness, service, location, rooms, value, business service, check-in/front desk, having a pool and being family-friendly. Two sample reviews from TripAdvisor are shown in Figure 4.2.

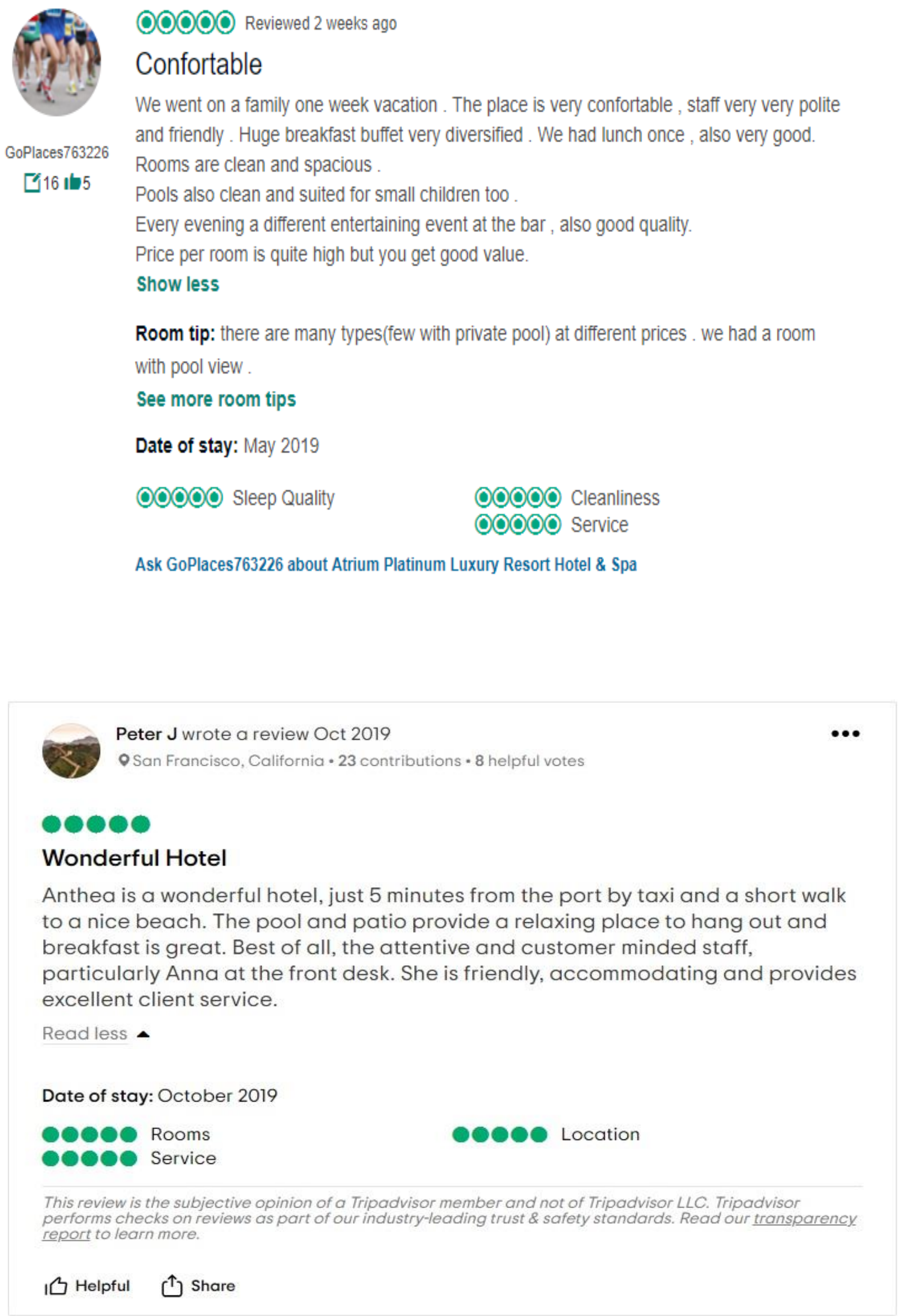


Figure 4.2. A typical TripAdvisor rating.

For the creation of the dependent variable time series, the average overall rating for each year of the period under consideration was used, reflecting the overall experience of the clients and their satisfaction degree, expressed as a number from 1 to 5. In the concrete escalation, 1 represents the minimum satisfaction experienced by the client, while 5 is the maximum rating corresponding to an amazing experience. The metric incorporates the overall perception of the client for a hotel, blending individual opinions for each one of several subcategories of satisfaction, such as location, cleanness, service, etc. The initial hotels' dataset size, was restricted to complete records of data on the TripAdvisor platform for the years under consideration.

4.2 Explanatory variables: financial characteristics

This thesis aims to correlate the hotels' ratings on TripAdvisor, representing customers' satisfaction, with their end-of-the- economic year reported financial data. Considering the need for reliable and homogenized data the initial significantly larger sample was decreased to comply with the following:

- ✓ Hotels' legal entities to report financial data in compliance with the Greek Accounting Rules in a comprehensible manner for all 10 fiscal years from 2007 to 2016.
- ✓ The companies to have income from operating the hotel and not from renting the premises or other sources.
- ✓ The entities to have continuous operation for the 2007 – 2016 period.
- ✓ Mergers and acquisitions cases were excluded from the sample.
- ✓ Legal entities operating more than one hotel, if separate data for each one was not provided, were excluded.

The final sample includes 554 Greek hotels falling under 3-star, 4-star, and 5-star levels for the period 2007 to 2016 (5.540 observations), representing 12% of the 3-star, 4-star, and 5-star hotel segment of Greece at the end of 2016. For each hotel, balance sheet information on fixed assets, total assets, equity and reserves, and income statement information on revenue and administrative and sales expenses were obtained. Furthermore, information on whether the hotel has a pool, and whether it defines or not itself as a family hotel was collected through TripAdvisor. Through the National Statistical Authority of Greece, national-level survey estimates of yearly average spending by international tourists per night and per visit were sourced. Finally, the number of beds per hotel was collected by the Greek Hotel Association.

Considering the regression analysis, the choice of financial variables in the present thesis follows Pnevmatikoudi and Stavrinoudis (2016) who, in a systematic review, identified 10 categories of financial variables previously applied in hotel studies. According to them:

"A study ought to concurrently incorporate multiple indicators, both financial and non-financial. By doing so, researchers can comprehensively understand the multifaceted aspects of hotel performance in their analyses. Assessing hotel performance is an intricate process with multiple dimensions, offering various angles of approach. This assessment is critically important for hotel businesses."

In addition to the latter, Pnevmatikoudi and Stavrinoudis (2016) mentioned that “in the global scientific literature, a vast array of variables, measures, and indicators have been identified and documented for the evaluation and measurement of hotel performance. Within this context, the financial aspect was initially explored, being the most recognized dimension of performance."

The sample was stratified into different categories to scout for more detailed insights. Firstly, hotels are divided into categories according to their star classification (3,4,5-star). Secondly, they are divided per the number of beds they operate (Small up to 50% of the distribution, and Large over 50% of the distribution) Thirdly, hotels are bundled in accordance with the amount of total assets they possess (once again, Small up to 50% of the distribution, and Large over 50% of the distribution). Stratification is conducted to control the differentiation in the behaviors of various hotels' aggregations. Meaningful results may emerge out of this stratification analysis resulting in important implications for hoteling entrepreneurs, hotel managers, and tourism policy makers. Whether satisfaction is higher in 3-, 4- or 5-star hotels, in large hotels with an important magnitude of total or fixed assets and a higher number of beds, or on smaller more versatile, and agile is a piece of crucial information to have when scheduling hotel investments, hoteling business transformation and hotel operation.

Moving to the selection of the explanatory variable, the first to be included into the model is logged Total Assets. "Total assets" is a magnitude measure capturing the financial size of the infrastructure and operational assets, involved in the offered by the hotel experience. It incorporates the total financial commitment of the business to its effort to provide services. The log-transformed data is used in the model as the concrete type of transformation offers several advantages. One assumption of OLS regression is that there is a linear relationship between the independent and dependent variables. In such cases, taking the logarithm can linearize the relationship. Continuously, "total assets squared and logged" are added to the equation.

The next variable to be added to the model is, the logged ratio of turnover per bed. The latter is an efficiency and activity tracking ratio indicating how much money each bed of the hotel generates per year. The more the volume per bed produced, the better the hotel uses its resources. The ratio, known in the hotel industry as RevPAB (Revenue Per Available Bed) is close to the RevPAR (Revenue Per Available Room) ratio approach which is widely used. The difference is that RevPAB accounts even for the use of a room from only one person. As a substantial number of hotels in high touristic destinations in Greece operate relying on bed selling contracts instead of rooms and evaluate their performance relying on bed occupancy terms, it was preferred to the RevPAR ratio for the purpose of this study.

The next variable is the turnover per fixed assets ratio. The ratio is a utilization metric of fixed assets, showing their ability to produce sales. The higher the ratio, the better the utilization of each euro invested in fixed assets, such as premises and recreational facilities. Common examples of fixed assets that provide long-term economic benefits include land, buildings and facilities, machinery, and equipment. Therefore, the fixed asset turnover ratio determines if a company's purchases of fixed assets – i.e. capital expenditures (Capex) – are being used effectively or not. Given how costly fixed asset creation or purchase can be and how– Capex decisions must be made carefully, the importance of the ratio is pronounced.

Rating, in its log transformation, is the next variable to count for. It is expressed as the hotel's yearly average rating on TripAdvisor. It resumes the overall perception of satisfaction a client received from his stay in a hotel. As it is already mentioned, logged values offer a series of advantages to the model interpretation justifying the use of such a transformation.

Administrative expenses are the expenses that are not attributable to the direct production or delivery of the services of a hotel. "Administrative expenses to turnover" is an efficiency ratio that measures how well a company can manage its overheads and generate sales. In other words, the administrative expenses to sales ratio measures how much administrative expenses are needed to generate a euro of turnover. Administrative expenses usually incorporate costs such as higher management compensation, HR and procurement costs as well as accounting and finance department costs. The ratio (winsorized at the 1/99 percentile) is added as a variable to account for short-term investment in guest experience. The hypothesis behind this, states, that better management, including people and procedures, may cost more per euro of sales. On the other hand, better management results in an upgraded guest experience, as the hotels operates smoothly, from reservations to all other services provided, avoiding unpleasant surprises for the visitor. In a well-managed hotel, inventory is always refreshed, reservations and appointments for services like sporting facilities and spa treatments are always on time, and room cleaning, and hygiene are at a state-of-the-art level. Furthermore, personnel turnover is minimized, leaving space for learning curves and the creation of added value services to the clientele, like out-of-premises destination experiences. Winsorized data are used to limit extreme values in the dataset. Outliers outside the 1st and 99th percentiles were replaced by the last closest values. It is well known that an extreme value can disproportionately influence the estimated regression line, leading to a model that doesn't represent the central tendency of the data well. By winsorizing the data, the undue influence of outliers is mitigated. Moreover, the linear model assumes that the residuals are normally distributed. Outliers can disturb this normal distribution. By limiting extreme values, winsorizing technique, can make the distribution of residuals more normal. In addition to that, winsorizing can help in addressing the issue of heteroscedasticity by stabilizing the variance across the dataset and increase the robustness of the results.

The next variable in use is the ratio of sales expenses to turnover, winzorized at the 1/99 percentile. The ratio is constructed by dividing the costs of selling by the total value of sales – and then multiplying the result by 100. Most of the hotels of the sample in "Sales Expenses", incorporate mainly sales salaries and commissions created by different sources of clientele generators, such as tour operators and agencies as well as online channels such as Booking.com, Expedia.com, etc. Moreover, internet presence, channelling, and booking tools costs, promotional initiatives, online advertising like Google AdWords, and all forms of "presents" and pro bono services to the clients are included in the numerator. The sales expenses/turnover ratio is an efficiency ratio. The lower the ratio the more efficient the hotel operates, in terms of sales production. Seen from a different angle, the ratio can be read as a short-term investment of the hotel to the overall satisfaction of the client. From the moment the client starts researching a hotel in a destination, to the booking decision and the "element of positively being surprised" on the spot, due to treatments and unexpected pro bono services, sales operation can sculpt the overall satisfaction perception of the client. Under this angle, more sales expenses may result in higher satisfaction.

The next variable to be added to the regression model is the first difference (D) or the year-on-year change of Fixed Assets logged. Taking the first difference of a time series is the process of subtracting the value of the series at the previous year from its value at the current year. The goal is to capture the change in the series from one period to the next. Generally, this method is commonly used to transform a non-stationary time series into a stationary one, making it more suitable for OLS regression. In our case, on top of the latter, the amount invested in renovations, refurbishments, or expansions emerges from the deductions. As renovation and refurbishment or expansion of the services provided (for example the addition of a SPA facility) is an add-up to satisfaction, it is of interest to this study.

Next, the first difference (d) or the year-on-year change logged of the administrative as well as sales expenses are added to the regression. The assumption is that a year-on-year increase in one or both metrics under consideration may result in higher satisfaction. More money invested in more efficient operations will have an impact on the perceived satisfaction level of the guest.

The next add-in component of the model is a macro-sector variable, the average amount a visitor spends per travel in Greece, calculated on a yearly basis. The variable is added in attempting to smooth distortion from changes in spending behavior, related to general economic conditions and/or other sector induced reasons such as destinations' competition air traveling prices, etc. Moreover, the average amount a visitor spends per overnight stay in Greece yearly is added to the model, accounting for the same reasons and on top of that to eliminate inherent spending behaviour interpretants as well.

Finally, two dummy variables, the hotel "having or not a pool", and "being family friendly" (mined from TripAdvisor) were added to the model to enhance its explanatory force and to check for statistical significance.

In a nutshell, as the present thesis focuses on hotel investment, both long- and short-term, it draws on variables from three categories, namely, sales (Kim et al., 2013), cost (Assaf and Barros, 2011), and growth (Han, 2012). A sales variable, revenue/beds, is first included as one of the most popular measures of hotel performance (Pnevmatikoudi and Stavrinoudis, 2016). Fixed and total assets are used as a proxy for long-term investment. Sales expenses and administration, as a proxy for short-term investment in guest experience. Lastly, among the primary independent variables, measures of change (growth) are added to the investment variables accounting for changes from one year to the next. The set of variables used in the regression analysis of the thesis is presented in the following table:

Table 4.1. Variables used in the regression analysis.

Variable name	*Variable abbreviation*	*Formula*	*Description/indication/utilization*
Rating	Rating	Hotel's yearly average rating on TripAdvisor	Dependent variable
Stars	3,4,5 star	N/A	Indicates hotel's star category.
Number of Beds (stratification)	BEDS	number of beds per hotel	Size metric in terms of population
Total Assets (stratification)	TA	Total assets extracted from the balance sheet	Size metric in terms of assets used
Efficiency (stratification)	Bc_h	SFA computation	Indicates how efficiently a hotel uses its inputs to create output
Own capital	OC	Equity plus reserves	the addition of equity injected and accumulated reserves
Fixed Assets	FA	The total amount invested in fixed assets provided by balance sheets (undepreciated values)	A metric of size and luxury
Turnover	T	The yearly amount of sales	A metric of activity and efficiency
Total Assets	Lta	Total Assets (log)	size
Fixed Assets (Δln)	lFA_d	Fixed Assets year-on-year log change	A metric capturing expansion and/or improvement in terms of refurbishment or renovation
Turnover / Beds (ln)	TurnBeds_l	turnover/number of beds (log)	A metric of activity. Indicates how much money is earned per bed yearly for each hotel
Number of Ratings (ln)	lRatingcount	Number of ratings (log)	How many ratings does a hotel receive in a given year?
Sale Expenses / Turnover	SalesTI_w	Sales Expenses / Turnover (winsorised 1/99)	Cost efficiency indicator (includes promotion, room amenities, and other costs to sales)
Admin Expenses / Turnover	CTI_w	Administrative Expenses / Turnover (winsorised 1/99)	Cost efficiency ratio (includes administrative and managerial costs to sales)

Total Assets	LTAsq	Total Assets (log)	Size metric
Turnover / Fixed Assets	TurnFA_w	turnover/fixed assets (winsorised	fixed assets utilization metric
Pool dummy	Pool	Y/N	The qualitative variable indicates whether a hotel has or not a pool
Family Friendly	FamFriendly	Y/N	The qualitative variable indicates whether a hotel self-identifies as
Average spending per person	AvSpendpp	The average amount a visitor spends per travel in Greece is calculated on a yearly basis	Macro- Sector average
Average spending per night	AvSpendpn	The average amount a visitor spends per overnight stay in Greece is calculated on a yearly basis	Macro- Sector average
Administrative Expenses / Turnover year-on-year log change	CTI_dl	Administrative Expenses / Turnover year-on-year log change	The year-on-year increase in the admin costs
Sale Expenses / Turnover (Δln)	SalesTI_dl	Sales Expenses / Turnover year-on-year log change	The year-on-year increase in the sales costs

Notes: L=log; sq=squared; d=first difference; w=winsorised at the 1/99 percentile

The following table presents the Pearson correlation coefficients between the variables in use:

Table 4.2. Correlation matrix

	BEDS	*STARS*	*Rating*	*FIXEDASSETS*	*TURNOVER*	*lTA*	*lFA_d*	*TurnBeds_l*	*lRatingcount*	*SalesTI_w*	*CTI_w*	*lTAsq*	*TurnFA_w*	*Pool*	*FamFriendly*	*AvSpendpp*	*AvSpendpn*	*CTI_dl*	*SalesTI_dl*	*bc_h*
BEDS	1,000																			
STARS	0,409	1,000																		
Rating	-0,037	0,138	1,000																	
FIXEDASSETS	0,504	0,296	0,040	1,000																
TURNOVER	0,559	0,338	0,092	0,703	1,000															
lTA	0,708	0,549	0,038	0,585	0,673	1,000														
lFA_d	0,030	0,036	0,019	0,072	0,074	0,098	1,000													
TurnBeds_l	0,170	0,417	0,126	0,296	0,510	0,568	0,062	1,000												
lRatingcount	0,305	0,315	0,267	0,206	0,316	0,366	0,017	0,321	1,000											
SalesTI_w	-0,250	-0,181	0,062	-0,124	-0,208	-0,318	-0,019	-0,379	-0,116	1,000										
CTI_w	-0,139	-0,093	-0,020	-0,044	-0,136	-0,176	-0,021	-0,301	-0,124	0,344	1,000									
lTAsq	0,723	0,551	0,041	0,625	0,702	0,998	0,097	0,563	0,368	-0,312	-0,171	1,000								
TurnFA_w	-0,077	-0,023	-0,055	-0,127	-0,005	-0,187	-0,165	0,236	0,093	-0,101	-0,095	-0,186	1,000							
Pool	0,008	-0,063	-0,096	-0,044	-0,050	-0,004	0,008	0,006	-0,011	0,017	-0,004	-0,008	0,047	1,000						
FamFriendly	-0,022	-0,040	0,006	-0,010	-0,014	-0,023	-0,022	-0,009	-0,040	-0,023	0,000	-0,023	-0,001	0,087	1,000					
AvSpendpp	0,006	0,009	-0,051	-0,023	-0,049	-0,021	-0,046	-0,086	-0,221	0,076	0,030	-0,021	-0,078	0,001	-0,003	1,000				
AvSpendpn	-0,005	-0,004	0,043	-0,010	0,000	-0,006	-0,017	-0,008	0,045	-0,019	-0,015	-0,007	-0,004	-0,007	0,004	0,409	1,000			
CTI_dl	-0,013	-0,043	-0,031	-0,029	-0,034	-0,026	-0,003	-0,086	-0,025	0,055	0,242	-0,027	-0,023	-0,009	0,033	-0,072	-0,054	1,000		
SalesTI_dl	0,001	-0,022	-0,041	0,014	0,006	-0,005	0,016	-0,061	-0,027	0,148	0,128	-0,005	0,009	-0,007	0,000	-0,078	-0,094	0,258	1,000	
bc_h	0,167	0,087	0,024	-0,057	0,214	0,128	-0,006	0,414	0,198	-0,436	-0,623	0,120	0,421	0,043	0,051	-0,120	0,010	-0,136	-0,048	1

Table 4.2 presents the correlation matrix between the variables of the regression model and the linear relationship between the pairs of the variables. The value of the coefficient can range between -1 and 1 with a coefficient close to 1 implying a strong positive linear relationship, a coefficient close to -1 implying a strong negative linear relationship, and a coefficient close to 0 implying little to no linear relationship. From inspecting the table, one can conclude the following for the different pairs of variables:

1. BEDS and lTA pair present a correlation coefficient of 0.708 indicating a strong positive relationship. As the number of beds increases, the "total assets" of a hotel also tend to increase.

2. FIXED ASSETS and TURNOVER pair have a coefficient of 0.703, there's a strong positive relationship between fixed assets and turnover. Bigger and/or more luxurious hotels produce a higher turnover.

3. BEDS and TURNOVER pair: displays a correlation of 0.559 showing a moderate to strong positive relationship. It implies that hotels with more beds tend to have higher turnover. The interesting finding is that the correlation is not as strong as it was observed in the fixed assets and turnover correlation. This is happening since a lot of Greek hotels, with a higher number of beds, sell cheaper as they usually operate under tour operators' enforcements and low hotel standards. Another possible explanation could be that a non-linear relationship holds.

4. For the SalesTI_w and CTI_w pair a coefficient of 0.344 is calculated indicating, a moderate positive correlation between sales expenses/turnover and administrative expenses/turnover.

5. bc_h and CTI_w pair exhibits a correlation of -0.623 suggesting a strong negative relationship between these two variables. As the efficiency of a hotel increases, administrative expenses/turnover tends to decrease, and vice versa. The latter is expected as efficient hotels tend to decrease administrative costs over turnover.

6. TurnBeds_l and bc pair, showcase a positive correlation of 0.414 between turnover/beds and technical efficiency indicating a moderate relationship confirming that more efficient hotels tend to generate higher per bed turnover.

7. STARS and Rating pair presents an interesting finding. Even though one might expect a high correlation between the stars of a hotel and its rating on TripAdvisor, with 4- and 5-star hotels to be correlated with higher ratings, the correlation is only 0.138, suggesting a weak positive relationship. In any case, correlation does not always imply causation i.e., a high correlation between two variables does not mean that one causes the other. Moreover, the practical significance of a correlation depends on the context of the data. A more thorough analysis of the relationship between these two variables is presented in the part of the thesis, discussing the results of the OLS model.

8. BEDS and FIXEDASSETS have a coefficient of 0.504 suggesting a moderate positive correlation between the two variables. This implies that hotels with more beds also have a higher value of fixed assets, potentially because, larger hotels require more infrastructure or amenities.

9. STARS and TurnBeds_l exhibit a correlation of 0.417 indicating a moderate positive relationship. Hotels with more stars present higher values for the ratio turnover/beds.

10. Rating and lRatingcount showcases a coefficient of 0.267 implying a weak to a moderate positive relationship. Establishments with higher ratings might receive more ratings in general, perhaps because popular or well-liked establishments are more frequently rated by customers.

11. TURNOVER and lFA_d pair present a coefficient of 0.074, meaning that a very weak positive relationship between turnover and money invested in fixed assets on a year-to-year basis holds.

12. AvSpendpp and AvSpendpn with a coefficient of 0.409, suggests a positive relationship. It's conceivable that these two variables represent average spending per person for different contexts, and if spending is high in one category, it tends to be high in the other.

13. Finally, Rating and FIXEDASSETS pair have a correlation of 0.040 is very close to zero, suggesting almost no linear relationship between a hotel's rating and its fixed assets value in this dataset.

4.3 Explanatory variables: Technical Efficiency

In the context of Stochastic Frontier Analysis, a set of output variables against input variables can be used to construct a production frontier and measure the efficiency of hotels. In the present study outputs (current assets and turnover) are compared against inputs (fixed assets, own capital, and sales and administrative expenses) to determine how efficiently hotels of the data set are utilizing their resources to generate outcomes.

The first output variable selected is the total of current assets, calculated as the substruction of fixed assets out of total assets, reflecting the accumulation of other types of assets such as cash and cash equivalent, that can be used for the enhancement of operation, services, and soft infrastructure over the next year. Total Assets (TA) is provided as a direct value, from the balance sheet. Total Assets represent the total value of everything a hotel owns and controls that can be used for its operations. Fixed Assets (FA) provided from the balance sheet, is the total amount invested in fixed assets after accounting for depreciation. Fixed assets encompass property, plant, and equipment, which are essential for the hotel's day-to-day operations. It reflects the size of the infrastructure, including the building, machinery, and other non-liquid assets. Current Assets (CA) represent assets that are expected to be used up within a year. This can include cash, accounts receivable, inventory, and other short-term assets. In a hotel context, it may also represent prepaid expenses or short-term investments. All above mentioned are essential facilitators of the hotel operations and performance.

Turnover (T) is the second output variable, extracted directly from P & L accounts. Turnover reflects the total revenue or sales made by the hotel each year. It's a primary activity metric, measuring the hotel’s efficiency in attracting and serving guests.

Input variables used in the technical efficiency specification are (i) fixed assets, ii) own capital, iii) administrative expenses, and iv) sale expenses, all associated with each hotel "i" in year "t" of the data set.

Own Capital (OC) is the addition of Equity plus reserves. Own Capital is a representation of the hotel's internal financing. It indicates the total financial strength from both injected equity and accumulated reserves over time. It can serve as an indicator of the hotel's financial health and stability. Both metrics are collected from balance sheets.

Administrative Expenses (C) are collected directly from P & L accounts. They encompass costs related to the general management and administration of the hotel. They can include salaries of managerial staff, utilities for administrative offices, and other overhead costs. It serves as a measure of how much the hotel spends on its non-operational, and administrative functions.

Sales Expenses (Sales) are provided directly from P & L accounts. These are the costs directly associated with generating sales, such as marketing, advertising, and the cost of amenities or services offered to guests. It can indicate how much the hotel is investing in attracting customers and enhancing the guest experience.

Table 4.3. Variables used in the SFA analysis.

Variable name	*Variable abbreviation*	*Formula*	*Description/indication/utilization*
Total Assets	TA	The total assets of a company at the end of the economic year	Size metric in terms of assets used
Fixed Assets	FA	The total amount invested in fixed assets (after depreciation)	A metric of the size of the infrastructure
Current Assets (output)	CA	Total Assets – Fixed Assets	Current assets of the hotel at the end of each economic year
Turnover (output)	T	The yearly succeeded amount of sales	A metric of activity and efficiency
Own capital (input)	OC	Equity plus reserves	the addition of equity injected and accumulated reserves
Administrative expenses (input)	C	The total of administrative expenses	The year total of management and administration costs
Sales expenses (input)	Sales	The total of sales expenses	The year total of sales costs including promotion and services offered costs

4.4 Descriptive statistics

The present chapter presents the descriptive statistics of the full sample as well as statistics for various stratifications of the dataset under examination. A brief presentation of the metrics presented in the tables provided on the course of the present chapter follows: N is the number of observations in the sample, mean is t the average value of the variable, p50 is the median (or 50th percentile) value i.e., the middle value when all data points are arranged in order, sd is the standard deviation, which measures variability, min is the smallest value observed, max is the largest value observed. Medians are used to understand the typical hotel (especially if data is skewed), and standard deviations to see how spread out the data is. Min and max values provide the range of information. For dummy variables like Pool and FamFriendly, the mean represents the proportion of hotels with that feature (since 0 = No and 1 = Yes).

Table 4.4. Descriptive statistics for the full sample

	Observation	*Mean*	*Median*	*SD*	*Min*	*Max*
Number of Beds	5,530	222.65	137.00	244.10	13.00	1,444.00
Stars	5,540	3.79	4.00	0.70	3.00	5.00
Rating	5540	4,09	4,24	0,72	1	5
Fixed Assets	5,525	13,000,00	2,590,206	44,900,00	0.00	561,000,00
Turnover	5,540	3,011,824	941,593	6,524,246	-193,620	65,500,000
Total Assets (ln)	5,525	15.25	15.07	1.43	9.99	20.16
Fixed Assets	4,661	0.02	-0.02	0.27	-8.75	3.29
Turnover / Beds	5,439	8.87	8.81	1.08	1.39	13.39
Number of	5,540	2.42	2.40	1.49	0.00	6.50
Sale Expenses /	3,798	0.18	0.08	0.23	0.00	1.23
Admin Expenses	5,102	0.19	0.12	0.23	0.01	1.40
Total Assets	5,525	234.50	227.20	44.83	99.78	406.44
Turnover / Fixed	5,521	0.47	0.33	0.56	0.00	4.35
Pool dummy	6,193	0.28	0.00	0.45	0.00	1.00
Family Friendly	6,193	0.00	0.00	0.06	0.00	1.00
Average	4,288	607.62	639.47	47.53	514.10	653.31
Average	4,288	70.58	70.38	2.28	66.96	73.93
Cost / Income	4,230	0.02	0.00	0.61	-4.75	5.33
Sale Expenses /	3,033	0.02	0.00	0.60	-5.04	6.32

NOTES: The table reports summary statistics for the variables used in the analysis. “ln” denotes the natural logarithm, “Δln" denotes the logarithmic first difference. All monetary values are expressed in Euros.

A presentation and statistical interpretation of the above tabulated descriptive statistics for each of the variables follows:

Considering the "Number of Beds" the table presents that data includes observations for 553 hotels for 10 years each. These observations present the following characteristics:

- Mean: On average, sample hotels have around 222.65 beds.
- Median: Half of these hotels have 137 beds or fewer, and half have more.
- Standard Deviation: The number of beds varies by about 244.10 beds around the mean.
- Min: The smallest hotel in the sample has 13 beds.
- Max: The largest hotel in the sample has 1,444 beds.

The hotels in the above-mentioned sample range from 3 to 5 stars. Data on fixed assets is available for 5,525 hotels. On average, hotels have fixed assets valued at 13,000,000 euros. The middle hotel, when ranked by fixed assets, has assets valued at 2,590,206 euros. According to the standard deviation metric, the value of fixed assets varies significantly across hotels, with a standard deviation of 44,900,000. The range is considerable, with some hotels having no fixed assets (assets being depreciated) and others having as much as 561,000,000.

Turnover data shows that the average turnover of the sample's hotels is 3,011,824 euro and the median turnover is 941,593 euro. There's a widespread turnover value, indicated by the large standard deviation of 6,524,246 euros. The minimum turnover value is -193,620 and the highest turnover reported is 65,500,000 euro.

The year-on-year first difference of Fixed assets presents a slight increase of 0.02. The median value is -0.02, suggesting that more than half of the hotels have a decrease. The standard deviation is 0.27, indicating the amount of variability. Finally, changes in fixed assets range from a large decrease (-8.75) to an increase of 3.29.

TurnBeds_l (Log of the ratio of turnover/number of beds) indicates revenue efficiency per bed. The average logged ratio is 8.87. The median logged ratio is 8.81. A standard deviation of 1.08 shows moderate variability.

lRatingcount (Number of ratings, logged) shows that the average log-transformed rating count is 2.42. The median is close to the mean at 2.40, suggesting a somewhat symmetric distribution.

SalesTI_w (Sales Expenses / Turnover ratio, winsorized at 1/99: the average ratio is 0.18, meaning that for every unit of turnover, 0.18 is spent on sales expenses. The median value is 0.08, indicating skewness in the distribution.

CTI_w (Administrative Expenses / Turnover, winsorized at 1/99) indicates that on average, 0.19 of turnover is spent on administrative expenses.

LTAsq (Logged Total Assets squared) informs that the average squared log of total assets is 234.50.

TurnFA_w (Turnover/Fixed Assets, winsorized at 1/99) reveals that the average ratio is 0.47, indicating the turnover generated for each unit of fixed assets.

FamFriendly: Data indicates very few are family-friendly, with a mean close to 0.

AvSpendpp (Average amount spent per travel) indicates that visitors, on average, spent 607.62 per travel in Greece for the years 2007 to 2016. This figure encompasses all potential expenses a visitor might incur during their travel, such as accommodation, food, transportation, and entertainment.

AvSpendpn (Average amount spent per overnight stay) Visitors spend an average of 70.58 per overnight stay for the decade under investigation.

CTI_dl (Administrative expenses/turnover YoY log change): the average year-on-year logarithmic change in the administrative expenses to turnover ratio is a constant slight increase of 0.02. This indicates that, on average, hotels are seeing a minor rise in their administrative expenses relative to their turnover. Potential implications of this rise might be:

- A potential increase in overhead costs.
- Possible inefficiencies creeping into the administration of some hotels.
- A stagnant or reduced turnover, making the administrative expenses proportionally higher.

SalesTI_dl (Sales Expenses / Turnover ratio YoY log change): the year-on-year log change in the ratio of sales expenses to turnover is also 0.02. This suggests that the average hotel is experiencing a slight increase on a year basis in sales expenses compared to its turnover. Possible reasons for this could be:

- Increased investment in marketing and promotional activities to attract customers.
- A shift in sales strategy that initially has higher costs.
- External factors leading to higher sales costs, such as commission fees to third-party booking platforms or increased competition require more aggressive sales efforts.

Both previous variables (CTI_dl and SalesTI_dl) seem to have small average year-on-year log changes, suggesting that for the average hotel in the sample, there isn't a drastic annual change in these expense-to-turnover ratios.

Lta (log of Total Assets): the assets of the hotels in the sample, when logged, average around a value of 15.25, indicating that the distribution is somewhat skewed (since the median is 15.07). The range of the logged total assets is between 9.99 and 20.16.

Data for the variable "Pool" showcase that the average of 0.28 indicates that 28% of the hotels in the sample have a pool. Given that the max is 1 (indicating "has a pool") the min is 0 (indicating "does not have a pool"), and the median is 0, this confirms that less than half of the hotels have a pool. Table 4.5. provides summary descriptive statistics for hotels segmented by star category: 3-star, 4-star, and 5-star hotels.

Table 4.5. Summary statistics by hotel star ratings

	Observations	*Mean*	*Median*	*SD*	*Min*	*Max*
Panel A. 3-star hotels						
Number of Beds	2,056	127.67	102.00	96.05	13.00	858.00
Stars	2,061	3.00	3.00	0.00	3.00	3.00
Rating	2061	3,99	4,05	0,81	1	5
Fixed Assets	2,061	2,731,272	1,302,398	8845006	16304	145,000,000
Turnover	2,061	904,948	423,001	2772743	0.00	50,900,000
Lta	2,061	14.45	14.41	0.92	11.98	19.02
lFA_d	1,686	0.01	-0.03	0.21	-3.40	3.25
TurnB~_l	2,021	8.35	8.31	0.93	1.39	12.20
lRatin~t	2,061	1.93	1.79	1.33	0.00	6.15
SalesT~w	1,243	0.23	0.11	0.26	0.00	1.23
CTI_w	1,871	0.21	0.13	0.25	0.01	1.40
lTAsq	2,061	209.62	207.76	27.29	143.53	361.89
TurnFA_w	2,061	0.49	0.33	0.60	0.00	4.35
Pool	2,061	0.22	0.00	0.41	0.00	1.00
FamFri~y	2,061	0.00	0.00	0.06	0.00	1.00
AvSpen~p	1,621	607.21	639.47	47.72	514.10	653.31
AvSpen~n	1,621	70.59	70.38	2.29	66.96	73.93
CTI_dl	1,501	0.06	0.02	0.69	-4.44	4.95
Sales~dl	954	0.04	0.02	0.64	-4.64	4.92
Panel B. 4-star hotels						
Number of Beds	2,560	223.65	144.00	234.52	13.00	1444.00
Stars	2,565	4.00	4.00	0.00	4.00	4.00
Rating	2565	4.10	4.24	0.71	1	5
Fixed Assets	2,565	9,522,632	3150013	24,800,000	0.00	409,000,000
Turnover	2,565	3,050,077	1196788	6,950,329	-193,620	65,500,000
lTA	2,565	15.35	15.26	1.30	9.99	19.90
lFA_d	2,185	0.02	-0.02	0.30	-8.75	3.29
TurnB~_l	2,507	9.01	8.99	1.03	1.49	13.39
lRatin~t	2,565	2.50	2.48	1.46	0.00	6.50
SalesT~w	1,777	0.17	0.07	0.24	0.00	1.23
CTI_w	2,365	0.19	0.12	0.22	0.01	1.40
lTAsq	2,565	237.23	232.97	40.39	99.78	396.04
TurnFA_w	2,562	0.47	0.35	0.49	0.00	4.35
Pool	2,565	0.28	0.00	0.45	0.00	1.00
FamFri~y	2,565	0.00	0.00	0.00	0.00	0.00
AvSpen~p	1,961	607.75	639.47	47.44	514.10	653.31
AvSpen~n	1,961	70.58	70.38	2.28	66.96	73.93
CTI_dl	1,975	0.01	-0.01	0.61	-4.75	5.33
Sales~dl	1,417	0.01	0.00	0.66	-5.04	6.32
Panel C. 5-star hotels						
Number of Beds	914	433.51	299.50	348.00	14.00	1353.00
Stars	914	5.00	5.00	0.00	5.00	5.00
Rating	914	4.29	4.39	0.49	1	5

Fixed Assets	899	46,700,000	14,000,000	95,100,000	0.00	561,000,000
Turnover	914	7,655,320	4,839,002	8,580,202	0.00	62,400,000
lTA	899	16.79	16.69	1.36	10.79	20.16
lFA_d	790	0.03	-0.01	0.28	-2.58	2.98
TurnB~_l	911	9.62	9.57	0.95	4.76	12.53
lRatin~t	914	3.31	3.37	1.44	0.00	6.46
SalesT~w	778	0.11	0.06	0.13	0.00	0.80
CTI_w	866	0.15	0.10	0.19	0.02	1.40
lTAsq	899	283.78	278.54	46.09	116.44	406.44
TurnFA_w	898	0.45	0.28	0.63	0.00	4.35
Pool	914	0.09	0.00	0.29	0.00	1.00
FamFri~y	914	0.00	0.00	0.00	0.00	0.00
AvSpen~p	704	608.47	639.47	47.25	514.10	653.31
AvSpen~n	704	70.56	70.38	2.27	66.96	73.93
CTI_dl	754	-0.02	-0.02	0.44	-3.75	2.78
Sales~dl	662	0.01	0.00	0.39	-1.86	2.24

NOTES: The table reports summary statistics for the variables used in the analysis. “ln” denotes the natural logarithm, “Δln" denotes the logarithmic first difference. All monetary values are expressed in Euros.

The key findings for each star’s category are:

Panel A. 3-star hotels

The average 3-star hotel in this sample has around 128 beds, but there's wide variation as the standard deviation is 96.05, ranging from small establishments with 13 beds to larger ones with up to 858 beds. On average, these hotels have fixed assets worth around €2.73 million. The wide disparity is evident, ranging from a modest €16,304 to a high of €145 million. The average turnover stands at about €905,000. However, some hotels have reported no turnover, while the highest goes up to €50.9 million. Roughly 22% of 3-star hotels have a pool facility, given that the mean is 0.22. Nearly no 3-star hotel in the sample is categorized as family-friendly since the mean and median are both zero.

Panel B. 4-star hotels

The average 4-star hotel has more capacity than the 3-star hotel, with approximately 224 beds. The range is from 13 beds to a massive 1,444 beds. These hotels have significantly higher fixed assets than the 3-star hotels, with an average value of around €9.52 million. The most endowed hotel in this category boasts assets worth €409 million. 4-star hotels, on average, bring in a turnover of about €3.05 million, with the range extending from a negative turnover (due to accounting adjustments, a no-turnover indication) of -€193,620 to a high of €65.5 million. 28% of 4-star hotels have a pool, indicating that this amenity is more common in this category compared to 3-star establishments.

Panel C. 5-star hotels

5-star hotels have an average of 434 beds, almost double the capacity of 4-star hotels. The range is 14 to 1,353 beds. These luxury hotels hold significant value in fixed assets. On average, a 5-star hotel has assets worth €46.7 million, with the top player having assets worth a staggering €561 million. The average turnover for 5-star hotels stands at about €7.66 million. The metric hit a high of €62.4 million. Surprisingly, only 9% of 5-star hotels in this sample have a pool since most 5-star establishments are city hotels.

Considering Table 4., it is concluded that capacity between the three categories varies significantly, with 5-star hotels having far more beds than the other two categories. At the same time, an important difference in the number of beds is held between 3-star and 4-star hotel categories, with 3-star hotels having half the beds of the next-level star category. Considering the lTA (Total Assets log) variable, 3-star hotels exhibit a mean of 14.45, 4-star hotels a mean of 15.35, and 5-star hotels a mean of 16.79. As expected, 5-star hotels, on average, have higher total assets than 3- and 4-star hotels, indicating a clear distinction between these categories in terms of size. Moreover, 5-star hotels have a higher average rating than the two other categories. 4-star hotels have the second-best rating score, while 3-star hotels succeeded in the lowest achievement. rating count (Number of rating log) variable presents for 5-star hotels a mean of 3.31, for 4-star hotels a mean of 2.50, and 3-star hotels a mean of 1.93. It could be assumed that 5-star hotels have, on average, received more ratings than the other two categories due to higher popularity and/or higher number of tourists serviced.

Considering the lFA_d (Year-on-year first difference of logged Fixed Assets) variable, 5-star hotels exhibit a mean of 0.03, 4-star hotels a mean of 0.02, and 3-star hotels a mean of 0.01. It is obvious that 5-star hotels, on average, experienced a more substantial growth in fixed assets from the previous year compared to 4-star and 3-star hotels. In terms of the TurnBeds_l (Log of Turnover/Number of beds ratio) variable, 5-star hotels exhibit a mean of 9.62, 4-star hotels a mean of 9.01, and 3-star hotels a mean of 8.35. In that context, 5-star hotels, on average, generate more turnover per bed than 4- and 3-star hotels. This might suggest higher efficiency or better pricing power for 5-star hotels. It is important to mention that both, 5-star hotels and 4-star hotels generate more per bed turnover compared to the full sample's 8.87 average.

SalesTI_w (Winsorised Sales Expenses / Turnover ratio) variable, counts a mean of 0.11 for 5-star hotels, a mean of 0.17 for 4-star hotels, and a mean of 0.23 for 3-star hotels. 5-star hotels spend a smaller portion of their turnover on sales expenses, while 4-star hotels spend a significantly higher percentage, and 3-star hotels spend even more. A plausible explanation is that usually, 5-star and 4-star hotels have a higher number of beds compared to 3-star hotels, resulting in economies of scale. A second explanation could be that higher-star category hotels usually establish more professional and effective managerial solutions and marketing tactics resulting in incrementally lower costs. The same pattern holds for variable CTI_w (Winsorised

Administrative Expenses / Turnover ratio) indicating that 5-star hotels have a lower administrative expense to turnover ratio as 5-star hotels have a mean of 0.15, 4-star hotels a mean of 0.19 and 3-star hotels a mean of 0.21. Once again higher category hotels seem to deal more efficiently with their administrative costs, probably due to economies of scale or better management.

Consequently, for the TurnFA_w (Winsorised Turnover/Fixed Assets ratio) variable, 5-star hotels exhibit a mean of 0.45, 4-star hotels a mean of 0.47, and 3-star hotels a mean of 0.49. Interestingly, the lower category succeeds slightly more turnover against their fixed assets, probably due to the lower value of the denominator.

The following variable, Pool (Hotel has a pool), calculates a mean of 0.09 for 5-star hotels, a mean of 0.28 for 4-star hotels, and a mean of 0.22 for 3-star hotels, clearly showing that a higher percentage of medium hotels have pools compared to both large and small hotels.

