Dynamic pricing strategies:

Evidence from European hotels

Abstract

How much do hoteliers actually make use of dynamic pricing strategies? We collected data on the price of a single room booked in advance (from three months to a single day), from almost 1000 hotels in eight European capital cities. Pricing strategies were analyzed by means of descriptive statistics, box plots and econometric panel data techniques. The empirical results show that the inter-temporal pricing structure primarily depends on the type of customer, the star rating and the number of suppliers with available rooms.

Keywords: hotel industry, dynamic pricing, customer heterogeneity

1. Introduction

The nature of hotel rooms as a perishable asset is prompting hoteliers to maximize their revenue by trying to achieve optimal dynamic prices using different strategies. Conversely, customers can strategically change their purchase plans in order to pay as little as possible. In this context, heterogeneity among hotels and customers plays a key role. The best form of inter-temporal pricing strategies depends on the composition of the customer population alongside factors such as customer valuations and patience, as highlighted by Su (2007). In particular, if *high-valuation customers* have a proportionally lower degree of patience while *low-valuation customers* are often sufficiently patient to wait for sales, setting last-minute promotional low prices is preferable. Otherwise, strategic waiting by high-value customers would need to be discouraged by setting increasing price dynamics.

The rapid growth of the Internet has had a massive impact on the hotel industry; however, hospitality literature lacks published articles that examine the trend and the variability of prices in online markets (Tso and Law, 2005). Unlike those methods in decline (such as agency, fax and phone), the Internet encourages greater price scrutiny since the relevant information is both easier to obtain and transparent, given that any party can compare the prices of several alternatives with very little effort. This is bound to have an impact on how hotel operators set room prices since they too can easily obtain this information and rapidly respond accordingly.

This work aims to provide some evidence on the actual behavior of operators in the hotel industry. Do hoteliers really make use of dynamic pricing strategies? If yes, do we observe increasing or decreasing price trends when approaching the check-in date? From the customer perspective, how should they react to the seller's pricing strategies? What are the main drivers behind the price trend structure?

The empirical analysis is carried out on a sample of almost 1000 hotels distributed around different European capital cities. The idea was to observe the evolution of prices on a predefined booking day in order to verify the extent of price variability and the significance of any trend. Moreover, we investigate the presence of alternative pricing policies in relation to different hotel characteristics, dynamic competition in the overall booking period and potential customers. With respect to the latter, we collected data for different types of booking days, in particular, an intraweek day - Tuesday night - which is usually characterized by business travelers, as confirmed by the manager of an important international chain, and a weekend day - Saturday night - more in line with a leisure trip.

The remainder of this paper is organized as follows. Section 2 provides the literature review. Section 3 describes the conceptual framework, defining the hypothesis that will be tested in the empirical analysis. Section 4) starts by clarifying the sources of data (4.1), thereafter presenting both descriptive data analysis (4.2) and a more structured empirical model based on panel data econometric techniques (4.3). Results of the panel data analysis are shown in Section 4.4. Section 5 provides a comprehensive discussion of the conceptual hypotheses in light of the results obtained. Finally, section 6 offers concluding remarks and directions for future research.

2. Literature review

We analyze both theoretical and empirical contributions in literature concerning revenue management in hotels and other service industries characterized by finite inventories (such as, for example, airline seats).

From the theoretical standpoint, Gallego and Ryzin (1994), starting from a general demand function, proved that in some situations it is possible to estimate the exact optimal policy as a function of the room stock and the length of the time horizon. Gurion (1995), under the assumption of an aggregate non-linear demand function, offered an optimal market strategic segmentation pricing strategy by dividing the market into *n* segments to maximize profit. The segmentation is confirmed in this field by other authors and proved by more recent analyses. Badinelli (2000) discussed a model appropriate for small hotels to determine the optimal solution based on the number of vacancies and also on time and revealed/hidden market prices. A serious problem and limitation in applied analysis is that hotels tend not to release any data on vacancies. Zhao and Zheng (2000) describe the relation between price and time as a non-homogeneous Poisson process. As a result, if the willingness of a customer to pay a premium for the product does not increase over time, the optimal price decreases over time for a given inventory level, as in the fashion retail market. However, in the hospitality field, the customer accepts paying a premium price. Moreover, Su (2007) takes into account that the customer population is heterogeneous along two dimensions: they attribute different values to the product and have different degrees of patience. This is a crucial turning point in literature: his theoretical model delineates that "when high-value customers are proportionally less patient, markdown pricing policies are effective because high-value customers would still buy early at a high price while low-value customers are willing to wait." To the contrary, "when high-value customers are more patient than low-value customers, prices should increase over time in order to discourage inefficient waiting." To precisely understand the variables that underpin price, Qu *et al* (2002) offer a simultaneous equations model suggesting that ‘‘hotel room price level’’ and ‘‘tourist arrivals’’ are significant factors driving demand for hotel rooms. Equally important, ‘‘hotel room quantity demanded’’, ‘‘room occupancy rate’’, ‘‘last period's room price’’, and ‘‘labor cost" significantly concur to form the final price. Another relevant aspect of this issue is that dynamic pricing strategies depend on competition between hotels. Albeniz and Talluri (2010) demonstrate that hotels with few rooms tend to sell their rooms more frequently at a discounted price, whereas large hotels are less likely to sell all their rooms, but tend to charge the full price.

Empirical studies on price dynamics mostly concern airline markets. In a time series of 650,000 flights for which up to 13 fares were available, Piga and Bachis (2006) identified through descriptive statistics that fares do not increase monotonically and moreover observed a higher volatility of fares in the four weeks preceding the departure date. Noone and Mattila (2008) focused their analysis on the effect of hotel price presentation on websites and demonstrated, with the ANOVA technique in a sample of 107 people, that the more clearly the price is identified on the hotel’s website, the more the customer is inclined to buy. They also highlighted a customer tendency to prefer decreasing rates during the period booked, even if the total amount to pay is the same, and this “bizarre” result needs further investigation.

Analyses of the hotel industry are more concerned with the booking method and other price discrimination strategies, not necessarily linked to variability over time. This set of papers refers to the difference in booking the same service from different countries, channels or websites. Yelkur and Da Costa (2001) and Chung and Law (2003) studied the performance of hotel websites according to information on facilities, customer contact, reservations, surrounding area, and management of websites. They demonstrated that the bigger the company or the quality (in terms of star rating), the better the website in terms of information and effectiveness. Piga and Bachis (2005) and Tso and Law (2005) found a significant difference in the amount of money one has to spend to obtain the same service from different distribution channels, local travel agents (traditional channels) vs. website or different countries (UK vs. Europe). This result is coherent with that obtained in the airline market by Brunger (2010), who demonstrated that in the same market those who book through traditional agencies pay more. This is the so called *internet price* effect. Ellison and Fisher (2005) argued that only one market place, for instance only one website selling a product, allows higher earnings either for the customer or for the company, as in the case of eBay.

3. The conceptual framework

Critical discussions from previous studies give evidence of theoretical models concerning the process of adjusting prices dynamically in the context of limited capacity. Empirical studies on price dynamics prevalently concern the airline industry, while there is a lack of empirical evidence in the context of hotels. In this section we build a number of theoretical hypotheses that will be verified through an empirical analysis, related to the presence in the hotel industry of an *online* dynamic pricing policy and a price discrimination strategy among heterogeneous customers.

First, the hotel industry is an ideal field of application for the theory of revenue management (as a general reference, we can refer to Talluri and Van Ryzin, 2004). Usual price differentiation criteria are the physical attributes concerning service provision (facilities and amenities), reputation (star classification and brand affiliation), site-specific attributes (location, local attractions, climate, beach). At the same time, the growth of the Internet has increased the hotels' opportunities to discriminate clients, applying different prices on the online market over time (Tso and Law, 2005; Kannan and Kopalle, 2001). However, this opportunity may be seized differently across hoteliers, and there are no studies yet that quantify the diffusion of dynamic pricing in practice. Moreover, the theoretical solution to the revenue management problem faced by each hotel can be affected by several factors such as the number of rooms that can be sold as well as their qualitative positioning on the market.

Hypothesis 1 (H1): hoteliers make use of online dynamic pricing strategies. The extent of the inter-temporal price variability depends on hotel characteristics such as star category and size.

In the preceding literature review we discussed Su's (2007) model. He develops a model where the seller adjusts prices dynamically in order to maximize revenue, demonstrating that customer composition and behavior are the main drivers of the optimal pricing strategies. The prices should decrease over time when high-valuation customers are quite impatient "to buy early at higher prices" and low-valuation customers are "sufficiently patient to wait for sales". When waiting costs are high, customers will not wait to buy; when waiting costs are low customers will delay their purchases in order to benefit from promotions and discounts. On the other side, prices should increase over time when "high-valuation customers are more patient than low-valuation customers: this discourages inefficient waiting and also captures surplus from high-valuation customers who miss the promotional prices".

H2: inter-temporal hotel pricing strategies depend on the composition and behavior of the customer population.

In order to empirically test H2, we consider that hotel clientele can be classified into different categories based on the interest in a specific hotel and on the waiting patience level. These are: business clients, usually keen on a specific hotel and conditioned by restrictive business appointments, and leisure clients, with variable interest in a specific hotel but mainly characterized by less impatience than business people. Thus, we deem there is a strong correlation between the composition of customer population and the type of booking date. During working days, business clients are predominant; they are prevalently high-valuation customers with a low degree of patience and are willing to book as soon as they find out they need to travel, even if prices are high. So, the hoteliers want to exploit their high willingness to pay independently from the period of booking. Nevertheless they might reduce prices at last minute in order to fill eventual empty rooms by attracting the more elastic (also leisure) clientele. During the weekend, the customer population is mainly composed by leisure clients. Among these, low-valuation customers tend to search for early offers to save more, comparable to what happens when booking low-cost flights for example.

H2a: when the booking date is a working day we expect a dynamic decreasing pricing strategy.

H2b: when the booking date is on the weekend we expect a dynamic increasing pricing strategy.