SalesTI_dl (Sales Expenses / Turnover ratio year-on-year log change) variable indicates that on average, 5-star and 4-star hotels experienced a moderate year-on-year increase in their sales expenses compared to 3-star hotels which succeeded an even lower increase.

CTI_dl (Administrative expenses/turnover year-on-year log change) variable, 5-star hotels exhibit a moderate decline in their year-on-year administrative expenses (-0.1), while 4-star hotels exhibit a slight increase (0.1). 3-star hotels present a more substantial increase in their y-o-y expense costs (0.06). This again might hint at greater efficiency gains, cost-cutting measures, and/or more substantial economies of scale implied in the larger hotels compared to their smaller counterparts.

In conclusion, the stratification by hotel star category (3–5 stars) shows that each hotel stratum has different financial profiles and attracts different customer satisfaction ratings. Overall, concerning star categorization of hotels it is obvious that the average number of beds increases with the star rating as 3-star hotels have an average of 127.67 beds, 4-star hotels have 223.65, and 5-star hotels have 433.51. Moreover, the value of fixed assets and total assets (log) is higher in 5-star hotels, indicating larger and more expensive properties. Similarly, Turnover also rises with the hotel star category. 5-star hotels have the highest average turnover (€7,655,320), followed by 4-star (€3,050,077), while 3-star hotels have the lowest (€904,948). Turnover per Bed is highest in 5-star hotels (9.62), followed by 4-star (9.01) and 3-star hotels (8.35). Turnover to Fixed Assets is relatively consistent across all categories, with 3-star hotels having a slightly higher mean (0.49) compared to 4-star (0.47) and 5-star (0.45) hotels. 5-star hotels enjoy the highest average log rating count (3.31), followed by 4-star (2.50) and 3-star hotels (1.93). Additionally, average rating improves with star classification: 3-star hotels have an average rating of approximately 3.99, 4-star hotels have around 4.10, and 5-star hotels have around 4.29.

Sales Expenses to Turnover and Administrative Expenses to Turnover ratios tend to be lower in higher-rated hotels, suggesting better cost management. The year-on-year log changes for administrative expenses to turnover showcase a moderate decrease for 5-star hotels and an increase for 3-star hotels. The sales Expenses to Turnover ratio does not show a consistent pattern across star categories. The presence of a pool is more common in 4-star hotels (28%) compared to 3-star (22%) and 5-star hotels (9%). Table 4.6. contains summary statistics for the variables under examination, categorized by the size of the assets (Large, Medium, Small).

Table 4.6. Summary statistics by total assets Large - Small

Total Assets. Large

Variables	*N*	*Mean*	*Median*	*SD*	*Min.*	*Max.*
BEDS	2762	344.5836	242	290.8528	14	1444
STARS	2762	4.12	4	0.67	3	5
Rating	2762	4.10	4.23	0.65	1	5
FIXEDA~S	2762	2.48E+07	6809650	6.13E+07	175174	5.61E+08
TURNOVER	2762	5466273	2491876	8541550	-37488	6.55E+07
lTA	2762	16.34603	16.01422	1.089903	15.0739	20.16028
lFA_d	2401	0.035908	-0.01561	0.230387	-2.58434	2.9801
TurnB~_l	2739	9.326799	9.248939	0.981308	2.1228	13.39302
lRatin~t	2762	2.794898	2.772589	1.522561	0	6.504288
SalesT~w	2188	0.117793	0.057794	0.168091	0.000363	1.228157
CTI_w	2592	0.155646	0.105461	0.188654	0.0075	1.401676
lTAsq	2762	268.3802	256.4553	37.0039	227.2226	406.437
TurnFA_w	2762	0.393463	0.300079	0.419717	0	4.353887
Pool	2762	0.223027	0	0.416352	0	1
FamFri~y	2762	0	0	0	0	0
AvSpen~p	2143	607.3958	639.4698	47.6082	514.0972	653.3084
AvSpen~n	2143	70.58874	70.37794	2.289301	66.95984	73.93074
CTI_dl	2222	0.012596	-0.00627	0.574385	-4.74546	5.331987
Sales~dl	1824	0.015214	0.002013	0.569375	-3.26009	6.320682

Total Assets. Small

Variables	*N*	*Mean*	*Median*	*SD*	*Min.*	*Max.*
BEDS	2752	101.2155	87	70.76164	13	1400
STARS	2762	3.46	3	0.55	3	5
Rating	2762	4.08	4.25	0.79	1	5
FIXEDA~S	2762	1222342	1088257	779689	0	6297740
TURNOVER	2762	548429.6	377850.5	541060	-193620	3782111
lTA	2762	14.14788	14.26323	0.682964	9.989023	15.07292

lFA_d	2259	0.000904	-0.02868	0.300258	-8.749	3.293752
TurnB~_l	2684	8.380519	8.391113	0.943401	1.390452	11.76964
lRatin~t	2762	2.046605	2.079442	1.348332	0	6.464588
SalesT~w	1597	0.255753	0.135672	0.285292	0.000363	1.228157
CTI_w	2494	0.233801	0.144834	0.256962	0.0075	1.401676
lTAsq	2762	200.6286	203.4397	18.8739	99.78059	227.1929
TurnFA_w	2758	0.55154	0.383817	0.656911	0	4.353887
Pool	2762	0.225199	0	0.417789	0	1
FamFri~y	2762	0.002534	0	0.050288	0	1
AvSpen~p	2129	607.9498	639.4698	47.41712	514.0972	653.3084
AvSpen~n	2129	70.56851	70.37794	2.27525	66.95984	73.93074
CTI_dl	1994	0.034899	0.007345	0.654744	-4.27864	4.953347
Sales~dl	1205	0.020687	0.008574	0.654177	-5.03677	5.340404

Considering Table 3, it is observed that the average number of "BEDS" is roughly 345 for large hotels in terms of total assets (with a wide range from 14 to 1444) against 101 for Small hotels (again with almost the same range, 14 – 1444), indicating the significant difference in the magnitude of bed capacity between the two categories. On average, for Large Hotels, fixed assets value is significantly high (around 24.8 million), indicating larger operations and investments, while the average turnover is about 5.47 million, with a wide range, reflecting significant business activity. Small hotels exhibit an average fixed assets value much lower (around 1.22 million) and an average turnover significantly lower at around 548,430, reflecting a smaller operational scale.

Considering the lTA (Total Assets log) variable, the average log of total assets for Large Hotels is about 16.35. The average log of total assets for Small Hotels is about 14.15, showing these hotels are smaller in scale compared to the large hotels. Hotels with more beds belong (as an average) to a significantly higher star category (4,12 for Large against 3,45 for Small).

Moreover, they tend to have a higher average rating than Small hotels. Moreover, the lRatingcount (Number of ratings log) variable presents for Large hotels a mean of 2,79, and Small hotels a mean of 2,04. Consequently, Large hotels have, on average, received more ratings than small hotels, suggesting they might be more popular or have served more customers.

Taking into account the lFA_d (Year-on-year first difference of logged Fixed Assets) variable, Large hotels exhibit a mean of 0.035, and Small hotels a mean of 0.0009 indicating that Large hotels, on average, experienced a more substantial growth in fixed assets from the previous year compared to small-sized hotels. In terms of the TurnBeds_l (Log of Turnover/Number of beds ratio) variable, Large hotels exhibit a mean of 9.30 and Small hotels a mean of 8.30. In that context, it is clear that Large hotels, on average, generate more turnover per bed than small hotels. This might suggest higher efficiency or pricing power for larger hotels. ales Expenses / Turnover ratio variable counts a mean of 0.12 for Large hotels and a mean of 0.25 for Small hotels. The latter is interpreted as Larger hotels spending a smaller portion of their turnover on sales expenses, implying economies of scale or more efficient marketing strategies. Small hotels spend a significantly higher percentage on sales expenses, as fixed sales overheads cannot be subdivided into a higher turnover due to limited size. Similarly, larger hotels have a lower administrative expense-to-turnover ratio, further indicating economies of scale for the category (0.15 against 0.23). Considering the Turnover/Fixed Assets ratio variable, Large hotels exhibit a mean of 0.39, and Small hotels a mean of 0.55. Small hotels have a higher turnover relative to their fixed assets than larger hotels. This might suggest that smaller hotels are maximizing the use of their assets more than larger hotels, or they have fewer fixed assets relative to their turnover. Another explanation could be that smaller hotels incorporate more boutiques and luxury hotels succeed with higher RevPaR. For the following variable, Pool (Hotel has a pool), a mean of 0.223 for Large hotels, and a mean of 0.225 for Small hotels is calculated, showing that both categories present the same percentage of units incorporating a pool. Whether the hotel is identified as Family-friendly presents very low means for the two hotel categories, affirming that very few hotels are flagged as family-friendly, with small hotels having a slightly higher proportion than large ones.

Sales Expenses / Turnover ratio year-on-year log change variable indicates that on average, larger hotels experienced a more moderate increase in their sales expenses relative to turnover compared to small hotels. This may indicate that larger hotels are becoming more efficient or scaling back their sales initiatives more than smaller hotels.

According to the Administrative expenses/turnover year-on-year log change variable, like the SalesTI_dl trend, larger hotels have seen a less substantial increase in their administrative expenses relative to turnover. This again might hint at greater efficiency gains, cost-cutting measures, or more substantial economies of scale implied in the larger hotels compared to their smaller counterparts.

Overall insights suggest that larger hotels tend to have received more ratings, potentially indicating higher popularity or larger customer bases. Moreover, they tend to have a higher average rating than the Small category, indicating they are generally more upscale establishments, potentially offering more amenities and services. Smaller hotels have higher turnover relative to their fixed assets, indicating efficient utilization or a different asset structure. Both categories present the same number of units including a pool in their amenities. larger hotels tend to have lower expenses in proportion to their turnover, suggesting economies of scale. Larger hotels, on average, have higher bed capacity and total assets than Small hotels. Stratifying the sample by Total Assets, hotels with more beds belong (as an average) to a higher star category than hotels with fewer beds. Large hotels, on average, experienced a more substantial growth in fixed assets from the previous year compared to small-sized hotels. On the other hand, Larger hotels spend a smaller portion of their turnover on sales expenses than their Small counterparts, implying economies of scale or more efficient marketing strategies. Similarly, larger hotels have a lower administrative expense-to-turnover ratio, further indicating economies of scale for the category. A slightly higher percentage of Small hotels have pools compared to large hotels. Very few hotels are flagged as family-friendly, with small hotels having a niglet higher proportion than Large. larger hotels experienced a moderate increase in their sales and administrative expenses relative to turnover compared to small hotels. This may indicate that larger hotels are becoming more efficient or scaling back their sales initiatives more than smaller hotels, exhibiting greater efficiency gains, cost-cutting measures, or more substantial economies of scale implied in the larger hotels compared to their smaller counterparts. Table 4.7 presents summary statistics for the variables under examination considering a "by the number of beds (Large, Small)" categorization.

Table 4.7. Summary statistics by beds Large. Small

Beds. Large

Variables	*N*	*Mean*	*Median*	*SD*	*Min.*	*Max.*
BEDS	2752	370.4251	258	274.1536	138	1444
STARS	2752	4.02	4	0.71	3	5
Rating	2752	4.02	4.17	0.69	1	5
FIXEDA~S	2752	2.35E+07	5670262	6.09E+07	0	5.61E+08
TURNOVER	2752	5121517	2179846	8215984	0	6.55E+07
lTA	2752	16.06268	15.80037	1.333212	9.989023	20.16028
lFA_d	2363	0.025632	-0.01831	0.252414	-2.87519	3.293752
TurnB~_l	2720	8.98665	8.970585	0.966661	2.1228	12.2488
lRatin~t	2752	2.795752	2.772589	1.547641	0	6.504288
SalesT~w	2093	0.122965	0.053788	0.189526	0.000363	1.228157
CTI_w	2562	0.149815	0.104617	0.176026	0.0075	1.401676
lTAsq	2752	259.7864	249.6517	43.82472	99.78059	406.437
TurnFA_w	2749	0.48682	0.334027	0.607034	0	4.353887
Pool	2752	0.226381	0	0.418565	0	1
FamFri~y	2752	0	0	0	0	0
AvSpen~p	2097	608.1961	639.4698	47.30213	514.0972	653.3084
AvSpen~n	2097	70.56717	70.37794	2.276044	66.95984	73.93074
CTI_dl	2169	0.010196	-0.00752	0.568906	-4.74546	5.331987
Sales~dl	1732	0.006406	0.000618	0.527202	-3.26009	3.755597

Beds. Small

Variables	*N*	*Mean*	*Median*	*SD*	*Min.*	*Max.*
BEDS	2740	75.41	75	32.61	13	136
STARS	2740	3.57	3	0.62	3	5
Rating	2740	4.16	4.33	0.76	1	5
FIXEDA~S	2725	2579774	1284359	1.06E+07	0	4.09E+08
TURNOVER	2740	919807.5	407573	3054034	-193620	6.55E+07
lTA	2725	14.43306	14.41284	0.9758	11.27885	19.90085
lFA_d	2256	0.011279	-0.02621	0.282474	-8.749	3.249714

TurnB~_l	2681	8.741728	8.625653	1.174673	1.390452	13.39302
lRatin~t	2740	2.05036	2.079442	1.319225	0	6.148468
SalesT~w	1681	0.238665	0.1237	0.265966	0.000363	1.228157
CTI_w	2496	0.237858	0.147399	0.264244	0.0075	1.401676
lTAsq	2725	209.2651	207.73	28.50195	127.2124	396.0437
TurnFA_w	2724	0.453886	0.32443	0.49856	0	4.353887
Pool	2740	0.217153	0	0.412383	0	1
FamFri~y	2740	0.002555	0	0.050489	0	1
AvSpen~p	2151	607.1749	639.4698	47.69511	514.0972	653.3084
AvSpen~n	2151	70.59039	70.37794	2.288268	66.95984	73.93074
CTI_dl	2025	0.037387	0.011101	0.657088	-4.27864	4.953347
Sales~dl	1282	0.032818	0.013788	0.696962	-5.03677	6.320682

The descriptive statistics stratification of the full sample of hotels, according to the number of beds, in two distinctive categories Large and Small, present similar patterns and characteristics as the per total assets stratification of the sample. According to Table 4. Large hotels have significantly more beds, with a mean of approximately 370 compared to 75 beds for Small hotels. Moreover, Hotels with more beds fall in the higher star category, on average (around 4.02) compared to small hotels (around 3,56 on average), indicating potentially more luxurious accommodations or additional amenities. Interestingly, small hotels, while having fewer ratings ("lRatingcount"), boast the highest average customer rating (around 4.02 against 4,15 for Large), suggesting that the quality of service, intimacy, or specific niches that these hotels serve, has a positive impact on guest satisfaction. Moreover, it shows that the quality of service could be a significant differentiator for these hotels.

Large hotels in terms of beds showcase the highest total assets and fixed assets, indicative of substantial investments in their property and facilities. The total assets (log) mean for large hotels is approximately 16.06, compared to 14.43 for small hotels. The year-on-year first difference of fixed assets ('lFA_d') shows minimal growth across the two categories, with small hotels slightly leading (0.025). This finding could imply a general industry trend of cautious capital expenditure. the larger number of beds hotels, present imperceptibly higher per bed turnover ratio of 8.9, with Small hotels ratio amounting to 8,74. As the "TurnBeds_l" (turnover per bed) indicates how efficiently the hotels are generating revenue based on their bed capacity, Large hotels seem to have only a slightly higher efficiency, suggesting that while large hotels leverage scale, small hotels might emphasize cost-effective operations or higher charges per stay.

Sales expenses to turnover ratios are highest for small hotels, suggesting either higher costs of operations per guest or more significant investments in marketing and sales activities. Administrative expenses to turnover ratios ('CTI_w') are also highest for small hotels, potentially indicating less operational efficiency or higher relative fixed costs in smaller operations. The 'TurnFA_w' metric, representing turnover efficiency relative to fixed assets, is higher for large hotels (0,49 compared to 0,45 or Small), suggesting they manage to utilize their investments more effectively to generate revenue.

The presence of a pool ('Pool') is relatively consistent across all hotel sizes, slightly favoring large hotels. Family friendliness ('FamFriendly') is negligible in large hotels and very low in small ones, indicating that this might not be a prevalent marketing point or amenity concentration.

Both 'CTI_dl' and 'SalesTI_dl' represent the year-on-year changes (log) in expense-to-turnover ratios, providing insight into operational adjustments. Small hotels exhibit higher ratios, particularly in sales expense changes, which could be due to reactive strategies in dynamic market conditions.

In summary, Larger hotels as categorized by the number of beds they operate, have more beds, higher star ratings, and significantly higher fixed assets, indicating more extensive facilities and potentially more diverse services. Smaller hotels, while more modest in scale and investment, achieve higher TripAdvisor ratings, possibly due to their ability to provide more personalized service and potentially benefit from their boutique size and personalized services. Larger hotels appear to be more operationally efficient, with lower sales and administrative expenses relative to turnover. Smaller hotels, while less efficient in this regard, might be investing more in attracting and serving guests on a relative basis.

Considering the three previous stratifications, by star category (Table 2), by total assets (Table 3), and by number of beds (Table 4) one can identify several similarities and a potential grouping among these categorizations. Larger by total assets hotels, having at the same time on average more beds, and belonging to the 5-star category (henceforth LB5), present some common characteristics. More precisely, these hotels have higher turnover and fixed assets than any other stratification. Furthermore, it is observed that LB5 hotels have received more ratings, potentially indicating higher popularity or larger customer bases and they tend to have a higher average rating than the two other categories, indicating more professional management. A surprising exemption holds for Small hotels in terms of beds, experiencing higher customer ratings, as they probably succeed in providing quality service and/or personalized experiences (on a boutique hotel context). In addition, LB5 hotels generate more turnover per bed than lower-star categories with fewer total assets and beds. This might suggest higher efficiency or pricing power. Furthermore, LB5 hotels present lower expenses in proportion to their turnover, suggesting economies of scale or more efficient marketing strategies. Similarly, LB5 hotels have a lower administrative expense-to-turnover ratio compared to lower hotel categories. Finally, LB5 hotels, on average, experienced a more substantial growth in fixed assets from the previous year compared to all other hotels in the sample. A final astonishing finding is that Small hotels in terms of Total Assets and terms of the number of beds have a higher turnover relative to their fixed assets than larger hotels. This might suggest that smaller hotels are maximizing the use of their assets more than larger hotels, or they have fewer fixed assets relative to their turnover. Another explanation could be that smaller hotels incorporate more boutique and luxury hotels succeeding higher RevPaR or they present a different asset structure.

In summary, the hotel industry's landscape, as depicted through the examined metrics, illustrates that while large hotels hold substantial assets and traditional star-based prestige, small hotels often outperform in investment efficiency, customer satisfaction, and potentially adaptive strategies. Nonetheless, the operational costs and challenges for smaller venues are evident, reflecting the competitive and diverse nature of the hospitality sector. Table 4.8. presents year-by-year full sample summary descriptive statistics.

Table 4.8. Summary statistics by year

YEAR 2007

Variables	*Observations*	*Mean*	*Median*	*SD*	*Min.*	*Max.*
BEDS	387	239.3282	147	256.6838	13	1400
STARS	387	3.81137	4	0.692727	3	5
Rating	387	3.83	4	0.91	1	5
FIXEDA~S	387	1.02E+07	2212086	2.98E+07	0	3.29E+08
TURNOVER	387	2960319	1015524	5785522	0	4.14E+07
lTA	387	15.11408	14.97646	1.414589	10.22437	19.7455
lFA_d	0	.	.	.	.	.
TurnB~_l	378	8.871848	8.894584	1.03085	3.846642	12.65077
lRatin~t	387	1.444498	1.386294	1.135536	0	4.330733
SalesT~w	276	0.1641197	0.0603875	0.240702	0.0003633	1.228157
CTI_w	353	0.1765219	0.1036812	0.212057	0.0074998	1.401676
lTAsq	387	230.4312	224.2945	43.90112	104.5378	389.8846

TurnFA_w	386	0.539126	0.3975253	0.506582	0	4.353887
Pool	387	0.245478	0	0.430927	0	1
FamFri~y	387	0	0	0	0	0
AvSpen~p	0	.	.	.	.	.
AvSpen~n	0	.	.	.	.	.
CTI_dl	0	.	.	.	.	.
Sales~dl	0	.	.	.	.	.
YEAR	2008					
stats	*Observations*	*Mean*	*Median*	*SD*	*Min*	*Max*
BEDS	410	232.6561	141.5	254.2276	13	1444
STARS	410	3.82	4	0.69	3	5
Rating	410	3.80204	4	0.949718	1	5
FIXEDA~S	409	1.22E+07	2716337	4.03E+07	0	4.42E+08
TURNOVER	410	3086499	1037608	6083474	0	4.26E+07
lTA	409	15.24624	15.09321	1.423372	9.989023	19.93279
lFA_d	336	0.1676039	0.1137519	0.31385	-0.760267	3.293752
TurnB~_l	402	8.928311	8.995337	1.059422	3.016351	12.84185
lRatin~t	410	1.501353	1.386294	1.188899	0	5.293305
SalesT~w	287	0.183633	0.0737532	0.248236	0.0003633	1.228157
CTI_w	367	0.1966187	0.1130878	0.246086	0.0074998	1.401676
lTAsq	409	234.4688	227.805	44.4964	99.78059	397.3163
TurnFA_w	408	0.4586474	0.3436775	0.438394	0	4.353887
Pool	410	0.2365854	0	0.425505	0	1
FamFri~y	410	0	0	0	0	0
AvSpen~p	0	.	.	.	.	.
AvSpen~n	0	.	.	.	.	.
CTI_dl	300	0.1129411	0.0230465	0.601087	-2.083995	4.09457
Sales~dl	222	0.0934532	0.0320745	0.629223	-2.333373	2.816166
YEAR	2009					
Variables	*Observations*	*Mean*	*Median*	*SD*	*Min*	*Max*

BEDS	457	228.2735	140	246.3238	13	1444
STARS	457	3.818381	4	0.703883	3	5
Rating	457	3.90	4	0.86	1	5
FIXEDA~S	456	1.24E+07	2749823	4.05E+07	48599	4.65E+08
TURNOVER	457	2848403	1069252	5570007	-193620	3.94E+07
lTA	456	15.29675	15.15036	1.390669	11.73517	20.02871
lFA_d	362	0.0231085	-0.032491	0.309111	-1.369006	2.9801
TurnB~_l	450	8.896995	8.912508	0.9917	4.827202	12.88477
lRatin~t	457	1.704743	1.609438	1.271516	0	5.389072
SalesT~w	329	0.1740204	0.0701174	0.237777	0.0003633	1.228157
CTI_w	411	0.1972008	0.1104905	0.251866	0.0074998	1.401676
lTAsq	456	235.9201	229.5335	43.71424	137.7143	401.1494
TurnFA_w	456	0.4420184	0.3112715	0.481344	0	4.353887
Pool	457	0.2319475	0	0.422538	0	1
FamFri~y	457	0.0021882	0	0.046778	0	1
AvSpen~p	0	.	.	.	.	.
AvSpen~n	0	.	.	.	.	.
CTI_dl	324	-0.011948	0.03299	0.601966	-4.44212	1.8583
Sales~dl	252	0.0245174	0.0418188	0.611534	-2.626988	2.237068

YEAR	2010					
Variables	*Observations*	*Mean*	*Median*	*SD*	*Min.*	*Max.*
BEDS	518	227.9807	142	247.1123	13	1444
STARS	518	3.80888	4	0.709263	3	5
Rating	518	3.97	4.03	0.77	1	5
FIXEDA~S	517	1.37E+07	2616850	4.79E+07	0	4.72E+08
TURNOVER	518	2541017	962188.5	4909768	-32388	3.74E+07
lTA	517	15.24665	15.06288	1.457131	10.79062	20.07134
lFA_d	420	-0.021579	-0.038395	0.181688	-2.432	0.801018
TurnB~_l	512	8.755632	8.763556	1.060625	2.993229	12.83136
lRatin~t	518	1.933876	1.94591	1.276449	0	5.541264

SalesT~w	374	0.2093727	0.076183	0.27848	0.0003633	1.228157
CTI_w	471	0.2115281	0.1219542	0.267106	0.0074998	1.401676
lTAsq	517	234.5796	226.8903	45.74831	116.4374	402.8588
TurnFA_w	516	0.4214653	0.2992229	0.469896	0	4.353887
Pool	518	0.2316602	0	0.422301	0	1
FamFri~y	518	0	0	0	0	0
AvSpen~p	518	640.4312	640.4312	0	640.4312	640.4312
AvSpen~n	518	68.56585	68.56585	0	68.56585	68.56585
CTI_dl	379	0.0580365	0.0640783	0.6394	-2.597187	5.331987
Sales~dl	292	0.0748863	0.0936439	0.694953	-5.036769	3.638514

YEAR	2011					
Variables	*Observations*	*Mean*	*Median*	*SD*	*Min.*	*Max.*
BEDS	585	221.1812	136	243.2576	13	1444
STARS	585	3.801709	4	0.705893	3	5
Rating	585	4.14	4.25	0.66	1	5
FIXEDA~S	583	1.26E+07	2561316	4.32E+07	48599	4.50E+08
TURNOVER	585	2737655	833328	5656092	0	4.13E+07
lTA	583	15.20584	15.04328	1.426899	11.73566	20.00638
lFA_d	492	-0.024951	-0.041456	0.11789	-0.710444	1.147755
TurnB~_l	576	8.800598	8.757722	1.07318	1.390452	12.93193
lRatin~t	585	2.280301	2.302585	1.373818	0	5.777652
SalesT~w	438	0.1874157	0.0825053	0.24354	0.0003633	1.228157
CTI_w	537	0.2024925	0.118406	0.251905	0.0074998	1.401676
lTAsq	583	233.2501	226.3002	44.76189	137.7257	400.2552
TurnFA_w	583	0.4439744	0.3180358	0.515118	0	4.353887
Pool	585	0.225641	0	0.418362	0	1
FamFri~y	585	0.0017094	0	0.041345	0	1
AvSpen~p	585	639.4698	639.4698	0	639.4698	639.4698
AvSpen~n	585	69.57777	69.57777	0	69.57777	69.57777
CTI_dl	441	-0.046372	-0.053783	0.441892	-2.696613	2.455353

Sales~dl	348	-0.039695	-0.038259	0.604289	-4.644823	2.707703

YEAR	2012					
Variables	*Observations*	*Mean*	*Median*	*SD*	*Min*	*Max*
BEDS	632	217.7991	136	240.596	13	1444
STARS	634	3.77918	4	0.70776	3	5
Rating	634	4.17	4.33	0.65	1	5
FIXEDA~S	632	1.21E+07	2635977	4.10E+07	48599	4.43E+08
TURNOVER	634	2510915	747947.5	5568460	-9917	4.90E+07
lTA	632	15.21945	15.07449	1.402924	11.80171	19.96261
lFA_d	568	0.0378672	0.0031137	0.133982	-0.57946	0.746749
TurnB~_l	620	8.672203	8.629889	1.146927	1.491655	13.10211
lRatin~t	634	2.423246	2.302585	1.4403	0	5.934894
SalesT~w	478	0.2035909	0.0853483	0.267792	0.0003633	1.228157
CTI_w	587	0.2168068	0.1221411	0.278346	0.0074998	1.401676
lTAsq	632	233.5967	227.2401	43.98741	139.2805	398.5058
TurnFA_w	632	0.3868657	0.2794546	0.493038	0	4.353887
Pool	634	0.214511	0	0.410807	0	1
FamFri~y	634	0.0015773	0	0.039715	0	1
AvSpen~p	634	646.0346	646.0346	0	646.0346	646.0346
AvSpen~n	634	71.13941	71.13941	0	71.13941	71.13941
CTI_dl	518	0.0200175	0.0072718	0.473303	-2.415915	4.940768
Sales~dl	417	0.0319705	0.0160704	0.518887	-2.205753	3.864934

YEAR	2013					
Variables	*Observations*	*Mean*	*Median*	*SD*	*Min.*	*Max.*
BEDS	629	219.0095	136	241.4803	13	1444
STARS	631	3.784469	4	0.704507	3	5
Rating	631	4.19	4.32	0.65	1	5
FIXEDA~S	629	1.25E+07	2592181	4.22E+07	25220	4.49E+08
TURNOVER	631	2960988	870736	6507290	-20804	5.08E+07
lTA	629	15.24711	15.08534	1.401101	11.79317	19.97587

lFA_d	610	-0.012841	-0.028468	0.132693	-1.570259	0.870463
TurnB~_l	616	8.831901	8.740481	1.042014	4.356148	13.138
lRatin~t	631	2.705476	2.639057	1.423586	0	6.272877
SalesT~w	483	0.1712599	0.083399	0.218001	0.0003633	1.228157
CTI_w	583	0.1766989	0.1041331	0.22871	0.0074998	1.401676
lTAsq	629	234.4344	227.5674	44.03785	139.079	399.0355
TurnFA_w	629	0.4515108	0.3180718	0.574428	0	4.353887
Pool	631	0.2139461	0	0.410415	0	1
FamFri~y	631	0.0015848	0	0.039809	0	1
AvSpen~p	631	653.3084	653.3084	0	653.3084	653.3084
AvSpen~n	631	73.05424	73.05424	0	73.05424	73.05424
CTI_dl	553	-0.172386	-0.125432	0.5405	-4.278641	2.070231
Sales~dl	449	-0.09188	-0.095783	0.635294	-2.540644	6.320682
YEAR	2014					
Variables	*Observations*	*Mean*	*Median*	*SD*	*Min.*	*Max.*
BEDS	638	217.4138	135.5	240.3356	13	1444
STARS	640	3.779688	4	0.703152	3	5
Rating	640	4.18	4.32	0.59	1	5
FIXEDA~S	638	1.30E+07	2572256	4.47E+07	30408	5.54E+08
TURNOVER	640	3420761	989617	7986582	0	6.24E+07
lTA	638	15.26578	15.08685	1.425358	11.86073	20.14646
lFA_d	619	0.0135543	-0.0247	0.269162	-3.404669	1.562557
TurnB~_l	630	8.946988	8.853725	1.08467	2.1228	13.29733
lRatin~t	640	2.904705	2.833213	1.436495	0	6.393591
SalesT~w	382	0.1585758	0.0770986	0.213504	0.0003633	1.228157
CTI_w	596	0.1876412	0.1338796	0.19388	0.0074998	1.401676
lTAsq	638	235.0726	227.6132	44.88501	140.677	405.8798
TurnFA_w	638	0.5204706	0.362951	0.631814	0	4.353887
Pool	640	0.2140625	0	0.410491	0	1
FamFri~y	640	0.0015625	0	0.039529	0	1

AvSpen~p	640	590.2404	590.2404	0	590.2404	590.2404
AvSpen~n	640	70.37794	70.37794	0	70.37794	70.37794
CTI_dl	560	0.2131384	0.0142629	0.938965	-3.022568	4.66293
Sales~dl	353	0.0073679	-0.003785	0.633461	-2.921724	4.922865
YEAR	2015					
Variables	*Observations*	*Mean*	*Median*	*SD*	*Min*	*Max*
BEDS	639	217.4351	136	240.2011	13	1444
STARS	641	3.778471	4	0.703277	3	5
Rating	641	4.22	4.33	0.59	1	5
FIXEDA~S	639	1.45E+07	2515137	5.02E+07	1350	5.49E+08
TURNOVER	641	3385838	1008203	7516379	0	5.98E+07
lTA	639	15.27146	15.09045	1.454879	11.27885	20.15242
lFA_d	628	0.005182	-0.018256	0.41824	-8.749002	3.249714
TurnB~_l	630	8.958722	8.871658	1.105505	4.432527	13.30161
lRatin~t	641	3.091878	3.091043	1.444717	0	6.464588
SalesT~w	373	0.1452558	0.0714854	0.189636	0.0003633	1.228157
CTI_w	599	0.1862206	0.1402828	0.175863	0.0074998	1.401676
lTAsq	639	235.331	227.7217	45.92596	127.2124	406.1202
TurnFA_w	639	0.52742	0.362859	0.644138	0	4.353887
Pool	641	0.2184087	0	0.413489	0	1
FamFri~y	641	0.0015601	0	0.039498	0	1
AvSpen~p	641	579.6403	579.6403	0	579.6403	579.6403
AvSpen~n	641	73.93074	73.93074	0	73.93074	73.93074
CTI_dl	575	0.0293267	0.0260354	0.572903	-4.745463	3.050335
Sales~dl	343	-0.017045	0.0074534	0.4818	-2.554022	2.67482
YEAR	2016					
Variables	*Observations*	*Mean*	*Median*	*SD*	*Min*	*Max*
BEDS	635	217.9181	135	240.8825	13	1444
STARS	637	3.77551	4	0.70365	3	5
Rating	637	4.21	4.36	0.59	1.63	5

FIXEDA~S	635	1.57E+07	2636079	5.69E+07	0	5.61E+08
TURNOVER	637	3508621	1015578	7781565	-37488	6.55E+07
lTA	635	15.31398	15.10503	1.458526	11.95837	20.16028
lFA_d	626	0.0312167	-0.018434	0.306646	-2.584344	2.951557
TurnB~_l	625	8.998583	8.924546	1.119921	3.955082	13.39302
lRatin~t	637	3.188493	3.135494	1.40993	0	6.504288
SalesT~w	378	0.1490191	0.0754333	0.186346	0.0003633	1.228157
CTI_w	598	0.1832642	0.1395651	0.162806	0.0074998	1.401676
lTAsq	635	236.642	228.162	46.246	143.0027	406.437
TurnFA_w	634	0.5332274	0.3621361	0.648102	0	4.353887
Pool	637	0.2197802	0	0.414423	0	1
FamFri~y	637	0.0015699	0	0.039621	0	1
AvSpen~p	637	514.0972	514.0972	0	514.0972	514.0972
AvSpen~n	637	66.95984	66.95984	0	66.95984	66.95984
CTI_dl	580	0.0265234	0.0121995	0.451302	-3.576725	4.953347
Sales~dl	357	0.1375535	0.05388	0.603809	-3.240405	4.549133

By examining Table 4.8., one can observe that the average rating score on TripAdvisor increased from 3.83 in the year 2007 to 4,2 in the year 2016, following a steady slight year-by-year increase. The finding may indicate that through the decade under investigation, hotels of the sample ameliorate their services and/or infrastructure, resulting in higher rating scores. Another finding is the increase in the average number of ratings over the years, from 1,44 to 3,18. The latter can be explained by the vast recognition and accelerated use of TripAdvisor over the concrete decade. All other variables don't present any statistically significant findings. Average turnover amounts to 3.960.319 in 2006, drops to 2,515,912 in 2012, and rebounds to 3,508,621 in 2016. Turnover seems to follow the country's receipts trend over the decade under examination. Finally, a consistent decrease in the average number of beds from 2007 to 2016 is detected, stemming from the fact that while some hotels are excluded from the sample, there are no new entries, as new-built doesn't meet the following prerequisite: in order a hotel to be included into the sample, it must exhibit ratings in TripAdvisor and financial statements for all the years of the decade under research.

Chapter 5. Econometric Methods

In this Chapter, the methods used for the analysis of the data are presented, namely, panel data models and the stochastic frontier analysis.

5.1 Panel data methods

Panel data sets, also known as longitudinal data or cross-sectional time-series data, have a cross-section (N) and time (T) dimension. They constitute several observations over time on several cross-sectional units such as individuals, firms, and countries allowing researchers to better explore the dynamics of change. According to Baltes and Nesselroade (1979), longitudinal data and techniques involve "a variety of methods connected by the idea that the entity under investigation is observed repeatedly as it exists and evolves". These methods have been applied in different research disciplines. Frees (2004), posits that they have been developed because of increased data availability and computing power to empirical researchers.

The term panel data was first introduced by Lazarsfeld and Fiske in their study of the relationship between radio advertising and product sales and the effect the first has on the latter, where they proposed to interview a "panel" of consumers over time. Toon (2000) acknowledges Engel's 1857 budget surveys as one of the earliest applications of longitudinal data. In the concrete survey, Engel studies the expenditure pattern from the same set of subjects over a period. The aim was to study expenditure on food as a function of income. Panel data modeling and estimation techniques were developed in the second half of the twentieth century. A widely used technique used to analyze panel data is Pooled GLS regression which estimates the relationship between a dependent variable and one or more independent variables.

Early applications of panel data analyzing techniques are those of Kuh (1959), Johnson (1960), Mundlak (1961), and Hoch (1962) who used the fixed effect models, and Balestra and Nerlove (1966) and Wallace and Hussain (1969) who used the random effect models. These techniques are explained later in the text.

5.1.1 Panel Data Modelling Techniques

Panel data analysis offers several advantages. First and foremost, inferences are drawn using a large sample. According to Baltagi (2005), more complex relationships can be modeled, for example, temporal changes in cross-sections can be analyzed. One of the most important advantages, however, is that panel data modeling allows for the control of heterogeneity in the sample. A standard approach to model the relationship between Y (the dependent variable) and X, a set of explanatory variables, is given below:

$$y_{it} = \beta' x_{it} + \varepsilon_{it} \tag{1}$$

where ε_{it} is the stochastic error term which considers the variation in the expected value of Y which cannot be explained by the X's. This is known as the pooled-GLS model, where GLS stands for generalized least squares.

For instance, in a tourism demand model, Y can stand for the number of arrivals to a particular destination while the X's includes factors affecting demand, such as income in the home country, relative prices, marketing expenditure, transportation cost, and so on. The X's or explanatory variables can be included in the model so long as they are observable and measurable. There are, however, factors such as culture and other unique characteristics of the individuals or groups under study that are not observable or measurable, but which influence the outcome of expected Y (Seetaram and Petit, 2012). These factors are referred to as the unobserved heterogeneity and are not directly part of the equation above. The effect is incorporated in one of the β's, should they be correlated with the respective X or otherwise included in the error term ε_{it}. As a result, in the first case, the estimated β will not reveal the true effect of the variation in X on expected Y. By modeling the relationship between X and Y, using the panel data techniques, the researcher can separate the effect of this unobserved heterogeneity from that of β.

Consider the Equation below, where Y_{it} is determined by three variables.

$$y_{it} = \beta_1 x_{1it} + \beta_2 x_{2it} + \beta_3 x_{3it} + \varepsilon_{it} \tag{2}$$

Omitting X_3 from the model will reduce the previous equation to the equation below, where the effect of X_3 is soaked by the ε_{it}. The actual model estimated is.

$$y_{it} = \beta_1 x_{1it} + \beta_2 x_{2it} + \beta_3 x_{3it} + u_{it} \tag{3}$$

where $u_{it} = x_{3t} + \varepsilon_{it}$

In this case, the residuals include the effect of x_{3t} and will display patterns leading to the conclusion that the model may be suffering from serial correlation (Green, 1999). The implications for the estimated β will depend on whether the x_{it} are correlated with x_{3t}. If so, the estimated coefficients will be biased. If on the other hand, the covariance between x_{3t} and the x_{it} is equal to zero, estimated β 's will be unbiased and consistent. According to Wooldridge (2002), the panel data modeling technique offers an effective solution to this problem. The inclusion of fixed effects namely μ_i in the previous equation will solve the problem as it will absorb the effect of x_{3t}. This solution is also applicable when measurement errors are present in the data.

Taking the example of the tourism model, Seetaram (2010) explains the complexities that arise when faced with the computation of airfare elasticities. Airfare data are often plagued with measurement errors which arise mainly because of the wide array of airfares and travel class categories that are prevalent in the market. This makes the task of the researcher complex as often no choice is left but to use an average airfare to represent the transportation cost to the destination. Average airfare is not always a good representation of actual airfare. Suppose that airfare, x_{3t} is measured with errors such that the actual variable which is included in the model is $x_{3t}^{*} = x_{3t} + v_t$. The model estimated is given by:

$$y_t = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_{3t}^{*} + \eta_{it} \tag{4}$$

where $\eta_{it} = \varepsilon_{it} + v_t$ (Eq. 2) is the stochastic error term.

The $cov(\eta_i, x_{3t}^{*}) = -\beta\delta_v^2$, where δ_v^2 is the variance of the measurement error (Green, 1999). This violates the crucial assumption of non-correlation between any explanatory variable and the residuals of the equation (Green, 1999). Therefore, all the Ordinary Least Squares (OLS) estimators, will be biased and inconsistent. If instead the relationship between Y and the explanatory variables are modeled using the format of the (Eq. 2) equation, the measurement errors, v_t will be absorbed by the unobserved heterogeneity μ_i leaving ε_{it} free from its effect.

5.1.2 The Fixed Effects Estimator

There are two ways of modeling μ_i, the unobserved heterogeneity, under the fixed and random effects. The choice between these two ways depends on, whether μ_i is correlated to any of the other explanatory variables of the model (Wooldridge, 2002).

The fixed effects method assumes that the heterogeneity in the model is time-invariant and specific to the individual group. The slopes are fixed but the intercepts vary for each cross-section. This is equivalent to adding a dummy variable for each cross section which is why fixed effects are also referred to as the Least Square Dummy Variable (LSDV) method. The slopes are treated as constant across groups and across time. It is, however, possible to allow the slope to vary across groups, across time, or both (Hsaio, 2003). The rationale behind this modeling approach is that since μi accounts for time-invariant characteristics of the group, it removes the pernicious effect of time-invariant omitted variables (Allison, 2005). Fixed effect is often chosen as a precaution against omitted variable bias. Fixed effects are useful for controlling unobserved heterogeneity that remains constant over time.

The Fixed effects technique explores the relationship between explanatory and dependent variables within one individual group. For each group, the variations of all variables from their mean values are considered and the estimated coefficients are also known as the within estimates. This can be a limitation of the Fixed effect method as cross-entity variations are ignored. Moreover, only the effect of variables with sufficient variability can be analyzed. Fixed effect models may additionally include an error component that changes over time but not for each unit τ_t.τ_t is treated as a constant in the model. This is known as "time-fixed effects" and a model with both time and entity-fixed effects is known as a "two-way fixed effect" and takes the following form:

$$y_{it} = x'_{it}\beta + \mu_i + \tau_t + \varepsilon_{it} \tag{5}$$

where τ_i are the time-fixed effects. In the present thesis, yearly dummies are used to control the general hotel market in Greece over the period.

In the fixed effects model, the primary goal is to control the effects of unobserved individual-specific factors that could potentially bias the estimates of the relationships between independent and dependent variables. In a panel dataset, each entity (individual, firm, etc.) has its unique characteristics that might affect the dependent variable. These characteristics are often referred to as entity-specific effects or individual effects. Examples could include unobservable managerial skills, cultural differences, or unmeasured preferences. The fixed effects model assumes that these entity-specific effects are time-invariant, meaning they don't change over the periods covered by the data. To account for these unobserved effects, the fixed effects model includes entity-specific dummy variables in the regression equation. A dummy variable takes the value of 1 if the observation belongs to a specific entity and 0 otherwise. By including these dummy variables, the model essentially subtracts out the entity-specific constant effect from each entity's observations. The fixed effects model focuses on the within-entity variation, essentially comparing how the variables of interest change within each entity over time. The variation across different entities is controlled by the entity-specific dummy variables. Including entity-specific dummy variables means that the intercept (constant term) of the regression equation is also entity-specific. This adjustment ensures that the estimated relationships are specific to each entity. The fixed effects model allows for valid statistical inference, even when the unobserved individual-specific effects are correlated with the independent variables. This is contrary to simpler approaches like pooled GLS, which may yield biased results in the presence of unobserved heterogeneity. The fixed effects model provides several advantages, including the ability to control unobserved heterogeneity and to analyze how changes within entities over time are associated with changes in the dependent variable. However, it does not capture time-varying unobserved effects, and it requires a certain level of within-entity variation to estimate the model effectively.