The star-rating system is a well known standard to rate hotels. The main components of the rating are the levels of service provision and quality attributes. Literature gives evidence that high star ratings are a quality signal supporting premium pricing (Abrate et al., 2010; Bul, 1994; Israeli, 2002; Wu, 1999). We argue that the highest star rankings give hotels the opportunity to maintain a more stable pricing policy over time when the general price trend is decreasing, and to take more advantage than average when the general price trend is increasing.

H3: the highest star ratings are a quality signal supporting not only a static but also a dynamic premium price.

The literature emphasizes the importance of the number of available rooms in defining dynamic pricing policies (Gallego and Ryzin, 1994; Badinelli, 2000). We do not have data concerning the evolution of the number of vacancies for each hotel, but can observe the number of hotels with available bookings within a certain city. Such information can be used to control for aggregate demand shocks at the city level. If there is a positive demand shock, measured by a drop in the number of hotels with available rooms, customers who have not yet booked will face more limited choice opportunities and may be induced to pay higher prices for the remaining hotels.

H4: Within a certain city, prices are higher when fewer hotels with a similar star rating have availability, and vice versa.

4. The empirical analysis

4.1. Data collection

From the start of this research project in September 2009, price information was collected through the Venere.com website, a search engine that included all the main hotels in Europe at that point in time. The availability of "no frills" prices facilitates comparison between hotels. Furthermore, to reduce the risk of obtaining biased data, and to show such a difference, a sample test with the official website of each hotel was provided.

As a first and necessary step, we checked whether the room prices between the hotel's website and other web search engines (such as booking.com, expedia.com) were similar in order to strengthen our decision to collect data using only Venere.com. In our test sample, conducted on hotels in Rome, we confirmed the findings of O'Connor (2002, 2003), who argued that "upscale hotel brands were more likely to quote more expensive prices on their own websites than on other channels". In fact, the level of prices was generally higher on the specific hotels' websites. However, the price trend, which is the focus of our study, showed very similar patterns for both channels.

Supposing a higher volatility of prices in the period closest to the check-in date, as Piga and Bachis (2006) suggested in the comparable airline field, we collected rates for hotel stays of respectively, 1, 2, 4, 7, 15, 22, 30, 45, 60, and 90 days from the date of query. The main reason to do so was to satisfy the need to identify the evolution of prices in the three months before the booking date.

The dataset includes information on two different queries (01/12/2009 and 30/01/2010), each in a time series of ten time periods. Both queries concern a standard single room. We do recognize that other extensions are possible considering that many hotels offer upgrades to the basic service (such as, for example, king size or particular view), but these specific attributes complicate the accuracy of comparability across hotels. The single room was preferred to the double room because it appears more suitable to investigate at the same time both the business and the leisure target. In particular, the first query is for a room booked on a Tuesday (higher proportion of business travelers), the second for a room booked on a Saturday (mostly leisure travelers).

The booking dates were chosen to avoid vacation periods which might alter the composition of customers if also favoring mid-week leisure travelers. We consider in our dataset several central and west European capitals, characterized by both business and tourist trade: Amsterdam, Berlin, London, Madrid, Paris, Prague, Rome and Vienna. The total number of hotels analyzed in our sample is 916 - at least one hundred hotels were randomly selected from each city. For each hotel, detailed information about the city, category (star rating), number of rooms, zone (city centre vs. suburb) and period was acquired. In the present setting we do not have information on the percentage of rooms already booked (load factor) at the time the prices were retrieved. This problem makes it difficult to exactly distinguish the factors behind the temporal price dispersion. Nevertheless, the size of our dataset enables us to address several reasons for the trend observed.

It is important to highlight that sometimes hotels had no room availability on Venere.com in one or more of the periods considered, making it impossible to record the corresponding price. This unavailability may be due to demand shocks at the city level or to a particular strategy adopted by the single hotel. In fact, hospitality managers may decide to ration the availability of rooms at several points in time. In this case, the booking unavailability may reveal only a partially exploited load factor. To solve this problem we first use "an ethical treatment of missing data" (Barnard and Lan, 2008), as follows: when one or two consecutive missing values were present in the 10 rates of each time series we imputed the average value of the previous and next real value in order not to change the underlying trend, which as the core of our analysis needs to be unbiased. To the contrary, when more than two consecutive missing values were present we decided to delete the observation because in the latter case our assumptions alone would have affected over 20% of our prices. By doing so, we reduced the number of hotels from 916 to 755 for the first booking date, and from 916 to 562 for the second booking date. This approach was used for the descriptive analysis in Section 4.2.

An alternative approach was to consider missing data as a meaningful source of information, interpreting booking unavailability as a signal of fully booked hotels. In this case, such information can be used to test whether the price levels are correlated with the number of comparable hotels with available rooms (H4). The panel data model developed in Section 4.3 allows studying this effect.

4.2. Descriptive data analysis

In this section, we investigate to what extent hoteliers make use of dynamic pricing strategies by means of descriptive statistics concerning price levels, price variability over the time horizon, and box plots referring to specific cities. Thereafter, we describe the price trend by means of linear regressions for each hotel and by clustering the results into specific categories.

[INSERT TABLE 1 AND TABLE 2 HERE]

Table 1 and Table 2 show the mean prices across all hotels by different booking days, split into the two periods considered. Even when using such a highly aggregate measure, some interesting features arise. For the first group (Table 1), the mean price decreases in the period considered. The "advance purchase discount theory" (Gale and Holmes, 1993) does not stand up in this case. The second group (Table 2) presents an opposite trend: even in a rather flat manner (tending to be U-shaped) the trend goes up. The best time to book for a weekend seems to be around three weeks before the date of the query. The period closest to the date of query is risky also because the amount of available hotels decreases, implicating a limited choice for the customer.

The trends in raw data are not large, however they are confirmed also by using a simple permutation test. The test takes the original data at the two extreme time points (90 days and 1 day before the booking date) and randomly reassigns the prices to the two dates. By counting the number of hotels that register a higher/lower price and by replicating this process 1000 times, a random distribution is obtained and represents the null hypothesis to be compared with the original data. In the case of intraweek date, the number of hotels with a lower price at the initial point is lower than the predicted value under the null hypothesis while in the weekend there is the opposite phenomenon.

[INSERT FIGURE 1 HERE]

[INSERT FIGURE 2 AND FIGURE 3 HERE]

Figure 2 and Figure 3 clearly show how the above described patterns continue to hold when we consider specific clusters. Each box in these figures provides a graphic summary of the distribution of prices for each booking day. We focus on the line inside each box, which represents the median of the distribution (the lower hinge in the box represents the 25th percentile while the top hinge the 75th percentile). It is evident that the trend is decreasing but the monotonic property is sometimes violated. For instance, in Figure 1 the median price available 15 days prior to the hotel stay is lower than the median price of the immediately preceding days. Interestingly, prices are consistent in the last week. In Figure 3, a decreasing variability in the period considered is also observed. In addition, even if there are black dots that represent values that are distant from the box, the trend is clearly decreasing in the period considered.

The second group of data, namely, the query referring to Saturday 30/01/2010, tells a completely different story.

[INSERT FIGURE 4 AND FIGURE 5 HERE]

In Figure 4, prices in London present an increasing trend without outliers and the variability of rates becomes higher and higher within the booking period. In Figure 5, rates in Prague present an upward U-shape. The choice of specific and different cities was meant to graphically highlight the opposite phenomenon. The subsequent analysis will investigate the significance of these trends.

The evidence presented so far suggests that in our sample we can identify opposing trends in the two groups taken into account, giving a first support to our hypothesis H2. Although the lack of sales data at each point in time does not permit us to fully understand the reasons behind price dispersion, our findings suggest a complex relationship between price, load factors, and leisure vs. business travelers.

Figure 6 shows the percentage of price changes observed between two consecutive observations. We can see that price adjustments are frequent over the entire time interval considered (3 months), providing support to hypothesis H1. If we exclude the last week preceding the accommodation, the share of hotels registering a price change is always higher than 30%, with peaks at over 60%. This share is slightly lower in the last week, but this is due to the fact that the observations are within much closer time intervals (3, 2 and 1 day). Indeed, another strong element supporting the diffusion of dynamic strategies appears to be the fact that around 20% of hotels change prices during the day immediately before the check-in date. If we had considered the share of hotels changing prices during the last week, we would have obtained a very high share: respectively, 46% and 71% for the two booking dates.

[INSERT FIGURE 6 HERE]

The analysis of price volatility allows us to understand the diffusion of dynamic pricing among operators as well as the relevance of price discounts over time. For this reason, we divided our data according to the coefficient of variation, a compound measure to analyze the standard deviation taking into account the mean of prices.

[INSERT TABLE 3 AND TABLE 4 HERE]

Table 3 and Table 4 illustrate the coefficient of variation among different hotels in the two periods analyzed. On the first booking date (Table 3), only 11% of times is there no change in rates during the overall period and half of the sample presents a coefficient of variation higher than 0.1. In the second group of data the variability is more marked (44.2 is the relative frequency of CV higher than 0.1).

Another step in our analysis consisted in dividing the data into different clusters in order to reveal some hidden trends. A simple linear regression for each hotel was conducted to disclose the overall trend. According to the results, we divided the hotel data into five clusters based on the price trend: decreasing by more than 20%, decreasing by 5-20%, increasing or decreasing at most by 5%, increasing by 5-10%, increasing by more than 20%.