5.1.3 The Random Effects Estimator

The random effects model is a regression analysis technique used to address unobserved heterogeneity or individual-specific characteristics that are assumed to be random and uncorrelated with the independent variables. Like the fixed effects model, the random effects model is commonly used in the context of panel data or longitudinal data where observations are collected on multiple entities (individuals, firms, countries, etc.) over multiple periods.

The random effects model focuses on capturing the average effects of individual-specific factors across the entire population of entities. Just like in the fixed effects model, each entity has its unique characteristics that might affect the dependent variable. These characteristics are referred to as entity-specific effects or individual effects. In the random effects model, it's assumed that these entity-specific effects are random and uncorrelated with the independent variables. This means that the factors contributing to the unobserved heterogeneity are assumed to be drawn from a larger population of entities. The random effects model accounts for these random effects by including an error term that captures the unobserved individual-specific factors. As it was mentioned, the error term is assumed to be uncorrelated with the independent variables. Estimation of the random effects model involves estimating the variance of the error term due to the random effects. This variance estimation allows the model to account for the variability of the dependent variable due to unobserved individual-specific factors. Unlike the fixed effects model, the random effects model treats the entity-specific effects as random disturbances that contribute to the overall variation. This pooling of individual-specific effects helps estimate the average relationships between the independent and dependent variables across the entire population of entities. The random effects model is more efficient than the fixed effects model when the assumption of uncorrelated and random individual-specific effects holds. This means that the random effects model can provide more precise parameter estimates. The random effects model offers several advantages, such as estimating average relationships across entities and requiring fewer parameters to estimate, compared to the fixed effects model. However, it assumes that the unobserved heterogeneity is random and not systematically related to the independent variables.

The Random effects model is given as:

$$Y_{it} = \beta X_{it} + a + u_{it} + \varepsilon_{it} \tag{6}$$

where α is the constant, u_{it} is the between-entity error, and ε_{it} the within-entity error.

The advantage of Random effect is that since variation across the sample is considered, it permits the study of time-invariant factors in the model (which is not possible under the fixed effects model). The Random effect method uses variations both within and between individuals, random effects methods typically have less sampling variability than fixed effects methods (Allison, 2005). The problem, however, is that all relevant measurable variables need to be included in the model, and since data on a few may not be available therefore leads to omitted variable bias in the model.

The choice between Fixed effect and Random effect depends on whether $\boldsymbol{\alpha} i$ is correlated to any of the other explanatory variables of the model (Wooldridge, 2002). When such a correlation exists, the fixed effect technique is superior. Otherwise, the random effect is more parsimonious and gives more efficient estimates (Wooldridge, 2002). A formal test for assessing the correlation between the unobserved heterogeneity and other explanatory variables is the Hausman (1978) specification test. In the tourism literature, the Fixed effect method has been more frequently applied since the groups under observation are often markets, or destinations that have characteristics that influence the other explanatory variables of the model.

5.1.4 Limitations of the Panel Data Methods

The first disadvantage of panel estimation techniques is their complexity of estimation and their large data requirements. Observing several individuals over a period of time usually results in data collection that can be tedious and expensive Baltagi (2005). From a statistical perspective, panel survey designs have some inherent disadvantages as noted in Kitamura (2000):

1. Respondents may find it cumbersome to regularly participate in the same survey, which results in increasing non-response.
2. Attrition or dropout rate from the sample can be high.
3. Over time the accuracy of data collection may decline. This is known as 'panel fatigue’.
4. The response of individuals may be influenced by their responses from previous participations.

These disadvantages can however be addressed although solutions do come at a cost. The solution proposed includes 'refreshing' the survey design and adding fresh participants at later stages. For a more in-depth analysis of attrition in panel data see Alderman et al. (2001), Fitzgerald, et al. (1998), and Uhrig et al. (2008).

5.1.5 Empirical specification

The panel random effects regression model is a common specification for analyzing panel data, where observations on multiple entities (i) over multiple periods (t) exist. The model includes both entity-specific effects (μ_i) and a time-varying effect (γM_t).

Concerning panel data regression model estimation in this study, initially, the influence of outliers is limited in the financial variables through winsorizing the 1st and 99th percentile. The testing approach is similar in design to (Papatheodorou et al., 2012). The dependent variable is average ratings per hotel per year and the model is estimated with a panel random effects regression model using Huber-White robust standard errors to account for potential heteroscedasticity and/or correlation within the error term (ε_{it}):

$$Y_{it} = \alpha + \beta X_{it} + \gamma M_t + \mu_i + \varepsilon_{it} \quad (7)$$

where i and t index the hotel category and years, respectively. Y_{it} represents the dependent variable for entity i at time t. It is the value under explanation and denotes the year-average hotel rating.

α is the intercept term, capturing the baseline or average level of the dependent variable when all the independent variables are zero.

X_{it} is a vector of hotel-level explanatory variables. βX_{it} are the coefficients associated with the independent variables X_{it}. They represent the change in the dependent variable for a unit change in the corresponding independent variable, holding other variables constant.

M_t is a vector of industry-level explanatory variables. This term captures the time-varying effect. It represents the impact of the variable M_t (time-varying characteristic) on the dependent variable Y_{it}.

μ_i denotes the random effects. This is the entity-specific effect, and it captures unobserved heterogeneity or individual-specific factors that influence the dependent variable and remain constant for each entity over time. This term accounts for differences between entities that are not explicitly included in the model.

ε_{it} is the stochastic error term. representing the unobservable factors that affect Yit but are not accounted for by the included variables. It also includes the random disturbances specific to each observation.

Huber-White robust standard errors are a method to estimate the standard errors of the coefficient estimates that are more robust to violations of assumptions, such as heteroscedasticity (unequal variances) and potential correlation within panels. These robust standard errors provide more reliable inferential results, particularly when the assumptions of classical standard errors are violated.

The panel random effects regression model is estimated using the statistical software, Stata. The Huber-White robust standard errors are applied to the estimated coefficients to produce more accurate and reliable hypothesis tests and confidence intervals.

For all variables, the logarithmic transformation is used to address issues related to nonlinearity, heteroscedasticity, and skewed distributions in the data, as well as to interpret the model coefficients in a more meaningful way. Inspection of the correlation matrix confirms there is no issue with multicollinearity in our specification. A Breusch-Pagan Lagrange Multiplier test was used to check whether the random effects specification is superior to a simple pooled GLS. The Breusch-Pagan Lagrange Multiplier test focuses on detecting the presence of heteroscedasticity, which is the unequal variability of residuals across different levels of the independent variables. Heteroskedasticity refers to a situation where the variance of the residuals is unequal over a range of measured values. If heteroskedasticity exists, the population used in the regression contains unequal variance, and the analysis results may be invalid. As regressions involving ordinary least squares (OLS) assume that the residuals are drawn from a population with constant variance, heteroskedasticity is a definite indication that the latter hypothesis is violated.

Regressions with time series variables involve two issues. First, one variable can influence another within a time lag. Second, if the variables are non-stationary, a spurious regression problem may occur. Considering the above, lagged values of the year are estimated, using Distributed Lag. In the research context, Lagged values are used in regression with time series variables to capture the temporal dependencies and dynamics present in the data. When dealing with time series data, observations are collected over time at regular intervals (e.g., days, months, years), and the order of observations matters. Time series data often exhibit patterns, trends, and relationships that evolve over time, and using lagged values in regression allows us to incorporate this temporal structure into the analysis. In this study, lagged values are used for the financial variables to mitigate endogeneity concerns on the relationship between financial values and customer satisfaction. Using lagged values rids the data of unwanted biases and auto-correlational effects which could weaken the regression results. However, this is only an effective estimation strategy if the lagged values do not themselves belong in the respective estimating equation, and if they are sufficiently correlated with the simultaneously determined explanatory variable (Reed, 2015).

Lagged variables come in several types:

- Distributed Lag (DL) variables are lagged values x_{t-k} of observed exogenous predictor variables x_t.
- Autoregressive (AR) variables are lagged values y_{t-k} of observed endogenous response variables y_t.
- Moving Average (MA) variables are lagged values e_{t-k} of unobserved stochastic innovations processes e_t.

5.2 Stochastic Frontier Analysis

The theory and the theoretical framework of the stochastic frontier analysis usually present the producers as successful optimizers. As such, the producers maximize their production function, minimize cost, and maximize their profits. Moreover, conventional econometric techniques, build on the base described above, to estimate production/cost/profit function parameters, they use regression techniques, where deviations of observed choices from optimal ones are modeled as statistical noise. In addition, econometric estimation techniques ought to allow for the fact that deviations of observed choices from optimal ones are considered to be a result of one or both of the following actors: Firstly, the failure to optimize (i.e. inefficiency) and secondly, due to random shocks. Stochastic Frontier Analysis is one such technique to model producer behavior.

Several factors influence the environment in which production takes place e.g., managerial decisions, degree of competitiveness, input and output quality, network characteristics, ownership form, regulation, etc. Econometric techniques interpret them using two different approaches: including them as variables in the production process and/or as control variables in a second-stage regression. The Stochastic Frontier Analysis produces efficiency estimates or efficiency scores of individual units at certain points in time. Thus, one could identify the units under examination that deviate from the calculated efficient frontier, interpret the cause of inefficiency, and propose corrective measures. Since efficiency scores vary across units, they can be related to a unit´s characteristics such as size, ownership, location, etc.

There are two major groups of methods related to the technique and methodology of the efficient frontier analysis. The first one consists of a parametric approach and the second one of non-parametric methods. More specifically, in parametric methods, there are techniques such as Stochastic Frontier Analysis (SFA), Thick Frontier Approach (TFA), and Distribution Free Approach (DFA). In these parametric methods, econometric theory is used to estimate pre-specified functional form, and inefficiency is modeled as an additional stochastic term. In The case of non-parametric methods, such as Data Envelopment Analysis (DEA) or Free Disposal Hull (FDH), the use of linear programming methods and techniques is applied to calculate an efficient frontier (deterministic) against which units are compared.

The methodology around Stochastic Frontier Analysis (SFA), first was developed by Aigner et al. (1977) and Meeusen & van Den Broeck (1977) and serves as a counterpoint to earlier DEA methods of estimating technical efficiency. In contrast to the DEA, which assumes that observed deviations from the frontier are the result of the inefficiency of the unit under consideration, SFA interprets observed deviations and the residuals. Moreover, a prominent distinction between these two techniques is that SFA includes statistical noise in the observed deviation from the estimated frontier. This allows the use of efficiency analysis in situations where it is not assumed with certainty that the 'output gap' between observed production and the optimal one is free of stochastic elements. The second distinction between these methods is that SFA takes the assumption of a clearly defined production technology—i.e., a parametric production function. In contrast to DEA, SFA relies heavily on this assumption of the production function to be utilized in the analysis of the data, while the DEA method avoids defining an explicit production function. This leads to a different interpretation of the results from these two methods—DEA estimates the convex hull of the technology set as the minimal enveloping frontier, while SFA estimates the parameters of the production function itself. These fundamental assumptions underpinning these two types of models make the estimation results of DEA and SFA difficult to compare. For example, in DEA, the estimated parameters for the enveloping hyperplanes serve as point estimates for the marginal productivities of the defined inputs for a given observation, but not as estimators of the parameters of the true production function. However, by construction, similar parameters in SFA serve both as estimators of the marginal productivities and the true production function parameters (Kuosmanen, et al., 2014). When these models are implemented in practice, for example by policymakers and regulators, these fundamental differences sometimes go unacknowledged (Greene, 2008). A commissioned study in energy regulation by Kuosmanen et. al (2013) found that in many countries not only are these differences unaccounted for, but sometimes they are completely ignored.

The stochastic frontier model retains some flexibility in allowing the actual frontier under estimation to be specified in various ways, usually according to the specific objectives of the researcher. In addition to production frontiers, the model lends itself readily to the estimation of cost frontiers (Schmidt & Lovell, 1979) and profit frontiers (Kumbhakar, 1987). These formulations differ from the simple production model by their choice of output and inputs and the impact and interpretation of inefficiency. For the most part, though, the inner workings of these extensions to cost frontiers are almost identical to the production side models we discuss in this thesis.

The stochastic frontier model (Aigner et al., 1977 & Meeusen et al., 1977) begins with an assumption of some production function, which relates the input vector to a single, nonnegative output vector. In applied work, a clear majority of the research utilizes the familiar first-degree flexible Cobb-Douglas or second-degree flexible transcendental logarithmic production function (abbreviated translogarithm), as noted by Greene (2008, p.98). Oftentimes, the choice of production technology is down to researcher preference, with studies on the correct functional forms for analysis being quite scarce. This avenue of research links closely to the broader discussion of justifiable functional forms in econometric analysis and will not be pursued for this thesis. A noted problem with the more flexible production functions is the fact that these specifications may yield estimates that do not satisfy the basic axioms of production (Kumbhakar & Lovell, 2000). More specifically, a more flexible functional form may create problems in the econometric estimation, if the function needs additional constraints on the estimated parameters to satisfy, e.g., monotonicity and concavity. A thorough survey of this issue was presented by Gong and Sickles (1992), who investigated the sensitivity of SFA results to the choice of functional form and concluded that the choice of correct production technology is imperative to obtain unbiased results.

The cornerstone for analyzing and a comprehensive introduction to stochastic efficient frontier is the book written by Kumbhakar and Lovell (2000), where they made a thorough analysis of the theoretical foundations and the mathematic analysis of both technical and economic efficiency. Moreover, Cornwell and Schmidt (2008), introduce a complete mathematical interpretation of the Stochastic Frontier analysis and the estimation of the efficiency using different types of estimations (e.g., panel data up to models with time-varying inefficiency). They advocate that "...panel data are advantageous as they permit less stringent assumptions or enhanced precision within a specific set of assumptions, compared to what a single cross-section would allow. The majority of research conducted on employing panel data for efficiency estimation has highlighted the potential for relaxed assumptions and more adaptable models."

Considering the latter, only a few researchers have tried to study efficiency and effectiveness using a panel data model in the tourist sector, such as Yu and Lee (2009) and Hsieh et al. (2010). As they have said, the finding of a panel data analysis has the potential to reveal the growth in the hotel's performance over time, leading to a more robust result.

A process of production in classical economics is defined as the usage of material and immaterial resources for making goods and services. Furthermore, a company, commonly referred to as a production unit, uses a set of resources (inputs) to produce a set of goods and services (outputs). Consider a company, that uses K inputs, indexed $k = 1, 2, \ldots, K$, to produce

M outputs, indexed $m = 1, 2, \dots, M$. Input and output bundles can be presented in a vector form as:

$$x = (x_1, x_2, \dots, x_K) \quad (8)$$

$$y = (y_1, y_2, \dots, y_M) \quad (9)$$

The production process can be defined as transforming of an input vector x into an output vector y. Technological limits of production are usually described as a set of pairs of input and output vectors, which are possible in the sense that a company can produce an output vector by using a given input vector. This set of input and output pairs is well known as a production possibility set denoted as PPS:

$$PPS = \{x, y: xcanproducey\} \quad (10)$$

The set of feasible outputs for an input vector can be defined as:

$$P(x) = \{y: (x, y) \in PPS\} \quad (11)$$

This set includes all output vectors y, which are feasible for a given input vector x, in other words, it represents the set of output values y that can be produced given an input value x while adhering to the constraints and possibilities outlined by the Production Possibility Set (PPS). The definition of efficiency of a company's activity strictly depends on the goal of this activity. The most widely used goals of a company are maximization of the output vector given by a fixed input vector (output-oriented) and minimization of the input vector given by a fixed output vector (input-oriented).

Efficiency, measured based on these production-oriented approaches, is called technical. There are several alternative goal specifications: revenue maximization, profit maximization, and others. Further in this chapter, we will consider output-oriented production.

An output vector is called technically efficient if, and only if (Koopmans's definition):

$$y_{eff}: y_{eff} \in P(x) \rightarrow \nabla_{y'>y} y' \notin P(x) \quad (12)$$

where:

y_{eff}: is a function that calculates the technically efficient output (y_eff) for a given input level (x).

$P(x)$: is the set of feasible output values associated with the input level x. In other words, it represents all the possible output values that can be produced using the given input level x while staying within the production possibilities set (PPS).

→: arrow indicates the mapping or relationship between the input value x and the calculated efficient output value y_eff.

$\nabla_{y'>y} y'$: represents a selection process. The symbol ∇ is used to denote a selection operator, selecting from a set of alternative output values (y') that are greater than the current output value y.

$y' \notin P(x)$: indicates that the alternative output value y' must not belong to the set $P(x)$, which represents the set of feasible output values for the given input level x. In other words, the selected alternative output value must not be attainable with the current input level x while staying within the production possibilities set. Concluding, technical efficiency means that given an input vector, there are no feasible output vectors exceeding y_{eff} in any component.

Expanding this concept to all feasible sets of input vectors, a production possibility frontier is defined as a function:

$$f(x) = \{y : y \in P(x), \nabla_{y'>y} y' \notin P(x)\} \tag{13}$$

In the case of a single output production process, the production possibility frontier can be presented as:

$$f(x) = max_y P(x) \tag{14}$$

Koopmans's definition of technically efficient output vectors is very general and can be applied to outputs of different nature. A more practically convenient definition of the technical efficiency of output vector y was presented by Debreu and Farrell (*1957, The measurement of productive efficiency. Journal of the Royal Statistical Society Series A: Statistics in Society, 120(3), 253-281*):

$$TE(x,y)=\left[sup_{\theta}\{\theta:\theta_y \leq f(x)\}\right]^{-1} \tag{15}$$

This definition is closely related to a distance function, introduced in Shephard's works on multi-output production. The main difference with Koopmans's definition is in the direction of output vector increasing. Koopmans's definition allows the increasing of any component of y, while the Debreu-Farrell definition considers only the equiproportional (radial) increase of y. Later the Debreu-Farrell definition was extended by Luenberger, D. G. (*1992, New optimality principles for economic efficiency and equilibrium. Journal of optimization theory and applications, 75, 221-264*.) and Chambers, R. G., Chung, Y., & Färe, R. (*1996, Benefit and distance functions. Journal of economic theory, 70(2), 407-419*), who introduced a directional technology distance function.

In the following graph, there is an illustration of the different definitions of technical efficiency. OA presents an arbitrary directional distance, OB presents Koopmans's (closest) distance, and OC presents Debreu-Farrell's (radial) distance. All discussed features can be extended to more general definitions of technical efficiency. According to the Debreu-Farrell definition, values of technical efficiency should satisfy the following properties:

$0 \leq TE(x,y) \leq 1$ indicating that technical efficiency (TE) values are bounded between 0 and 1. A TE value of 0 indicates complete inefficiency and a TE value of 1 indicates perfect efficiency, $TE(x, y_{eff}) = 1$ stating that when the output is y_{eff} (technically efficient output) for a given input level x, the technical efficiency is 1, $TE(x,y)\ is\ non-decreasinginy$ securing that as the actual output y increases, the technical efficiency value should not decrease. In other words, producing more output from the same input levels should not lead to lower efficiency. This property reflects the idea that utilizing inputs more effectively should lead to higher efficiency, $TE(x, \lambda y) = \lambda TE(x, y)$ interpreting the scaling of output. It states that if the output is scaled by a factor λ, then the technical efficiency value should also be scaled by the same factor λ. This is in line with the concept that efficiency should be proportional to the level of output achieved. If the output is doubled (λ = 2), the efficiency should also double.

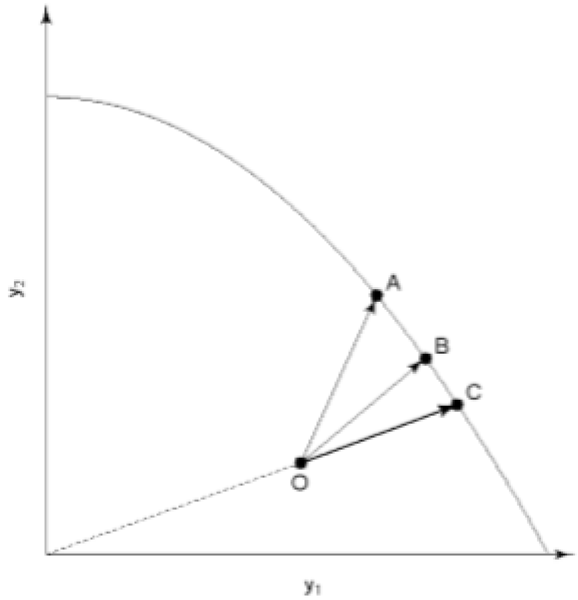


Figure 5.1. Graphical illustration of different definitions of technical efficiency

As a result, the value of technical efficiency equals one (1) for a company located on the production possibility frontier (produced a maximum possible vector of outputs given by its input vector). Companies producing less than the maximum possible outputs, compared to their inputs, are classified as inefficient. The Debreu-Farrell definition of technical efficiency can be presented in the form of equation:

$$y = f(x) \cdot TE(x, y) \tag{16}$$

So, given x and y, tasks of construction of production frontier $f(x)$ and technical efficiency $TE(x, y)$ are dual to each other. The above-mentioned model combines a production function with technical efficiency to analyze the performance of hotels in this thesis case. y: represents the observed or actual output of a production unit. $f(x)$ represents the production function, where the inputs x are transformed into output y. $TE(x, y)$ is the technical efficiency term, which quantifies how efficiently a production unit is using its inputs to produce a given output. It's a value between 0 and 1, where 1 represents perfect efficiency and values below 1 represent inefficiencies. The equation $y = f(x) \cdot TE(x, y)$ decomposes the observed output into two components: one that represents what the output would be if the production unit were perfectly efficient, and the other that accounts for the deviation from perfect efficiency due to inefficiencies or random factors.

For estimation purposes, the technical efficiency term is usually transformed as:

$$TE(x, y)=\exp(-u), u \geq 0 \tag{17}$$

After this transformation properties for technical efficiency values are satisfied automatically. The term u is an inverse to the technical efficiency value, so it is frequently noticed as an inefficiency term. In that context $exp(-u)$ represents the exponential function raised to the power of negative u helping to describe a growth or decay process. In this context, it's used to transform the inefficiency factor u into the efficiency measure.

The previous equation can be presented as:

$$y = f(x) \cdot \exp(-u) \tag{18}$$

The equation accounts for both the systematic process of transforming inputs into outputs (as described by $f(x)$) and the impact of inefficiency (as captured by $exp(-u)$). The term $exp(-u)$ attenuates the output based on the level of inefficiency: the higher the inefficiency (larger u), the lower the output, reflecting the reduction in the effectiveness of input utilization.

The model assumes that the production frontier $f(x)$ is deterministic. This assumption ignores the fact that the production of a company can be affected by random disturbances. The presence of these random disturbances in practice is widely acknowledged and considered as a background to be interpreted. Random disturbances are usually explained by the influence of a large set of factors, generated both from the company's internal and external environment. Introducing the random disturbances *v* into the last formula, we consider a classical stochastic frontier (SF) model:

$$y = f(x) \cdot \exp(v) \cdot \exp(-u) \tag{19}$$

The latter SFA equation combines the systematic impact of the input-output relationship (as described by $f(x)$), the influence of random variations ($exp(v)$), and the effect of technical inefficiency ($exp(-u)$). This equation allows for the separation of systematic inefficiency from random variations in observed output.

For the econometric estimation of this model, we assume that we have a sample of n companies, indexed $i = 1, 2, \ldots, N$. Values of output (yi) and input (xi) vectors are available for each company, while values of random disturbances (vi) and inefficiencies (ui) are not observable. Supposing that the production possibility frontier $f(x)$ is common for all companies in the sample and depends on a vector of parameters β, we receive a cross-sectional specification of the stochastic frontier model:

$$y_i = f(x_i, \beta) \cdot \exp(v_i) \cdot \exp(-u_i) \tag{20}$$

When a production process is described only by one output ($M = 1$), the specification of the above-presented model represents a standard econometric model, in which parameters can be estimated. This approach is frequently used in cases when the single-output assumption is appropriate for a real production process or when production outputs can be aggregated. The model is frequently presented in the logarithmic form, which is more convenient in practice:

$$\ln y_i = \ln f(x_i, \beta) + v_i - u_i \tag{21}$$

Models with multiple outputs ($M > 1$) production require a transformation to become econometrically estimable. A popular transformation utilizes the property (homogeneity of degree 1 in outputs) of technical efficiency.

There is a set of widely known theoretical production functions including the Cobb-Douglas function, translog function, Diewert function, and CES (constant elasticity of substitution) function. The Cobb-Douglass function is one of the simplest forms:

$$\ln f(x_i, \beta) = \beta_0 + \sum_{j=1}^{K} \beta_j \ln x_{ji} \tag{22}$$

assuming that the relationship between the dependent variable $ln(f(x_i,\beta))$ and the independent variables $ln(x_{ji})$ is linear (after taking logarithms). Moreover, the Cobb-Douglass function helps to model relationships where percentage changes in the dependent variable are associated with percentage changes in the independent variables.

All elasticities of substitution between inputs in the Cobb-Douglas function are equal to one (1).

The translog production function is more flexible in terms of elasticity substitution:

$$\ln f(x_i,\beta) = \beta_0 + \sum_{j=1}^{K} \beta_j \ln x_{ji} + \sum_{j=1}^{K}\sum_{k=1}^{K} \beta_{ij} \ln x_{ji} \ln x_{ki} \tag{23}$$

The elasticity of substitution in the translogarithm production function is not fixed to one (1) but can be estimated.

In the concrete study, an efficiency measurement is calculated, and hotels are divided into two categories, one with high and one with low efficiency. Technical efficiency is estimated for all hotels in the sample using the following translog specification:

$$\begin{aligned}\ln y_{Mit} = a_0 &+ \sum_{m=1}^{M-1} a_m \ln\left(\frac{y_{mit}}{y_{Mit}}\right) + 0.5 \sum_{m=1}^{M-1}\sum_{n=1}^{M-1} a_{mn} \ln\left(\frac{y_{mit}}{y_{Mit}}\right) \ln\left(\frac{y_{mit}}{y_{Mit}}\right) \\ &+ \sum_{k=1}^{K} \beta_k \ln x_{kit} + 0.5 \sum_{k=1}^{K}\sum_{l=1}^{K} \beta_{kl} \ln x_{kit} \ln x_{lit} \\ &+ \sum_{k=1}^{K}\sum_{m=1}^{M-1} \delta_{km} \ln x_{kit} \ln\left(\frac{y_{mit}}{y_{Mit}}\right) + \varepsilon_{it}\end{aligned} \tag{24}$$

where i,t index the hotels and years, respectively. The hotels are assumed to produce two outputs while using four inputs.

Considering the equation, ln y_{Mit}: This represents the natural logarithm of the observed output (value added ($y1$) and sales revenue ($y2$)) for hotel "i" in year "t".

$a_0, a_m,\ a_{mn}$: These are parameters that need to be estimated from the data. They are associated with the logarithmic transformations of output ratios, where "a_0" represents the intercept, "a_m," and "a_{mn}" are coefficients for the logarithmic ratios of outputs for different categories "m" (M-1 categories are considered).

$\ln\left(\frac{y_{mit}}{y_{Mit}}\right)\ln\left(\frac{y_{mit}}{y_{Mit}}\right)$: These are the logarithmic ratios of output "y_mit" for category "m" relative to the total output "y_Mit" of hotel "i" in year "t." These ratios capture how output in each category contributes to the total output.

$\beta_k,\ \beta_{kl}$: These are parameters associated with the logarithmic transformations of input variables. "β_k" represents the coefficient for the logarithmic transformation of input "k," and "β_{kl}" represents interaction terms between input "k" and input "l".

x_{kit}: These are the input variables used, namely: (i) fixed assets, ii) equity, iii) administrative expenses, and iv) sales expenses associated with hotel "i" in year "t".

δ_{km}: These are parameters associated with the interaction between input "k" and the logarithmic ratio of output in category "m" It captures how inputs relate to output categories.

ε_{it}: is the error term that accounts for unobserved factors and random variations that affect the observed output. The stochastic frontier analysis (SFA) assumes that the residual in the above specification is made up of these two components; $\varepsilon_{it} = v_{it} - u_{it}$ where $v_{it} \sim N(0, \sigma_v^2)$ reflects the standard stochastic error term including random variations and unobservable factors that affect the data and $u_{it} \sim N^+(\mu, \sigma_u^2)$ which reflects systematic or non-random part of the error interpreted as the efficiency score which is the quantity of interest for our purposes. Estimation of the model was conducted in Stata using maximum likelihood.

Parameters (a's, β's, δ's) are estimated from the data, and the resulting efficiency scores can be used to categorize hotels into high and low-efficiency groups based on their relative performance in terms of utilizing inputs to generate outputs.

The specification is in line with similar hotel technical efficiency literature, see Brida et al., (2012), Delotto et al., (2014), and Pulina et al., (2010) among others. However, it is the first time that total expenses are decomposed into administrative (reflecting managerial quality) and sales expenses reflecting promotion effort.

Chapter 6 - Empirical results and discussion

This chapter presents the empirical results of the thesis in a series of tables. The findings are first presented for the full sample and then for the stratified sub-groupings.

Table 6.1. Regression results for the full sample

Variables/Models	*(1)*	*(2)*	*(3)*	*(4)*
Lta	-0.382*	-0.424**	-0.549***	-0.537***
	(0.201)	(0.199)	(0.199)	(0.200)
lTAsq	0.0103*	0.0113*	0.0148**	0.0146**
	(0.00612)	(0.00606)	(0.00607)	(0.00608)
TurnBeds_l	0.107***	0.118***	0.134***	0.138***
	(0.0251)	(0.0232)	(0.0243)	(0.0246)
TurnFA_w	-0.0729*	-0.0760*	-0.0781*	-0.0734*
	(0.0400)	(0.0411)	(0.0418)	(0.0414)
lRatingcount	0.104***	0.104***	0.108***	0.0993***
	(0.0142)	(0.0154)	(0.0154)	(0.0163)
CTI_w	-0.0617	-0.0602	-0.00946	-0.0126
	(0.109)	(0.115)	(0.131)	(0.131)
SalesTI_w	0.303***	0.298***	0.273***	0.276***
	(0.0818)	(0.0798)	(0.0833)	(0.0837)
lFA_d		0.0702**	0.0748**	0.0722**
		(0.0336)	(0.0360)	(0.0360)
CTI_dl			-0.0269	-0.0225
			(0.0276)	(0.0275)
SalesTI_dl			0.00867	0.00886
			(0.0225)	(0.0221)
Pool	-0.163***	-0.170***	-0.175***	-0.175***
	(0.0503)	(0.0518)	(0.0526)	(0.0527)
FamFriendly	0.693***	0.947***	0.968***	0.938***
	(0.0586)	(0.0629)	(0.0668)	(0.0686)
AvSpendpn	0.00660	0.00644	0.00536	0.00885**
	(0.00433)	(0.00410)	(0.00420)	(0.00433)
AvSpendpp	2.51e-05	2.31e-05	0.000120	-0.000698*
	(0.000249)	(0.000249)	(0.000249)	(0.000368)
2011.YEAR				0.114***
				(0.0364)
2012.YEAR				0.104***
				(0.0351)
2013.YEAR				0.0738*
				(0.0388)

2014.YEAR				-0.000277
				(0.0257)
2015o.YEAR				-
2016o.YEAR				-
Constant	5.866***	6.211***	7.139***	7.212***
Observations	2,843	2,645	2,521	2,521
No. hotels	540	536	520	520
R-squared	0.1333	0.1474	0.1509	0.1538

Notes: Heteroskedasticity robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6.1. presents the results of regression analysis across four different models (1, 2, 3, and 4) using the predetermined explanatory variables related to hotel characteristics and performance. Different models (1 through 4) are constructed by progressively adding variables.

Model (1) is the basic model with a set of core explanatory variables. Model (2) introduces L.lFA_d, the year-on-year first difference of a hotel's Fixed Assets. Model (3) adds two more variables - L.CTI_dl (the administrative expenses/turnover year-on-year log change) and L.SalesTI_dl (the Sales Expenses / Turnover ratio year-on-year log change). Model (4) further expands on Model (3) by introducing year dummy variables for the years 2011 to 2014, capturing the effect of specific years on the TripAdvisor rating and adjusting for international or national trends.

Model (1) uses the largest sample of hotels and observations, Model (2) has slightly fewer data while, Models (3) & (4) both use slightly fewer observations compared to Model (2). Considering Goodness-of-fit (R-squared), Model (1) has an R-squared value of 0.1333, indicating that it explains about 11.78% of the variability in the dependent variable. Model (2) slightly improves on this with an R-squared of 0.1474, while Model (3) has an R-squared value of 0.1509. Model (4) has the highest R-squared value of 0.1538, suggesting that, among the four models, it does the best job explaining the variation in TripAdvisor ratings using the given variables. Standard Errors associated with the coefficients are presented in brackets, a row after the variable in Table 6.1. Larger standard errors indicate greater uncertainty in the coefficient estimates. Smaller standard errors suggest more precise estimates.

Taking into consideration the first variable of the models, Lta (Total Assets, log), it presents negative and significant coefficients across all models, ($p < 0.1$ in Model 1, $p < 0.05$ in Model 2, and $p < 0.01$ in Models 3 and 4), indicating that as the total assets of a hotel increase, the dependent variable (TripAdvisor rating score) decreases, holding other factors constant. The latter suggests that larger hotels (in terms of assets) tend to have lower ratings. Consequently, as the hotel's magnitude in terms of infrastructure and operational metrics increases the customer's satisfaction decreases. The descending direction of the curve can be assumed to be explained by the sense a customer takes of a faceless establishment when he visits a sprawled hotel with multiple (sometimes identical) compounds.

lTAsq (Total Assets, log squared), is positive and significant across all models, ($p < 0.1$ in Models 1 and 2, $p < 0.05$ in Models 3 and 4), indicating a nonlinear U-shaped relation (quadratic) with total assets. This squared term typically is used to capture non-linear effects. In this context, while initially, the dependent variable decreases with an increase in assets (as indicated by the negative Lta coefficient), it eventually starts increasing at higher levels of assets. In other words, as hotels grow very large, the negative effect on ratings becomes less pronounced, beginning to lessen or possibly reverse. The latter probably suggests that there's an optimal size for a hotel in terms of assets.

TurnBeds_l (log of turnover/number of beds) is positive and significant at a 1% level across all models, indicating that hotels with higher turnover/bed enjoy higher customer satisfaction. The latter can be interpreted as hotels with higher standards usually charging higher prices and/or enjoying higher occupancy per year.

TurnFA_w (turnover/fixed assets, winsorized) is negative but not significant at any level in all models, implying that the concrete variable might not be a strong predictor.

lRatingcount (number of ratings, log) variable is positive and significant at a 1% level across all models, suggesting that hotels with more ratings generally have higher TripAdvisor ratings. A plausible explanation is that people tend to rate when they have a good perception of the hotel, they stayed and skipped rating when they had a non-superb or average experience.

CTI_w (Administrative Expenses/Turnover) is not statistically significant in any model, indicating a weak or no relationship.

The sales Expenses / Turnover (SalesTI_w) variable showcases a positive relation and high significance across all models, meaning that higher sales expenses relative to turnover are associated with higher ratings. Sales expenses incorporate marketing expenses, commissions for sales to tour operators and agencies advertisements, and other related expenses. A possible interpretation of the concrete finding could be that taking care of the purchasing experience of the client (having a highly functional site and reservation tool) as well as of the on-site total experience (better amenities, tailored-made pampering, etc) results in higher satisfaction and better rating.

lFA_d (year-on-year first difference of Fixed Assets) variable is positive and significant in models (2) to (4), suggesting that hotels that increased their fixed assets from the previous year tend to have higher ratings. The increase could reflect either a renovation or an extension in capacity or facilities. So capital investments for infrastructure investment result in higher customer contentment, as visitors enjoy an increased value received/price paid ratio, sometimes being surprised by amelioration of the general status of the hotel.

Year-on-year log change in Administrative Expenses / Turnover (CTI_dl) and Sales Expenses / Turnover (SalesTI_dl) - both variables are not statistically significant, suggesting a weak or no relationship with ratings.

The pool (Presence of a pool) variable is negative and significant ($p < 0.01$), across all models. Hotels with a pool have a lower value of the dependent variable compared to those without. While the presence of a pool is positively assessed by tourists, a series of factors involved with the "pool experience" including hygiene issues, safety issues, crowdedness, availability of sunbeds, food & beverage service, changing and WC facilities, provision of pool towels, pool maintenance, water purity and clarity, music selection and loudness can alternate the client's perception and resulting to decreased ratings.

The next variable, FamFriendly (Family-friendly) is positive and significant at a 1% level in all models, suggesting family-friendly hotels tend to have a higher average value of ratings on TripAdvisor. Family-friendly hotels have significantly higher ratings which indicates that amenities or services catering to families can positively impact a hotel's overall rating. Family-friendly is expected to have en-suite rooms lodging, cradles and baby coats, playgrounds, babysitting on demand, programs specially designed for children, kids' breakfast, and menu, etc. These facilities are boosting client's gratification.

AvSpendpn (Average spending per night) becomes significant in model (4) with a positive coefficient. AvSpendpp (Average spending per travel) is negative and significant only in the model (4). The results are mixed: average spending per night is positively related to ratings and significant in Model 4 ($p < 0.05$), while average spending per trip is negatively related and significant in Model 4 ($p < 0.1$ with relatively low values. The amount a visitor spends per travel or per overnight stay in Greece seems to have a minor impact on the ratings, with per night spending having a more substantial positive relation to ratings.

Finally, Year Dummies (2011, 2012, 2013, and 2014), capturing year-specific effects are positively correlated and significant for 2011, 2012, and 2013 in Model (4), suggesting higher ratings in those years compared to the base year. 2014 is not statistically significant. More specifically, the coefficients suggest that, compared to the reference year (2007), the average rating on TripAdvisor was higher in 2011, 2012, and 2013 and nearly unchanged in 2014.

The Constant represents the average (predicted) TripAdvisor rating when all independent variables are zero. It is positive and statistically significant across all models. However, in this context, it doesn't have a meaningful interpretation given that many of these variables can't realistically be zero.

The model's fit as represented by R-squared ranges from 0.1333 to 0.1538 on the 4 models, indicating the proportion of the variance in the TripAdvisor rating that is explained from the independent variables. An R-squared value of 0 means that the dependent variable cannot be predicted from the independent variables, while an R-squared value of 1 indicates that the dependent variable can be perfectly predicted from the independent variables. Consequently, the models explain between approximately 13.33% to 15.38% of the variance in the dependent variable. R-squared increases from Model 1 to Model 4, indicating that the additional variables in the later models explain more of the variance in the dependent variable. Although relatively low in explanatory capacity, the results suggest that factors such as the total assets, turnover per bed, presence of a pool, and whether a hotel is family-friendly, have significant impacts on customer satisfaction.

In summary, several variables like the size of a hotel (Lta, LTAsq), turnover per bed (TurnBeds_l), turnover relative to fixed assets (TurnFA_w), number of ratings (lRatingcount), family-friendly nature (FamFriendly), having a pool (Pool) and yearly effects show consistent statistical significance across the 4 models, suggesting they are important factors influencing a hotel's average yearly rating on TripAdvisor. Being family-friendly, the number of ratings, turnover per bed, and sales expenses emerge to have a positive relationship, while having a pool and total assets (up to a point) relate negatively. Assets and financial metrics play a role in hotel ratings, but so do amenities and services such as the hotel having a pool being family-friendly. Yearly effects might be impacting ratings, suggesting that external factors affect ratings during these years.

The models can be enhanced further, given the relatively low R-squared values. It's essential to remember that correlation doesn't imply causation all the time, so while these variables are associated with TripAdvisor ratings, they might not necessarily cause the ratings to change. Finally, the change in sign and significance of some variables across models suggests potential multicollinearity.

Table 6.2. Regression results for the full sample with one-year lagged regressors.

Variables/ Model	*(1)*	*(2)*	*(3)*	*(4)*
L.lTA	-0.575***	-0.602***	-0.582***	-0.574***
	(0.211)	(0.200)	(0.219)	(0.219)
L.lTAsq	0.0164**	0.0170***	0.0162**	0.0160**
	(0.00641)	(0.00613)	(0.00667)	(0.00666)
L.TurnBeds_l	0.109***	0.126***	0.132***	0.138***
	(0.0256)	(0.0258)	(0.0299)	(0.0301)
L.TurnFA_w	-0.0513	-0.0536	-0.0519	-0.0467
	(0.0390)	(0.0395)	(0.0419)	(0.0423)
L.lRatingcount	0.0731***	0.0789***	0.0769***	0.0697***
	(0.0114)	(0.0120)	(0.0123)	(0.0131)
L.CTI_w	0.0470	0.162	0.157	0.165
	(0.135)	(0.101)	(0.101)	(0.101)
L.SalesTI_w	0.187*	0.254***	0.246***	0.242**
	(0.0958)	(0.0880)	(0.0946)	(0.0955)
L.lFA_d		0.0185	0.0183	0.0396
		(0.0428)	(0.0441)	(0.0436)
L.CTI_dl			0.00767	0.00571
			(0.0212)	(0.0207)
L.SalesTI_dl			0.0155	0.0154
			(0.0168)	(0.0167)
Pool	-0.149***	-0.137***	-0.137**	-0.138**
	(0.0525)	(0.0517)	(0.0538)	(0.0538)
FamFriendly	0.588***	0.614***	0.611***	0.609***
	(0.0647)	(0.0589)	(0.0645)	(0.0666)
AvSpendpn	0.0107**	0.00976**	0.00934**	0.0145***
	(0.00433)	(0.00420)	(0.00432)	(0.00456)
AvSpendpp	-0.000167	-7.36e-05	-7.59e-05	-0.00106***
	(0.000263)	(0.000263)	(0.000273)	(0.000402)
2011.YEAR				0.154***
				(0.0362)
2012.YEAR				0.117***
				(0.0347)
2013.YEAR				0.0928**
				(0.0392)
2014.YEAR				0.0348
				(0.0256)
2015.YEAR				-
2016.YEAR				-
Constant	7.301***	7.390***	7.237***	7.312***

N	2,713	2,483	2,366	2,366
No. hotels	536	524	499	499
R-squared	0.1178	0.1334	0.1305	0.1345

Notes: Heteroskedasticity robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

The dependence of a variable Y (dependent variable) on another variable (s) X (explanatory variables) is rarely instantaneous and often, Y responds to X with a lapse of time. Such a lapse of time is called a lag. The latter could occur when the explanatory variable has a causal effect on the response variable, but the causal effect occurs gradually and/or manifests in changes to the response, later. By using a lagged response variable in the model, we account for autocorrelation in the response variable. Table 6.2 presents the results of the same exercise with the same sample, conducted with one-year lagged values. This suggests that the independent variables in this second regression effort are based on data from the prior year rather than the current year. Once again, the table provides the results of regression analyses with four different models (columns (1) to (4)). The dependent variable across all models is the average yearly rating on TripAdvisor for hotels for the years 2007 to 2016. Each coefficient represents the change in the average yearly rating for a unit change in the predictor variable, assuming all other predictors are held constant. Almost all variables show a change in their coefficient values between Table 6.1 and Table 6.2. Some changes are minor, while others are more substantial. Some variables show a change in their significance level. Moreover, the standard errors for some variables have changed. This is expected since the data set and model structure have been modified. Furthermore, as it is expected, the number of observations and hotels has changed, indicating that the dataset for Table 6.2 (with lagged regressors) is slightly smaller.