[INSERT TABLE 5 AND TABLE 6 HERE]

Table 5 and Table 6 show the distribution of hotels by star rating and size (rows) and price trends (columns) for each set of data. In this research, one and two stars are bundled together, as are five and five luxury stars (5L). In the first set of data (Tuesday, 01/12/2009), one, two and three star hotels show a percentage of times when prices decrease more than others. The median for these types of hotels decreases over the entire period. The highest quality category presents a more stable trend. In over 38% of 5/5L star hotels, there is no significant price variation within the overall period. To the contrary, in the group referring to the second booking date (Saturday, 30/01/2010), four and five/5L stars present a stable or increasing trend. For instance, more than one quarter of 5/5L star hotels show a price increase between 5% and 10% from the price displayed three months before. Generally, in this group of data, as we anticipated earlier, the volatility is less pronounced. Regarding size, the smallest hotels, which are hotels under the median of 53 rooms, seem to show a less pronounced volatility, especially when considering the second booking dates where more than 40% of hotels reveal no, or only a minimal, variation. However, a statistical test showed that there is no significant difference in dynamic price variability (as measured by the coefficient of variation) between the two groups of hotels defined according to size (p-value = 0.27).

Finally, we show a Principal Component Analysis (PCA) of the difference of prices, standardized with the following approach: $(\ln (p_{t+1}-p_t))/p_t$.

The presence of interval data on different scales, grades from 1 to 10 and percentiles from 1 to 100, required a standardization of the variables because PCA is sensitive to the relative scaling of the original variables. Therefore to weight equally hotel with high and low prices to get the trend we divided by the original prices and we had a new table with all the standardized variations.

PCA uses Euclidean Distance calculated from the *n* variables as the measure of dissimilarity among the m objects and derives the best possible *k* dimensional ($k < n$) representation of the Euclidean distances among objects. The first *k* components display as much as possible of the variation among objects.

In the graph presented we can see that there are clearly two trends: One at the end, from one week to the day before the date of the check-in and the other from three months before until one week before the check-in. This means that hoteliers decide to substantially change the dynamic pricing strategy the last week. It should be mentioned that this technique allows to see correlations among the strategy adopted and that in this case was used only for the intra-week day.

There are two velar dimensions that stand out among the others in terms of correlation. The power of this graph is to present with only two components all the 9 variations of prices across time (T60, T45, T30, T22, T15, T7, T4, T2, T1).

We are losing one observation because calculating the differences among booking dates the first observation disappears.

4.3. Panel data analysis

The panel data analysis allows strengthening the results by exploiting the whole sample. At the same time, with this approach we can take into account the specific characteristics of each hotel (or group of hotels) affecting price levels and, eventually, variation.

Let *i* be the subscript indicating a hotel in our sample. For each hotel, the price proposal (*P*) was observed with reference to different check-in dates (subscript d = 0, 1) and progressively at 10 different time distances from the date of the query (90, 60, 45, 30, 22, 15, 7, 4, 2, 1 day(s) from the date of query). Let t = 0 to 9 be the subscript indicating each progressive observation (independently from the actual number of days of advance booking). The structure of the panel is unbalanced in the dimension *d*, since in some cases the hotel was not available for booking.

The model can be written as follows:

$$\ln P_{idt} = \alpha + \beta D_d + \gamma Z_i + \omega t D_d S_i + \tau A_{idt} + \mu_i + \varepsilon_{idt}$$

The dependent variable P has been transformed in logarithm, in order to interpret estimated coefficient in terms of percentage impact on price. More precisely, the percentage effect of a dummy variable should be computed according to the formula: 100{exp[β – Var(β)/2] – 1}, where β and Var(β) are the estimated coefficient and related variance for the dummy (Kennedy, 1981). However, for small values of β, the coefficient directly yields the effect with an acceptable degree of approximation. D indicates the dummy variables referring to the two different check-in dates; because of collinearity with the constant term (α), β is a vector composed of a number of parameters that is equal to the number of booking days minus one (in our case one). Z indicates hotel specific characteristics that affect price levels (for example, city and zone dummies, hotel star rating): these variables are intended to explain cross-sectional price variability. S indicates hotel specific characteristics which, along with the type of booking date (D), can affect the trend of dynamic prices. The variables in S may or may not coincide with Z. The interaction between time, D and S allows estimating specific trends for each desired set of combinations of hotel characteristics and booking date. Finally, it is possible to control for aggregate demand shocks within a certain area, defined as the share of hotels with booking availability within a certain group of competing hotels (A). The vector of parameters γ, ω, τ estimates the effects attached to the respective set of explanatory variables.

The error term is composed of the panel specific effect (u_i) and the random noise (ϵ_{idt}). The panel specific effect may be random or fixed; in the latter case, the estimation would drop the term γZ_i. We used a random effect model, which assumes that u_i is normally distributed and uncorrelated with the regressors. To test the validity of such assumption we compared the results with those obtained with a fixed effect model. On the basis of a Hausman test, the hypothesis that coefficients are not significantly different under the two alternative specifications can not be rejected (p-value = 0.8191 in the base model and p-value = 0.2319 in the model accounting for city-specific demand shocks).

4.4. Results from the panel data analysis

Table 7 (first column) shows the results of the estimation carried out using the dataset without missing values (thus, coherent with the descriptive statistics section) and without measuring the city specific demand shocks. The constant term indicates the average price (in logarithm) for the particular case corresponding to the booking date D_0 (30.01.10 - Saturday) of a two star hotel located in Amsterdam. The city dummy variables are mostly significant and capture the average price differences between the cities (not surprisingly, Paris is the most expensive while Prague is the most economic). The star rating dummies are also strongly significant and measure the average price premium that one has to pay with respect to a two star hotel. One can retrieve the average price estimation for each particular "group" of hotels using the appropriate dummy coefficients. By way of an example, the average price for a 3 star hotel located in London, with a weekend booking date (30-01-10) observed 15 days before the check-in date (t = 5) can be computed as follows: exp[3.938 + 0.211 + 0.355 + 5*0.00286] = 91.68.

The fit of the model, as expected, is low when it comes to within variation, since the latter is explained exclusively by means of the price trend, which is imposed as fixed among the (four) groups of hotels identified. Nevertheless, the estimated trends are strongly significant. The four specific trends estimated are generated from the interaction between the two types of booking dates (D) and two clusters of hotels identified by the star category: the first (Z_{SH}) defined as "high star rating hotels" (at least 4 stars) and the second defined as "low star rating hotels" (less than or equal to 3 stars).

[INSERT TABLE 7 HERE]

First, we find a significant negative trend when considering the case of working days, confirming hypothesis H2a, suggesting that, from the customer perspective, booking at the last minute is more convenient. This savings opportunity seems to be associated especially with the lowest star category. On average, for low star rating hotels, the estimated effect implies a 1.34% price decrease on each of the 9 succeeding observations (in Table 7, the estimated effect is given by the coefficient associated to the interaction $t * D_1 * Z_{SL}$), yielding an overall reduction of around 12% from the starting price. This decrease is only partially confirmed (less than 2% overall) in high star rating hotels (coefficient associated to the interaction $t * D_1 * Z_{SH}$).

The weekend data shows an opposite trend, thus booking in advance is better, with slightly more marked evidence for high star rating hotels (0.48% with respect to 0.29% price increase on each succeeding observation). Thus, hypotheses H2b and H3 also find support from the analysis.

An additional possible explanation of the opposite trend between the first and the second booking date is offered by observing the opposite level of prices at the beginning of the booking period. In fact, the first booking date is characterized by a higher level at the beginning (+ 2.53%, as implied by the coefficient relating to the booking date dummy), while the second booking date, which presents an increasing trend, has a lower starting point.

Table 7 (second column) shows the results of the estimation carried out using the full dataset and including the effect of city-specific demand shocks. The latter was computed by grouping hotels according to city and to the two star categories clusters (high and low star rating categories). The value of the variable ranges between 0.20 and 1, the latter corresponding to when all hotels are available. The estimated parameter has the expected negative and significant sign, indicating that prices increase when there are fewer competitors with available rooms (hypothesis H4). In particular, for example, if half the hotels can be booked (A = 0.5), we can expect, ceteris paribus, a 7.3% price increase with respect to when no competitors are fully booked. To obtain the quantitative effect, it is sufficient to simulate the change in the predicted value of price due to a change in the competition variable. Another way is to compute the price elasticity with respect to the competition variable, which is equal to $\frac{\partial \ln P}{\partial \ln A} = \frac{\partial \ln P}{\partial A} * A = \tau * A$.

5. Discussion

We believe the results of the empirical analysis validate our conceptual framework and contribute to enhancing understanding of dynamic pricing practices in the hotel industry. Nonetheless, we are aware of some limitations related to the dataset. In the following, we discuss the hypotheses in light of the empirical evidence and the critical aspects that may need further investigation.

H1. Diffusion of dynamic pricing strategies.

In line with revenue management theory, we find that the large majority of hotels use some form of dynamic pricing. For each hotel, we observe the price proposals in ten different periods of time, from three months before the check-in date onwards. In less than 10 percent of cases the price never changes. In almost half the sample, the coefficient of variation is higher than 0.10, indicating that, on average, we can expect price levels to deviate from the mean more than (+ or –) 10 percent. From the customer perspective, this suggests important potential savings if choosing the best period to book, even if it is worth mentioning that price volatility is certainly lower with respect to the case of the airline industry. Moreover, price adjustments are frequent over the entire time interval considered, not only in the days and weeks immediately before the check-in date. This suggests to future research that it may be worth increasing the number of records, reducing the time distance between consecutive observations over the whole period.

H2: Dynamic pricing and composition of customer population.

Results show statistically significant opposing price trends in the two sets of data analyzed (a mid-week and a weekend day). We argue that the different trends depend on the different nature of the customer population, mainly business in the first case where the query date is a Tuesday and mostly leisure in the second where the query date is a Saturday. Even if we must be cautious in generalizing our results, since they derive from only two sets of data, we find interesting interpretations in light of relevant studies in the field.

Mandelbaum, the director of research of PKF consulting, a hotel-industry research agency, states that “generally, prices get lower closer to your check-in date as the hotel looks to fill empty rooms”. This assertion is confirmed only for the set of data referring to a working day, the group that we define as mostly business travelers. On the other hand, in the second group of data of mostly leisure travelers, the trend increases overall, discouraging last minute booking. We interpret this evidence in light of Su's concept (2007), indicating two specific inter-temporal pricing strategies dividing the population according to different product valuations and degrees of patience. With a customer population mainly composed by business travelers, we can attribute to high-value customers a lower degree of patience. Instead, with a significant number of leisure travelers, as on the second date selected in our investigation, low-value customers with budget constraints may be keener to search for early promotions and avoid the risk of facing higher prices when booking last minute.