Taking into consideration the Goodness-of-fit in Table 6.2 (R-squared), Model (1) has an R-squared value of 0.1178, indicating that it explains about 11.78% of the variability in the dependent variable. Model (2) slightly improves on this with an R-squared of 0.1334, Model

(3) has a similar R-squared value of 0.1305 while Model (4) has the highest R-squared value of 0.1345, suggesting that, among the four models, it does the best job explaining the variation in TripAdvisor ratings using the given variables. R-squared values of Table 6.1 and Table 6.2 are very close indicating that one is expecting no dramatic changes in the outcome.

Overall, introducing one-year lagged regressors has changed the model dynamics, which is reflected in different coefficient values, significance levels, and model fits. This is typical when introducing lagged variables, as they account for the impact of past values on the current outcome.

Moving to the interpretation of the variables and considering L.lTA (Total Assets - log): For every unit increase in the log of total assets, there's a decrease of between -0.574 to -0.602 in the average yearly rating, holding other variables constant. This relationship is statistically significant at the 1% level across all models. Consequently, lagged values intensify the findings of Table 6.61, exhibiting a reverse relationship between the magnitude of total assets and rating on TripAdvisor. Once again, as was the case in Table 6.1, the Total Assets - log squared variable presents a positive coefficient, indicating a quadratic relationship with total assets. This suggests that as total assets increase, the decline in ratings becomes less severe or may even reverse at higher levels.

Log of ratio turnover/number of beds) indicates that a unit increase in the log ratio of turnover to the number of beds leads to an increase in ratings between 11% to 14%, statistically significant at the 1% level. Results are consistent with previous findings in Table 6.1, showcasing that hotels with higher turnover per bed tend to have higher TripAdvisor ratings.

TurnFA_w (turnover/fixed assets, winsorized) is negative and becomes significant at a 1% level in all four models of the lagged values regression version in Table 6.2, implying that as the turnover to fixed assets ratio increases, the customer's satisfaction as expressed by TripAdvisor's rating tends to decrease. The latter implies that the larger the hotel's premises the visitor's satisfaction slightly declines. The finding may be reflective of a faceless establishment feel when a customer visits a sprawled hotel with multiple (sometimes similar) compounds, where the value of personal touch may be lost. Most of the time, such types of hotels and resorts are overcrowded. Especially during high season, service is standard or below standard, and infrastructure (such as pool areas) is poorly cleaned. Considering that Greece has a large number of this type of hotels, defined as "sea and sun" accommodation, especially in the touristic developed islands such as Rhodes, Crete, Kos, and Corfu, the arguments seem plausible. The type of hotel that is located near a sandy beach, receives clientele through the major tour operating channels such as TUI, it refers to itself as a "touristic village" where, all compounds and rooms look alike, and they keep an average level of services. Having the "sun and sea" as a given, they compete on a price basis which leads to squeezing the offered quality of services.

L.lRatingcount (Number of ratings - log): Indicates a positive relationship between the number of ratings and the average yearly rating. Hotels with more ratings (in log terms) tend to have a statistically higher rating, significant at the 1% level. The finding in Table 6.62 using lagged values is in line with the outcome of Table 6.1.

L.CTI_w and L.SalesTI_w: These variables indicate expenses as a fraction of turnover. While L.CTI_w is not significant in any model, L.SalesTI_w is significant in models 1-4, indicating that as the ratio of Sales Expenses to Turnover increases, the TripAdvisor rating might also increase. Results are in accordance with those in Table 6.1.

Considering lFA_d (year-on-year first difference of Fixed Assets), in contrast with no lagged values models, no significant relationship is observed. An interpretation of the finding is, that the lag effect was already considered in the year-on-year calculation, and therefore, taking the lag biases the results.

Considering log change in Administrative Expenses / Turnover and Sales Expenses / Turnover neither shows a statistically significant relationship with the TripAdvisor rating in the models they're included. The results are in consonant with that of Table 6.1.

Pool and FamFriendly: Hotels with a pool have significantly lower ratings, while family-friendly hotels have significantly higher ratings. More precisely, hotels having a pool have ratings between -14% to -15% lower than those not having a pool, significant at the 1% level. The negative relationship is significant across all models. Findings are in full accordance with no-lagged values regression models in Table 6.61. Family-friendly hotels have between 59% to 61% higher ratings than those that are not family-friendly, significant at the 1% level. The findings are in line with the outcome of Table 6.1.

AvSpendpn and AvSpendpp represent average spending by visitors per day and trip. Both show varied significance across models. There's a positive relationship between the average amount spent per overnight stay and the rating, significant at the 5% and 1% levels for models 1-3 and model 4 respectively. Lagged values operated as an amplifier of significance across the four models in Table 6.1. Meanwhile, the relationship between the average amount spent per travel and the rating is negative in the model (4) and significant at the 1% level. Again, results are in line with no-lagged regressions.

Yearly Dummies (2011. YEAR to 2014. YEAR) show the difference in the average yearly rating for the specified year, compared to the base year, i.e., 2007. There's a positive and significant relationship for the years 2011 to 2013, implying ratings in these years were higher relative to 2007, in consistency with results in Table 6.1.

Overall, the main findings for no-lagged regressions hold for lagged data regressed for the variables under consideration. Comparing the two tables we conclude that almost all variables show a change in their coefficient values between them. Some changes are minor, while others are more substantial. For example, the coefficient for Lta in the model (1) of the first table is -0.382 and in the second table is -0.575. Some variables show a change in their significance level. For instance, in the first table, AvSpendpn showcases a stronger significance in all four models in Table 6.2 in comparison to Table 6.1. Moreover, the standard errors for some variables have changed. This is expected since the data set and model structure have been modified. Finally, the change in R-squared values between the two sets of regressions signifies that the models' ability to explain the variance in the dependent variable has shifted when lagged regressors were incorporated. The difference in R-squared values could also suggest that the lagged values of some variables might be more relevant or provide additional information in explaining the dependent variable than their contemporaneous values. It's essential to bear in mind, though, that a higher R-squared doesn't necessarily mean a better model.

Table 6.3. presents the stratified by star category (3, 4, and 5-star) regression results across the four different models (1, 2, 3, and 4).

Table 6.3. Regression results stratified by hotel star ranking.

	3 stars	4 stars	5 stars	3 stars	4 stars	5 stars	3 stars	4 stars	5 stars	3 stars	4 stars	5 stars
Variable/Model	*(25)*	*(26)*	*(27)*	*(28)*	*(29)*	*(30)*	*(31)*	*(32)*	*(33)*	*(34)*	*(35)*	*(36)*
lTA	-0.349	-0.824***	-0.213	-0.192	-0.871***	-0.0542	-0.518	-0.937***	-0.00365	-0.504	-0.919***	0.00136
	(0.472)	(0.316)	(0.306)	(0.536)	(0.286)	(0.288)	(0.498)	(0.298)	(0.298)	(0.502)	(0.298)	(0.304)
lTAsq	0.00808	0.0232**	0.00463	0.00279	0.0240***	-0.000250	0.0127	0.0256***	-0.00159	0.0123	0.0251***	-0.00174
	(0.0153)	(0.00986)	(0.00907)	(0.0176)	(0.00892)	(0.00855)	(0.0161)	(0.00928)	(0.00877)	(0.0163)	(0.00928)	(0.00891)
TurnBeds_l	0.128***	0.0881**	0.114***	0.119***	0.120***	0.133***	0.130***	0.158***	0.139***	0.133***	0.160***	0.142***
	(0.0446)	(0.0372)	(0.0309)	(0.0416)	(0.0366)	(0.0304)	(0.0440)	(0.0397)	(0.0318)	(0.0441)	(0.0402)	(0.0319)
TurnFA_w	-0.290***	0.0147	-0.0610*	-0.282***	0.00281	-0.0647*	-0.305***	0.00435	-0.0608*	-0.303***	0.0108	-0.0582
	(0.0562)	(0.0434)	(0.0329)	(0.0594)	(0.0493)	(0.0346)	(0.0609)	(0.0454)	(0.0359)	(0.0618)	(0.0439)	(0.0367)
lRatingcount	0.111***	0.0912***	0.109***	0.0922***	0.0972***	0.128***	0.110***	0.0908***	0.128***	0.103***	0.0822***	0.127***
	(0.0281)	(0.0186)	(0.0231)	(0.0290)	(0.0203)	(0.0259)	(0.0293)	(0.0203)	(0.0258)	(0.0314)	(0.0211)	(0.0269)
CTI_w	0.126	-0.240	-0.0592	0.0478	-0.160	-0.0450	0.0455	-0.116	0.0153	0.0425	-0.110	0.00699
	(0.176)	(0.171)	(0.162)	(0.190)	(0.175)	(0.182)	(0.195)	(0.218)	(0.184)	(0.197)	(0.219)	(0.191)
SalesTI_w	0.191	0.411***	0.0746	0.278**	0.339***	0.0796	0.250*	0.312***	0.0759	0.257*	0.307**	0.0638
	(0.144)	(0.106)	(0.186)	(0.139)	(0.111)	(0.166)	(0.143)	(0.121)	(0.180)	(0.147)	(0.121)	(0.185)
lFA_d				0.178	0.0928**	-0.00150	0.165	0.0904**	0.00607	0.174	0.0840*	0.0119
				(0.145)	(0.0408)	(0.0390)	(0.158)	(0.0431)	(0.0399)	(0.161)	(0.0437)	(0.0422)
CTI_dl							-0.0342	-0.0195	-0.0248	-0.0332	-0.0137	-0.0216
							(0.0478)	(0.0438)	(0.0385)	(0.0478)	(0.0430)	(0.0376)
SalesTI_dl							0.0486	-0.0167	0.0222	0.0501	-0.0165	0.0203
							(0.0339)	(0.0338)	(0.0424)	(0.0345)	(0.0331)	(0.0422)
Pool	-0.181**	-0.124*	-0.200*	-0.197**	-0.141**	-0.192*	-0.182*	-0.165**	-0.183*	-0.183*	-0.166**	-0.182*
	(0.0910)	(0.0653)	(0.108)	(0.0983)	(0.0659)	(0.107)	(0.0999)	(0.0671)	(0.108)	(0.100)	(0.0671)	(0.108)
FamFriendly	0.753***			0.993***			1.008***			1.001***		

	(0.0967)			(0.111)			(0.119)			(0.123)		
AvSpendpn	0.0112	0.00590	0.00235	0.0138	0.00440	0.000915	0.00994	0.00348	0.00196	0.0136	0.00696	0.00542
	(0.00965)	(0.00617)	(0.00497)	(0.00940)	(0.00584)	(0.00504)	(0.00979)	(0.00585)	(0.00530)	(0.0101)	(0.00606)	(0.00580)
AvSpendpp	0.000239	-0.000328	0.000152	-1.53e-05	-0.000187	0.000274	0.000296	-0.000126	0.000194	-0.000303	-0.00125**	4.38e-06
	(0.000491)	(0.000375)	(0.000305)	(0.000484)	(0.000385)	(0.000306)	(0.000498)	(0.000384)	(0.000330)	(0.000709)	(0.000569)	(0.000512)
2011.YEAR										0.112	0.140**	0.0482
										(0.0749)	(0.0553)	(0.0515)
2012.YEAR										0.0529	0.154***	0.0276
										(0.0738)	(0.0534)	(0.0421)
2013.YEAR										0.0557	0.110*	-0.0154
										(0.0791)	(0.0584)	(0.0524)
2014.YEAR										0.0178	-0.0111	-0.0205
										(0.0576)	(0.0388)	(0.0312)
2015o.YEAR										-	-	-
2016o.YEAR										-	-	-
Constant	5.241	10.06***	4.896*	4.183	10.31***	3.375	6.788*	10.65***	2.811	6.716*	10.87***	2.613
N	936	1,325	582	842	1,240	563	801	1,170	550	801	1,170	550
No. hotels	196	248	96	193	247	96	184	241	95	184	241	95
R-squared	0.1321	0.118	0.2496	0.1374	0.1391	0.2731	0.1571	0.1409	0.2793	0.1573	0.1471	0.2809

Notes: Heteroskedasticity robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Considering firstly Total Assets (log) it is observed a negative relationship across all star categories, with 4-star hotels showing significance (***p<0.01) in all models. This suggests that as total assets increase, the rating on TripAdvisor decreases for 4-star hotels. Additionally, Total Assets (log) Squared present a positive relationship across all star categories, with 4-star hotels again showing significance (**p<0.05, ***p<0.01) in most models, possibly due to the quadratic relationship between total assets and the dependent variable, especially for 4-star hotels.

Concerning the Turnover per Bed (log) variable, a consistently positive and significant relationship across all star categories and models (***p<0.01) is highlighted by the regression results. The latter means that as the ratio of turnover to the number of beds increases, the rating on TripAdvisor also increases.

As far as it concerns Turnover/Fixed Assets have a negative relationship for 3 and 5-star hotels (significant at ***p<0.01 and *p<0.1 respectively) and no significant relationship for 4-star hotels. This suggests that an increase in the turnover to fixed assets ratio leads to a decrease in TripAdvisor's ratings for 3 and 5-star hotels.

Elaborating on the Number of Ratings (log) (lRatingcount) variable a positive and significant relationship across all star categories (***p<0.01) is emerging, indicating that more ratings correspond to an increase in customer satisfaction as it is expressed by the TripAdvisor score of ratings of the hotel. Administrative Expenses/Turnover variable has no significance observed across all star categories. In contrast, the Sales Expenses/Turnover variable presents a positive and significant relationship, especially for 4-star hotels (***p<0.01). Consequently, for 4-star hotels an increase in the sales expenses to turnover ratio corresponds to an increase in the ratings of TripAdvisor.

Year-on-Year First Difference of Fixed Assets (lFA_d) variable showcases a positive relationship for 4-star hotels in models 29, 32, and 35 (*p<0.1, **p<0.05).

Both, Administrative Expenses/Turnover Year-on-Year Log Change and Sales Expenses/Turnover Year-on-Year Log Change present no consistent significant relationship observed.

The Presence of a Pool is negatively related and significant across all star categories in all models at levels *p<0.1 and **p<0.05. As was already identified in the previous analysis, having a pool is associated with a decrease in the ratings of TripAdvisor.

Family Friendly (FamFriendly) variable is positive and significant for 3-star hotels across all relevant models (***p<0.01), indicating that family-friendly 3-star hotels tend to have a higher rating on TripAdvisor.

Average Spending per Travel and Average Spending per Night variables present no consistent significant relationship observed.

For Year Dummies (2011. YEAR to 2014. YEAR), coefficients capture year-specific effects, positive and significant for 4-star hotels in 2012 (***p<0.01) and 2013 (*p<0.1).

Finally, R-squared varies from 0.118 to 0.2809 across all models and star categories, indicating that the models explain between 11.8% and 28.09% of the variance in the dependent variable. Moreover, R-squared values show that the models explain more of the variance in the dependent variable for 5-star hotels, followed by 3-star and 4-star hotels.

Overall, 4-star hotels show significant relationships across more variables, suggesting that factors like Total Assets, Total Assets Squared, Turnover per Bed, Number of Ratings, Sales Expenses/Turnover, and certain years impact them more significantly. 3-star hotels have fewer significant variables, with the most prominent being Family-Friendly. 5-star hotels show fewer significant relationships, but the models have higher R-squared values, indicating that the included variables together explain more of the variation for 5-star hotels compared to 3-star and 4-star hotels.

Table 6.4. presents the stratified by total assets regression results across the four different models (1, 2, 3, and 4).

Table 6.4. Regression results stratified by hotel total assets.

	Small	Large	Small	Large	Small	Large	Small	Large
Variable/models	*(9)*	*(10)*	*(11)*	*(12)*	*(13)*	*(14)*	*(15)*	*(16)*
lTA	-1.040	-1.275*	-0.651	-1.240*	-0.964	-1.153*	-0.946	-1.136*
	(0.827)	(0.705)	(0.835)	(0.657)	(0.848)	(0.663)	(0.854)	(0.661)
lTAsq	0.0330	0.0355*	0.0190	0.0340*	0.0288	0.0317*	0.0284	0.0312*
	(0.0283)	(0.0200)	(0.0286)	(0.0187)	(0.0291)	(0.0188)	(0.0293)	(0.0188)
TurnBeds_l	0.115***	0.105***	0.118***	0.142***	0.139***	0.138***	0.144***	0.141***
	(0.0314)	(0.0391)	(0.0285)	(0.0373)	(0.0309)	(0.0379)	(0.0314)	(0.0382)
TurnFA_w	-0.101**	0.0681	-0.104**	0.0687	-0.107**	0.0727	-0.102**	0.0770
	(0.0449)	(0.0498)	(0.0438)	(0.0614)	(0.0456)	(0.0624)	(0.0453)	(0.0611)
lRatingcount	0.107***	0.101***	0.106***	0.104***	0.111***	0.0980***	0.103***	0.0928***
	(0.0177)	(0.0213)	(0.0187)	(0.0252)	(0.0188)	(0.0252)	(0.0198)	(0.0264)
CTI_w	-0.0627	0.0371	-0.0549	-0.0907	-0.00584	-0.00905	-0.00132	-0.0182
	(0.122)	(0.138)	(0.127)	(0.236)	(0.145)	(0.241)	(0.145)	(0.250)
SalesTI_w	0.301***	0.276	0.295***	0.249	0.270***	0.271	0.272***	0.277
	(0.0887)	(0.209)	(0.0870)	(0.230)	(0.0909)	(0.242)	(0.0912)	(0.243)
lFA_d			0.113**	0.0533	0.120**	0.0615	0.109**	0.0670
			(0.0469)	(0.0548)	(0.0492)	(0.0592)	(0.0493)	(0.0602)
CTI_dl					-0.0232	-0.0592	-0.0207	-0.0578
					(0.0329)	(0.0423)	(0.0327)	(0.0420)
SalesTI_dl					0.0152	-0.00591	0.0151	-0.00659
					(0.0293)	(0.0242)	(0.0287)	(0.0240)
Pool	-0.129**	-0.283***	-0.139**	-0.271***	-0.146**	-0.274***	-0.148**	-0.272***
	(0.0587)	(0.0899)	(0.0607)	(0.0887)	(0.0619)	(0.0916)	(0.0620)	(0.0912)
FamFriendly	0.682***		0.931***		0.957***		0.932***	

	(0.0689)		(0.0736)		(0.0774)		(0.0801)	
AvSpendpn	0.00693	0.00621	0.00752	0.00403	0.00601	0.00441	0.0108*	0.00459
	(0.00606)	(0.00445)	(0.00578)	(0.00473)	(0.00590)	(0.00493)	(0.00606)	(0.00537)
AvSpendpp	0.000155	-0.000158	7.82e-05	1.76e-05	0.000238	-0.000138	-0.000761	-0.000434
	(0.000342)	(0.000301)	(0.000348)	(0.000297)	(0.000347)	(0.000307)	(0.000496)	(0.000471)
2011.YEAR							0.155***	0.0147
							(0.0497)	(0.0425)
2012.YEAR							0.124***	0.0529
							(0.0478)	(0.0423)
2013.YEAR							0.0859*	0.0281
							(0.0520)	(0.0502)
2014.YEAR							0.00420	-0.0194
							(0.0362)	(0.0303)
2015.YEAR							-	-
2016.YEAR							-	-
Constant	10.46*	13.91**	7.765	13.43**	10.05	12.75**	10.06	12.73**
N	1,950	893	1,785	860	1,686	835	1,686	835
No. hotels	409	171	403	168	387	165	387	165
R-squared	0.1151	0.2209	0.1284	0.2435	0.136	0.2375	0.14	0.2389

Notes: Heteroskedasticity robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

While the coefficient of Total Assets (log) for Small Hotels is negative across all models but not statistically significant, for Large Hotels the negative coefficient is significant in all models ($p < 0.05$), indicating that an increase in total assets is associated with a decrease on the TripAdvisor rating. At the same time, Total Assets Squared present for Small Hotels a positive coefficient but not statistically significant, and for Large Hotels a positive and significant in all models ($p < 0.05$) coefficient, suggesting a nonlinear, U shape relationship between total assets and the TripAdvisor rating.

As far as it concerns the Turnover per Bed variable, both Small and Large Hotels show a positive and highly significant coefficient ($p < 0.001$), implying that an increase in turnover per bed correlates with an increase in TripAdvisor rating. Moreover, there is a significant U-shaped relationship between total assets and the dependent variable for Large hotels, which is not evident for Small hotels. The relationship is consistent across both categories, indicating that turnover per bed is a strong predictor for both Large and Small hotels.

In terms of the Turnover/Fixed Assets variable, for Small Hotels the coefficient is negative and significant ($p < 0.05$), while for Large Hotels the coefficient is positive but not significant. A positive and highly significant relationship ($p < 0.001$) is shown for both Small and Large Hotels, as far as it concerns the Number of Ratings. Administrative Expenses/Turnover variable has no consistent significance across both Small and Large Hotels.

Conversely, Sales Expenses/Turnover presents a positive and highly significant ($p < 0.001$) coefficient for Small Hotels. Large Hotels also have a positive but not consistently significant coefficient.

Year-on-year difference in Fixed Assets is positive and significant $p < 0.05$) for Small Hotels but not significant for Large Hotels.

Year-on-year log change in Administrative Expenses/Turnover and Sales Expenses/Turnover presents no consistent significance in either Small or Large Hotels.

Having the hotel, a Pool is for both Small and Large Hotels negative and significant ($p < 0.05$ for Small Hotels and $p < 0.001$ for Large Hotels). Having a pool is associated with lower values of the dependent variable for both, but more intensified for Large hotels.

Moving to the Family Friendly characterization, Small Hotels have a positive and highly significant ($p < 0.001$) TripAdvisor rating.

Average Spending per Travel is not significant in any model while Average Spending per Night is positive and significant for Small Hotels in Model 16 ($p < 0.05$).

Year Dummy Variables (2011. YEAR, 2012. YEAR, etc.) is significant and positive for Small Hotels for 2011.YEAR, 2012.YEAR, and 2013.YEAR in Model 15 ($p < 0.05$ or better) while for Large Hotels is not significant.

Both Small and Large Hotels have a significant constant in most models ($p < 0.05$ or better). Finally, the R-squared values indicate that the models explain more of the variance in the dependent variable for Large Hotels compared to Small Hotels.

In total, Large Hotels have more significant relationships across variables like Total Assets, Total Assets Squared, Turnover per Bed, and Number of Ratings. Small Hotels show significant coefficients for Turnover per Bed, Number of Ratings, Turnover/Fixed Assets, Sales Expenses/Turnover, and Family-Friendly. The R-squared values suggest that the models fit Large Hotels slightly better than Small Hotels. Total assets, turnover per bed, number of ratings, and the presence of a pool are consistently significant predictors across the two categories. The squared term of total assets indicates a potential non-linear relationship with the dependent variable, especially for Large hotels. Notably, family-friendly status significantly affects small hotels.

Table 6.5. presents the regression results stratified by the number of hotel beds into two categories: Large and Small hotels.

Table 6.5. Regression results stratified by hotel beds

Variable/models	Small *(17)*	Large *(18)*	Small *(19)*	Large *(20)*	Small *(21)*	Large *(22)*	Small *(23)*	Large *(24)*
lTA	-0.719**	0.284	-0.671**	0.154	-0.862**	0.350	-0.840**	0.365
	(0.333)	(0.700)	(0.327)	(0.568)	(0.339)	(0.527)	(0.343)	(0.528)
lTAsq	0.0225**	-0.00759	0.0202*	-0.00431	0.0260**	-0.0102	0.0253**	-0.0106
	(0.0108)	(0.0202)	(0.0106)	(0.0163)	(0.0111)	(0.0152)	(0.0112)	(0.0152)
TurnBeds_l	0.0847***	0.100*	0.101***	0.116**	0.114***	0.140***	0.118***	0.145***
	(0.0306)	(0.0532)	(0.0293)	(0.0489)	(0.0308)	(0.0514)	(0.0316)	(0.0520)
TurnFA_w	-0.110***	0.0951*	-0.114***	0.0894	-0.113***	0.0713	-0.107**	0.0692
	(0.0410)	(0.0569)	(0.0400)	(0.0670)	(0.0414)	(0.0676)	(0.0417)	(0.0661)
lRatingcount	0.111***	0.104***	0.106***	0.106***	0.112***	0.0995***	0.104***	0.0923***
	(0.0174)	(0.0215)	(0.0188)	(0.0243)	(0.0191)	(0.0234)	(0.0202)	(0.0240)
CTI_w	-0.138	0.126	-0.120	0.0857	-0.0691	0.0376	-0.0686	0.0300
	(0.123)	(0.130)	(0.128)	(0.199)	(0.149)	(0.237)	(0.149)	(0.240)
SalesTI_w	0.264***	0.315	0.268***	0.302	0.246***	0.145	0.251***	0.136
	(0.0900)	(0.209)	(0.0877)	(0.240)	(0.0915)	(0.268)	(0.0921)	(0.270)
lFA_d			0.0995**	0.0378	0.102**	0.0361	0.0946**	0.0342
			(0.0400)	(0.0652)	(0.0423)	(0.0682)	(0.0425)	(0.0704)
CTI_dl					-0.0159	-0.0596	-0.0118	-0.0601
					(0.0326)	(0.0460)	(0.0323)	(0.0442)
SalesTI_dl					0.00419	0.0277	0.00415	0.0286
					(0.0280)	(0.0301)	(0.0276)	(0.0297)
Pool	-0.125**	-0.230***	-0.138**	-0.228***	-0.141**	-0.244***	-0.144**	-0.242***
	(0.0618)	(0.0835)	(0.0644)	(0.0816)	(0.0660)	(0.0785)	(0.0663)	(0.0781)
FamFriendly	0.648***		0.902***		0.922***		0.898***	

	(0.0673)		(0.0726)		(0.0773)		(0.0796)	
AvSpendpn	0.00697	0.00437	0.00733	0.00556	0.00482	0.00799	0.00851	0.0115*
	(0.00550)	(0.00672)	(0.00532)	(0.00611)	(0.00543)	(0.00632)	(0.00569)	(0.00607)
AvSpendpp	5.86e-05	9.47e-05	3.61e-06	8.15e-05	0.000189	-5.11e-05	-0.000684	-0.000616
	(0.000311)	(0.000400)	(0.000319)	(0.000372)	(0.000321)	(0.000368)	(0.000471)	(0.000526)
2011.YEAR							0.129***	0.0712
							(0.0481)	(0.0468)
2012.YEAR							0.102**	0.0950**
							(0.0467)	(0.0464)
2013.YEAR							0.0813	0.0306
							(0.0496)	(0.0606)
2014.YEAR							-0.00672	0.00984
							(0.0330)	(0.0369)
2015.YEAR							-	-
2016.YEAR							-	-
Constant	8.370***	-0.185	8.041***	0.836	9.533***	-1.062	9.549***	-1.144
N	1,991	852	1,829	816	1,732	789	1,732	789
No. hotels	399	148	394	148	379	145	379	145
R-squared	0.1311	0.2024	0.1433	0.2073	0.1484	0.2209	0.1508	0.2251

Notes: Heteroskedasticity robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Considering table 6.5., it is observed a negative relationship between the number of beds (Large – Small hotels) and the dependent variable in all models (significant at 5% level) for Small Hotels and a positive relationship but not statistically significant across all models for Large Hotels. For Total Assets Squared (lTAsq), as far as it concerns Small Hotels, there's a positive relationship across all models. The relationship is significant at a 5% level in most models and at a 10% level in one. For Large Hotels, the relationship is negative but not significant across all models.

Concerning the Turnover per Bed variable, for both Small and Large hotels, there's a positive and significant relationship between turnover per bed and the TripAdvisor rating in all models.

Regarding the Turnover/Fixed Assets variable, Small Hotels present a negative and significant relationship across all models, while Large Hotels confirm a positive relationship, significant in one of the models.

For both categories, there's a positive and highly significant relationship in all models regarding the Number of Ratings.

Administrative Expenses/Turnover showcase no significant relationship for both Large and Small hotels in terms of beds, across all models.

The sales Expenses/Turnover variable is positively related and significant in all models for Small Hotels but not significant for Large Hotels.

Year-on-year difference in Fixed Assets is positive and significant at a 5% value for Small Hotels and not significant for Large Hotels.

Administrative Expenses/Turnover YoY Log Change and Sales Expenses/Turnover YoY Log Change variables are not significant for both categories of hotels under investigation. The same non-significant finding holds for the Sales Expenses/Turnover variable.

Having a pool showcases the same pattern as in previous stratifications of the sample, i.e., a negative and significant relationship with the dependent variable across all models. Being Family Friendly is observed to be similar in other categorizations with Small Hotels to present a strong positive and significant relationship with TripAdvisor's rating score.

Average Spending per Night and Travel shows no significant relationship with the dependent variable.

Regarding Year Dummy Variables, Small Hotels have a positive and significant relationship for 2011 and 2012, while other years aren't significant while Large Hotels have a positive relationship for 2012 and insignificant for other years.

Finally, for Small hotels, the models explain between 13.11% to 15.08% of the variance in the dependent variable, and for Large hotels, the models explain between 20.24% to 22.51% of the variance in the dependent variable, indicating a better fit for Large hotels compared to Small hotels.

In summary, Turnover per bed and number of ratings are consistently significant predictors for both Small and Large hotels. Large hotels' models generally have higher explanatory power (R-squared values) compared to Small hotels. While some variables like Turnover per Bed, Number of Ratings, and having a Pool show similar effects across both categories, there are distinct differences in how Total Assets, Turnover/Fixed Assets, and Family Friendly status impact Small and Large hotels. Regression models for Large hotels tend to have a higher explanatory power compared to Small hotels. While some variables like TurnBeds_l, lRatingcount, and Pool exhibit consistent behavior across both Small and Large hotels, others like lTA, TurnFA_w, and lFA_d show contrasting effects between the two categories.

Table 6.6. presents the regression results stratified by Total Assets and hotel star rating (3-4-5 stars hotels).

Table 6.6. Regression results stratified by Total Assets and hotel star rating

Variables /models	*Small & 3 stars (81)*	*Large & 3 stars (82)*	*Small & 4 stars (83)*	*Large & 4 stars (84)*	*Small & 5 stars (85)*	*Large & 5 stars (86)*	*Small & 3 stars (87)*	*Large & 3 stars (88)*	*Small & 4 stars (89)*	*Large & 4 stars (90)*	*Small & 5 stars (91)*	*Large & 5 stars (92)*	*Small & 3 stars (93)*	*Large & 3 stars (94)*	*Small & 4 stars (95)*	*Large & 4 stars (96)*	*Small & 5 stars (97)*	*Large & 5 stars (98)*
lTA	-0.733	-5.090	-1.696*	-3.805***	-0.715	0.290	-0.150	-5.119	-1.647	-3.245***	1.599	-0.0137	-0.595	-9.808*	-1.805*	-3.526***	1.328	0.138
	(1.202)	(6.821)	(0.973)	(1.214)	(2.246)	(0.758)	(1.236)	(6.997)	(1.002)	(1.149)	(2.316)	(0.806)	(1.233)	(5.632)	(1.024)	(1.194)	(2.855)	(0.808)
lTAsq	0.0221	0.146	0.0536	0.108***	0.0209	-0.00907	0.00190	0.146	0.0509	0.0908***	-0.0554	-0.000848	0.0159	0.279*	0.0554	0.0988***	-0.0464	-0.00500
	(0.0417)	(0.192)	(0.0334)	(0.0347)	(0.0762)	(0.0214)	(0.0428)	(0.197)	(0.0345)	(0.0330)	(0.0787)	(0.0227)	(0.0427)	(0.159)	(0.0353)	(0.0344)	(0.0968)	(0.0227)
TurnBeds_l	0.123***	0.0435	0.1000**	0.0783	0.221***	0.119***	0.113***	0.0406	0.125***	0.143**	0.233***	0.120***	0.121***	-0.0492	0.166***	0.127*	0.254***	0.122***
	(0.0472)	(0.0892)	(0.0443)	(0.0759)	(0.0546)	(0.0361)	(0.0424)	(0.0926)	(0.0421)	(0.0684)	(0.0543)	(0.0372)	(0.0458)	(0.0777)	(0.0506)	(0.0715)	(0.0473)	(0.0374)
TurnFA_w	-0.285***	-0.0267	-0.00399	0.240*	-0.137***	-0.00253	-0.274***	-0.0266	-0.0152	0.159	-0.127***	0.0256	-0.301***	-0.232	-0.00989	0.161	-0.132***	0.0312
	(0.0586)	(0.560)	(0.0486)	(0.136)	(0.0252)	(0.0343)	(0.0620)	(0.573)	(0.0522)	(0.131)	(0.0245)	(0.0518)	(0.0632)	(0.482)	(0.0503)	(0.135)	(0.0218)	(0.0548)
lRatingcount	0.105***	0.0698	0.0915***	0.0784**	0.0834*	0.101***	0.0848***	0.0652	0.102***	0.0607*	0.0858*	0.132***	0.104***	-0.0174	0.0954***	0.0513	0.0959***	0.130***
	(0.0298)	(0.147)	(0.0226)	(0.0319)	(0.0469)	(0.0271)	(0.0319)	(0.155)	(0.0241)	(0.0350)	(0.0470)	(0.0321)	(0.0322)	(0.102)	(0.0241)	(0.0353)	(0.0365)	(0.0322)
CTI_w	0.140	1.278	-0.201	-0.110	-0.466***	0.145	0.0649	1.243	-0.0908	-0.731	-0.421***	0.126	0.0502	0.774	-0.0509	-0.779	-0.358	0.186
	(0.180)	(0.829)	(0.185)	(0.332)	(0.145)	(0.122)	(0.199)	(0.916)	(0.179)	(0.482)	(0.153)	(0.142)	(0.204)	(0.582)	(0.225)	(0.549)	(0.237)	(0.140)
SalesTI_w	0.190	3.489***	0.416***	0.386	0.193	-0.0633	0.270*	3.524***	0.334***	0.387	0.150	-0.0105	0.235	4.039***	0.310**	0.466	0.186	-0.0632
	(0.149)	(1.003)	(0.115)	(0.267)	(0.171)	(0.305)	(0.145)	(1.081)	(0.120)	(0.300)	(0.158)	(0.327)	(0.147)	(0.612)	(0.131)	(0.302)	(0.147)	(0.365)
lFA_d							0.177	0.0503	0.0812*	0.0646	-0.0651	0.0547	0.172	0.352	0.0809	0.0298	-0.0643	0.0614
							(0.151)	(0.344)	(0.0487)	(0.100)	(0.114)	(0.0551)	(0.166)	(0.346)	(0.0502)	(0.117)	(0.122)	(0.0577)
CTI_dl													-0.0368	-0.0446	-0.00371	-0.0524	-0.0269	-0.0487
													(0.0487)	(0.166)	(0.0507)	(0.0697)	(0.0885)	(0.0480)
SalesTI_dl													0.0520	-0.299***	-0.0142	-0.0351	0.000734	0.0322
													(0.0354)	(0.0827)	(0.0443)	(0.0281)	(0.0895)	(0.0490)
Pool	-0.176*	-1.594***	-0.0720	-0.237**	-0.408	-0.205*	-0.186*	-1.620***	-0.0973	-0.236**	-0.407	-0.208*	-0.173*	-2.177***	-0.130*	-0.222**	-0.314	-0.208*
	(0.0941)	(0.455)	(0.0749)	(0.111)	(0.308)	(0.109)	(0.102)	(0.536)	(0.0756)	(0.108)	(0.304)	(0.110)	(0.103)	(0.324)	(0.0780)	(0.111)	(0.317)	(0.109)
FamFriendly	0.734***						0.972***						1.000***					
	(0.105)						(0.121)						(0.129)					
AvSpendpn	0.0113	0.0544**	0.00962	0.00716	0.0163*	0.00688	0.0157	0.0555**	0.0105	0.00631	0.0173*	0.00201	0.0132	0.0301	0.00901	0.00368	0.0200*	0.00352
	(0.0105)	(0.0261)	(0.00842)	(0.00789)	(0.00982)	(0.00678)	(0.0103)	(0.0260)	(0.00808)	(0.00861)	(0.00984)	(0.00693)	(0.0108)	(0.0386)	(0.00780)	(0.00855)	(0.0104)	(0.00715)
AvSpendpp	-0.000516	-0.00243	-0.00133*	-0.000813	-0.000163	-0.000423	-0.000721	-0.00254	-0.00150**	-0.00105	-0.000337	0.000133	-0.000386	-0.00472*	-0.00131*	-0.00110	-0.000433	7.66e-05
	(0.000710)	(0.00318)	(0.000696)	(0.000705)	(0.00104)	(0.000574)	(0.000711)	(0.00332)	(0.000737)	(0.000728)	(0.00104)	(0.000618)	(0.000743)	(0.00259)	(0.000735)	(0.000758)	(0.00120)	(0.000630)

2011.YEAR	0.128*	-0.0615	0.173**	0.00220	0.252*	0.0578	0.102	-0.0501	0.197***	0.0507	0.202	0.0195	0.119	-0.0444	0.183***	0.0262	0.183	0.00473
	(0.0748)	(0.232)	(0.0738)	(0.0725)	(0.136)	(0.0512)	(0.0741)	(0.221)	(0.0667)	(0.0726)	(0.147)	(0.0535)	(0.0777)	(0.208)	(0.0710)	(0.0738)	(0.145)	(0.0529)
2012.YEAR	0.0575	0.155	0.158**	0.0587	0.0510	0.0700	0.0627	0.161	0.196***	0.0818	0.0866	0.0236	0.0495	0.408**	0.187***	0.0816	0.0791	0.0186
	(0.0749)	(0.242)	(0.0685)	(0.0697)	(0.102)	(0.0469)	(0.0719)	(0.259)	(0.0644)	(0.0746)	(0.103)	(0.0495)	(0.0768)	(0.175)	(0.0690)	(0.0757)	(0.104)	(0.0512)
2013.YEAR	0.111	-0.0203	0.0855	0.0807	0.0663	-0.00983	0.0797	-0.0153	0.101	0.137	0.0929	-0.0416	0.0664	0.198	0.0937	0.126	0.0781	-0.0594
	(0.0763)	(0.403)	(0.0674)	(0.0903)	(0.0956)	(0.0595)	(0.0798)	(0.414)	(0.0669)	(0.0836)	(0.0906)	(0.0614)	(0.0827)	(0.267)	(0.0742)	(0.0848)	(0.0979)	(0.0637)
2014.YEAR	0.0448	-0.0179	-0.0316	0.0175	0.0886	-0.0200	0.0234	-0.0147	-0.0198	0.00588	0.0989	-0.0367	0.0251	0.0825	-0.0236	0.00972	0.0971	-0.0405
	(0.0607)	(0.106)	(0.0483)	(0.0523)	(0.0955)	(0.0337)	(0.0595)	(0.119)	(0.0488)	(0.0549)	(0.0946)	(0.0340)	(0.0609)	(0.122)	(0.0514)	(0.0561)	(0.0831)	(0.0342)
2015o.YEAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2016o.YEAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Constant	8.308	45.17	16.45**	36.52***	7.145	0.277	4.078	45.48	16.10**	31.59***	-10.48	2.924	7.458	91.14*	17.12**	34.49***	-8.811	1.474
Observations	901	35	925	400	124	458	807	35	858	382	120	443	767	34	802	368	117	433
Number of id	192	8	188	85	29	78	189	8	185	82	29	78	180	8	179	79	28	78
R-squared	0,1255	0,8686	0,1221	0,2124	0,3654	0,2653	0,127	0,8686	0,1488	0,2701	0,381	0,2763	0,1479	0,906	0,1529	0,2634	0,3853	0,2855

Notes: Heteroskedasticity robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

When stratified by total assets (categorized as Large or Small) and star ratings (3-star, 4-star, and 5-star hotels), the following results emerged (Table 6.6):

The total Assets (log) variable is negative across most categories, but it is only statistically significant in Large & 4-star, where we see coefficients like -3.805 with and $p<0.01$ level of significance. Total Assets (log) squared is included to capture any non-linear effects of size. The squared term is significant and positive once again in the Large & 4-star category, which might suggest that there's a non-linear relationship where ratings decrease up to a point with size but then start to increase after reaching a certain threshold.

Turnover/Number of Beds (log) his variable is consistently positive and significant across almost all categories, indicating that higher ratios of turnover to the number of beds are associated with higher ratings.

Turnover/Fixed Assets (winsorized) exhibits negative coefficients and high significance at 1% value in the categories of Small and 3-star and Small and 5-star hotels, implying that a higher turnover relative to fixed assets is associated with lower ratings. This may indicate that for these establishments, simply having a high turnover does not lead to higher ratings, and perhaps other factors like service quality are more important.

Number of Ratings exhibits a positive coefficient that is significant across almost all categories suggesting that a higher number of ratings is associated with higher hotel ratings.

Administrative Expenses / Turnover (winsorized) does not show a consistent pattern of significance or sign across categories.

The sales Expenses / Turnover ratio (winsorized) variable shows a significant positive association in "Large & 3-star" and "Small & 4-star" indicating that higher sales expenses relative to turnover are associated with a higher dependent variable rating. This might mean that these hotels are successfully investing in marketing or customer service, which is translating into higher perceived quality.

Year-on-Year First Difference of Fixed Assets - variable is not significant in all categories. Moreover, “Administrative Expenses / Turnover Year-on-Year Log Change” and “Sales Expenses / Turnover Ratio Year-on-Year Log Change" do not show a consistent pattern over the models.

Pool: The presence of a pool is negatively correlated with ratings in several categories, much more significantly and interestingly in larger 3-star hotels. Family Friendly is positively and significantly associated with ratings in the Small & 3-star category, indicating family-friendly hotels have higher ratings, likely due to the appeal to a broader customer base and the additional services and accommodations that cater to families.

Overall, Total assets had the strongest association in large 4-star hotels and the number of beds in small 5-star hotels. Fixed assets utilization had the largest negative coefficient in small 3-star hotels and the number of ratings was the largest positive coefficient for the two extreme categories (small & 3-star and large & 5-star). Cost-efficiency (SalesTI_w) had a large positive association with customer rating in large 3-star hotels but promotion costs (SalesTI_dl) had a small negative association. Having a swimming pool was most negatively associated with large 3-star hotels. Table 6.7 presents the regression results stratified by asset and star ranking (six categories).