The increasing price schedule appears more coherent with the airline sector. However, it is worth noting some major differences. Venere.com allows cancelling a reservation without any penalty within 72 hours and, moreover, it is at times also possible to reschedule/cancel a reservation later, by negotiating directly with the hotel. Therefore, at least in theory, a decreasing price trend may be risky for the hoteliers if customers behave strategically, booking in advance and rescheduling later. Such strategic behavior is probably more diffused among low-value leisure customers and this can be a further justification in proposing increasing price dynamics at the weekend.

H3: Dynamic pricing and star rating.

In a price decreasing scenario (mid-week), high star hotels present more consistent prices. Conversely, a more pronounced increase is shown when prices rise in the overall period, as on the second booking date. This result is in line with the idea of a dynamic premium price associated with quality. In other words, the different behaviours could be related to the different characteristics of hotels: low star categories tend to also capture a quantitatively broader segment of customers at the end of the period, while high star rating hotels protect their image without presenting intensive discounts in the period immediately preceding the query. Moreover, when the price increases overall, high star categories can increase their margins more than low star categories (i.e., the price differential between low and high star hotels tends to increase over time). This implies the higher willingness of last minute clients to pay for quality.

H4: Dynamic pricing and demand shocks.

One important limitation of the study is the unavailability of occupancy rate data for each hotel. However, we can approximate aggregate demand shocks at the city level by measuring the share of available hotels within a certain area. The results confirm the hypothesis that prices are significantly higher when fewer hotels with similar star ratings have availability. On one hand, this demonstrates the crucial role of demand shocks, even if they are approximated only at the aggregate level. On the other hand, it implies that hoteliers constantly monitor the local market, looking at competitors' availability and adjusting pricing decisions accordingly. This last consideration strenghtens the idea that hotels employ strategic behavior, since their use of dynamic pricing also depends on the strategies of other hoteliers.

6. Conclusions and future research

Through a better understanding of online pricing practices, hotel management room distributors can effectively manage the online market, leading to the capture of high-value customers. This research tries to bridge a gap in literature, investigating not only the customer perspective but also the relationship between online pricing strategies and the interest of hotel room suppliers. According to the empirical results, we demonstrated that over 90% of prices changed in the period considered, with an inter-temporal structure of the trend primarily depending on the type of customer (leisure or business) and on star rating.

Considering weekdays, when customers are prevalently *business* people, the lowest prices seem to appear in the period immediately preceding the hotel stay. Instead, on a weekend, when the number of *leisure customers* is predominant, prices tend to increase when approaching the check-in date. The presence of significant trends that are heterogeneous according to the booking period and hotel characteristics, such as the star rating, is confirmed by means of panel data techniques. In particular, both in the mid-week and in the weekend date, last minute booking is characterized by a higher price differential between high star and low star hotels. Furthermore, a variable approximating demand shocks at the city level was created and proved to significantly affect hotelier pricing dynamics. As expected, the price tends to increase when there is a scarcity of hotels available to book in a certain area. This suggests strategic behavior, with hotels adapting optimal prices according to competitor room availability.

The results in this paper can be extended in two broad directions. The first is to investigate the relation between price dynamics and occupancy rates. In this paper, the latter information is not available. It would be interesting to consider price decisions, with occupancy rates from three months before the date of query. Mannix (2008) suggested an historical database, owned by Hotelligence and Smith Travel Research, which contains historical information on some occupancy rates.

The second is to obtain data referring to more booking dates. On this occasion, Piga and Bachis (2006) used an electronic spider directly connected to the website of the data source, saving collection time and granting the opportunity to analyze a longer trend. Our analysis, in fact, has an impressive number of hotels but it could be beneficial to increase the number of booking dates to strengthen the results. It could also be interesting to test if in high season the gap in the price trend between midweek and weekend days becomes smaller.

List of Tables

Table 1. Descriptive statistics. Check-in date 01.12.09.

	N	Mean	St. Dev.
T90	755	102.5	59.4
T60	755	99.9	56.5
T45	755	98.5	56.6
T30	755	97.8	56.7
T22	755	96.3	56.7
T15	755	96.2	58.3
T7	755	97.2	60.7
T4	755	97.4	61.2
T2	755	98.2	62.7
T1	755	97.3	61.8

Table 2. Descriptive statistics. Check-in date 30.01.10.

	N	Mean	St. Dev
T90	562	107.0	63.1
T60	562	104.2	61.4
T45	562	102.9	55.7
T30	562	105.6	56.7
T22	562	102.5	63.6
T15	562	107.0	60.0
T7	562	108.1	62.1
T4	562	108.4	62.1
T2	562	108.5	61.8
T1	562	109.2	62.6

Table 3. Clusters based on coefficient of variation (CV). Check-in date 01.12.09.