Table 6.7. Regression results stratified by number of beds and star ranking

Variables/ models	*Small & 3 stars (99)*	*Large & 3 stars (100)*	*Small & 4 stars (101)*	*Large & 4 stars (102)*	*Small & 5 stars (103)*	*Large & 5 stars (104)*	*Small & 3 stars (105)*	*Large & 3 stars (106)*	*Small & 4 stars (107)*	*Large & 4 stars (108)*	*Small & 5 stars (109)*	*Large & 5 stars (110)*	*Small & 3 stars (111)*	*Large & 3 stars (112)*	*Small & 4 stars (113)*	*Large & 4 stars (114)*	*Small & 5 stars (115)*	*Large & 5 stars (116)*
lTA	-0.0110	4.654*	-0.998***	-0.864	-0.364	0.573	0.434	1.474	-1.064***	-0.549	0.109	0.356	0.0419	1.561	-1.126***	0.154	-0.0148	0.316
	(1.184)	(2.607)	(0.383)	(0.856)	(0.604)	(0.690)	(1.264)	(1.490)	(0.343)	(0.852)	(0.654)	(0.707)	(1.241)	(1.470)	(0.368)	(0.869)	(0.703)	(0.715)
lTAsq	-0.00351	-0.152*	0.0304**	0.0260	0.00956	-0.0153	-0.0187	-0.0629	0.0312***	0.0161	-0.00515	-0.00969	-0.00658	-0.0658	0.0328***	-0.00537	-0.000735	-0.00868
	(0.0413)	(0.0775)	(0.0125)	(0.0254)	(0.0191)	(0.0195)	(0.0441)	(0.0450)	(0.0111)	(0.0254)	(0.0209)	(0.0200)	(0.0434)	(0.0435)	(0.0120)	(0.0261)	(0.0226)	(0.0202)
TurnBeds_l	0.104**	0.711**	0.0565	0.0772	0.132**	0.0215	0.108**	0.790***	0.101**	0.120	0.143**	0.0388	0.111**	0.795***	0.118**	0.183**	0.146**	0.0491
	(0.0437)	(0.302)	(0.0476)	(0.0787)	(0.0591)	(0.0557)	(0.0436)	(0.258)	(0.0461)	(0.0820)	(0.0583)	(0.0603)	(0.0465)	(0.243)	(0.0502)	(0.0821)	(0.0681)	(0.0664)
TurnFA_w	-0.284***	-0.331	-0.0183	0.160***	-0.0988**	0.0145	-0.273***	-0.898	-0.0417	0.145**	-0.0895**	0.00667	-0.295***	-1.015	-0.0232	0.122	-0.0783*	0.00138
	(0.0581)	(0.887)	(0.0371)	(0.0596)	(0.0411)	(0.0302)	(0.0631)	(0.679)	(0.0384)	(0.0681)	(0.0412)	(0.0434)	(0.0634)	(0.663)	(0.0341)	(0.0752)	(0.0428)	(0.0449)
lRatingcount	0.0962***	0.0677	0.105***	0.0657**	0.110***	0.110***	0.0664**	0.207***	0.115***	0.0480	0.127***	0.143***	0.0871**	0.215***	0.111***	0.0425	0.128***	0.143***
	(0.0315)	(0.0705)	(0.0233)	(0.0293)	(0.0367)	(0.0337)	(0.0335)	(0.0689)	(0.0255)	(0.0323)	(0.0419)	(0.0370)	(0.0339)	(0.0701)	(0.0255)	(0.0308)	(0.0411)	(0.0364)
CTI_w	0.0957	0.293	-0.311*	0.00216	-0.338*	0.100	0.0404	0.238	-0.191	0.0649	-0.242	0.106	0.00424	0.223	-0.125	-0.354	-0.159	0.159
	(0.188)	(0.434)	(0.183)	(0.301)	(0.178)	(0.135)	(0.202)	(0.570)	(0.182)	(0.410)	(0.191)	(0.158)	(0.206)	(0.558)	(0.230)	(0.547)	(0.213)	(0.154)
SalesTI_w	0.204	1.021**	0.322***	0.610*	0.0199	0.182	0.296**	-0.0720	0.265**	0.570	0.0227	0.246	0.276*	-0.0846	0.228*	0.258	0.0657	0.0809
	(0.153)	(0.510)	(0.121)	(0.322)	(0.206)	(0.327)	(0.147)	(0.363)	(0.125)	(0.363)	(0.184)	(0.367)	(0.151)	(0.383)	(0.136)	(0.406)	(0.197)	(0.443)
lFA_d							0.137	0.118	0.0965**	0.0556	0.0239	0.00896	0.125	0.137	0.0962**	0.0376	0.0353	-0.00413
							(0.156)	(0.407)	(0.0464)	(0.111)	(0.0575)	(0.0582)	(0.172)	(0.416)	(0.0482)	(0.130)	(0.0674)	(0.0621)
CTI_dl													-0.0255	-0.198	-0.00295	-0.0546	0.00247	-0.0493
													(0.0488)	(0.203)	(0.0514)	(0.0685)	(0.0570)	(0.0466)
SalesTI_dl													0.0383	0.182*	-0.0146	-0.0151	-0.0657*	0.0787
													(0.0356)	(0.0986)	(0.0420)	(0.0336)	(0.0388)	(0.0601)
Pool	-0.173*	-0.163	-0.0724	-0.226*	-0.0634	-0.192*	-0.189*	-0.158	-0.0993	-0.230**	-0.0249	-0.196*	-0.177	-0.217*	-0.121	-0.227**	0.0199	-0.190*
	(0.0987)	(0.176)	(0.0783)	(0.116)	(0.0682)	(0.105)	(0.108)	(0.102)	(0.0792)	(0.114)	(0.0670)	(0.107)	(0.111)	(0.119)	(0.0834)	(0.108)	(0.101)	(0.106)
FamFriendly	0.711***						0.929***						0.953***					
	(0.107)						(0.124)						(0.132)					
AvSpendpn	0.0157	0.00280	0.00899	0.00775	-0.00582	0.0174**	0.0177*	0.00917	0.0110	0.00688	-0.00739	0.0117	0.0145	0.00993	0.00829	0.00632	-0.00673	0.0137*
	(0.00989)	(0.0411)	(0.00866)	(0.00768)	(0.00840)	(0.00725)	(0.0101)	(0.0367)	(0.00824)	(0.00828)	(0.00880)	(0.00748)	(0.0107)	(0.0353)	(0.00804)	(0.00809)	(0.00897)	(0.00777)

AvSpendpp	-0.000598	-0.000361	-0.00125*	-0.00125	0.000349	-0.000779	-0.000868	0.00121	-0.00150**	-0.00142*	0.000487	-0.000123	-0.000459	0.00132	-0.00127*	-0.00132*	0.000434	-0.000110
	(0.000729)	(0.00243)	(0.000697)	(0.000774)	(0.000764)	(0.000646)	(0.000713)	(0.00256)	(0.000746)	(0.000779)	(0.000802)	(0.000678)	(0.000745)	(0.00260)	(0.000745)	(0.000793)	(0.000835)	(0.000683)
2011 .YEAR	0.0926	0.249	0.137*	0.103	0.154	0.0629	0.0814	0.117	0.167**	0.142*	0.118	0.0136	0.0970	0.170	0.162**	0.0960	0.111	-0.00604
	(0.0767)	(0.236)	(0.0740)	(0.0768)	(0.0978)	(0.0519)	(0.0755)	(0.238)	(0.0673)	(0.0756)	(0.105)	(0.0525)	(0.0789)	(0.259)	(0.0715)	(0.0729)	(0.106)	(0.0487)
2012. .YEAR	0.0376	0.159	0.111	0.169**	0.0733	0.0511	0.0662	-0.127	0.153**	0.181**	0.0617	0.000408	0.0519	-0.217	0.140**	0.186**	0.0557	0.000336
	(0.0761)	(0.218)	(0.0681)	(0.0757)	(0.0715)	(0.0527)	(0.0749)	(0.183)	(0.0657)	(0.0757)	(0.0744)	(0.0527)	(0.0800)	(0.218)	(0.0706)	(0.0790)	(0.0731)	(0.0549)
2013 .YEAR	0.0806	0.158	0.0802	0.100	0.0569	-0.0204	0.0678	-0.103	0.105	0.141*	0.0640	-0.0599	0.0520	-0.173	0.0994	0.109	0.0463	-0.0746
	(0.0767)	(0.278)	(0.0694)	(0.0861)	(0.0672)	(0.0695)	(0.0777)	(0.291)	(0.0701)	(0.0800)	(0.0687)	(0.0705)	(0.0800)	(0.300)	(0.0770)	(0.0827)	(0.0759)	(0.0703)
2014. YEAR	0.0186	0.159	-0.0362	0.0126	0.0110	0.000672	0.00106	0.0683	-0.0238	0.0109	0.00709	-0.0182	-0.000451	0.0587	-0.0190	-0.00324	-0.0155	-0.0163
	(0.0607)	(0.184)	(0.0474)	(0.0534)	(0.0558)	(0.0420)	(0.0584)	(0.132)	(0.0480)	(0.0555)	(0.0572)	(0.0415)	(0.0604)	(0.137)	(0.0507)	(0.0586)	(0.0468)	(0.0431)
2015. YEAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2016. YEAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Constant	3.189	-37.79*	11.65***	10.34	6.407	-2.466	0.0144	-11.83	12.03***	7.685	2.458	-0.671	3.047	-12.59	12.50***	1.426	3.227	-0.521
Observations	846	90	911	414	234	348	758	84	845	395	226	337	719	82	794	376	219	331
Number of id	182	16	175	76	42	56	179	16	173	76	42	56	169	16	169	73	41	56
R-squared	0,119	0,5113	0,1231	0,1245	0,303	0,251	0,1187	0,4773	0,1504	0,1438	0,324	0,2547	0,1414	0,4926	0,1448	0,1729	0,327	0,268

Notes: Heteroskedasticity robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6.7. presents the regression results stratified by number of beds and star ranking (six categories).

The total Assets log present for Small & 4-star has a negative coefficient at $p < 0.01$ suggesting that, for these hotels, an increase in the total assets is associated with a significant decrease in the TripAdvisor score. Moreover, the Total Assets log squared for the same category, presents a positive coefficient at $p < 0.05$ indicating a nonlinear relationship where increases in the number of beds are initially associated with lower ratings, but after a certain point, further increases in beds are associated with higher ratings, possibly indicating economies of scale.

Log of turnover to number of beds ratio variable present for Small & 3-star a positive coefficient of 0.111 and significance of ($p < 0.05$) and Large & 3-star a positive coefficient of 0.795 and significance of ($p < 0.05$). Consequently, for both small and large 3-star hotels, higher turnover per bed is associated with better performance or ratings, suggesting efficient utilization of assets.

Turnover/Fixed Assets (winsorized) exhibits a positive coefficient and is statistically significant for Small 3-star at 1% value, while it is negative and statistically significant for Small 5-star hotels. This suggests that a higher turnover to fixed assets ratio is reversely related to TripAdvisor scores in Small 5-star hotels.

lRatingcount - Log of number of ratings is recorded as positive and significant in most of the sub-categories, suggesting that a higher number of ratings is associated with higher hotel ratings.

Administrative Expenses / Turnover (winsorised) does not show a consistent pattern of significance. The sales Expenses / Turnover ratio (winsorized) variable shows a significant positive association in "Small & 4-star" indicating that higher sales expenses relative to turnover are associated with a higher TripAdvisor score.

The year-on-year first difference of Fixed Assets is significant and positive for Small 4-star hotels, indicating that probably, increase in the fixed assets of the concrete category causes an increase in TripAdvisor rating. Investment in fixed assets, such as renovations or new facilities, could directly impact guest experiences in small 4-star hotels. Guests may perceive improvements as value additions, thereby increasing the average score rating on TripAdvisor.

Sales and administrative expenses to Turnover ratio, winsorized, do not present significance at any sub-category of the sample. The presence of a pool is negatively related and significant at a 10% level for Large in beds 5-star hotels. Being Family-friendly is statistically meaningful for Small in terms of beds & 3-star with a coefficient = 0.953* ($p < 0.05$). Family travelers seem to value amenities catered to children and family needs, and positive experiences of family travelers can lead to higher ratings on TripAdvisor.

Overall, increasing assets might improve TripAdvisor scores to a point, but the effect diminishes beyond a certain level, suggesting that there are optimal levels of investment for maximization of customer satisfaction as reflected in ratings. Additionally, features like being family-friendly or having a pool can be strategically emphasized or developed based on their positive association with customer ratings in specific hotel segments.

6.1 Results for Technical Efficiency

The following table presents summary statistics for Technical Efficiency (TE) scores estimated via Stochastic Frontier Analysis (SFA) for hotels. Technical efficiency scores close to 1 indicate high efficiency, while scores close to 0 suggest inefficiency. The table presents full sample descriptive statistics as well as statistics for four different categorizations: a) by Total Assets, b) by Beds, c) by year, and d) by star category.

Table 6.8. Summary statistics for Technical Efficiency

	Observations	*Mean*	*Median*	*SD*	*Min.*	*Max.*
Panel A. Full sample						
All hotels	3,666	0.622	0.647	0.151	0.051	0.916
Panel B. By size						
Large (TA)	1,150	0.653	0.679	0.136	0.060	0.892
Medium (TA)	1,855	0.616	0.639	0.149	0.051	0.916
Small (TA)	661	0.583	0.593	0.169	0.086	0.909
Large (Beds)	1,102	0.662	0.690	0.133	0.060	0.916
Medium (Beds)	1,801	0.619	0.640	0.149	0.107	0.909
Small (Beds)	761	0.569	0.596	0.161	0.051	0.889
Panel C. By Year						
2007	267	0.647	0.672	0.145	0.195	0.904
2008	272	0.631	0.662	0.149	0.060	0.886
2009	316	0.618	0.638	0.144	0.169	0.893
2010	364	0.595	0.631	0.155	0.086	0.893
2011	424	0.610	0.636	0.151	0.107	0.883
2012	465	0.587	0.595	0.154	0.125	0.899
2013	469	0.617	0.639	0.157	0.136	0.916
2014	368	0.639	0.663	0.152	0.051	0.907
2015	362	0.648	0.680	0.142	0.052	0.890
2016	359	0.646	0.675	0.137	0.198	0.909
Panel D. By Hotel Stars						
3-star	1,204	0.605	0.622	0.151	0.086	0.904
4-star	1,712	0.625	0.654	0.153	0.051	0.916
5-star	750	0.641	0.672	0.141	0.060	0.900

Panel A presents descriptive statistics for the full sample. Among all hotels, the average technical efficiency is 0.622, which indicates that, on average, hotels operate at approximately 62.2% of their potential efficiency. The efficiency scores range from a minimum of 0.051 (or 5.1%) to a maximum of 0.916 (or 91.6%). The median score is 0.647, and there's a standard deviation of 0.151. The latter indicates a moderate level of variability in the efficiency scores across all hotels.

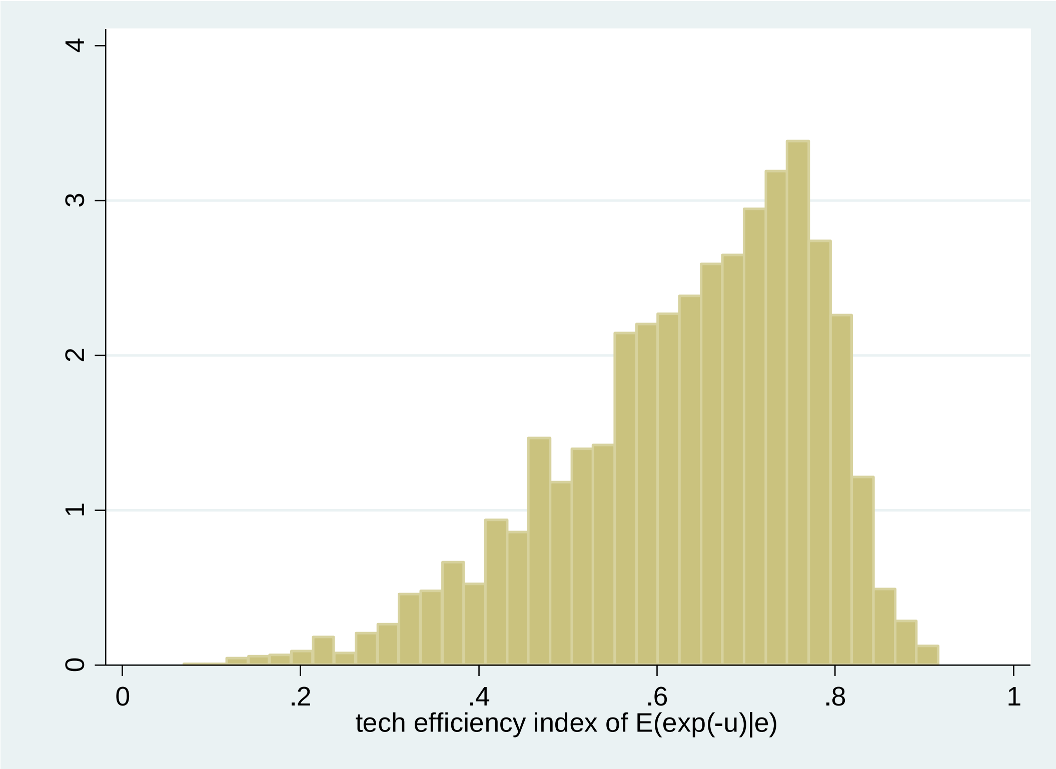


Figure 6.1. Dispersion of Technical Efficiency for the full sample

Panel B. presents by size stratification descriptive statistics in terms of Total Assets as well as in terms of the number of beds under operation. Large hotels, when categorized based on Total Assets, have an average efficiency score of 0.653. They range from 6% to 89.2% in their efficiency scores. Large hotels' efficiency is somewhat less variable than the entire sample, with a standard deviation of 0.136. Medium-sized hotels have an average efficiency score slightly lower than large hotels, standing at 0.616. Their efficiency scores range between 5.1% to 91.6%. Medium-sized hotels, on average, have a technical efficiency slightly below the overall sample mean. The smallest hotels (based on Total Assets) show the lowest average efficiency calculated at 0.583. Their efficiency spans from 8.6% to 90.9%. Furthermore, the variability in efficiency among small hotels is the highest with a standard deviation of 0.169.

When categorized by the number of beds, large hotels perform better with an average efficiency of 0.662, ranging from 6% to 91.6%. Medium hotels have an average score of 0.619 with efficiency scores spanning from 10.7% to 90.9%. Small hotels have an average efficiency score of 0.569, which is the lowest in this categorization, with scores ranging from 5.1% to 88.9%. Both size stratifications, by Total Assets and by bed showcase similar results among the three categories, large, medium, and small hotels ie., Large hotels have the best efficiency scores, outperforming the other two categories as well as the full sample mean efficiency. Medium hotels come second in a row, with their mean efficiency score being very close to the full sample's average. Lastly, Small hotels exhibit the worst out of the three categories efficiency score, being up to 19,7% less efficient than Large hotels. A plausible explanation could be economies of scale succeeded by larger hotels. Another interpretation is that smaller hotels tend to be family-operated, and lack professional management, a fact resulting in poorer management and operational results.

Table 6.8., Panel C., provides efficiency scores over 10 years. It is worthwhile to mention that the highest mean efficiency was in 2015 at 0.648, and the lowest was in 2012 at 0.587. Moreover, from 2007 to 2010, there was a gradual decline in average efficiency, which then slightly increased in 2011 but dipped again in 2012. Post-2012, there's an evident recovery with a peak in 2015, and it remains relatively stable till 2016. The standard deviation, which indicates the variability of technical efficiency among hotels, is relatively stable over the years. It peaks in 2013 at 0.157, suggesting a more diverse performance in that year. Although there doesn't seem to be a clear trend over the years; however, the efficiency levels seem to hover around the mid-0.6 range. Efficiency over the years seems to follow a parallel path with turnover averages, for the 2007 – 2016 decade, presented in Table 6.5.

Panel D., Table 6.8, provides a By Hotel Star stratification efficiency score. According to the table, 3-star hotels have an average efficiency score of 0.605, ranging from 8.6% to 90.4%. 4-star hotels exhibit an average efficiency score of 0.625, ranging from 5.1% to 91.6% while the

5-star hotels segment seems to be the most efficient on average, with a score of 0.641 and scores spanning from 6% to 90%. However, 4-star hotels have the highest maximum efficiency score, indicating that the best-performing 4-star hotels are as efficient as, if not more than, 5-star hotels. It is worthwhile to mention that all Star categories have champions outliers, performing at 90% of Technical Efficiency, and at the same time all categories, even 5-star hotels, have hotels that they do very poorly operating at a mediocre 5 – 6% of the efficiency range. This wide variability may indicate inconsistent standards or management practices within all-star categories.

Summing up, larger hotels, whether categorized by Total Assets or Beds, tend to be more technically efficient than their smaller counterparts. Moreover, there have been fluctuations in average technical efficiency over the years. The size and Star rating seem to have a positive correlation with efficiency, with 5-star hotels and Large hotels in terms either of Total Assets or beds being the most efficient on average. All-star categories, even 5-star hotels, have outliers, performing almost at the maximum or minimum efficiency range. Some 4-star hotels operate more efficiently than a few of their counterparts 5-star hotels. While there are yearly fluctuations, there isn't a clear upward or downward trend in technical efficiency over the years. Efficiency fluctuations over the years while following the average Greece's tourism receipts diagram of the same time, also reflect shifts in the industry's operations, economic conditions, or other exogenous factors affecting hotel performance.

The following table presents descriptive statistics of the hotels' sample for all variables of the OLS model used in the present thesis, stratified into two categories: hotels with high efficiency compared to those with low efficiency.

Table 6.9. Summary statistics by efficiency: Low, High

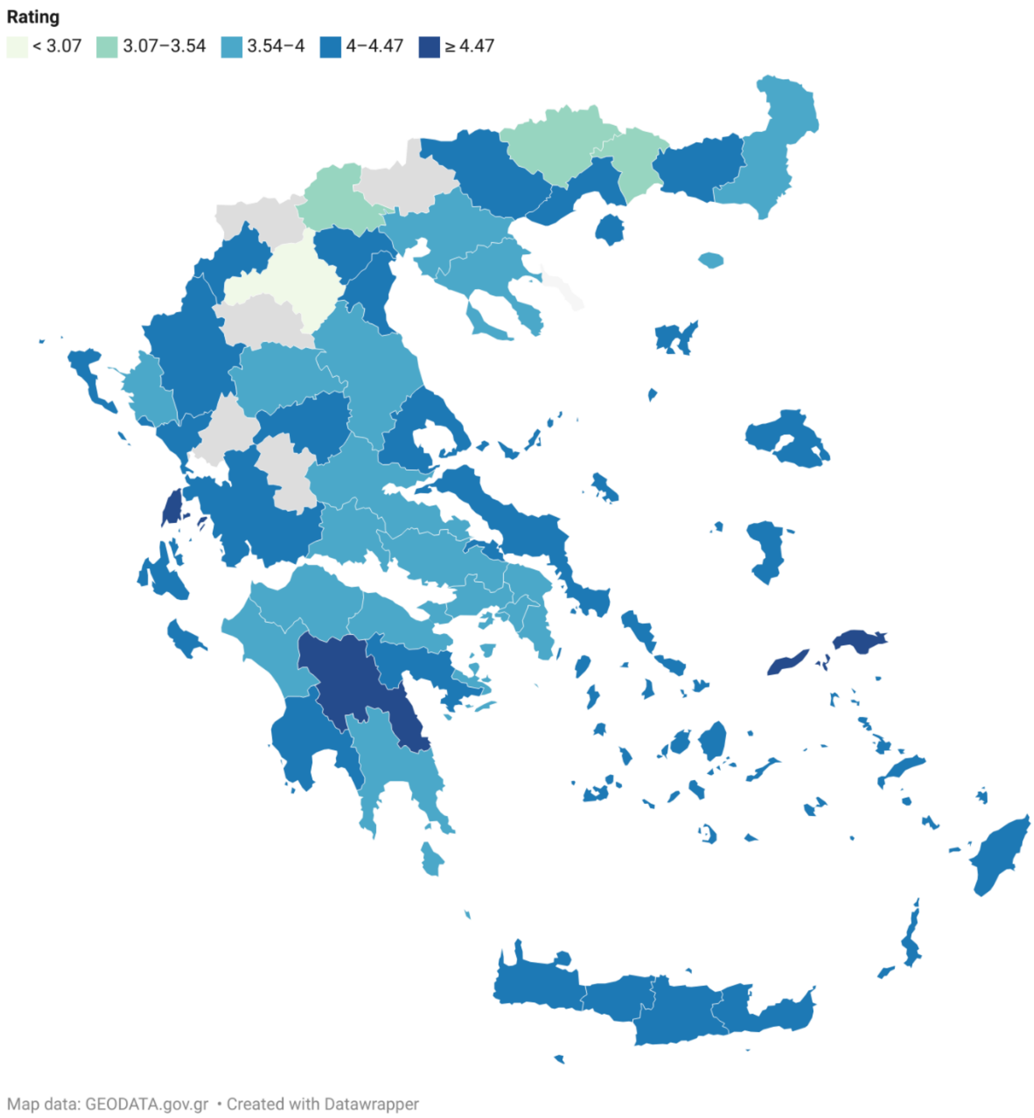
Efficiency, High	stats	N	mean	p50	sd	Min	max
-	BEDS	1833	286,3579	198	263,6326	16	1444
-	STARS	1833	3,932351	4	0,7146363	3	5
-	FIXEDA~S	1833	1,34E+07	3774515	3,57E+07	21638	4,09E+08
-	TURNOVER	1833	4652381	2119979	7630289	120	6,55E+07
-	lTA	1833	15,6486	15,57944	1,353182	11,98043	19,9584
-	lFA_d	1567	0,0223138	-0,0211306	0,2131	-2,875186	2,951557
-	TurnB~_l	1833	9,304677	9,239738	0,8901659	1,491655	13,39302
-	lRatin~t	1833	2,742403	2,70805	1,554438	0	6,504288
-	SalesT~w	1833	0,095333	0,048606	0,1397272	0,0003633	1,228157
-	CTI_w	1833	0,0784186	0,0694722	0,0581824	0,0074998	1,401676
-	lTAsq	1833	246,7089	242,7189	42,87072	143,5306	398,3378
-	TurnFA_w	1833	0,662757	0,4740872	0,6638334	0,0000557	4,353887
-	Pool	1833	0,2444081	0	0,429853	0	1
-	FamFri~y	1833	0,0021822	0	0,0466759	0	1
-	AvSpen~p	1388	608,3742	639,4698	47,4191	514,0972	653,3084
-	AvSpen~n	1388	70,63119	70,37794	2,297825	66,95984	73,93074
-	CTI_dl	1542	-0,0611845	-0,0249221	0,4564798	-3,567382	3,050335
-	Sales~dl	1513	-0,0038302	0,0009279	0,5771287	-3,260092	6,320682
-	Rating	1833	4,09888	4,234043	0,6713671	1	5
Efficiency, Low	stats	N	mean	p50	sd	Min	max
-	BEDS	1831	218,1436	130	257,7805	13	1353
-	STARS	1833	3,819967	4	0,7203336	3	5
-	FIXEDA~S	1833	1,99E+07	2984725	6,67E+07	70364	5,61E+08
-	TURNOVER	1833	2377715	780791	4981089	241	4,35E+07
-	lTA	1833	15,37477	15,11981	1,472366	12,03388	20,16028
-	lFA_d	1523	0,0264301	-0,0217323	0,2180583	-2,500572	2,9801
-	TurnB~_l	1831	8,715193	8,726381	1,006335	1,390452	11,87415
-	lRatin~t	1833	2,198855	2,197225	1,412207	0	6,150603

_	SalesT~w	1833	0,2460234	0,1325233	0,2713699	0,0003633	1,228157
_	CTI_w	1833	0,2158786	0,1637131	0,2024328	0,0074998	1,401676
_	lTAsq	1833	238,5502	228,6087	47,37731	144,8142	406,437
_	TurnFA_w	1833	0,2992238	0,2251812	0,3250456	0,0003119	4,353887
_	Pool	1833	0,2274959	0	0,4193298	0	1
_	FamFri~y	1833	0	0	0	0	0
_	AvSpen~p	1423	617,1719	640,4312	44,15436	514,0972	653,3084
_	AvSpen~n	1423	70,60134	70,37794	2,117481	66,95984	73,93074
_	CTI_dl	1499	0,0250269	0,0109235	0,5659766	-4,44212	3,694487
_	Sales~dl	1447	0,0368183	0,0106281	0,6300175	-5,036769	5,340404
_	Rating	1833	4,056418	4,2	0,7439847	1	5

Table 6.9. showcases that high-efficiency hotels have on average more beds than low-efficiency hotels (~286 vs. ~218). They also tend to have a higher average star rating (~3.93 vs. ~3.82). High-efficiency hotels, on average, have a higher turnover relative to their fixed assets, suggesting better asset utilization. Lower efficiency hotels have higher fixed assets on average, but considerably lower turnover. Furthermore, High-efficiency hotels, on average, have a higher turnover per bed ratio (9.30) than low-efficiency ones (8.71), suggesting better room revenue management. In addition to the latter, these hotels have lower sales expenses and administrative expense ratios indicating efficient expense management. A slightly higher percentage of high-efficiency hotels have pools, but very few are designated as family-friendly. High-efficiency hotels receive more reviews on average (2.74) than low-efficiency hotels (2.20), possibly indicating higher occupancy or guest engagement and popularity. At the same time, High-efficiency hotels received on average, only slightly better rating scores on TripAdvisor (4.09) than low-efficiency (4.05).

The next figure presents average TripAdvisor ratings per Greek region. Greece is administratively divided into 52 prefectures. Usually, its prefecture comprises a geographical formation with its limits to rivers, mountain ranges, coastlines, or the edges of valleys. Moreover, several island complexes were designated as prefectures.

Figure 6.1. Customer satisfaction rating

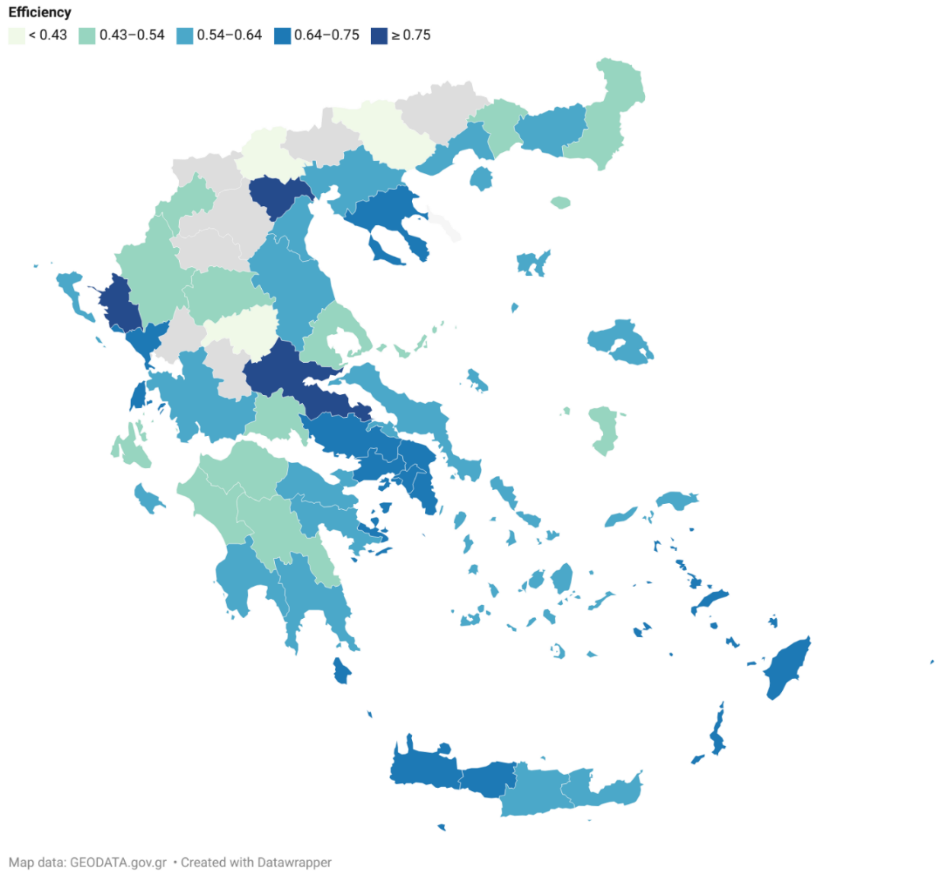


NOTES: The graph shows the average customer satisfaction rating obtained from TripAdvisor. The average rating is shown throughout the study.

Considering the above map, more touristic mature prefectures and especially hotels located on the islands tend to enjoy higher ratings. Consequently, Crete, Dodecanese (where Rhodes and Kos belong), Cyclades (the island complex incorporating Mykonos, Santorini, and Paros), Ionian Islands complex (containing Corfu, Zakynthos, and Lefkada), and the Northern islands (Thasos, Mytilene) as well as Messinia in Peloponnese exhibit average ratings between 4.0 and 4.47. Surprisely, Attica (the prefecture of Greece's capital, Athens) and Central Macedonia, the prefecture of the second biggest metropolis in Greece present lower average ratings, ranging from 3.54 to 4. A plausible explanation could be that there is a vast number of hotels concentrated in this two Metropolis, with an important part of them holding minimum service standards since they enjoy intensive demand and year-round occupancies. On the other hand, hotels on the islands compete with other destinations nationally and internationally to attract clientele, while they strive to extend holidays period. Consequently, over the years they have become more efficient, service-oriented, and mastered customer satisfaction. Finally, it is worthwhile to mention that the minimum rating is achieved by the less touristic developed, mostly mountainous, prefectures of Greece.

The next figure presents efficiency scores per Greek region.

Figure 6.2. Technical efficiency per region



NOTES: The graph plots the average technical efficiency estimates over the period of study.

A different situation holds when we observe efficiency scores per prefecture of the Greek territory. Attica (where Athens stands) and Central Macedonia, the prefecture of the second biggest metropolis in Greece, Salonica, along with Dodecanese, Chania, and Rethymno (half of Crete Island) and Lefkada succeeded the better efficiency scores in the period under examination. While Lefkada, Crete, and Dodecanese seem to excel due to managerial expertise, upper-scale hotels, and abundance of clientele, Athens, and Salonica areas, propelled by their capital cities, succeed with higher efficiency probably due to higher values on the nominator (turnover and total assets) rather than lower denominator making intensive demand being the crucial metric of the equation. Important is to mention that touristic mature prefectures generally exhibit higher efficiency in comparison to less developed regions.

Table 6.10. presents the regression results stratified by hotels' Stochastic Frontier Analysis (SFA) efficiency levels (Low and High).

Table 6.10. Regression results by efficiency

	Low	*High*	*Low*	*High*	*Low*	*High*	*Low*	*High*
Variable/model	*(37)*	*(38)*	*(39)*	*(40)*	*(41)*	*(42)*	*(43)*	*(44)*
lTA	-0.378*	-0.426	-0.488**	-0.273	-0.637***	-0.268	-0.640***	-0.239
	(0.220)	(0.348)	(0.222)	(0.363)	(0.217)	(0.383)	(0.217)	(0.400)
lTAsq	0.00941	0.0127	0.0123*	0.00830	0.0168**	0.00835	0.0169**	0.00740
	(0.00670)	(0.0105)	(0.00677)	(0.0109)	(0.00660)	(0.0115)	(0.00660)	(0.0120)
TurnBeds_l	0.121***	0.140***	0.132***	0.139***	0.138***	0.149***	0.143***	0.154***
	(0.0282)	(0.0449)	(0.0250)	(0.0453)	(0.0264)	(0.0469)	(0.0267)	(0.0475)
TurnFA_w	-0.0903**	-0.0712	-0.0930**	-0.0583	-0.0776*	-0.0541	-0.0775*	-0.0513
	(0.0433)	(0.0613)	(0.0461)	(0.0616)	(0.0443)	(0.0598)	(0.0441)	(0.0588)
lRatingcount	0.112***	0.107***	0.114***	0.102***	0.120***	0.101***	0.116***	0.0858***
	(0.0158)	(0.0272)	(0.0172)	(0.0295)	(0.0174)	(0.0300)	(0.0180)	(0.0317)
CTI_w	-0.0942	0.681**	-0.111	0.728**	-0.0460	0.472	-0.0514	0.527
	(0.113)	(0.303)	(0.121)	(0.316)	(0.136)	(0.709)	(0.136)	(0.701)
SalesTI_w	0.283***	0.485***	0.291***	0.400**	0.257***	0.485**	0.255***	0.531***
	(0.0867)	(0.169)	(0.0835)	(0.172)	(0.0880)	(0.198)	(0.0878)	(0.198)
lFA_d			0.129**	0.0252	0.134**	0.0261	0.143**	0.0157
			(0.0575)	(0.0464)	(0.0586)	(0.0520)	(0.0602)	(0.0534)
CTI_dl					-0.0384	-0.0182	-0.0332	-0.00661
					(0.0350)	(0.0494)	(0.0349)	(0.0477)
SalesTI_dl					0.0296	-0.0471	0.0303	-0.0487
					(0.0307)	(0.0336)	(0.0303)	(0.0342)
Pool	-0.178***	-0.0465	-0.188***	-0.0404	-0.207***	-0.0377	-0.208***	-0.0452
	(0.0538)	(0.0870)	(0.0555)	(0.0863)	(0.0568)	(0.0862)	(0.0569)	(0.0848)
FamFriendly		0.724***		0.986***		1.026***		0.962***

		(0.112)		(0.117)		(0.122)		(0.120)
AvSpendpn	0.00474	0.00460	0.00582	0.00172	0.00432	0.00303	0.00836	0.00467
	(0.00531)	(0.00785)	(0.00491)	(0.00815)	(0.00501)	(0.00859)	(0.00510)	(0.00874)
AvSpendpp	0.000208	-0.000158	0.000139	3.92e-06	0.000327	-0.000138	-0.000232	-0.00145*
	(0.000291)	(0.000468)	(0.000294)	(0.000461)	(0.000291)	(0.000502)	(0.000415)	(0.000818)
2011.YEAR							0.0904**	0.153**
							(0.0426)	(0.0758)
2012.YEAR							0.0687*	0.174**
							(0.0409)	(0.0686)
2013.YEAR							0.0306	0.163**
							(0.0461)	(0.0713)
2014.YEAR							-0.0349	0.0566
							(0.0316)	(0.0492)
2015o.YEAR							-	-
2016o.YEAR							-	-
Constant	5.918***	5.824**	6.782***	4.630	7.923***	4.428	7.950***	4.801
N	2,137	706	1,973	672	1,876	645	1,876	645
No. hotels	482	202	475	197	457	193	457	193
R-squared	0.1414	0.1528	0.159	0.1692	0.1674	0.175	0.1705	0.1863

Notes: Heteroskedasticity robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Total Assets (log) variable for Low Efficiency Hotels presents a negative significant coefficient across models (37-39-43) suggesting that an increase in the log of total assets is associated with a decrease in the average score rating. High-efficiency hotels also present negative coefficients but are not statistically significant, indicating no clear impact of total assets on TripAdvisor scores for these hotels. As the Total Assets (log) squared variable for Low-Efficiency hotels is positive and significant in some models (from * to **), it is assumed a nonlinear relationship where initial increases in assets may have diminishing positive scores on TripAdvisor after a certain point.

Log of the ratio turnover/number of beds variable, for both Low and High-Efficiency hotels, is consistently positive and highly significant across all models, implying that hotels with higher turnover per bed tend to have higher average score ratings, which could be indicative of more efficient operations, higher popularity or due to better service quality or experiences.

Low-efficiency hotels present a negative and significant in some models' relation when interpreting the Turnover/Fixed Assets (winsorized) variable, suggesting that a higher turnover to fixed assets ratio correlates with lower TripAdvisor scores, which might reflect disappointment of "overexploitation of the premises of the hotels for the concrete category. The concrete ratio does not have a clear influence on the TripAdvisor scores for highly efficient hotels.

The Number of ratings (log) is for both Low and High-Efficiency hotels positive and highly significant in all models, indicating that a greater number of ratings correlates with higher TripAdvisor scores for both low and high-efficiency hotels. This could suggest a virtuous cycle where higher ratings lead to more reviews, and more reviews could boost visibility and perceived trustworthiness.

Sales Expenses to Turnover Ratios are highly positive and significant across all models, for both, Low and high-efficiency hotels, suggesting that higher sales expenses relative to turnover are associated with higher TripAdvisor scores. The latter might indicate that spending on sales and marketing is effective in increasing the TripAdvisor score rating for these hotels.

Administrative/ Expenses to Turnover Ratio shows, only for High-Efficiency hotels significance in two models (38 and 40) at a 5% level, indicating that higher administrative expenses relative to turnover are associated with higher TripAdvisor scores for more efficient hotels, possibly reflecting better management or service quality.

The first difference of Fixed Assets is positive and significant in some models, for Low-Efficiency hotels indicating that recent investments in fixed assets have led to increased TripAdvisor scores for less efficient hotels.

Year-on-year changes for CTI_dl and SalesTI_dl, are not significant either for High or Low-efficiency hotels, suggesting that the year-on-year changes in these areas have not had a significant impact on TripAdvisor scores.

Having a Pool is negative and significant for low-efficiency hotels, suggesting that simply having a pool does not guarantee higher TripAdvisor scores among these hotels. The latter is somewhat counterintuitive as one might expect having a pool to be a positive feature. This could imply that in the context of less efficient hotels, a pool does not translate to higher ratings, possibly due to poor maintenance or other issues.

FamFriendly: Highly positive and significant for high-efficiency hotels, indicating that high-efficiency hotels account and invest in being family-friendly, acknowledging that is a strong positive factor for TripAdvisor scores.

Lastly, the significance of the year dummies in the models (2011 to 2014) for high-efficiency hotels indicates that some time-specific effects or trends impact the score ratings, which could be due to economic, industry-wide, or other external factors affecting the hospitality sector during those years.

The R-squared values are relatively low to moderate, suggesting that while the models do explain some variability in TripAdvisor scores, there is still a substantial amount of variation that is not captured by these variables.

In summary, for low-efficiency hotels, factors such as the number of ratings, the sales expenses to turnover ratio, and recent investments in fixed assets have a more pronounced impact on TripAdvisor scores. For high-efficiency hotels, being family-friendly and the administrative expenses to turnover ratio are more significant factors.

6.3 Overall assessment

In summary, taking into consideration the regression exercise results, for Total Assets log and Total Assets, log squared, it is observed that when total assets of a hotel increase, the TripAdvisor score rating decreases, holding other factors constant. The latter suggests that larger hotels (in terms of assets) tend to have lower ratings. By examining Total Assets, log squared) one realizes that it is positive and significant across all models, indicating a nonlinear U-shaped relation with total assets. this context, while initially the dependent variable decreases with an increase in assets, it eventually starts increasing at higher levels of assets. In other words, as hotels grow very large, the negative effect on ratings becomes less pronounced, beginning to lessen or possibly reverse. The descending part of the curve can be explained by the sense a customer takes of a faceless establishment when he visits a sprawled hotel with multiple (sometimes identical) compounds. Most of the time, such types of hotels and resorts are overcrowded, especially during high season, service is standard or below standard and amenities are poor. Considering that Greece has many of these types of hotels, defined as "sea and sun" accommodation, especially in the touristic developed islands, the arguments seem to stand. The concrete type of hotel usually receives clientele through the major tour operating channels such as TUI. Having the "sun and sea" as a given, they compete on a price basis which leads to squeezing the offered quality of services. Nevertheless, after a certain point, the total assets of the hotel appear to influence positively customer satisfaction, reflecting the upper-scale large hotels. When stratifying the sample by star categories, 4-star hotels show an intense and high significance (***$p<0.01$) in all models. The finding empowers previous observations, as it is a fact that many hotels described as “sea and sun village resorts or hotels” operate under a 4-star classification. Those hotels use tour operators as a source of clientele more intensively than the 5-star hotels. Often, but not always, they bear the 4-star classification as a lure of quality but at the same time they keep service low, and they crowd the places during summer. The same finding holds when stratifying by Total Assets the sample in Large and Small hotels, as Large Hotels present a negative coefficient, significant in all models ($p < 0.05$). The finding is verified by the stratification by “total assets (categorized as Large or Small) and star ratings (3-star, 4-star, 5-star hotels)”, where it holds for Large & 4-star. The finding is strong even in the Small & 4-star category, suggesting that it overruns horizontally the 4-star category independently of the size of the hotels. Finally, stratifying the sample by hotels' Stochastic Frontier Analysis (SFA) efficiency levels (Low and High), Total Assets (log) variable, presents a negative significant coefficient for Low-Efficiency Hotels which is becoming less negative with larger total assets, as indicated by the positive coefficient for Total Assets (log) squared. In a nutshell, usually low efficient, large in terms of Total Assets and beds, falling at the 4-star category hotels, when the total assets of the hotel increase, the TripAdvisor score rating decreases. The log of turnover/number of beds, is positive and significant at a 1% level across all models for the full sample, indicating that hotels with higher turnover/beds enjoy higher customer satisfaction. The finding is horizontal across all hotel star categories as well as all other stratifications of the sample used in this analysis. Thus, hotels that charge relatively more per bed tend to enjoy higher customer satisfaction rates, suggesting that hotel ratings are more driven by experience than the cost of the experience. Moreover, psychological factors may be a potential explanation, as the customer may pre-assume that "I pay more, so I receive higher quality service than less expensive hotels". Of course, the latter holds for less experienced travelers. A vice-versa interpretation suggests that higher satisfaction leads to higher ratings on TripAdvisor. The latter can be interpreted as hotels with higher standards usually charge higher prices and/or enjoy higher occupancy per year, indicative of more efficient operations.

Turnover/fixed assets is negative and becomes significant at the 1% level in all four models of the lagged values regression version in Table 6.2, implying that as the turnover to fixed assets ratio increases, the customer's satisfaction as expressed by TripAdvisor's rating tends to decrease. Turnover divided by fixed assets is a measure of fixed asset utilization. The higher the quotient the better utilization fixed assets have. It comes as a surprise that this high gearing optimization (which shows efficient entrepreneurship) has a negative relation to customer satisfaction, although not so statistically robust on the overall dataset application. Examining the breakdown analysis, as it is presented in Tables 6.3, 6.4, 6.5, and 6.6, it is obvious that this negative relation becomes statistically fully meaningful for small hotels, hotels with fewer beds, and 3-star hotels. The latter can be explained as reverse value-for-money sentiment, created by small—in terms of assets and the number of beds—and/or 3-star hotels, as it is perceived by the client. In other words, these hotels, small in size and operating on lower star rating standards give the sentiment that they charge a lot for the infrastructure they provide. A plausible explanation can be that due to constrained size and bed capacity, these types of hotels cannot achieve economies of scale and they live on the edge of break even. Consequently, they charge more than the acceptable value they offer to the client, to break even or to enjoy a marginal profit. The above holds for Small 5-star hotels too.

The number of ratings log variable is positive and significant, in the full sample and across all stratifications, suggesting that hotels with more ratings generally have higher TripAdvisor ratings. A plausible explanation is that people tend to rate when he has a good perception of the hotel they stayed in and skip rating when they had a non-superb or average experience. Hotel visitors tend to rate more frequently if they are happy with the experience they received. Moreover, it could suggest that a virtuous cycle holds, where higher ratings lead to more reviews, and more reviews could boost visibility and perceived trustworthiness. The association seems to be more intense in the 5-star category, a finding that enforces the previous interpretation, as 5-star hotels tend to offer better services and amenities.

Administrative Expenses/Turnover is not statistically significant in any model and stratification, indicating a weak or no relationship.

The sales Expenses / Turnover variable showcases a positive relation and high significance across all models in the full sample, meaning that higher sales expenses relative to turnover are associated with higher ratings. Sales expenses incorporate marketing expenses, commissions for sales to tour operators and agencies advertisements, and other related expenses. A possible interpretation of the concrete finding could be that taking care of the purchasing experience of the client (having a highly functional site and reservation tool) as well as of the on-site total experience (better amenities, tailored-made pampering, etc) results in higher satisfaction and better rating. Satisfaction seems to be related to selling expenses in small hotels in terms of assets and beds, and the 4-star category. These hotels are willing to pay to attract the clientele, showing managerial competence and this probably showcases a horizontal approach towards the customer, meaning better overall service. Furthermore, Sales Expenses to Turnover Ratios are highly positive and significant across all models, for both, Low and high-Efficiency hotels. The latter might indicate that spending on sales and marketing is effective in increasing the TripAdvisor score rating no matter how the hotel is positioned to market.

Year-on-year log change in Administrative Expenses / Turnover and Sales Expenses / Turnover does not have a statistical significance for the hypotheses under testing for the full sample as well as the sub-categories of the sample. Interestingly, the increase in sales expenses does not seem to be reflected in customer satisfaction. The finding must be further investigated.

The year-on-year first difference of the Fixed Assets variable is positive and significant in models (2) to (4) of the full sample, suggesting that hotels that increased their fixed assets from the previous year tend to have higher ratings. The increase could reflect either a renovation or an extension in capacity or facilities. So capital investments for infrastructure investment result in higher customer contentment, as visitors enjoy an increased "value received / price paid" ratio, sometimes being surprised by the amelioration of the general status of the hotel. This relationship is particularly evident in 4-star, low-efficiency, small hotels with fewer beds. These kinds of hotels positively surprise the most the customer when he visits a partially or fully renovated small hotel.

The pool (Presence of a pool) variable is negative and significant ($p < 0.01$), across all models, and all stratifications. While the presence of a pool is initially positively assessed by tourists, a series of factors involved with the "pool experience" can alternate the client's perception resulting in decreased ratings. These factors include hygiene issues, safety issues, crowdedness, availability of sunbeds, food & beverage service, changing and WC facilities, provision of pool towels, pool maintenance, water purity, and clarity, music selection, and loudness. Having a swimming pool is related to lower satisfaction, especially in large hotels in terms of assets and beds. The customer takes for granted the existence or not of a pool, information already acknowledged via the internet site. Par sequence, what the visitor rates are the quality of the pool experience he or she received and not the existence or not of a pool. The bigger a hotel is, the more difficult is to keep all previously stated prerequisites continuously on excellent status, especially during high season. A bad experience is reflected in rates. Interestingly, having a pool is negative and significant only for low-efficiency hotels. High-efficiency hotels look to be efficient in the way they operate and maintain their pools as well.

The next variable, FamFriendly (Family-friendly) is positive and significant at a 1% level in all models in the full sample, suggesting family-friendly hotels tend to have a higher average value of ratings on TripAdvisor. Family-friendly hotels have significantly higher ratings which indicates that amenities or services catering to families can positively impact a hotel's overall rating. Family-friendly is expected to have en-suite rooms lodging, cradles and baby coats, playgrounds, babysitting on demand, programs specially designed for children, kids' breakfast, and menu, etc. These facilities are boosting client's gratification. Family Friendly is significant, especially for 3-star hotels, and for small hotels in terms of total assets, across all relevant models (***p<0.01), indicating that these hotels tend to have a higher rating on TripAdvisor. In addition, being family-friendly positive and significant for high-efficiency hotels, once again verifying their high-gearing operational performance.

Finally, Year Dummies (2011, 2012, 2013, and 2014), capturing year-specific effects are positively correlated and significant for 2011, 2012, and 2013, suggesting higher ratings in those years compared to the base year. The significance of the year indicates that some time-specific effects or trends impact the score ratings, which could be due to economic, industry-wide, or other external factors affecting the hospitality sector during those years.

R-squared indicates that the models explain between approximately 13.33% to 15.38% of the variance in the dependent variable when the full sample is considered. R-squared increases from Model 1 to Model 4, indicating that the additional variables in the later models explain more of the variance in the dependent variable. The percentage climbs to 28,09% for 5-star hotels. The R-squared values are relatively moderate, suggesting that while the models do explain some variability in TripAdvisor scores, there is still a substantial amount of variation that is not captured by these variables. By the way, although relatively low in explanatory capacity, the results suggest that factors such as the total assets, turnover per bed, presence of a pool, and whether a hotel is family-friendly, have an impact on customer satisfaction. It is worthwhile to bear in mind that, satisfaction is a human sentiment and the conjunction of multiple psychological, and personal temperaments, state of the mind, and body conditions, is involved in the overall perception of a person for the experience he receives from his visit to a hotel, it is difficult to be fully measured and related to financial factors. Moreover, a set of other factors play an important role in the overall satisfaction sentiment a hotel visitor gets. Traveling experience is one of them. The more experienced a customer is, the better measure for shaping a rational judgment he or she has. Furthermore, the anticipation and the expectation he or she has for a hotel—created by the hotel's site and previous ratings and comments—and the discrepancy between expectations and reality, shape the final perception of satisfaction, making the client sometimes over or underreact to pleasant or bad surprises. Finally, gender, age, as well as reasons such as overall experience from the destination, or unforeseen events (such as an accident or a sickness/malaise) during staying at a hotel, may affect or distort final judgment. Consequently, a vast field for further research is open.

Chapter 7. Discussion & Conclusions

Customer satisfaction is a key goal in the hotel sector. While many service quality factors have been linked to customer satisfaction, the connection between hotels' financial performance and customer satisfaction remains largely unexplored. This thesis investigates the relationship between online ratings from TripAdvisor, representing a numerical version of customer satisfaction and hotel financial performance. Greece has been chosen as the primary research location for this study due to its prominent status as a major international tourist destination. The country boasts a significant number of hotels operating within its borders. Moreover, as previously established, tourism plays a pivotal role in contributing to the overall economy of Greece. Customer satisfaction is arguably an important performance measure in the hotel industry and achieving it is essential for maintaining a sustainable and high-quality influx of tourists to the country. The research involved collecting, correlating, and analyzing data from TripAdvisor ratings, financial statements, qualitative hotel features, and tourism sector information.

Financial data spanning from 2007 to 2016 on hotels in the 3-star, 4-star, and 5-star categories are employed for the purpose of the study. A few panel data models with robust standard errors were estimated, while accounting for the usual array of control variables typically found in tourism research. Then the models are saturated with hotel and time fixed effects, while also a battery of robustness checks is conducted, including a breakdown of the analysis by hotel size, star rating, technical efficiency, and prefecture. For the estimates of hotel technical efficiency, the study relied upon stochastic frontier analysis. For the creation of the dependent variable time series, the average overall rating for each year in the study period was utilized.

Results show that there is a significant reverse relation indicating that an increase in a hotel's total assets leads to a decrease in its TripAdvisor rating, all other factors being constant. This suggests that larger hotels, in terms of assets, generally receive lower ratings. However, the positive and significant coefficient for the squared log of Total Assets across all models indicates a non-linear, U-shaped relationship with total assets. Initially, ratings decline with an increase in assets, but then start to improve at higher asset levels. This implies that very large hotels eventually experience a lessening, or even a reversal, of the negative impact on ratings. The initial decline in ratings can be attributed to the impersonal nature customers often feel in expansive hotels with multiple, similar structures. These types of hotels and resorts are usually crowded, particularly in peak seasons, and tend to offer standard or substandard service and amenities. This is particularly relevant in Greece, where many 'sea and sun' accommodations fit this description, especially on tourist-heavy islands like Rhodes, Crete, Kos, and Corfu. These hotels, attract customers through major tour operators like TUI, typically present themselves as 'touristic villages', with uniform rooms and an average level of all-inclusive services. Their focus on competitive pricing often leads to a reduction in service quality. However, past a certain point, larger total assets in a hotel start positively influencing customer satisfaction, reflecting the higher quality of upscale large hotels. When the data is segmented by star ratings, 4-star hotels display the same pattern, of an increase in a hotel's total assets leading to a decrease in its TripAdvisor rating. This supports the observation that many 'sea and sun village resorts or hotels' are classified as 4-star, often relying more on tour operators for clientele than 5-star hotels. These hotels might bear the 4-star label as an indicator of quality but frequently provide average or low service levels and experience overcrowding during the summer and thus visitors rate them lower. This trend is also evident when stratifying Total Assets into Large and Small Hotels, with Large Hotels showing a negative coefficient significant across all models. This is further validated in the stratification by both total assets (categorized as Large or Small) and star ratings (3-star, 4-star, 5-star), particularly notable for Large & 4-star hotels. The pattern persists even in the Small & 4-star category, indicating a consistent trend across the 4-star category regardless of hotel size. Lastly, when categorizing hotels based on their Stochastic Frontier Analysis (SFA) efficiency levels (Low and High), a negative significant coefficient for Low-Efficiency Hotels holds, which becomes less negative as total assets increase, as shown by the positive coefficient for Total Assets (log) squared. Essentially, large hotels that are low in efficiency and classified as 4-star tend to see a decrease in their TripAdvisor ratings as their total assets increase.

The turnover per bed ratio is positively correlated and significant or the entire sample suggesting that hotels with a higher turnover per bed tend to receive higher customer satisfaction ratings. This trend is consistent across all hotel star categories and other sample stratifications used in this analysis. Therefore, it appears that hotels that charge relatively more per bed generally achieve higher customer satisfaction levels. This implies that hotel ratings are influenced more by the quality of the experience than its cost. Moreover, a possible psychological factor behind this could be the customer's perception that paying more equates to receiving superior quality service compared to less expensive hotels, though this might be more applicable to less experienced travellers. Conversely, this could also simply mean that higher satisfaction results in better TripAdvisor ratings. This can be interpreted as hotels with higher standards typically charging more and/or having higher occupancy rates annually, reflecting more efficient operations.

The ratio of turnover to fixed assets shows a negative correlation to customer satisfaction as rated on TripAdvisor. Surprisingly, this efficient use of assets (indicating effective entrepreneurship) has a negative impact on customer satisfaction, although this relationship is not as statistically robust across the entire dataset. Upon closer examination, this negative correlation is particularly significant for smaller hotels, hotels with fewer beds, and 3-star hotels. This could be interpreted as a perception of diminished value for money, for the above-mentioned hotel types. Essentially, these smaller hotels, with lower star ratings, may give the impression of overcharging for the facilities they provide. A likely explanation is that due to their limited size and bed capacity, such hotels are unable to achieve economies of scale and operate close to their break-even point. As a result, they might charge higher rates than what is perceived as reasonable by the guests to either break even or make a marginal profit. This observation is also applicable to small 5-star hotels.

The number of ratings is positive and significant in the full sample and across all stratifications, indicating that hotels with a greater number of ratings typically receive higher TripAdvisor ratings. This trend could be attributed to the tendency of guests to leave ratings when they have a positive impression of their stay, often refraining from rating in cases of ordinary or average experiences. Essentially, hotel visitors are more inclined to provide ratings if they are satisfied with their stay. This phenomenon might create a virtuous cycle where higher ratings lead to more reviews, which in turn can enhance the hotel's visibility and perceived reliability. This association is particularly strong in the 5-star category, supporting the idea that 5-star hotels, known for better services and amenities, tend to receive more frequent and positive ratings.

The Sales Expenses/Turnover ratio indicates that higher sales expenses relative to turnover correlate with higher ratings. One interpretation of this finding is that investing in the customer's purchasing experience (such as a user-friendly reservation system) and on-site experience (like enhanced amenities and personalized services) leads to greater satisfaction and better ratings. This relation is particularly apparent in smaller hotels in terms of assets and beds and in the 4-star category. Moreover, the ratio is significantly positive for both low and high-efficiency hotels, indicating that investment in sales and marketing effectively improves TripAdvisor ratings, regardless of the hotel's efficiency.

The year-on-year first difference of the Fixed Assets variable is positive and significant suggesting that hotels that have increased their fixed assets from the previous year generally receive higher ratings. Such increases could be due to renovations or expansions in capacity or facilities, leading to higher customer satisfaction as guests perceive better value for the price paid, sometimes pleasantly surprised by improvements in the hotel's overall condition. This trend is particularly pronounced in 4-star hotels, those with low efficiency, and smaller hotels with fewer beds, where renovations often lead to a notably enhanced guest experience.

The presence of a pool is negatively correlated to visitor satisfaction across all models and stratifications. While having a pool initially seems appealing to tourists, various aspects related to the pool experience can alter their perception, leading to lower ratings. These aspects include hygiene, safety, overcrowding, availability of sunbeds, food, and beverage services, changing and toilet facilities, provision of pool towels, pool maintenance, water quality, and issues with music. Large hotels, in terms of assets and beds, often see a negative impact on satisfaction due to challenges in maintaining excellent pool conditions, particularly during peak seasons. Interestingly, the negative impact of having a pool is only significant for low-efficiency hotels, suggesting that high-efficiency hotels are more capable of managing and maintaining their pool facilities effectively.

Hotels that identify themselves as Family-friendly, have a positively correlated and statistically significant relation with customer satisfaction. Such hotels, offering amenities and services tailored to families, like en-suite rooms, cribs and baby cots, playgrounds, babysitting services, special children's programs, kids' breakfasts, and menus, tend to enhance guest satisfaction. The Family-Friendly attribute is particularly significant in 3-star hotels and smaller hotels in terms of total assets. Furthermore, being family-friendly is positively associated with and significant for high-efficiency hotels, reinforcing their effective operational performance.

In terms of efficiency, larger hotels, defined either by Total Assets or the number of Beds, generally exhibit higher technical efficiency compared to smaller ones. Both the size of the hotel and its star category appear to positively influence efficiency, with 5-star and larger hotels (by Total Assets or beds) being the most efficient on average. However, outliers exist in all-star categories, including 5-star hotels, with some performing at either the highest or lowest ends of the efficiency spectrum. Interestingly, certain 4-star hotels demonstrate greater efficiency than some 5-star hotels. Despite yearly variations, there is no distinct overall trend in technical efficiency across the years.

Furthermore, more established tourist destinations, particularly those on the island complexes, tend to have higher hotel ratings. This includes Crete, the Dodecanese (which includes Rhodes and Kos), the Cyclades (encompassing islands like Mykonos, Santorini, and Paros), the Ionian Islands (including Corfu, Zakynthos, and Lefkada), and the Northern islands (such as Thasos and Mytilene), as well as Messinia in the Peloponnese, with average ratings ranging from 4.0 to 4.47. Surprisingly, Attica (home to Athens) and Central Macedonia, the region of Greece's second-largest city, Thessaloniki, have lower average ratings, between 3.54 and 4. A likely reason is the high concentration of hotels in these metropolitan areas, many of which offer minimal service standards due to steady demand and year-round occupancy. Conversely, island hotels, competing with other national and international destinations for guests, have become more efficient and service-oriented over the years, focusing on enhancing customer satisfaction and extending the holiday season. It is also noted that less touristic, predominantly mountainous regions of Greece tend to have the lowest ratings.

When considering efficiency scores per prefecture, Attica, and Central Macedonia, along with Dodecanese, Chania, Rethymno (parts of Crete), and Lefkada, achieve higher efficiency scores during the examined period. Lefkada, Crete, and the Dodecanese likely excel due to upscale hotels, abundant clientele, and accumulated managerial expertise. In contrast, Athens and Thessaloniki areas might achieve higher efficiency due to higher turnover over total assets rather than operational efficiency which varies a lot between its hotel establishments, with intensive demand playing a key role. Notably, more established tourist prefectures generally show higher efficiency compared to less developed regions.

7.1 Policy recommendations

Finally, the findings of the study could be useful to policymakers. The Greek state may encourage through incentives the upgrading of existing underperforming "sea and sun" large in terms of assets falling in the 3- and 4-star category hotels. As the size of the hotel and its star category appear to positively influence efficiency, with 5-star and larger hotels (by Total Assets or beds) being the most efficient on average, the policymakers can encourage and pipeline new investments in the hoteling sector to that direction. In addition to the latter, as there is a great gap in the quality provided by hotels in different regions reflected in the asymmetrical TripAdvisor ratings between destinations, special incentives for the upgrading of services can be in place from the state for the underperforming regions. The same boost could be provided by the government to the regions with relatively low-efficiency scores such as Peloponnese, Eastern Macedonia, and Epirus. Finally, the state could imply more frequent controls concerning the hygiene situation and maintenance of pools.

The thesis findings provide interesting reading to the hoteling sector management teams, in the design of services and hotel guest experience. An initial useful conclusion is that during the 2007 – 2016 period "sea and sun" accommodation in Greece was widely offered by large-scale, 4-star hotels, offering diminished level services on average, resulting in lower customer satisfaction. These hotels operate in collaboration with massive tour operators, having price as their edge to compete. Consequently, an upgrading of the services and amenities of these hotels will improve tourist satisfaction leading eventually to better pricing for these hotels. Moreover, Smaller hotels, with fewer beds, falling into the 3-star hotels category present a negative value-for-money perception to clients, seeming to overcharge. A suggestion for the concrete type of hotels is, instead of diminishing their prices, to enhance their amenities and pamper clients with tangible and intangible services. Furthermore, the study concluded that sales expenses over turnover are highly related to satisfaction. The argument holds especially for small hotels in terms of assets and beds and the 4-star category. Consequently, spending wisely on marketing expenses is a way to enjoy higher ratings and better reviews on the eWoM world resulting in generating clientele. Additionally, regularly audit and optimize sales and administrative expenses, fostering better guest engagement. The presence of a pool is negatively correlated to visitor satisfaction, especially in large hotels in terms of assets and beds. Interestingly, the negative impact of having a pool is only significant for low-efficiency hotels. In that sense, managers must be very conscious of issues related to hygiene, infrastructure maintenance, and service of pools. Finally, family-friendly hotels achieve higher average TripAdvisor ratings. Consequently, adding services and infrastructure for the easing of parents' worries and the pastime of children is a recommendation for existing and new hotels.

The findings of the study can be a useful consideration for the shaping of new investments in hotel infrastructure undertaken by entrepreneurs and investors in the future as well. The study claims that a specific category of hotels that are larger in terms of total assets, with more beds and falling into the 5-star category attract higher ratings on TripAdvisor and seem to satisfy guest's needs at a higher level. Moreover, large in terms of total assets only hotels, falling at the same time into the 3-star and 4-star category, showcase diminished customer satisfaction. Contrastingly, hotels with an average number of rooms, having a boutique size and falling on the 5-star category, seem to satisfy their clientele better. Consequently, shifting towards either investments in larger in terms of total assets, with more beds and falling into the 5-star category or in smaller in boutique size hotels, more delicate hotels, offering tailor-made services is a proposition for new investments in tourism lodging in Greece. Furthermore, to boost customer satisfaction, existing hotels need to invest in constant refurbishments and renovations. This is particularly true for 4-star hotels with a lower amount of fixed assets and a lower number of beds. Finally, an important number of hotels present a very low technical efficiency, leaving space for important improvements, managerially and infrastructurally wise. An investment opportunity arises from the previous findings, as investors and entrepreneurs can invest in underperforming existing hoteling infrastructure, proceed with infrastructure renovation, upgrade their services, and maybe brand them as well, which will lead to substantial future profits. Another opportunity arises from the study's findings concerning the regional technical efficiency of hotels. An important part of the Greek territory is still underdeveloped in hoteling infrastructure, encompassing hotels with average low efficiency and low visitor's satisfaction. As most of these regions host an abundance of natural beauty, important cultural and historical themes, and transport infrastructure is significantly improved lately, they possess dynamics of emerging destinations, to be explored.

7.2 Suggestions for further research

Several directions for future research are provided following this study. The concept of hotel performance is intricate and multifaceted, as depicted in international literature that employs a diverse range of measures for academic evaluation. Qualitative variables, such as the utilization of modern technologies for booking, operations, and customer service, along with their impact on efficiency and guest satisfaction, as well as factors like proximity to the beach, distance from the city center (for city hotels), and the presence of amenities like a restaurant or spa, could serve as characteristics for future research. Additionally, exploring differences between branded and non-branded hotels could be a valuable avenue for further analysis.

Beyond these considerations, overarching trends influencing the tourism landscape, such as the adoption of ESG initiatives, cultural and environmental sensitivity, integration with the local community, guest personalization, customization, narrative, and storytelling, are variables that may be challenging to quantify numerically but can exert a critical influence on customer satisfaction and the financial performance of hotels.

It's important to note that the present research focused on a Greek hotel sample, utilizing data spanning from 2007 to 2016. This period was relatively calm in terms of major geopolitical events affecting inbound tourism to Greece. The Greek economic crisis unfolded during 2008–2009, it did not drastically alter the tourism flow in terms of arrivals and receipts due to the fact that tourism demand is mostly exogenous driven. Moreover, internationally, events like the Lebanon war (2006), the global financial crisis (2007-2008), the Arab Spring (2010-2012), the Syrian civil war (started in 2011), the annexation of Crimea by Russia (2014), the rise of ISIS (2014), and the European refugee crisis (started in 2015) had some impact but not a drastic one on tourism flows in Greece.

In contrast, from the end of the investigated period until 2023, significant events had a profound impact on the tourism and hotel sectors. Notable occurrences include the COVID-19 pandemic (2019), wars in Ukraine (2021) and Gaza (2023), increases in fuel prices and inflation, climate disruptions, and the rise of Airbnb. These events are considered serious disruptors, often described metaphorically as "black swans," and understanding their impact is crucial for academic research as the frequency of such unexpected events seems to be on the rise.

7.3 Limitations of study

This section outlines the limitations of this thesis. First, TripAdvisor ratings as a measure of customer satisfaction may suffer from measurement and self-selection bias. The former arises when customer satisfaction is not accurately measured by the instrument, that is, the platform's rating system. The latter, and more detrimental bias, arises due to the voluntary nature of these reviews. Moreover, TripAdvisor ratings are subjective, reflecting individual guest experiences and expectations. They can vary widely based on factors unrelated to the hotel's objective qualities or financial performance, such as personal guest preferences or one-off incidents.

Secondly, there is no concrete evidence that the financial indices—used in the analysis as a proxy for various hotel financial performance indicators—accurately measure what they intend. Moreover, measurement errors or inconsistencies in the covariates might exist due to the bookkeeping practices of the hotels. Additionally, the selection of the final sample of hotels has led to the exclusion of hotels with poor bookkeeping practices which would, on average have lower TripAdvisor rankings—assuming wise bookkeeping is positively associated with better hotel service. This limits the representability of the sample and therefore the transportability of the results to hotels with good bookkeeping practices. Furthermore, information about amenities (like pools) may be challenging to obtain. Consequently, inconsistencies or inaccuracies in data can skew results.

Additionally, the analysis is limited in describing associations and does not deal with intermediate mechanisms that the customer experiences (i.e., hotel services) and cause the rating score. Accordingly, one cannot exclude the reverse causality scenario, that is, TripAdvisor ratings causing financial performance (Sayfuddin and Chen, 2021). Furthermore, the regression analysis may identify correlations between hotel ratings and explanatory variables, but it cannot conclusively establish causation. Some explanatory variables might be endogenous, i.e., they could be influenced by the hotel ratings themselves. For instance, hotels with higher ratings might attract more guests, leading to higher turnover. Another issue of the study is that combining financial figures with qualitative dummies (such as the presence of a pool) in the same model is challenging, as they represent fundamentally different types of data. In addition to the latter, the study might not account for temporal changes like seasonal variations, economic cycles, or shifts in consumer behaviour over time, which can affect both hotel performance and ratings. Another baffling issue is that TripAdvisor ratings can be subject to manipulation, such as fake reviews which can distort the true performance of a hotel. Another consideration is that findings from the sample of Greek hotels may not apply to hotels in other countries or regions due to cultural, economic, and market differences.

Concerning the methodology used, there is a possibility of multicollinearity, especially with financial variables, where one financial metric might be highly correlated with another, leading to difficulties in isolating the impact of individual variables. In addition to that, the model assumes that the individual-specific random effects (μi) are uncorrelated with the explanatory variables (Xit). If this assumption does not hold, it may lead to biased estimates. This is especially critical given the diverse nature of hotels in terms of size, location, and clientele. As was mentioned, the explanatory variables selected (such as number of beds, total assets, efficiency computed through SFA, own capital, fixed assets, and turnover) do not encompass all factors that influence TripAdvisor ratings. While Huber-White Robust standard errors are used to address potential heteroscedasticity and autocorrelation, they do not correct for potential biases in the coefficient estimates that could arise from omitted variables or endogeneity issues. Moreover, applying logarithmic transformations to variables aims to address issues like nonlinearity and skewed distributions. However, this transformation alters the interpretation of the coefficients, which can complicate the understanding of the relationships between variables. Finally, while random effects account for unobserved heterogeneity across entities, there may still be unaccounted-for factors that vary within entities over time, affecting the reliability of the model's predictions.

When using Stochastic Frontier Analysis (SFA) to calculate the technical efficiency of hotels limitations concerning the sample size, period, choice of inputs, and outputs: may hold. It is worthwhile to mention that in the service industry, like hotels, inputs and outputs might be interdependent. Standard SFA models may not adequately capture these interdependencies. Moreover, SFA typically focuses on operational efficiency and may not fully account for external factors such as economic conditions, regulatory changes, or competitive dynamics, which can also impact hotel performance. The translog model allows for a flexible representation of the production function, accommodating non-linear relationships and interactions among variables. However, this complexity also means that the model is more data-intensive and sensitive to data quality. Any inaccuracies in the data could significantly affect the results. Finally, the decomposition of the error term into a random component (v_it) and a non-random inefficiency component (u_it) is a key feature of SFA. The assumptions about the distribution of these components affect the efficiency scores. Any misspecification here can lead to biased estimates of technical efficiency.

Bibliography

Aakash, A., & Gupta Aggarwal, A. (2022). Assessment of hotel performance and guest satisfaction through eWOM: big data for better insights. International Journal of Hospitality & Tourism Administration, 23(2), 317-346.

Abrate, G., Capriello, A., & Fraquelli, G. (2011). When quality signals talk: Evidence from the Turin hotel industry. *Tourism Management*, *32*(4), 912-921.

Aigner, D., Lovell, C. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of econometrics*, *6*(1), 21-37.

Aissa, S. B., & Goaied, M. (2016). Determinants of tourism hotel market effciency. *International Journal of Culture, Tourism and Hospitality Research*.

Akan, P. (1995). Dimensions of service quality: a study in Istanbul. *Managing Service Quality: An International Journal*, *5*(6), 39-43.

Alam, M. S., & Paramati, S. R. (2017). The dynamic role of tourism investment on tourism development and CO2 emissions. *Annals of Tourism Research*, *66*, 213-215.

Albaladejo, I. P., González-Martínez, M. I., & Martínez-García, M. P. (2016). Nonconstant reputation effect in a dynamic tourism demand model for Spain. *Tourism Management*, *53*, 132-139.

Alderman, H., Behrman, J. R., Kohler, H. P., Maluccio, J. A., & Watkins, S. C. (2001). Attrition in longitudinal household survey data: some tests for three developing-country samples. *Demographic research*, *5*, 79-124.

Allison, P. D. (2005). *Fixed effects regression methods for longitudinal data using SAS*. Sas Institute.

Amin, S. B., Kabir, F. A., & Khan, F. (2020). Tourism and energy nexus in selected South Asian countries: a panel study. *Current Issues in Tourism*, *23*(16), 1963-1967.

Anderson, R. I., Fish, M., Xia, Y., & Michello, F. (1999). Measuring efficiency in the hotel industry: A stochastic frontier approach. *International journal of hospitality Management*, *18*(1), 45-57.

Anderson, R. I., Fok, R., & Scott, J. (2000). Hotel industry efficiency: An advanced linear programming examination. *American Business Review*, *18*(1), 40.

Assaf, A. G. (2012). Benchmarking the Asia Pacific tourism industry: A Bayesian combination of DEA and stochastic frontier. Tourism Management, 33(5), 1122-1127.

Assaf, A. G. and Barros, C. (2011). Performance analysis of the Gulf hotel industry: A Malmquist index with bias correction. International Journal of Hospitality Management, 30(4):819–826.

Assaf, A. G., & Agbola, F. W. (2011). Modelling the performance of Australian hotels: a DEA double bootstrap approach. *Tourism economics*, *17*(1), 73-89.

Assaf, A. G., & Tsionas, M. G. (2019). A review of research into performance modeling in tourism research-Launching the Annals of Tourism Research curated collection on performance modeling in tourism research. Annals of Tourism Research, 76, 266-277.

Assaf, A. G., A. Josiassen, K. L. Cvelbar (2012) Does triple bottom line reporting improve hotel performance? International journal of hospitality management 31, 596-600.

Assaf, A. G., Oh, H., & Tsionas, M. (2017). Bayesian approach for the measurement of tourism performance: A case of stochastic frontier models. Journal of Travel Research, 56(2), 172-186.

Assaf, A., & Cvelbar, K. L. (2011). Privatization, market competition, international attractiveness, management tenure and hotel performance: Evidence from Slovenia. *International Journal of Hospitality Management*, *30*(2), 391-397.

Assaf, A., Barros, C. P., & Josiassen, A. (2010). Hotel efficiency: A bootstrapped metafrontier approach. *International Journal of Hospitality Management*, *29*(3), 468-475.

Baik, B., Chae, J., Choi, S., & Farber, D. B. (2013). Changes in operational efficiency and firm performance: A frontier analysis approach. *Contemporary Accounting Research*, *30*(3), 996-1026.

Baka, V. (2016). The becoming of user-generated reviews: Looking at the past to understand the future of managing reputation in the travel sector. Tourism Management, 53, 148–162.

Bakar, N. A., & Rosbi, S. (2020). Effect of Coronavirus disease (COVID-19) to tourism industry. *International Journal of Advanced Engineering Research and Science*, *7*(4).

Balestra, P., & Nerlove, M. (1966). Pooling cross section and time series data in the estimation of a dynamic model: The demand for natural gas. *Econometrica: Journal of the Econometric Society*, 585-612.

Baltes, P. B., & Nesselroade, J. R. (1979). History and rationale of longitudinal research. In *Longitudinal research in the study of behavior and development* (pp. 1-39). Academic Press.

Banerjee, S., & Chua, A. Y. K. (2016). In search of patterns among travellers' hotel ratings in TripAdvisor. *Tourism Management*, 53, 125–131.

Barros, C. A. P., & Santos, C. A. (2006). The measurement of efficiency in Portuguese hotels using data envelopment analysis. *Journal of Hospitality & Tourism Research*, *30*(3), 378-400.

Barros, C. P. (2005). Evaluating the efficiency of a small hotel chain with a Malmquist productivity index. *International Journal of tourism research*, *7*(3), 173-184.

Barros, C. P., & Alves, F. P. (2004). Productivity in the tourism industry. *International Advances in Economic Research*, *10*(3), 215-225.

Barros, C. P., & Dieke, P. U. (2008). Technical efficiency of African hotels. *International Journal of Hospitality Management*, *27*(3), 438-447.

Barros, C. P., & Mascarenhas, M. J. (2005). Technical and allocative efficiency in a chain of small hotels. *International journal of hospitality management*, *24*(3), 415-436.

Barros, C. P., Botti, L., Peypoch, N., Robinot, E., & Solonandrasana, B. (2011). Performance of French destinations: Tourism attraction perspectives. *Tourism Management*, *32*(1), 141-146.

Barros, C. P., Peypoch, N., & Solonandrasana, B. (2009). Efficiency and productivity growth in hotel industry. *International Journal of Tourism Research*, *11*(4), 389-402.

Bell, R. A., & Morey, R. C. (1995). Increasing the efficiency of corporate travel management through macro benchmarking. *Journal of travel research*, *33*(3), 11-20.

Blalock, C. (2004). First-quarter results reaffirm recovery. *Hotel & Motel Management*, *219*(11), 1-2.

Bolat, T., & Yılmaz, Ö. (2009). The relationship between outsourcing and organizational performance: is it myth or reality for the hotel sector?. *International journal of contemporary hospitality management*, *21*(1), 7-23.

Botti, L., Briec, W., & Cliquet, G. (2009). Plural forms versus franchise and company-owned systems: A DEA approach of hotel chain performance. *Omega*, *37*(3), 566-578.

Brau, R., Lanza, A., & Pigliaru, F. (2007). How fast are small tourism countries growing? Evidence from the data for 1980–2003. *Tourism Economics*, *13*(4), 603-613.

Brida, J. G., Garrido, N., Deidda, M., & Pulina, M. (2012). Exploring the dynamics of the efficiency in the Italian hospitality sector. A regional case study. *Expert Systems with Applications*, *39*(10), 9064-9071.

Briggs, S., Sutherland, J., & Drummond, S. (2007). Are hotels serving quality? An exploratory study of service quality in the Scottish hotel sector. *Tourism management*, *28*(4), 1006-1019.

Brotherton, B., & Mooney, S. (1992). Yield management—progress and prospects. *International Journal of Hospitality Management*, *11*(1), 23-32.

Brown, J. R., & Ragsdale, C. T. (2002). The competitive market efficiency of hotel brands: An application of data envelopment analysis. *Journal of Hospitality & Tourism Research*, *26*(4), 332-360.

Bulchand-Gidumal, J., Melián-González, S., & González López-Valcárcel, B. (2011). Improving hotel ratings by offering free Wi-Fi. Journal of Hospitality and Tourism Technology, 2(3), 235-245.

Callan, R. J. (1998). Attributional analysis of customers' hotel selection criteria by UK grading scheme categories. *Journal of Travel Research*, *36*(3), 20-34.

Callan, R. J. (1999). Augmenting the range of hotel selection attributes: Perceptions of managers and customers of three-to five-star UK hotels compared. *Tourism and Hospitality Research*, *1*(3), 205-230.

Callan, R. J., & Bowman, L. (2000). Selecting a hotel and determining salient quality attributes: a preliminary study of mature British travellers. *International Journal of tourism research*, *2*(2), 97-118.

Casalo, L. V., Flavian, C., Guinaliu, M., & Ekinci, Y. (2015). Do online hotel rating schemes influence booking behaviors. *International Journal of Hospitality Management*, *49*, 28-36.

Chaabouni, S. (2019). China's regional tourism efficiency: A two-stage double bootstrap data envelopment analysis. *Journal of destination marketing & management*, *11*, 183-191.

Chang, H. H., & Sokol, D. D. (2022). How incumbents respond to competition from innovative disruptors in the sharing economy—The impact of Airbnb on hotel performance. *Strategic Management Journal*, *43*(3), 425-446.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European journal of operational research*, *2*(6), 429-444.

Charnes, A., Cooper, W. W., Golany, B., Seiford, L., & Stutz, J. (1985). *Foundations of Data Envelopment Analysis for Pareto-Koopmans Efficient Empirical Productions Functions* (No. CCS-504). Texas Univ at Austin center for cybernetic studies.

Charnes, A., Cooper, W. W., Huang, Z. M., & Sun, D. B. (1990). Polyhedral cone-ratio DEA models with an illustrative application to large commercial banks. *Journal of econometrics*, *46*(1-2), 73-91.

Chen, C. F. (2007). Applying the stochastic frontier approach to measure hotel managerial efficiency in Taiwan. *Tourism Management*, *28*(3), 696-702.

Chen, F. H., Hsu, T. S., & Tzeng, G. H. (2011). A balanced scorecard approach to establish a performance evaluation and relationship model for hot spring hotels based on a hybrid MCDM model combining DEMATEL and ANP. *International Journal of Hospitality Management*, 30(4), 908-932.

Chen, M. H. (2010). The economy, tourism growth and corporate performance in the Taiwanese hotel industry. *Tourism Management*, *31*(5), 665-675.

Chen, M. H., & Kim, W. G. (2010). Hotel valuation in China: A case study of a state-owned hotel. *Cornell Hospitality Quarterly*, *51*(3), 429-445.

Chen, T. H. (2009) Performance measurement of an enterprise and business units with an application to a Taiwanese hotel chain. International journal of hospitality management 28, 415-422.

Choi, T. Y. and Chu, R. (2001). Determinants of hotel guests' satisfaction and repeat patronage in the Hong Kong hotel industry. *International Journal of Hospitality Management*, 20(3):277–297.

Choi, T. Y., & Chu, R. (2000). Levels of satisfaction among Asian and Western travellers. *International Journal of Quality & Reliability Management*, *17*(2), 116-132.

Choi, T. Y., & Chu, R. K. (1999). Consumer perceptions of the quality of services in three hotel categories in Hong Kong. Journal of Vacation Marketing, 5(2), 176-189.

Christos Alexakis, Michael Dowling, Vasileios Pappas, Manimozhi Ramachandiran, Favianos Sklavos, (2021). Do hotel financial factors influence satisfaction? Annals of Tourism Research, Volume 90.

Corne, A. (2015). Benchmarking and tourism efficiency in France. *Tourism Management*, *51*, 91-95.

Cornwell, C., & Schmidt, P. (2008). Stochastic frontier analysis and efficiency estimation. In *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice* (pp. 697-726). Berlin, Heidelberg: Springer Berlin Heidelberg.

Cracolici, M. F., Nijkamp, P., & Cuffaro, M. (2007). Efficiency and productivity of Italian tourist destinations: a quantitative estimation based on data envelopment analysis and the Malmquist method. In *Advances in modern tourism research: Economic perspectives* (pp. 325-343). Heidelberg: Physica-Verlag HD.

Crouch, G. I. (1994). The study of international tourism demand: A review of findings. *Journal of Travel research*, *33*(1), 12-23.

Crouch, G. I., & Shaw, R. N. (1992). International tourism demand: a meta-analytical integration of research findings (pp. 175-207). Mansell Publishing.

Cser, K., & Ohuchi, A. (2008). World practices of hotel classification systems. *Asia Pacific Journal of Tourism Research*, *13*(4), 379-398.

Danziger, S., Israeli, A., & Bekerman, M. (2006). The relative role of strategic assets in determining customer perceptions of hotel room price. International Journal of Hospitality Management, 25(1), 129-145.

Davutyan, N. (2007). Measuring the quality of hospitality at Antalya. International Journal of Tourism Research, 9(1), 51-57.

De Pelsmacker, P., Van Tilburg, S., & Holthof, C. (2018). Digital marketing strategies, online reviews and hotel performance. International Journal of Hospitality Management, 72, 47-55.

Denizci Guillet, B., & Tasci, A. D. (2010). Travelers' takes on hotel—Restaurant co-branding: Insights for China. *Journal of Hospitality & Tourism Research*, *34*(2), 143-163.

Dioko, L. A., So, S. I. A., & Harrill, R. (2013). Hotel category switching behavior—Evidence of mobility, stasis or loyalty. *International Journal of Hospitality Management*, *34*, 234-244.

Donaghy, K., McMahon, U., & McDowell, D. (1995). Yield management: an overview. *International journal of hospitality management*, *14*(2), 139-150.

Dritsakis, N. (2004). Tourism as a long-run economic growth factor: an empirical investigation for Greece using causality analysis. *Tourism economics*, *10*(3), 305-316.

Dwyer, L., & Kim, C. (2003). Destination competitiveness: determinants and indicators. *Current issues in tourism*, *6*(5), 369-414.

Eleftheriou K., Sambracos E. (2019). Tourism–growth nexus and spatial spillovers: Evidence from Greece. Tourism Economics, 25(2) 297–302.

Espino-Rodríguez, T. F., & Padrón-Robaina, V. (2005). A resource-based view of outsourcing and its implications for organizational performance in the hotel sector. *Tourism management*, *26*(5), 707-721.

Eugenio-Martin, J. L., Martín Morales, N., & Scarpa, R. (2004). Tourism and economic growth in Latin American countries: A panel data approach.

Eugenio-Martin, J. L., Martín-Morales, N., & Sinclair, M. T. (2008). The role of economic development in tourism demand. *Tourism Economics*, *14*(4), 673-690.

Fang, B., Ye, Q., & Law, R. (2016). Effect of sharing economy on tourism industry employment. *Annals of tourism research*, *57*, 264-267.

Fernández, M. C. L., & Bedia, A. M. S. (2004). Is the hotel classification system a good indicator of hotel quality?: An application in Spain. *Tourism Management*, *25*(6), 771-775.

Fernández, M. C. L., & Bedia, A. M. S. (2005). Applying SERVQUAL to diagnose hotel sector in a tourist destination. Journal of Quality Assurance in Hospitality & Tourism, 6(1-2), 9-24.

Fernández-Barcala, M., González-Díaz, M., & Prieto-Rodriguezb, J. (2009). Factors influencing guests' hotel quality appraisals. European Journal of Tourism Research, 2(1), 25-40.

Fitzgerald, J., Gottschalk, P., & Moffitt, R. (1998). *An analysis of sample attrition in panel data: The Michigan Panel Study of Income Dynamics* (No. t0220). National Bureau of Economic Research.

Frees, E. W. (2004). *Longitudinal and panel data: analysis and applications in the social sciences*. Cambridge University Press.

Fu, Y., Li, D., & Li, N. (2011, August). Hotel performance evaluation based on cross-efficiency DEA models. In *2011 International Conference on Management and Service Science* (pp. 1-4). IEEE.

Fuentes, R. (2011). Efficiency of travel agencies: A case study of Alicante, Spain. *Tourism management*, *32*(1), 75-87.

Garin-Munoz, T., & Montero-Martín, L. F. (2007). Tourism in the Balearic Islands: A dynamic model for international demand using panel data. *Tourism management*, *28*(5), 1224-1235.

Garrigós-Simón, F. J., Palacios Marqués, D., & Narangajavana, Y. (2005). Competitive strategies and performance in Spanish hospitality firms. *International Journal of Contemporary Hospitality Management*, *17*(1), 22-38.

Geetha, M., Singha, P., & Sinha, S. (2017). Relationship between customer sentiment and online customer ratings for hotels—An empirical analysis. Tourism Management, 61, 43–54.

Gong, B. H., & Sickles, R. C. (1992). Finite sample evidence on the performance of stochastic frontiers and data envelopment analysis using panel data. *Journal of Econometrics*, *51*(1-2), 259-284.

Gonzalez, P., & Moral, P. (1995). An analysis of the international tourism demand in Spain. *International Journal of forecasting*, *11*(2), 233-251.

Gössling, S., Hall, C. M., & Andersson, A. C. (2018). The manager's dilemma: a conceptualization of online review manipulation strategies. Current Issues in Tourism, 21(5), 484-503.

Gössling, S., Scott, D., & Hall, C. M. (2020). Pandemics, tourism and global change: a rapid assessment of COVID-19. Journal of sustainable tourism, 29(1), 1-20.

Green, W. (1999). Econometric Analysis Third Edition Prentice-Hall.

Grönroos, C. (2000). Relationship Marketing. *Handbook of relationship marketing*, *95*.

Guetat, H., Jarboui, S., & Boujelbene, Y. (2015). Evaluation of hotel industry performance and corporate governance: A stochastic frontier analysis. Tourism Management Perspectives, 15, 128-136.

Gunduz, L., & Hatemi-J, A. (2005). Is the tourism-led growth hypothesis valid for Turkey? *Applied Economics Letters*, *12*(8), 499-504.

Han, H. (2012). The relationship among corporate culture, strategic orientation, and financial performance. Cornell Hospitality Quarterly, 53(3):207–219.

Hao, F., Xiao, Q., & Chon, K. (2020). COVID-19 and China's hotel industry: Impacts, a disaster management framework, and post-pandemic agenda. *International journal of hospitality management*, *90*, 102636.

Harris, P. J., & Mongiello, M. (2001). Key performance indicators in European hotel properties: general managers' choices and company profiles. *International Journal of Contemporary Hospitality Management*, *13*(3), 120-128.

Hathroubi, S., Peypoch, N., & Robinot, E. (2014). Technical efficiency and environmental management: The Tunisian case. *Journal of Hospitality and Tourism Management*, *21*, 27-33.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the Econometric Society*, 1251-1271.

Hensens, W. (2015). The future of hotel ratings. *Journal of Tourism Futures*, *1*(1), 69-73.

Hoch, I. (1962). Estimation of production function parameters combining time-series and cross-section data. *Econometrica: journal of the Econometric Society*, 34-53.

Hsiao, C. (2014). *Analysis of panel data* (No. 54). Cambridge University Press.

Hsieh, L. F., Wang, L. H., Huang, Y. C., & Chen, A. (2010). An efficiency and effectiveness model for international tourist hotels in Taiwan. *The Service Industries Journal*, *30*(13), 2183-2199.

Hua, N., O'Neil, J., & Mattila, A. S. (2010). More Marketing Expenditures, Better Hotel Financial Performance?. *Journal of Hospitality Financial Management*, *16*(1), 4.

Huang, Y., Mesak, H. I., Hsu, M. K., & Qu, H. (2012). Dynamic efficiency assessment of the Chinese hotel industry. *Journal of Business Research*, *65*(1), 59-67.

Hwang, S. N., & Chang, T. Y. (2003). Using data envelopment analysis to measure hotel managerial efficiency change in Taiwan. *Tourism management*, *24*(4), 357-369.

Ibrahim, M. A. M. A. (2011). The Determinants of International Tourism Demand for Egypt: Panel Data Evidence. *European Journal of Economics, Finance and Administrative Sciences*, (30).

Israeli, A. A. (2002). Star rating and corporate affiliation: their influence on room price and performance of hotels in Israel. International Journal of Hospitality Management, 21(4), 405-424.

Israeli, A. A., & Uriely, N. (2000). The impact of star ratings and corporate affiliation on hotel room prices in Israel. *Tourism and Hospitality Research*, *2*(1), 27-36.

Jambor, A., & Leitão, N. (2017). Economic Growth and sustainable development: evidence from Central and Eastern Europe. *International Journal of Energy Economics and Policy*, *7*(5), 171-177.

Jiang, Y., & Wen, J. (2020). Effects of COVID-19 on hotel marketing and management: a perspective article. *International journal of contemporary hospitality management*, *32*(8), 2563-2573.

Johns, N., Howcroft, B., & Drake, L. (1997). The use of data envelopment analysis to monitor hotel productivity. *Progress in tourism and hospitality research*, *3*(2), 119-127.

Johnson, E. B. (1960). Standard Operating Procedure for the Pool Critical Assembly (No. ORNL-2449). Oak Ridge National Lab., Tenn.

Kaul, R. N. (1985). *Dynamics of tourism: A trilogy* (Vol. 1). Sterling.

Kavoura, A., & Borges, M. T. T. (2016). Understanding online communities on social networks via the notion of imagined communities: The case of Tripdvisor. International Journal of Web Based Communities, 12(3), 238–261.

Keum, K. (2011). International tourism and trade flows: A causality analysis using panel data. *Tourism Economics*, *17*(5), 949-962.

Khadaroo, J., & Seetanah, B. (2007). Transport infrastructure and tourism development. *Annals of tourism research*, *34*(4), 1021-1032.

Khoshnevis-Yazdi, S., & Khanalizadeh, B. (2017). Tourism demand: A panel data approach. *Current Issues in Tourism*, *20*(8), 787-800.

Kim, H. J., & Chen, M. H. (2006). Tourism expansion and economic development: The case of Taiwan. *Tourism management*, *27*(5), 925-933.

Kim, W. G., & Park, S. A. (2017). Social media review rating versus traditional customer satisfaction: which one has more incremental predictive power in explaining hotel performance?. International journal of contemporary hospitality management, 29(2), 784-802.

Kim, W. G., Cho, M., and Brymer, R. A. (2013). Determinants affecting comprehensive property-level hotel performance: The moderating role of hotel type. International Journal of Hospitality Management, 34:404–412.

Kitamura, R. (2000). Longitudinal methods. *Handbook of transport modelling*, *50*, 113-129.

Knutson, B. J. (1988). Frequent travelers: Making them happy and bringing them back. *Cornell Hotel and Restaurant Administration Quarterly*, *29*(1), 82-87.

Kozak, M., & Rimmington, M. (1999). Measuring tourist destination competitiveness: conceptual considerations and empirical findings. *International journal of hospitality management*, *18*(3), 273-283.

Kuh, E. (1959). The validity of cross-sectionally estimated behavior equations in time series applications. Econometrica: Journal of the Econometric Society, 197-214.

Kumbhakar, S. C. (1987). The specification of technical and allocative inefficiency in stochastic production and profit frontiers. *Journal of econometrics*, *34*(3), 335-348.

Kumbhakar, S. C., & Lovell, C. K. (2000). Stochastic production frontier. *Cambridge University Press. Kumbhakar, SC, & Sarkar, S.(2003). Deregulation, ownership and productivity growth in the banking industry: Evidence from India. Journal of Money Credit and Banking*, *35*(3), 403424.

Kuosmanen, T., Johnson, A., & Saastamoinen, A. (2015). Stochastic nonparametric approach to efficiency analysis: A unified framework. In *Data envelopment analysis* (pp. 191-244). Springer, Boston, MA.

Lado-Sestayo, R., & Fernández-Castro, Á. S. (2019). The impact of tourist destinations on hotel efficiency: A data envelopment analysis approach. European Journal of Operational Research, 272(2), 674-686.

Lazarsfeld, P., & Fiske, M. (1938). The "panel" as a new tool for measuring opinion. *Public Opinion Quarterly*, *2*(4), 596-612.

Ledesma-Rodríguez, F. J., Navarro-Ibáñez, M., & Pérez-Rodríguez, J. V. (2001). Panel data and tourism: a case study of Tenerife. *Tourism Economics*, *7*(1), 75-88.

Lee, M., & Youn, S. (2009). Electronic word of mouth (eWOM). International Journal of Advertising, 28(3), 473–499.

Lee, Y. H., Lu, L. T., & Sung, A. D. (2012). A measure to the operational performance of international hotels in Taiwan: DEA and Malmquist approach. *Review of Economics & Finance*, *2*, 73-83.

Leitão, N. C. (2010). Does trade help to explain tourism demand? The case of Portugal. *Theoretical and Applied Economics*, 63-74.

Li, R., Guo, Q., Wu, D., Yin, H., Zhang, H., & Zhu, T. (2014). Spatial characteristics of development efficiency for urban tourism in eastern China: A case study of six coastal urban agglomerations. *Journal of Geographical Sciences*, *24*, 1175-1197.

Lim, C. (1997). Review of international tourism demand models. *Annals of tourism research*, *24*(4), 835-849.

Lim, C. (1999). A meta-analytic review of international tourism demand. *Journal of Travel Research*, *37*(3), 273-284.

Lopez-Fernandez, M. C., Serrano-Bedia, A. M., & Gómez-López, R. (2016). Determinants of innovation decision in small and medium-sized family enterprises. *Journal of Small Business and Enterprise Development*, *23*(2), 408-427.

Luo, H., Yang, Y., & Law, R. (2014). How to achieve a high efficiency level of the hotel industry? *International Journal of Contemporary Hospitality Management*.

Manasakis, C., Apostolakis, A., & Datseris, G. (2013). Using data envelopment analysis to measure hotel efficiency in Crete. *International Journal of Contemporary Hospitality Management*.

Mansourirad, E. (2013). A categorical fuzzy DEA method to evaluate efficiency of hotels based on stars rating. *Applied Mathematical Sciences*, *7*(73), 3625-3628.

Martin-Fuentes, E. (2016). Are guests of the same opinion as the hotel star-rate classification system? *Journal of Hospitality and Tourism Management*, 29, 126-134.

Martins, L. F., Gan, Y., & Ferreira-Lopes, A. (2017). An empirical analysis of the influence of macroeconomic determinants on World tourism demand. *Tourism management*, *61*, 248-260.

Mayzlin, D., Dover, Y., & Chevalier, J. (2014). Promotional reviews: An empirical investigation of online review manipulation. American Economic Review, 104(8), 2421-2455.

Meeusen, W., & van Den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International economic review*, 435-444.

Mellinas, J. P. and Nicolau, J. L. (2020). Let’s hook up fast! Hotel reviews and Wi-Fi flaws. *Annals of Tourism Research*, 80:102842.

Miller, G. (2001). The development of indicators for sustainable tourism: results of a Delphi survey of tourism researchers. *Tourism management*, *22*(4), 351-362.

Milman, A., & Pizam, A. (1995). The role of awareness and familiarity with a destination: The central Florida case. *Journal of travel research*, *33*(3), 21-27.

Min, H., Min, H., & Joo, S. J. (2008). A data envelopment analysis-based balanced scorecard for measuring the comparative efficiency of Korean luxury hotels. *International journal of quality & reliability management*, *25*(4), 349-365.

Minazzi, R. (2010). Hotel classification systems: a comparison of international case studies. *Acta Universitatis Danubius. Œconomica*, *6*(4), 64-86.

Mitra, S. K. (2019). Is tourism-led growth hypothesis still valid?. *International Journal of Tourism Research*, *21*(5), 615-624.

Morey, R. C., & Dittman, D. A. (1995). Evalatn a hotel GM's performance: A case study in benchmarking. *Cornell Hotel and Restaurant Administration Quarterly*, *36*(5), 30-35.

Mudambi, S., & Schuff, D. (2010). What Makes a Helpful Review? A Study of Customer Reviews on Amazon.com. Retrieved August 27, 2019, Retrieved from www.TripAdvisor.com/results.cfm

Mundlak, Y. (1961). Aggregation over time in distributed lag models. International Economic Review, 2(2), 154-163.

Narayan, P. K., Sharma, S. S., & Bannigidadmath, D. (2013). Does tourism predict macroeconomic performance in Pacific Island countries? *Economic Modelling*, *33*, 780-786.

Nasution, H. N., & Mavondo, F. T. (2008). Customer value in the hotel industry: What managers believe they deliver and what customer experience. *International Journal of Hospitality Management*, *27*(2), 204-213.

Naudé, W. A., & Saayman, A. (2005). Determinants of tourist arrivals in Africa: a panel data regression analysis. *Tourism Economics*, *11*(3), 365-391.

Nicolau, J. L., & Sellers, R. (2010). The quality of quality awards: diminishing information asymmetries in a hotel chain. *Journal of Business Research*, *63*(8), 832-839.

Nie, C., Zheng, Z. E., & Sarkar, S. (2022). Competing with the sharing economy: incumbents' reaction on review manipulation. MIS Quarterly, 46(3).

Núñez-Serrano, J. A., Turrion, J., & Velázquez, F. J. (2014). Are stars a good indicator of hotel quality? Assymetric information and regulatory heterogeneity in Spain. *Tourism management*, *42*, 77-87.

O' Cass, A., & Grace, D. (2004). Exploring consumer experiences with a service brand. *Journal of Product & Brand Management*, *13*(4), 257-268.

Öğüt, H., & Onur Taş, B. K. (2012). The influence of internet customer reviews on online sales and prices in hotel industry. *The Service Industries Journal*, *32*(2), 197-214.

Oliveira, R., Pedro, M. I., & Marques, R. C. (2013). Efficiency and its determinants in Portuguese hotels in the Algarve. *Tourism Management, 36*, 641-649.

Oliver, R. L. (1977). Effect of expectation and disconfirmation on postexposure product evaluations: An alternative interpretation. *Journal of applied psychology*, *62*(4), 480.

Papatheodorou, A., Lei, Z., & Apostolakis, A. (2012). Hedonic price analysis. Handbook of research methods in tourism: Quantitative and qualitative approaches, 170-182.

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A conceptual model of service quality and its implications for future research. *Journal of marketing*, *49*(4), 41-50.

Park, T. J., Jang, H. J., & Kim, H. I. (2011). Effects of management performance with outsourcing environment in hotel industry. *The Journal of the Korea Contents Association*, *11*(10), 479-487.

Parte-Esteban, L., & Alberca-Oliver, P. (2015). Determinants of technical efficiency in the Spanish hotel industry: regional and corporate performance factors. *Current Issues in Tourism*, *18*(4), 391-411.

Pérez-Granja, U., & Inchausti-Sintes, F. (2023). On the analysis of efficiency in the hotel sector: Does tourism specialization matter? Tourism Economics, 29(1), 92-115.

Peypoch, N. (2007). On measuring tourism productivity. *Asia Pacific Journal of Tourism Research*, *12*(3), 237-244.

Peypoch, N., & Solonandrasana, B. (2008). Aggregate efficiency and productivity analysis in the tourism industry. *Tourism Economics*, *14*(1), 45-56.

Pien, Y. H. (2001). The study on operational style and efficiency of international tourist hotels in Taiwan (Doctoral dissertation, Chaoyang University of Technology).

Pnevmatikoudi, K. and Stavrinoudis, T. (2016). Classification of hotel performance measurement indicators presented in international scientific research. European Journal of Tourism Research, 12:82–98.

Prideaux, B., Thompson, M., & Pabel, A. (2020). Lessons from COVID-19 can prepare global tourism for the economic transformation needed to combat climate change. *Tourism Geographies*, *22*(3), 667-678.

Pulina, M., Detotto, C., & Paba, A. (2010). An investigation into the relationship between size and efficiency of the Italian hospitality sector: A window DEA approach. *European journal of operational research*, *204*(3), 613-620.

Qu, H., Ryan, B., & Chu, R. (2000). The importance of hotel attributes in contributing to travelers' satisfaction in the Hong Kong hotel industry. *Journal of Quality Assurance in Hospitality & Tourism*, *1*(3), 65-83.

Radojevic, T., Stanisic, N., and Stanic, N. (2017). Inside the rating scores: A multilevel analysis of the factors influencing customer satisfaction in the hotel industry. *Cornell Hospitality Quarterly*, 58(2):134–164.

Ramos, C. M., & Rodrigues, P. M. (2013). The importance of ICT for tourism demand: A dynamic panel data analysis. In *Quantitative methods in tourism economics* (pp. 97-111). Physica, Heidelberg.

Reed, W. R. (2015). On the practice of lagging variables to avoid simultaneity. *Oxford Bulletin of Economics and Statistics*, *77*(6), 897-905.

Reynolds, D. (2003). Hospitality-productivity assessment: Using data-envelopment analysis. *Cornell Hotel and Restaurant Administration Quarterly*, *44*(2), 130-137.

Reynolds, D., & Thompson, G. M. (2007). Multiunit restaurant productivity assessment using three-phase data envelopment analysis. *International Journal of Hospitality Management*, *26*(1), 20-32.

Rhee, H. T., & Yang, S. B. (2015). Does hotel attribute importance differ by hotel? Focusing on hotel star-classifications and customers' overall ratings. *Computers in Human Behavior,* 50, 576-587.

Ritchie, J. B., & Crouch, G. I. (2003). The competitive destination: A sustainable tourism perspective. Cabi.

Sainaghi, R. (2010). A meta-analysis of hotel performance. Continental or worldwide style? *Tourism Review*.

Sanjeev, G. M. (2007). Measuring efficiency of the hotel and restaurant sector: the case of India. *International Journal of Contemporary Hospitality Management*, *19*(5), 378-387.

Santana-Gallego, M., Ledesma-Rodríguez, F. J., & Pérez-Rodríguez, J. V. (2016). International trade and tourism flows: An extension of the gravity model. *Economic Modelling*, *52*, 1026-1033.

Santana-Gallego, M., Ledesma-Rodríguez, F., & Pérez-Rodríguez, J. V. (2011). Tourism and trade in small island regions: The case of the Canary islands. *Tourism Economics*, *17*(1), 107-125.

Sardo, F., Serrasqueiro, Z., & Alves, H. (2018). On the relationship between intellectual capital and financial performance: A panel data analysis on SME hotels. *International Journal of Hospitality Management*, *75*, 67-74.

Sayfuddin, A. T. M., & Chen, Y. (2021). The signaling and reputational effects of customer ratings on hotel revenues: Evidence from TripAdvisor. International Journal of Hospitality Management, 99, 103065.

Schmidt, P., & Lovell, C. K. (1979). Estimating technical and allocative inefficiency relative to stochastic production and cost frontiers. *Journal of econometrics*, *9*(3), 343-366.

Schuckert, M., Liu, X., & Law, R. (2016). Insights into suspicious online ratings: direct evidence from TripAdvisor. Asia Pacific Journal of Tourism Research, 21(3), 259-272.

Seddighi, H. R., & Theocharous, A. L. (2002). A model of tourism destination choice: a theoretical and empirical analysis. *Tourism management*, *23*(5), 475-487.

Seetanah, B., Durbarry, R., & Ragodoo, J. N. (2010). Using the panel cointegration approach to analyse the determinants of tourism demand in South Africa. *Tourism Economics*, *16*(3), 715-729.

Seetaram, N. (2010). Use of dynamic panel cointegration approach to model international arrivals to Australia. *Journal of Travel Research*, *49*(4), 414-422.

Seetaram, N., & Petit, S. (2012). Panel data analysis in Tourism Research.

Sequeira, T. N., & Maçãs Nunes, P. (2008). Does tourism influence economic growth? A dynamic panel data approach. *Applied Economics*, *40*(18), 2431-2441.

Serrasqueiro, Z., & Nunes, P. M. (2014). Financing behaviour of Portuguese SMEs in hotel industry. *International Journal of Hospitality Management*, *43*, 98-107.

Shaheen, K., Zaman, K., Batool, R., Khurshid, M. A., Aamir, A., Shoukry, A. M., & Gani, S. (2019). Dynamic linkages between tourism, energy, environment, and economic growth: evidence from top 10 tourism-induced countries. *Environmental Science and Pollution Research*, *26*(30), 31273-31283.

Shang, J. K., Wang, F. C., & Hung, W. T. (2010). A stochastic DEA study of hotel efficiency. *Applied Economics, 42*(19), 2505-2518.

Sigala, M. (2004). Using data envelopment analysis for measuring and benchmarking productivity in the hotel sector. *Journal of travel & tourism marketing*, *16*(2-3), 39-60.

Sigala, M. (2020). Tourism and COVID-19: Impacts and implications for advancing and resetting industry and research. *Journal of business research*, *117*, 312-321.

Simar, L., & Wilson, P. W. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management science*, *44*(1), 49-61.

Simar, L., & Wilson, P. W. (2000). A general methodology for bootstrapping in non-parametric frontier models. *Journal of applied statistics*, *27*(6), 779-802.

Singh, R., Charan, P., & Chattopadhyay, M. (2020). Evaluating the Hotel Industry Performance Using Efficiency and Effectiveness Measures. International Journal of Hospitality & Tourism Administration, 1-24.

Song, H., & Li, G. (2008). Tourism demand modelling and forecasting—A review of recent research. Tourism management, 29(2), 203-220.

Song, H., Wong, K. K., & Chon, K. K. (2003). Modelling and forecasting the demand for Hong Kong tourism. *International Journal of Hospitality Management*, *22*(4), 435-451.

Su, C. S., & Sun, L. H. (2007). Taiwan's hotel rating system: A service quality perspective. *Cornell Hotel and Restaurant Administration Quarterly*, *48*(4), 392-401.

Surugiu, C., Leitão, N. C., & Surugiu, M. R. (2011). A panel data modelling of international tourism demand: Evidences for Romania. *Economic research-Ekonomska istraživanja*, *24*(1), 134-145.

Talias, M. (2018). Global governance conceptualization and the case of hotel classification. International Journal of Hospitality Management, 72, 132-139.

Tanford, S., Raab, C., & Kim, Y. S. (2012). Determinants of customer loyalty and purchasing behavior for full-service and limited-service hotels. *International Journal of Hospitality Management*, *31*(2), 319-328.

Thompson, R. G., Langemeier, L. N., Lee, C. T., Lee, E., & Thrall, R. M. (1990). The role of multiplier bounds in efficiency analysis with application to Kansas farming. *Journal of econometrics*, *46*(1-2), 93-108.

Thompson, R. G., Singleton Jr, F. D., Thrall, R. M., & Smith, B. A. (1986). Comparative site evaluations for locating a high-energy physics lab in Texas. *interfaces*, *16*(6), 35-49.

Torres, E. N., Fu, X., & Lehto, X. (2014). Examining key drivers of customer delight in a hotel experience: A cross-cultural perspective. *International Journal of Hospitality Management*, *36*, 255-262.

Tosun, C. (1999). Towards a typology of community participation in the tourism development process. *Anatolia*, *10*(2), 113-134.

Tsai, H. (2009). Star-Rated Hotel Productivity in China: A Provincial Analysis Using the DEA Cross-Efficiency Evaluation Approach. *Journal of China Tourism Research*, *5*(3), 243-258.

Tsaur, S. H., & Tsai, C. W. (1999). Cost structure for international tourist hotels in Taiwan. *Asia Pacific Journal of Tourism Research*, *4*(2), 41-48.

Tsionas, E. G., & Assaf, A. G. (2014). Short-run and long-run performance of international tourism: Evidence from Bayesian dynamic models. *Tourism Management*, *42*, 22-36.

Tuominen, P. (2011). The influence of TripAdvisor consumer-generated travel reviews on hotel performance.

Uhrig, S. C. (2008). *The nature and causes of attrition in the British Household Panel Study* (No. 2008-05). ISER Working Paper Series.

Uhrig, S. N. (2011). Understanding panel conditioning: an examination of social desirability bias in self-reported height and weight in panel surveys using experimental data. *Longitudinal and Life Course Studies*, *3*(1), 120-136.

United Nations World Tourism Organization (UNWTO), Blomberg-Nygard, A., & Anderson, C. K. (2016). United Nations world tourism organization study on online guest reviews and hotel classification systems: an integrated approach. *Service Science*, *8*(2), 139-151.

Van Der Wiele, T., Van Iwaarden, J., Dale, B. G., & Williams, R. (2006). A comparison of five modern improvement approaches. *International Journal of Productivity and Quality Management*, *1*(4), 363-378.

Wallace, T. D., & Hussain, A. (1969). The use of error components models in combining cross section with time series data. Econometrica: Journal of the Econometric Society, 55-72.

Wang, F. C., Hung, W. T., & Shang, J. K. (2006). Measuring the cost efficiency of international tourist hotels in Taiwan. *Tourism Economics*, *12*(1), 65-85.

Wang, Y., Lee, W., & Wong, C. (2007). Productivity and efficiency analysis of international tourist hotels in Taiwan: An application of the stochastic frontier approach. Taiwan Economic Review, 35(1), 55.

Wassenaar, D. J., & Stafford, E. R. (1991). The lodging index: an economic indicator for the hotel/motel industry. *Journal of Travel Research*, *30*(1), 18-21.

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

Wöber, K. W. (2002). *Benchmarking in tourism and hospitality industries: The selection of benchmarking partners*. CABI.

Wooldridge, J. M. (2002). Econometric analysis of cross section and panel data MIT Press. *Cambridge, MA*, *108*.

Wu, J., Cheng, X., & Liao, S. S. (2020). Tourism forecast combination using the stochastic frontier analysis technique. Tourism Economics, 26(7), 1086-1107.

Wu, J., Tsai, H., & Zhou, Z. (2011). Improving efficiency in international tourist hotels in Taipei using a non-radial DEA model. *International Journal of Contemporary Hospitality Management*, *23*(1), 66-83.

Xiang, Z., Schwartz, Z., Gerdes, J. H., & Uysal, M. (2015). What can big data and text analytics tell us about hotel guest experience and satisfaction? International Journal of Hospitality Management, 44, 120–130.

Xie, K. L., So, K. K. F., & Wang, W. (2017). Joint effects of management responses and online reviews on hotel financial performance: A data-analytics approach. International Journal of Hospitality Management, 62, 101-110.

Xie, K. L., Zhang, Z., & Zhang, Z. (2014). The business value of online consumer reviews and management response to hotel performance. International Journal of Hospitality Management, 43, 1–12.

Xie, K. L., Zhang, Z., Zhang, Z., Singh, A., & Lee, S. K. (2016). Effects of managerial response on consumer eWOM and hotel performance: Evidence from TripAdvisor. International Journal of Contemporary Hospitality Management, 28(9), 2013-2034.

Xie, K., Kwok, L., & Wang, W. (2017). Monetizing managerial responses on TripAdvisor: Performance implications across hotel classes. Cornell Hospitality Quarterly, 58(3), 240-252.

Xu, Y., Zhang, Z., Law, R., & Zhang, Z. (2020). Effects of online reviews and managerial responses from a review manipulation perspective. Current Issues in Tourism, 23(17), 2207-2222.

Yan, C. H. (1997). *The performance evaluation for the operational efficiency of international tourist hotel industry in Taiwan* (Doctoral dissertation, Chinese Culture University).

Yang, C., & Lu, W. M. (2006). Performance benchmarking for Taiwan's international tourist hotels. *INFOR: Information Systems and Operational Research*, *44*(3), 229-245.

Yang, Y. (2012). Agglomeration density and tourism development in China: An empirical research based on dynamic panel data model. *Tourism Management*, *33*(6), 1347-1359.

Yang, Y., Park, S., & Hu, X. (2018). Electronic word of mouth and hotel performance: A meta-analysis. Tourism management, 67, 248-260.

Yoo, K. H., Sigala, M., & Gretzel, U. (2016). Exploring TripAdvisor. Open tourism: open innovation, crowdsourcing and co-creation challenging the tourism industry, 239-255.

Yu, M. M., & Lee, B. C. (2009). Efficiency and effectiveness of service business: Evidence from international tourist hotels in Taiwan. *Tourism Management*, *30*(4), 571-580.

Zeithaml, V. A., Bitner, M. J., & Gremler, D. D. (2006). Services marketing: Integrating customer focus across the firm. McGraw-Hill/Irwin.

Zhou, Z., Huang, Y., & Hsu, M. K. (2008). Using data envelopment analysis to evaluate efficiency: An exploratory study of the Chinese hotel industry. *Journal of Quality Assurance in Hospitality & Tourism*, *9*(3), 240-256.

APPENDIX A: Demand and supply for Hotel Services in Greece & the importance of tourism to the Greek economy

A. Demand for Hotel Services

This appendix aims to present the basic metrics constituting the shape, magnitude, and texture of the Greek tourism and hoteling sector in terms of demand and supply. To capture all effects and changes taking place during the period of the study, from 2007 to 2016, the figures, and metrics of the touristic and hoteling sector of the immediate next year, 2017, are presented and elaborated. The presentation is divided into two parts, the first exhibits data concerning touristic demand, and the second presents information related to hoteling supply.

The global financial crisis of 2007-2008, followed by the Greek debt crisis, initially led to a decline in tourist arrivals in the period from 2007 to 2012. Economic uncertainty and negative media coverage impacted Greece's image as a tourist destination. In the years to follow the economic turbulence of 2007–2008, a trend for budget travel occurred internationally, with tourists seeking more affordable options, benefiting from low-cost accommodations and services. In this context, Greece was a considerable option as an economic crisis and its effects were squishing costs and accommodation asking prices. Post-2012, there was a significant rebound for the tourism sector, with record-breaking numbers of inbound tourists. Beyond the value-for-money prices and the eternal sea, sun, and cultural product of the country, the Greek government and tourism organizations launched campaigns to promote Greece as a safe and attractive destination, countering negative perceptions from the economic crisis. Furthermore, instability in neighbouring regions in the Middle East and North Africa potentially led some tourists to choose Greece as a perceived safer alternative. Finally, the economic growth observed in the countries of origin of incoming tourists, and the expansion and upgrading of hotel capacity propelled touristic demand for Greece in the eleven years of 2007-2017.

While European countries remained primary sources of tourists, there was an increase in visitors from new markets, including Asia and North America. Interest in Greece's rich cultural and historical heritage, including visits to archaeological sites and museums, grew during this period. Meanwhile, investments in tourism infrastructure, including upgrades to airports, ports, and transportation networks, were made to accommodate the growing number of tourists. The hospitality sector saw growth and diversification, with a rise in boutique hotels, luxury resorts, and private villa rentals. International arrivals for 2017 in Greece stood at 27.2 million, converted to 88 million overnight stays in accommodation establishments. Arrivals increased by 9.7% compared to 2016 (24.8 million). The 2017 performance had been the fifth consecutive year-on-year increase which reveals the resilience and competitiveness of the Greek Tourism industry at that time. The next table presents the number of arrivals and the number of beds available in Greece from 1960 to 2017. The last column presents the ratio of arrivals per bed.

Table A.1. International Arrivals in Greece Non-Residents - Arrivals/ Beds (1960-2017)

Year	*Arrivals*	*Beds*	*Arrivals/Beds*
1960	400,000	55,000	7.3
1970	1,455,000	119,000	12.2
1980	5,271,000	278,000	19
1990	9,310,000	423,660	22
2000	13,567,453	586,372	23.1
2010*	15,007,490	763,407	19.7
2011*	16,427,247	763,668	21.5
2012*	15,517,621	771,271	20.1
2013*	17,919,580	773,445	23.2
2014*	22,033,463	780,721	28.2
2015*	23,599,455	784,315	30.1
2016*	24,799,300	788,553	31.4
2017*	27,194,200	806,045	33.7

** Non-Residents Arrivals*

Source: ELSTAT/ BoG / HCG.

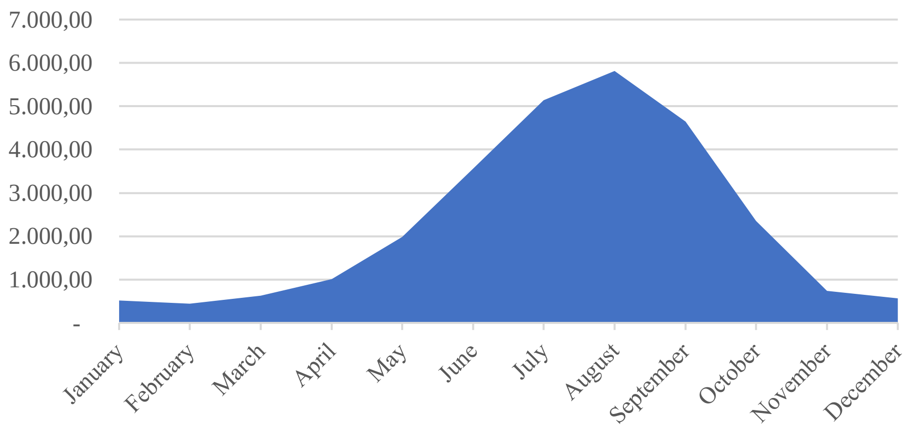
Considering the above table, one can easily observe that all three sizes, arrivals, number of beds under disposal and arrivals/beds, increased dramatically from 1960 up to 2017. It is important to mention that the arrivals per available bed ratio exhibits a constant rise, indicating that additional beds are fully "absorbed" by increasing demand. For 2017 the rate of arrivals per bed stood at 33.7 arrivals per hotel bed. Table A.2 and Graph A.1. present the monthly share of international arrivals per month in Greece for 2017.

Table A.2. Monthly International Arrivals (2017)

Month	*Arrivals*	*Share*
January	520.4	2%
February	444.4	2%
March	627.6	2%
April	1,009.70	4%
May	1,982.10	7%
June	3,356.30	12%
July	5,141.60	19%
August	5,813.40	21%
September	4,639.70	17%
October	2,355.80	9%
November	740.5	3%
December	562.7	2%
Total	27,194.20	100%

in thousands, Source: BoG

Graph A.1. Arrivals by month in m. Year 2017



Source: BoG

Examining the previous Table A.2. and Graph A.1., it is evident that a seasonality pattern clearly holds, with 85% of all arrivals occurring during the six months of May to October. The pattern is ongoing, characterising Greek tourism, from its infantry, to nowadays (the percentage was 87% for 2023). The next Table A.3. showcases the country of origin of the international arrivals in Greece for 2017.

Table A.3. International Arrivals by Country of Origin

In thousands	*2014*	*2015*	*2016*	*2017*	*% of total tourist arrivals -2017*
EE-28	13,249	14,974	17,217	18,583	68%
Germany	2,459	2,810	3,139	3,706	14%
UK	2,090	2,397	2,895	3,002	11%
Italy	1,118	1,355	1,387	1,441	5%
France	1,463	1,522	1,314	1,420	5%
Romania	543	540	1,026	1,149	4%
Cyprus	448	470	652	632	2%
Other*	7,218	8,690	6,804	7,233	27%
Other of which	8,784	8,625	7,583	8,611	32%
Russia	1,250	513	595	589	2%
USA	592	750	779	865	3%
Australia	183	183	169	324	1%
Canada	146	182	153	198	1%
China	47	55	150	175	1%
TOTAL	22,033	23,599	24,799	27,194	100%

Source: BoG

Table A.4. Figures for major markets for 2017

Market	*Overnight stay*		*Receipts*		*Spending & Length of Stay*		
	in thousands	*% of total visits*	*€ million*	*% of total visits*	*Cost per Visit (€)*	*Cost per O/N Stay (€)*	*Average Length of Stay*
Germany	37,637	17.90%	2,553	18.00%	624	68	9.2
UK	26,552	12.70%	2,065	14.50%	637	78	8.2
France	12,268	5.80%	994	7.00%	577	81	7.1
Italy	12,042	5.70%	753	5.30%	453	63	7.2
US	9,549	4.60%	814	5.70%	573	85	6.7
Russia	5,885	2.80%	418	2.90%	605	71	8.5
Total of all Markets	209,855	100.00%	14,202	100.00%	458	68	6.8

Source: BoG

Germany and the UK furnishes to Greece the most visitors constantly over the years comparingly to other countries of the world. Specifically, from Germany and the UK respectively, 3.7 and 3.0 million tourists arrived in Greece in 2017, accumulating 25% of the total share of international arrivals (the ratio will increase to 31% or 2023). Italy and France follow, presenting cumulatively another 10 % of the total arrivals for 2017 (the same percentage holds for 2023). Overall, 68% of total arrivals in Greece originated from countries of the EU in 2017 (a 60% ratio is recorded in 2023). The next table presents overnight stays, receipts spending & length of stay for the major markets of incoming tourism for 2017 for Greece.

Table A.5. Tourism Receipts (2007-2017)

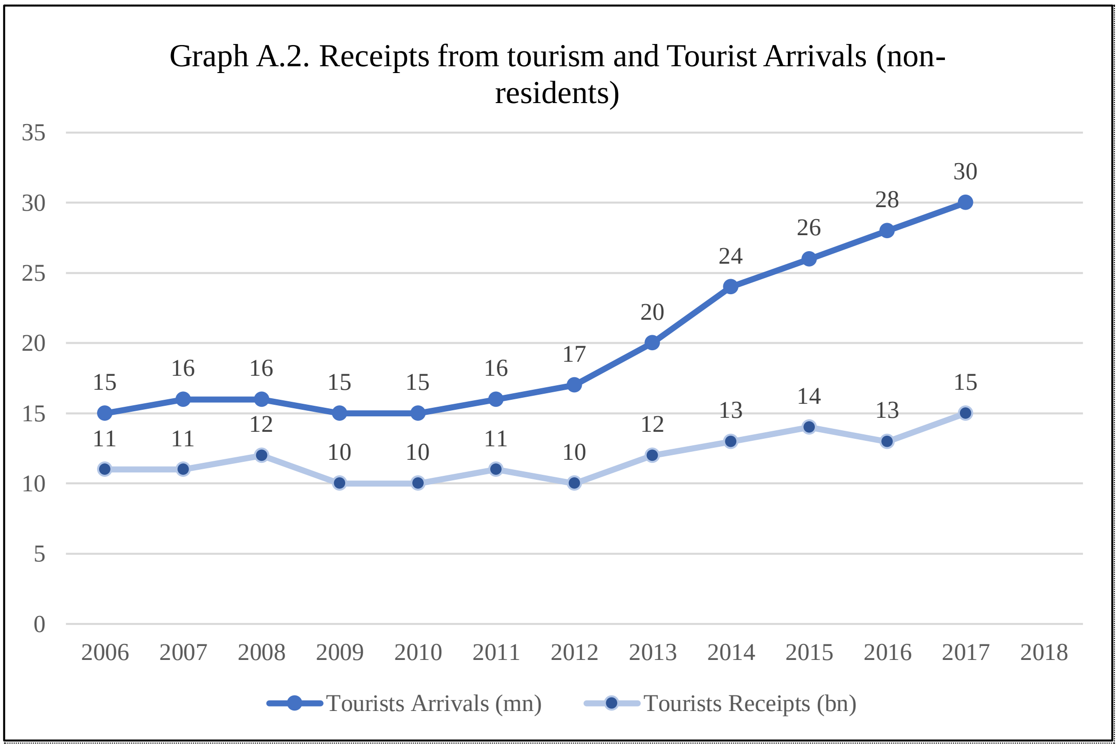
Year	*Tourism Receipts (€mil.)*	*Change*
2007	11,319.20	-
2008	11,635.90	2.80%
2009	10,400.30	-10.60%
2010	9,611.30	-7.60%
2011	10,504.70	9.30%
2012	10,442.50	-4.60%
2013	12,152.20	16.40%
2014	13,393.10	10.20%
2015	14,125.80	5.50%
2016	13,206.70	-6.50%
2017	14,595.80	10.50%

Source: Bog

Data from the previous Table A.4. justify that the major inbound touristic markets for Greece, are the larger and more economically strong countries of Europe. In terms of overnight stay, 42,1% are pipelined by Germany. UK, France, and Italy, while 44,8% of total receipts are generated by the same four European countries. Moreover, visitors from Germany. UK, France, and Italy spend above the country's average per trip (see Table A.6.) and UK and French tourists spend much higher than the country's average in terms of cost per overnight stay. Finally, all four countries exhibit a higher length of stay per person than Greece's average.The next table showcases the tourism receipts for the years 200- to 2017 and the y-o-y change.

According to data presented in Table A.5., receipts directly related to tourism amounted to €14.6 billion in 2017 which constitutes an increase of 22.4% compared to 2007. Although there were fluctuations in the number of total receipts over these 11 years under research, it is obvious that an overall rising trend exists from the year 2011 to 2017.

The next graph A.2. contrasts the receipts from tourism and arrivals for non-residents.



Source: Bog

In a nutshell, the previous graph shows that receipts tend to follow the upward trend of arrivals but not on a more relaxed slope. The latter indicates that more tourists provided fewer receipts per arrival to the Greek tourist industry over the years, implying a diminishing per capita expenditure.

The next table A.6. exhibits the average per capita spending for the 2007 – 2017 period.

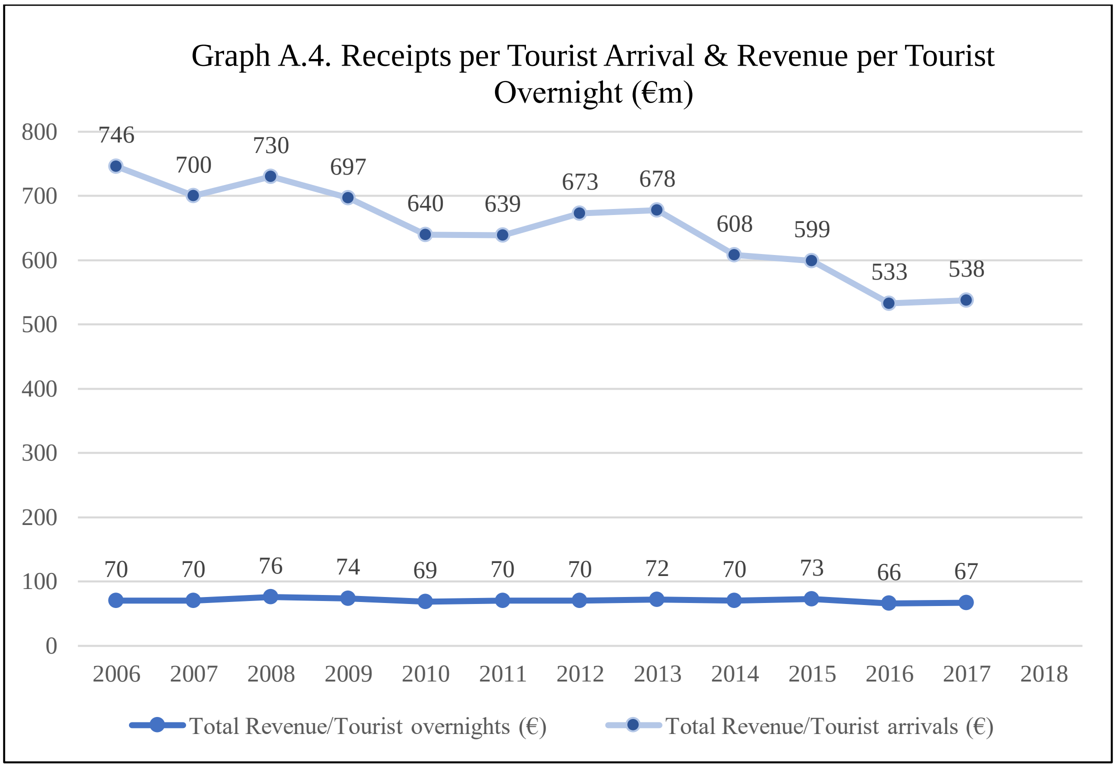
Table A.6. Average Per Capita Spending (APCS) 2007-2017

Year	Arrivals	Revenue (in € mill)	APCS (€)
2007	16,165,283	11,319	700
2008	15,938,806	11,636	730
2009	14,914,534	10,400	697
2010	15,007,490	9,611	640
2011	16,427,247	10,505	640
2012	15,517,621	10,443	673
2013	17,919,580	12,152	678
2014	22,033,463	13,393	608
2015	23,599,455	14,126	599
2016	24,799,300	13,207	533
2017	27,194,200	14,596	537

Source: ELSTAT / BoG

Data from the table indicate that the average per capita spending in Greece has shrunk over the 2007–2017 period by 23%, reflecting mostly the shrinking of the prices paid by the tour operators to the Greek hoteliers as well as the differentiated mix of tourists visiting Greece (more visitors from the Balkans and China, having less available budget). Another factor explaining the diminished expenditure per person is competition between lower star classes. 1- and 2-star hotels, as well as most 3-star hotels having no quality advantage to compete with, they use pricing to attract customers.

Graph A.4. contrasts the receipts per tourist arrival & revenue per tourist overnight, showcasing that over the 2007–2017 period, while the spending per overnight remains almost unchanged, the receipt per travel shrunk by 23%. The latter is explained, among other reasons, by the diminishing length of stay of tourists in Greece over the concrete period.



Source: ELSTAT / BoG

The next Table A.7. presents the cost per overnight Stay by Greek region for 2017.

Table A.7. Cost per overnight Stay 2017

REGION	*2017 (€)*
Eastern Macedonia & Thrace	52
Attica	71
Northern Aegean	52
Western Greece	56
Western Macedonia	52
Epirus	59
Thessaly	58
Ionian Islands	71
Central Macedonia	45
Create	81

Southern Aegean	79
Peloponnese	59
Central Greece	56
Greece Average	*68*

Source: ELSTAT / BoG

The average Cost per Stay for the total of the 13 regions of Greece for 2017 is €68.00. As shown on the above Table A.7. above, the index varies significantly among regions: The regions of Crete and Southern Aegean had the highest rates in 2017 (€81 and €79 respectively). On the other hand, the lowest rates refer to the regions of Central Macedonia (€45) and Central Greece (€56). The regions close to the Greece average of €68 for 2017 are the regions of Attica (€71) and the Ionian Islands (€71).

The following two Tables A.8 and A.9 present respectively the arrivals as well as the overnight stays and their relevant total shares, for both, international and domestic visitors per Greek region in 2017.

Table A.8. Arrivals (international & domestic by region) 2017

Region	*Arrivals (in thousands)*	*%*
Central Macedonia	7.262	23.4%
Southern Aegean	5.841	18.8%
Attica	5.137	16.6%
Create	4.806	15.5%
Ionian Islands	2.966	9.6%
Eastern Macedonia and Thrace	1.349	4.3%
Peloponnese	727	2.3%
Epirus	713	2.3%
Thessaly	694	2,20%
Western Greece	563	1.8%
Central Greece	376	1.2%
Northern Aegean	364	1.2%
Western Macedonia	222	0.7%
Total	31.021	100%

Source: ELSTAT / BoG

Table A.9. Overnight stays per region in 2017

Regions ranked	*Stays (in thousands)*	*% of total Stays*
1. Southern Aegean	46.210	22.0
2. Central Macedonia	40.782	19.4
3. Crete	40.271	19.2
4. Attica	29.437	14.0
5. Ionian Islands	24.944	11.9
6. Eastern Macedonia & Thrace	5.421	2.6
7. Peloponnese	5.214	2.5
8. Thessaly	5.027	2.4
9. Epirus	3.643	1.7
10. Northern Aegean	3.217	1.5
11. Western Greece	2.819	1.3
12. Central Greece	2.013	1.0
13. Western Macedonia	859	0.4
Total	209.855	100%

Source: INSETE

From the above tables is evident that 86,5% of the total overnight stays and 83,7% of the total arrivals are observed in 5 out of the 13 regions of Greece, namely Southern Aegean, Central Macedonia, Crete, Attica, and the Ionian islands. These happen to be the most tourist-developed areas of the country.

Concerning the average length of overnight stays (Table A.10), a slightly different condition holds regarding the regional performance. The regions of Northern Aegean, Crete, the Ionian Islands, the Southern Aegean, and Thessaly have the highest numbers of overnight stays in 2017 (8.8, 8.4, 8.4, 7.9 and 7.3 respectively). On the other hand, the lowest stay durations refer to the regions of Western Macedonia (3.9) and of Eastern Macedonia & Thrace. For 2017, the average stay duration per visit to the Greek regions has been estimated to be 6.8 overnight stays as opposed to 6.7 in 2016. The average stay length was slightly higher, amounting to 7.8 stays in 2022 according to the data of the Bank of Greece for the specific year.

Table A.10. Average Length of Overnight Stays

Region	*2017*
Eastern Macedonia & Thrace	4.0
Attica	5.7
Northern Aegean	8.8
Western Greece	5.0
Western Macedonia	3.9
Epirus	5.1
Thessaly	7.3
Ionian Islands	8.4
Central Macedonia	5.6
Crete	8.4
Southern Aegean	7.9
Peloponnese	7.2
Central Greece	5.4
Greece Average	*6.8*

Source: ELSTAT / BoG

The next Table A.11 exhibits receipts per Greek region for 2017.

Table A.11. Receipts per region 2017

Region ranking	*Receipts (in million)*	*% of total Receipts*
1. Southern Aegean	3.653	25.7
2. Crete	3.260	23.0
3. Attica	2.083	14.7
4. Central Macedonia	1.852	13.0
5. Ionian Islands	1.775	12.5
6. Peloponnese	307	2.2
7. Thessaly	290	2.0
8. Eastern Macedonia & Thrace	282	2.0
9. Epirus	216	1.5
10. Northern Aegean	167	1.2
11. Western Greece	159	1.1
12. Central Greece	113	0.8
13. Western Macedonia	45	0.3
Total	14.202	100%

Source: INSETE

Greece in 2017 has received a total of €14.202 million from inbound tourism. The five regions that have received the bulk (88.9%) of the total amount are the regions of Southern Aegean, Crete, Attica, Central Macedonia, and the Ionian Islands. The rest of the regions (Eastern Macedonia & Thrace, Peloponnese, Epirus, Thessaly, Western Greece, Central Greece, Northern Aegean, and Western Macedonia) received in 2017, only a total number of €1.579 million (11.1%). Finally, the next table A.12 and Graph A.5 presents a comparative presentation of arrivals, overnight stays, and receipts by region, 2017.

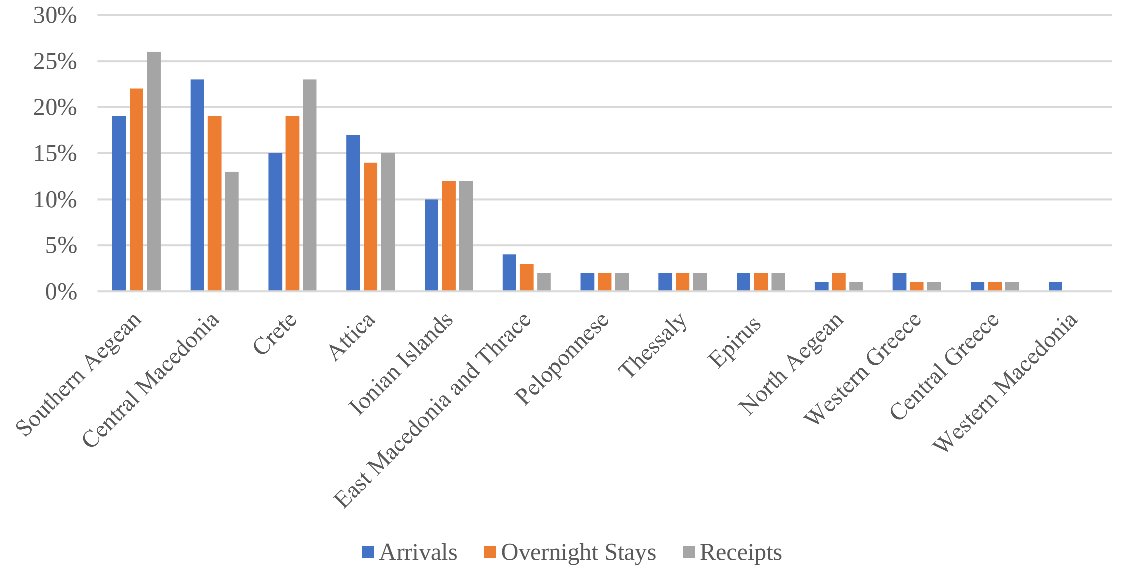
Table A.12. Distribution of arrivals, overnight stays, and receipts by region, 2017

Regions	*Arrivals*	*Overnight Stays*	*Receipts*
Southern Aegean	19%	22%	26%
Central Macedonia	23%	19%	13%
Crete	15%	19%	23%
Attica	17%	14%	15%
Ionian Islands	10%	12%	12%
East Macedonia and Thrace	4%	3%	2%
Peloponnese	2%	2%	2%
Thessaly	2%	2%	2%
Epirus	2%	2%	2%
North Aegean	1%	2%	1%
Western Greece	2%	1%	1%
Central Greece	1%	1%	1%
Western Macedonia	1%	0%	0%

Source: INSETE

Comparing the columns of the previous Table A.12. it is concluded that the two more efficient regions of Greece in terms of producing receipts are Southern Aegean and Crete. They are followed by Attica and Central Macedonia. Moreover, Southern Aegean and Crete succeed in the second and the third ranking in terms of arrivals per region, and the first and second placement in terms of overnight stays. As an overall picture, Southern Aegean and Crete are the most efficient regions in terms of producing receipts, as well as in terms of generating income per trip and per overnight stay against all other regions. The next Graph A.5. visualizes the percentages on the totals of Greece for arrivals, overnight stays, and receipts per region, exhibiting the overall dominance of Southern Aegean, Crete, Attica, the Ionian islands, and Central Macedonia. The Southern Aegean, Crete, and Ionian islands are developed touristic destinations relying mainly on the sun and sea competitive advantage, in combination with cultural monuments and unparalleled natural beauty. They attract tourists mainly from the major European markets such as the UK, Germany, France, and especially the Ionian islands Italy and they exhibit a highly seasonal operation (May to October). Attica is the capital of Greece accommodating both tourists and business visitors, with major archaeological sites and museums located on its territory. It attracts visitors from all over the world. Finally, Centra Macedonia hosts the capital of Northern Greece, Salonica, and Chalkidiki, a 3-peninsula territory with breathtaking beaches. It attracts tourists from neighboring countries such as Bulgaria, Romania, and Serbia as well as from Russia and Ukraine. As Salonica is a city and business destination, the seasonality effect is smoother than the islanding part of Greece.

Graph A.5. % of Total -Arrivals, Stays and Receipts by region 2017



Source: INSETE

Considering domestic tourism, according to data from ELSTAT and BoG, Greeks made 5,3 million vacation trips to their own homeland in 2017. Domestic tourism expenditure for trips of more than 1 overnight stay, for the same year, was € 1,4 million, while the spending per trip per person was 264 euro. It is worthwhile to mention that according to the same sources, the domestic tourist expenditure for 2008 was € 3,9 million, thus a decrease of 63.9% took place over the 2008 – 2017 decade due to the ranging economic crisis.

Table A.13. Domestic tourism flow

	2014	*2015*	*2016*	*2017*
Trips	5,340,163	4,841,525	4,590,484	5,296,499
Expenditure (€)	1,352,466,146	1,264,125,934	1,286,735,621	1,398,365,311
Spend per person	253	261	280	264

Source: BoG/ELSTAT

Summing up, in 2017, Greece welcomed 27.2 million international visitors, translating to 88 million overnight stays in various accommodations. This marked a 9.7% increase from 2016's 24.8 million visitors, continuing a five-year trend of growth. Tourism-related revenue hit €14.6 billion in 2017, up 22.4% from 2007. Despite fluctuations over these 11 years, a general upward trend is evident from 2011 to 2017. However, the average spending per visitor in Greece decreased between 2007 and 2017 from 700 euros to 537 respectively. This is largely due to lower prices offered by tour operators to Greek hotels and the changing mix of tourists, with more visitors originating from the Balkans and China who tend to have smaller budgets. Furthermore, competition among lower-star hotels, which rely on pricing rather than quality to attract guests, also contributes to this trend.

Concluding, arrivals and the arrivals per available bed ratio have increased constantly from the 60s up to 2017. Notably, the ratio of arrivals per available bed has consistently risen, indicating that the increasing demand is effectively met by the additional beds. In 2017, this ratio reached 33.7 arrivals per bed. A seasonality pattern holds, with 85% of all arrivals occurring during the six months of May to October. Germany and the UK furnishes the most visitors constantly over the years. Specifically, from Germany and the UK respectively, 3.7 and 3.0 million tourists arrived in Greece in 2017, accumulating 25% of the total share of international arrivals. Italy and France follow, presenting cumulatively another 10 % of the total arrivals for 2017.

The average per capita spending in Greece has shrunk over the 2007–2017 period, reflecting mostly the diminishing prices paid by the tour operators to the Greek hoteliers as well as the differentiated mix of tourists visiting Greece (more visitors from the Balkans and China, having less available budget). Another factor explaining the diminished expenditure per person is competition between lower star classes. 1 and 2-star hotels, as well as many 3-star hotels having no quality advantage to compete with, they use pricing to attract customers.

Considering 2017, 83.7% of the total arrivals, 86.5% of the total overnight stays, and 89% of the total receipts were achieved by 5 out of the 13 regions of Greece, namely Southern Aegean, Central Macedonia, Crete, Attica, and the Ionian islands. These happen to be the most tourist-developed areas of the country. The average Cost per Stay for the total of the 13 regions of Greece for 2017 is €68.00, with the regions of Crete and Southern Aegean having the highest rates (€81 and €79, respectively).

B. <u>Supply of Hotel Services</u>

According to the following Table B.1., during 2017 in total, 9.783 hotel units were in operation in Greece, with a total capacity of 414,127 rooms and 806,045 beds. 5-star Hotels accounted for 19% of the share of bed capacity for the concrete year, while 4-star hotels held 26.2% and 3-star hotels held 18.6% of the total bed capacity. Most of the hotel units operating in Greece in 2017, had 2-star classification (3,900 units or 39.9%). Comparing 2007 to 2017 the number of hotels grew by 6.26% in total while the number of beds grew by 15%. Consequently, larger hotels were constructed on average during the concrete 11-year period.

Table B.1. Hotel capacity of Greece 2007 - 2017, Units, Rooms, Beds

	2007	*2008*	*2009*	*2010*	*2011*	*2012*	*2013*	*2014*	*2015*	*2016*	*2017*	*% 2007-2017*
Hotels	9207	9385	9559	9732	9621	9661	9677	9745	9757	9730	9783	6,26%
Rooms	367992	375067	383008	397660	396475	399037	401322	404779	406250	407146	414127	12,54%
Beds	700933	715857	732279	763407	761964	767756	773445	780721	784315	788553	806045	15,00%

Source: Hellenic Chamber of Hotels

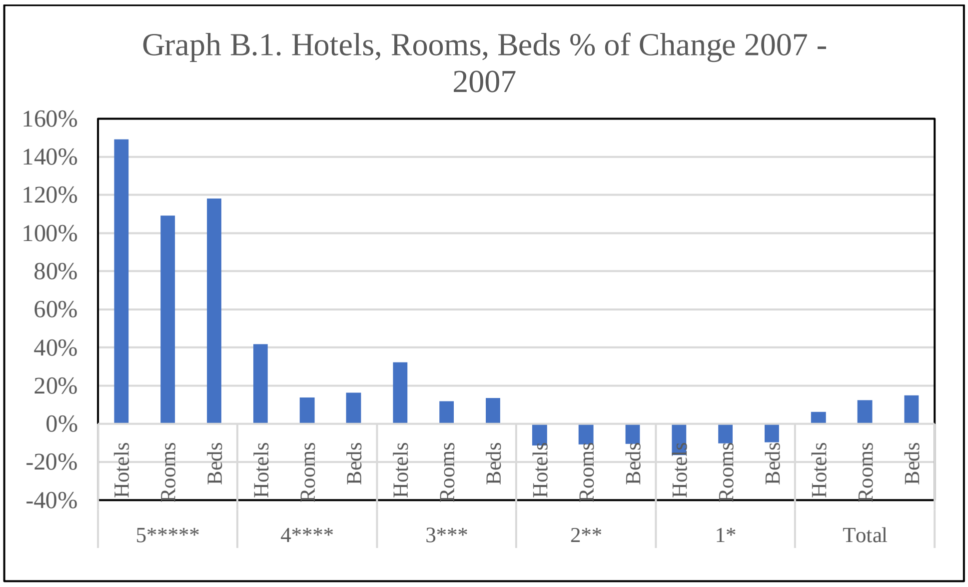
The next Table B.2 presents the hotel units, rooms, and beds stratified by star category and by year, for 2007 to 2017 as well as the percentage change in the metrics.

Table B.2 Hotel Units, Rooms, and Beds stratified by star category and by year

	5-star			4-star			3-star			2-star			1-star			Total		
Year	*Hotels*	*Rooms*	*Beds*	*Hotels*	*Rooms*	*Beds*	*Hotels*	*Rooms*	*Beds*	*Hotels*	*Rooms*	*Beds*	*Hotels*	*Rooms*	*Beds*	*Hotels*	*Rooms*	*Beds*
2007	199	35.782	70.198	1.048	94.737	181.476	1.900	85.920	163.729	4.403	121.589	228.404	1.657	29.964	57.126	9.207	367.992	700.933
2008	230	39.614	78.464	1.102	95.790	183.900	2.058	89.142	169.941	4.387	120.935	227.146	1.608	29.586	56.406	9.385	375.067	715.857
2009	280	46.186	91.770	1.164	97.432	187.494	2.179	89.749	171.202	4.368	120.733	226.707	1.568	28.908	55.106	9.559	383.008	732.279
2010	312	51.100	102.429	1.234	101.837	196.862	2.268	92.847	177.923	4.349	122.645	230.358	1.569	29.231	55.835	9.732	397.660	763.407
2011	332	53.580	107.921	1.232	99.981	193.169	2.282	93.865	180.057	4.262	120.386	226.102	1.513	28.663	54.715	9.621	396.475	761.964
2012	343	55.209	111.675	1.245	100.241	193.687	2.295	94.166	180.783	4.268	120.597	226.573	1.510	28.824	55.038	9.661	399.037	767.756
2013	361	57.878	117.555	1.277	100.289	194.010	2.358	95.674	183.722	4.203	119.157	223.932	1.478	28.334	54.226	9.677	401.332	773.445
2014	375	59.581	120.902	1.298	101.334	196.402	2.402	96.742	185.922	4.198	118.929	223.562	1.472	28.193	53.933	9.745	404.779	780.721
2015	419	63.297	128.672	1.340	102.690	199.088	2.436	96.308	185.081	4.110	116.015	218.143	1.452	27.890	53.331	9.757	406.200	784.315
2016	444	67.407	137.210	1.412	104.562	203.203	2.472	96.033	185.560	3.990	111.842	210.365	1.412	27.302	52.215	9.730	407.146	788.553
2017	496	74.884	153.132	1.485	107.805	211.064	2.515	96.129	186.056	3.900	108.383	204.193	1.387	26.926	51.600	9.783	414.127	806.045
% Change 2007 - 2017	149%	109%	118%	42%	14%	16%	32%	12%	14%	-11%	-11%	-11%	-16%	-10%	-10%	6%	13%	15%

Source: ELSTAT

The total number of 5-star hotel units increased by +149% from 2007 to 2017, while the corresponding number of 1-star and 2-star hotels decreased by 11% and 16% respectively. Three-star hotel units show a moderate increase of 16% during the same time. Correspondingly, the total number of rooms and beds for 5-star hotel units increased by +109% and 118% respectively from 2007 to 2017, while the corresponding number of 1-star and 2-star metrics for rooms and beds decreased by 10% and 11% respectively. Three-star hotel rooms increased by 12% and beds by 14%. The following graph visualizes the percentage changes in the number of hotels, beds, and rooms for the period 2007 to 2017.



Source: Hotels Chamber of Greece

Table B.3 presents the geographic distribution among the regions of Greece in terms of hotel units, rooms, and beds for 2017.

Table B.3. Geographical Distribution of Hotels in Greece (2016)

Region		*5**	*4**	*3**	*2**	*1**	*Total*
Eastern Macedonia and Thrace	Units	11	29	96	168	75	379
	Beds	2,073	3,850	5,958	7,245	2,198	21,324
Attica	Units	34	103	136	267	109	649
	Beds	12,565	15,203	12,257	15,062	3,935	59,022
Northern Aegean	Units	6	32	125	177	42	382
	Beds	1,595	2,862	8,115	8,132	1,240	21,944
Western Greece	Units	4	40	95	105	25	269
	Beds	3,106	4,175	5,867	5,111	562	18,821
Western Macedonia	Units	3	18	65	36	4	126
	Beds	137	879	3,659	1,274	141	6,090
Epirus	Units	11	89	149	133	29	411
	Beds	1,756	3,977	5,474	4,675	949	16,831
Thessaly	Units	26	117	132	216	77	568
	Beds	2,545	6,728	7,454	9,523	2,851	29,101
Ionian Islands	Units	28	117	221	490	74	930
	Beds	9,349	21,625	27,514	30,939	2,923	92,350
Central Macedonia	Units	48	100	269	361	417	1,195
	Beds	17,459	15,711	21,134	18,765	16,719	89,788
Crete	Units	97	249	361	656	205	1,568
	Beds	38,274	49,429	32,014	41,548	10,251	171,516
South Aegean	Units	144	356	484	870	214	2,068
	Beds	40,742	65,064	37,709	45,234	6,498	195,247
Peloponnese	Units	22	120	210	233	75	660
	Beds	6,137	7,958	11,223	10,538	1,758	37,614
Central Greece	Units	10	42	129	278	66	525
	Beds	1,472	5,742	7,182	12,319	2,190	28,905
Total	Units	444	1,412	2,472	3,990	1,412	9,730
	Beds	137,210	203,203	185,560	210,365	52,215	788,553

Source: Hotels Chamber of Greece

Considering Table B.3 presenting the geographical distribution by hotel classification, most of the bed capacity, in 2017, was in Southern Aegean (205,073 or 25.4%) followed by Crete (174,275 or 21.6%). They were followed by the Ionian islands with 93,440 beds or 11.6% of the total, Central Macedonia with 90,727 beds, or 11.25% of total capacity, and Attica region with 59,878 beds with a share of 7.42% out of the total. As was the case for arrivals and receipts, the predominant share of bed capacity is accumulated in Southern Aegean, Crete, Ionian islands, Central Macedonia, and Attica, indicating that these are the most touristic developed regions of Greece. These five regions host the most 5-star hotels with Southern Aegean (144) and Crete (97) outperforming the other three of the five predominant regions with 110 5-star units in total. The situation holds the same concerning 4-star hotels, with Southern Aegean and Crete accommodating 43% of the total hotels in the category.

Table B.4. Hotel occupancy by month for 2017

Month	*Occupancy 2017*
January	
February	22.10%
March	21.80%
April	25.50%
May	47.10%
June	66.80%
July	75.20%
August	76.80%
September	66.30%
October	40.70%
November	19.90%
December	22.90%
Total	52.80%

Source: ELSTAT

Considering the occupancy ratios presented in Table B.4, it is observed, that the average year-round occupancy for Greece is 52.80% for 2017 in comparison to 60.2% for Turkey according to data from STR Global, to 69% for Italy according to Istat (Italian statistical authority) and 68.2% according to data sourced by the “Institut National des Statistics et des etudes economiques". The occupancy rate in Greece follows the seasonal effect of arrivals, with July and August exhibiting occupancy ratios of 75.20%

and 76.80%, June and September succeeding occupancy of 66.8% and 66.30% respectively, while May and October showcase occupancy ratios of 47.10% and 40.70% for 2017. For the rest six months of the year, the average occupancy ratio is around 21%. The next Table B.5 exhibits the annual average occupancy rate in hotels and resorts for the period of 2012 to 2017 per region.

Table B.5. Annual Average Occupancy Rate in Hotels and Resorts (2012-2017) per region

Regions	*2012*	*2013*	*2014*	*2015*	*2016*	*2017*
Southern Aegean	54.20%	53.80%	55.30%	59.00%	57.60%	62.00%
Attica	33.60%	37.40%	45.10%	47.10%	47.20%	50.50%
Crete	57.30%	60.70%	62.00%	61.90%	65.20%	66.80%
Central Macedonia	42.20%	43.40%	44.40%	45.60%	46.30%	49.10%
Ionian Islands	53.90%	57.60%	54.20%	60.00%	62.10%	63.70%
Peloponnese	23.20%	25.30%	27.30%	27.40%	30.50%	32.20%
Central Greece	20.70%	21.10%	21.90%	23.20%	25.60%	27.90%
Thessaly	25.30%	26.60%	29.50%	29.60%	30.20%	31.80%
Western Greece	27.70%	29.10%	31.70%	33.50%	35.10%	36.40%
Eastern Macedonia and Thrace	29.80%	27.50%	34.40%	36.30%	35.50%	35.50%
Northern Aegean	29.90%	32.90%	35.90%	38.90%	33.40%	36.80%
Epirus	25.20%	26.40%	27.40%	26.40%	27.90%	30.60%
Western Macedonia	17.20%	18.70%	17.80%	17.00%	15.90%	17.20%
Total	43.20%	45.20%	47.30%	49.10%	50.10%	52.80%

Source: ELSTAT

Once again, in terms of occupancy by region the five "region champions" exhibit the higher ratios succeeding the following percentages for 2017, Southern Aegean 62%, Crete 66.8%, Ionian Islands 63.7%, Central Macedonia 49.10%, and Attica 50.5%. The six of the rest Greek regions show occupancy of less than 36.4%, with Western Macedonia showcasing the worst performance with an occupancy ratio of 17.23%.

Greek Hotel Groups and International Chain Hotels in Greece

Table B.6 includes the largest Greek Group Hotels operating more than 1,000 beds in 2017. The largest Greek Hotel Groups (19 in total) control more than 138 Hotel Units and approximately 70,000 beds. Daskalantonakis Group controls the Grecotel brand, incorporating city, and resort hotel units with a total capacity of 11,842 beds. Mitsis Group was also developed into a key player in the industry with 17 units and 5,440 rooms throughout the country.

Table B.6. Greek Hotel Groups

Group / Chain	Number of Hotels	Capacity in beds (room number is provided where bed number is not available)
Grecotel (Daskalantonakis Group) (www.grecotel.gr)	27	11,381
Mitsis Group (www.mitsishotels.com)	17	5,440 rooms
Aldemar (www.aldemarhotels.com)	8	5,500
Vasilakis - Esperia Hotels (www.esperia-hotels.gr)	7	4,000
Mantonakis – Helios Hotels (www.helioshotels.gr)	7	3,300
Divani (www.divanis.com)	7	2,700
Gregoriadis – G Hotels (www.g-hotels.gr)	6	3,400
Hadjilazarou – H Hotels (www.hhotels.gr)	6	2,000 rooms

Sbokos - Sbokos Hotels (www.sbokoshotels.gr)	6	3,000
Koulouvatos Enterprises – Amalia Hotels (www.amalia.gr)	6	1,700
Aegean Star Hotels (www.aegeanstar.com)	5	3,000
Kypriotis (www.kipriotis.gr)	5	4,000
Fokas – Electra Hotels (www.electrahotels.gr)	5	1,800
Mamidakis (www.bluegr.com)	5	1,300
Aquila Hotels & Resorts (www.aquilahotels.com)	4	2,100
Ikos Resorts (www.ikosresorts.com)	4	1,500
Metaxas – Maris Hotels (www.maris.gr)	3	2,200
Chandris (www.chandris.gr)	3	1,800
Capsi Eleni (www.capsishotel.gr)	2	1,000

Source: ICAP Group SA, web

In terms of foreign Hotel Groups, it is estimated that 150 hotel units belong to or are operated by foreign Hotel Groups. TUI has developed a solid presence controlling 41 hotel units throughout Greece in 2017. At the same time, it had a direct collaboration with over 2,000 independent Greek hotels. In addition, in 2017, Starwood operated under 6 different brands and 23 hotel units in various regions in Greece. Table B.7 presents the main foreign hotel groups and chains in Greece for 2017. Notably, foreign investors and international operators since 2014 have started to exhibit an accelerated interest in the Greek market. The trend became strong in the years to follow and continues until 2023.

Table B.7. Foreign Hotel Groups and Chains in Greece

Group / Chain	Brands	Number of Hotels/Units
TUI Group (www.tuigroup.com)	TUI	41
Atlantica Hotels (www.atlanticahotels.com)	Atlantica	29
Starwood (www.starwoodhotels.com)	Sheraton The Westin Luxury Collection Design Hotels Marriott Autograph Collection	23
Louis Hotels (www.louishotels.com)	Louis	10
Best Western (www.bestwestern.com)	Best Western	8
Labranda Hotels & Resorts (www.labranda.com)	Labranda	7
Wyndham Hotel Group (www.wyndhamhotels.com)	Wyndham Ramada	6
Intercontinental Hotel Group (IHG) (www.ihg.com)	InterContinental Holiday Inn Crowne Plaza	4
Accor (www.accorhotels.com)	Novotel Sofitel Ibis	3
AKS Hotels	AKS	3

(www.akshotels.com)		
Sentido Hotels	Sentido Hotels	3
Thomas Cook (www.thomascookgroup.com)	Casa Cook	2
Rezidor Group (www.rezidor.com)	Radisson Blue	2
Iberostar (www.iberostar.com)	Iberostar	2
Sunwing Hotels	Sunwing	2
Robinson	Robinson Club	2
Hilton Hotels International (www.hilton.com)	Hilton	1
Hyatt (www.hyatt.com)	Hyatt Regency Grand Hyatt	2

Source: ICAP Group SA, web

B.1. Greece: Hotel Industry Investment Activity

The rising demand for tourism led to significant enhancements and expansions in Greece's hotel infrastructure over the past five years. This trend has not only improved the capacity and quality of services in existing hotels but also attracted potential investors. Consequently, this interest has spurred the growth of Greek hotels and drawn well-known international hotel chains and management companies to the Greek hospitality sector. According to the "Developments in Key Figures for the Greek Hotel Industry 2018 Survey" conducted by the Hellenic Chamber of Hotels, investment in the hotel sector for the construction of new and for renovation/repair of existing hotel rooms in 2017 was estimated at € 1,541 million. A breakdown analysis of new constructions and renovation of hotels, stratified by star category follows in Table B.8 and Table B.9.

Table B.8. New hotels

		Total	5*	4*	3*	2*&1*
New rooms	2017	9,317	1,474	3,173	2,157	2,513
Construction costs (m€)	2017	779	179	326	171	102
DOMESTIC Value Added (m€)	2017	493	105	203	113	71

Source: MHTE, HCH-INSETE

Table B.9. Hotel Renovations

		Total	5*	4*	3*	2*& 1*
RENOVATION COSTS (m€)	2017	762	243	243	143	134 134
DOMESTIC Value Added (m€)	2017	388	108	121	79	81

Source: MHTE, HCH-INSETE

B.2. The importance of tourism to the Greek economy

According to The Centre of Planning and Economic Research (KEPE) (2014) Study for Tourism, the resulting tourism multiplier for the Greek tourism economy amounts to 2.65, which means that for every €1.0 from tourism activity, an additional €1.65 is generated from indirect and induced economic activity, and therefore the GDP increases by €2.65 in total. According to the IOBE (2012) relevant study, every €1.0 created by tourism activity generates additional indirect and induced economic activity of €1.2 and therefore generates a total of €2.2 in GDP. That is, the multiplier of tourism activity amounts to 2.2. Therefore, the multiplier of the tourism activity ranges between 2.2 (according to IOBE) and 2.65 (based on KEPE multipliers). Furthermore, according to a sub-sector breakdown presented at the KEPE 2014 study, the multiplier for accommodation stands at 2,5. The following Table B.10 presents the distribution of tourism revenues among the different touristic sub-sectors as well as the applied multiplier as calculated by KEPE.

Table B.10. distribution of tourism revenues and applied multiplier

Category	*Distribution of Tourism Revenues %*	*Multiplier*
Accommodation	45.30%	2.50
F & B	18.00%	2.50
Maritime transport	9.00%	2.41
Road transport	7.10%	3.25
Air transport	5.40%	2.98
Retail	4.90%	3.69
Entertainment	3.80%	1.90
Travel agencies	3.70%	3.68
Car rental	1.80%	1.39
Conferences	1.00%	4.13
Weighted average		2.65

Source: KEPE, IOBE, INSETE

According to *ELSTAT*, the Greek GDP amounted to €180,218 million in 2017. According to SETE Intelligence, the direct contribution of the tourism industry to the

Greek GDP was estimated to be 10.3% that year, while its indirect contribution was estimated to be 22.6% if its multiplier effect is considered (multiplier circa 2.2). Consequently, tourism is a key pillar of the Greek economy.

Table B.11. The contribution of tourism to GDP

Expenditure Category	*2017 (m€)*
Inbound tourism expenditure	14,203
Cruise tourism expenditure	428
Cruise company expenditure	158
Air transport	1,468
Maritime Transport	90
Domestic tourism	1,398
Investments	1,294
Direct tourism impact	19,039€
As % GDP	10.60%
IOBE multiplier	2.2
Direct and indirect	41,885€
As % GDP	23.20%
KEPE multiplier	2.65
Direct and indirect result	50,452€
As % GDP	28.00%
GDP 2017	180,218

Source: KEPE, IOBE, INSETE, ELSTAT

It is worthwhile to mention that the economy of three island regions of Greece is entirely dependent on tourism since tourism's contribution to regional GDP is 47.2% in Crete, 71.2% in the Ionian Islands, and 97.1% in the South Aegean. Furthermore, these regions have among the highest per capita GDPs in the country.

Table B.12. The contribution of tourism to the balance of payments

	2017 (m €)
Balance of Trade	-19,834
Travel Receipts (incl. Cruise)	14,630
As % deficit of Balance of Trade	*74%*
Estimate of Revenues from Transport	1,716
Travel receipts & Transports/ Balance of Trade	82%
Exports of Goods	28,040
Exports of Goods except Ships & Oil	20,051
Travel Receipts/Exports of Goods	52%
Travel Receipts & Transports/Export of Goods	58%
Travel Receipts/Exports of goods except Ships and oil	73%
Travel Receipts & Transport/Export of Goods except Ships & Oil	82%

Source: BoG-INSETE

Table B.15 exhibits that for 2017, the travel receipts covered 74.0% of the balance of trade in Greece. The percentage is mounting to 82% if revenues from air transport, cruises, and others are also included. Furthermore, travel receipts in 2017 represent 52.0% of Greece's export income, a percentage increased to 58% if receipts from the air and marine transport from inbound tourism are included in the computation. Moreover, excluding exports of fuel and ships, receipts from tourism represent 73.0% in 2017 of receipts from Greece's exports of all other products.

Finally, tourism is also an important contributor to employment. In the period 2007 - 2017 it has contributed significantly to reducing unemployment. More specifically, according to the "Employment in Tourism and the Other Sectors of the Greek Economy, 2008-2017, June 2018" prepared by INSETE, from 2008 to 2017, there was an upward trend in tourism employment, in contrast to the overall job decline in other industries of Greece during this timeframe. The recovery in tourism was particularly notable from the first quarter of 2014, outpacing other sectors by exhibiting substantial increases (17.9% in Q3 of 2014 over 2013, 8.4% in 2015 over 2014, 4.4% in 2016 over 2015, and 4.6% in 2017 over 2016). By the third quarter of 2017, tourism employment reached a peak of 398.7 thousand individuals. Moreover, between 2008 and 2017, the tourism sector's employment experienced a Compounded Annual Growth Rate (CAGR) of +0.9%, in contrast to a -2.5% decrease in other sectors. Each sector saw a decline from 2008 to 2012 (tourism by -4.1%, other sectors by -5.5%) and an uptick from 2013 to 2017 (tourism by +7.8%, other sectors by +1.1%).

C. <u>Summary</u>

Between 2007 and 2012, the global financial crisis and the Greek debt crisis led to a drop in tourist arrivals in Greece, negatively affected by economic uncertainty and adverse media portrayals. However, this period also saw a rise in budget travel globally, with tourists seeking more affordable options. Greece, with its reduced costs due to the economic crisis, emerged as a favourable destination. Post-2012, Greece's tourism industry experienced a remarkable recovery, achieving record inbound tourist numbers. This resurgence was supported by Greece's value-for-money appeal, its enduring attractions of sea, sun, and culture, and promotional campaigns by the Greek government and tourism bodies that rebranded Greece as a safe and appealing destination.

Between 2007 and 2017, Germany and the UK consistently remained the top sources of tourists to Greece, contributing significantly to international arrivals. In 2017, tourists from Germany and the UK made up 25% of all international visitors. Italy and France followed, together accounting for another 10% of total arrivals for 2017. Overall, in 2017, 68% of Greece's tourists came from EU countries. Most overnight stays and tourism revenue in Greece were generated by visitors from Germany, the UK, France, and Italy. Tourists originating from these countries not only spent more per trip than the country's average but also had longer stays in Greece.

While tourist arrivals in Greece have increased, the total revenue from these tourists has not grown at the same rate, suggesting a decrease in per capita spending. In 2017, the average cost per stay across Greece's 13 regions was €68.00, with significant regional variations: Crete and the Southern Aegean had the highest rates (€81 and €79, respectively), while Central Macedonia and Central Greece were at the lower end (€45 and €56, respectively). Attica and the Ionian Islands were close to the national average.

Tourism in Greece is highly seasonal, with 85% of arrivals concentrated between May and October, a pattern that has persisted over the years. Five regions – the Southern Aegean, Central Macedonia, Crete, Attica, and the Ionian Islands – accounted for

86.5% of overnight stays and 83.7% of total arrivals. These regions are the most developed tourist areas in Greece.

Regarding the length of stay, the Northern Aegean, Crete, the Ionian Islands, the Southern Aegean, and Thessaly had the highest numbers in 2017. In contrast, Western Macedonia and Eastern Macedonia & Thrace had the shortest stays. The average stay duration was 6.8 nights in 2017, slightly up from 6.7 in 2016, and increased to 7.8 nights in 2022.

In 2017, Greece earned €14.202 million from inbound tourism, with 88.9% of this revenue generated by the Southern Aegean, Crete, Attica, Central Macedonia, and the Ionian Islands. The remaining regions collectively earned €1.579 million (11.1%). The Southern Aegean and Crete were the most efficient regions in terms of revenue generation, followed by Attica and Central Macedonia.

Domestic tourism also played a role, with Greeks making 5.3 million trips within the country in 2017, spending €1.4 billion. This marked a significant decrease from the €3.9 billion spent in 2008, reflecting the impact of the economic crisis on domestic tourism expenditure over the decade.

In 2017, Greece had 9,783 operational hotel units with a total of 414,127 rooms and 806,045 beds. Five-star hotels represented 19% of the bed capacity, four-star hotels 26.2%, and three-star hotels 18.6%. Most hotels were two-star establishments, comprising 39.9% of the total. From 2007 to 2017, the number of hotels in Greece grew by 6.26%, and bed capacity increased by 15%, indicating a trend towards larger hotels.

Significantly, the number of five-star hotel units rose by 149% between 2007 and 2017, while one-star and two-star hotels decreased by 11% and 16% respectively. The bed and room capacity in five-star hotels also surged dramatically during this period. Most of the bed capacity in 2017 was concentrated in the Southern Aegean, Crete, Ionian Islands, Central Macedonia, and Attica, with these regions hosting the majority of five-star and four-star hotels.

The average occupancy rate for Greek hotels in 2017 was 52.80%, lower than in neighbouring countries like Turkey, Italy, and France. Occupancy rates were highest in July and August and significantly lower in the off-season months. The regions with the highest occupancy rates were the Southern Aegean, Crete, and the Ionian Islands.

Regarding hotel groups, the largest Greek hotel groups, totalling 19, control over 138 units and around 70,000 beds. Foreign hotel groups also had a substantial presence, with TUI controlling 41 units and collaborating with over 2,000 independent Greek hotels, and Starwood operating 23 units under six different brands in Greece in 2017.

In 2017, travel receipts reached 14.6 billion, accounting for 74.0% of Greece's trade balance. Travel receipts also represented 52.0% of Greece's export income. KEPE's study highlights a 2.5 multiplier specifically for accommodation services. Tourism's impact is especially pronounced in three Greek island regions, where it constitutes a major part of the regional GDP: 47.2% in Crete, 71.2% in the Ionian Islands, and 97.1% in the South Aegean. These regions also feature some of the highest per capita GDPs in Greece.

Tourism also plays a vital role in employment. Between 2008 and 2017, the tourism sector's employment grew annually by 0.9%, while other sectors decreased by 2.5%.

While both sectors experienced declines from 2008 to 2012, tourism rebounded significantly from 2013 to 2017, contrasting with the modest recovery in other industries. According to INSETE's data by the third quarter of 2017, tourism employment reached 398.7 thousand people.

Appendix references

Internet sources

Bank of Greece, (BoG) www.bankofgreece.gr

Hellenic Statistical Authority (ELSTAT) www.elstat.gr.

Hellenic Chamber of Hotels (HCH) https://www.grhotels.gr

Publications

Arbitrage Real Estate Advisors: Greek Hospitality Market edition December 2018.

Aris Ikkos & Serafeim Koutsos, Greece-and-Competition-2013-2017, 2018, 2018, INSETE.

Aris Ikkos & Serafeim Koutsos. Balance of travel services by region 2017, July 2018, INSETE.

Aris Ikkos & Serafeim Koutsos, The contribution of Tourism to the Greek economy in 2017, May 2018 (INSETE).

Aris Ikkos & Serafeim Koutsos (2018), The Contribution of Tourism to the Greek Economy in 2017 (INSETE).

Aris Ikkos & Serafeim Koutsos (2019), The contribution of Tourism to the Greek economy in 2018 (INSETE).

Development of figures and indicators of inbound tourism in Greece from its main source markets, 2016–2018, Synopsis, October 2019, INSETE.

Foundation for Economic and Industrial Research (IOBE) (2012), The impact of tourism on the Greek economy.

Long-term employment trends (1983-2016) in Tourism and the other sectors of the Greek economy, May 2017, INSETE.

Nikos Karavitis. The contribution of tourism to the economy and tax revenues, the Case of Greece, 2010-2016, September 2018, INSETE (provided only in the Greek language).

The Hellenic Chamber of Hotels Survey," Developments in Key Figures for the Greek Hotel Industry 2018").