	Absolute frequency	Relative frequency	Cumulative Frequency
Stable	82	10.9	10.9
CV less than or equal to 0.1	355	47.0	57.9
CV higher than 0.1 and less than or equal to 0.25	277	36.7	94.6
CV higher than or equal to 0.25	41	5.4	100.0
Total	755	100.0	

Table 4. Clusters based on CV. Check-in date 30.01.10.

	Absolute frequency	Relative frequency	Cumulative frequency
Stable	40	7.1	7.1
CV less than or equal to 0.1	274	48.8	55.9
CV higher than 0.1 and less than or equal to 0.25	228	40.6	96.4
CV higher than or equal to 0.25	20	3.6	100.0
Total	562	100.0	

Table 5. Clusters based on price variations. Check-in date 01.12.09.

		Decreasing by more than 20%	Decreasing by 5-20%	Increasing or decreasing at most by 5%	Increasing by 5-10%	Increasing by more than 20%	Total
Stars	1, 2	28	28	30	15	8	109
		25.7%	25.7%	27.5%	13.8%	7.3%	100.0%
	3	71	81	75	16	29	272
		26.1%	29.8%	27.6%	5.9%	10.7%	100.0%
	4	71	77	82	49	32	311
		22.8%	24.8%	26.4%	15.8%	10.3%	100.0%
	5, 5L	5	15	24	8	11	63
		7.9%	23.8%	38.1%	12.7%	17.5%	100.0%
Size	Small	92	89	100	32	32	345
		26.7%	25.8%	28.9%	9.3%	9.3%	100.0%
	Big	83	112	111	56	48	410
		20.2%	27.3%	27.1%	13.7%	11.7%	100.0%
Total		175	201	211	88	80	755
		23.2%	26.6%	27.9%	11.7%	10.6%	100.0%

Table 6. Clusters based on price variations. Check-in date 30.01.10.

		Decreasing by more than 20%	Decreasing by 5-20%	Increasing or decreasing at most by 5%	Increasing by 5-10%	Increasing by more than 20%	Total
Stars	1, 2	4	12	20	4	11	51
		7.8%	23.5%	39.2%	7.8%	21.6%	100.0%
	3	19	43	76	25	33	196
		9.7%	21.9%	38.8%	12.8%	16.8%	100.0%
	4	20	41	97	52	54	264
		7.6%	15.5%	36.7%	19.7%	20.5%	100.0%
	5,5L	2	7	23	14	5	51
		3.9%	13.7%	45.1%	27.5%	9.8%	100.0%
Size	Small	21	54	99	33	38	245
		8.6%	22.0%	40.4%	13.5%	15.5%	100.0%
	Big	24	49	117	62	65	317
		7.5%	15.5%	36.9%	19.5%	20.5%	100.0%
Total		45	103	216	95	103	562
		8.0%	18.3%	38.4%	16.9%	18.3%	100.0%

Table 7. Random-effect GLS regression (dependent variable is lnP)

Regressor	Coefficient (Standard error)	Coefficient (Standard error)
Constant (i)	3.938***	4.065***
	(0.0378)	(0.0364)
Booking date dummy		
D_1 (01.12.09 - Tuesday)	0.0253***	0.0475***
	(0.00461)	(0.00490)
City dummies		
Z_{C1} (Berlin)	-0.172***	-0.198***
	(0.0420)	(0.0390)
Z_{C2} (London)	0.211***	0.175***
	(0.0469)	(0.0391)
Z_{C3} (Madrid)	-0.00686	-0.0176
	(0.0419)	(0.0393)
Z_{C4} (Paris)	0.573***	0.519***
	(0.0411)	(0.0366)
Z_{C5} (Prague)	-0.493***	-0.537***
	(0.0405)	(0.0371)
Z_{C6} (Rome)	0.144***	0.122***
	(0.0413)	(0.0391)
Z_{C7} (Wien)	-0.116***	-0.140***
	(0.0422)	(0.0393)
Hotel star rating		
Z_{S3} (Star rating = 3)	0.355***	0.361***
	(0.0334)	(0.0296)
Z_{S4} (Star rating = 4)	0.693***	0.728***
	(0.0343)	(0.0306)
Z_{S5} (Star rating = 5)	1.406***	1.411***
	(0.0477)	(0.0440)
Time effects		
$t * D_1 * Z_{SL}$	-0.0134***	-0.0171***
	(0.0007)	(0.0008)
$t * D_0 * Z_{SL}$	0.00286***	0.00336***
	(0.0008)	(0.0008)
$t * D_1 * Z_{SH}$	-0.00205***	-0.00527***
	(0.0007)	(0.0008)
$t * D_0 * Z_{SH}$	0.00477***	0.00420***
	(0.0008)	(0.0008)
Share of available hotels		-0.1415***
		(0.0147)
Observations	13,170	15,082
Number of hotels	780	916
R-square: overall	0.6332	0.6352
between	0.6774	0.6850
within	0.0384	0.0502

(i) the constant term indicates the average price (in logarithm) in the booking date D_0 (30.01.10 - Saturday) of a two-star hotel located in Amsterdam.

List of Figures

Figure 1. Permutation test. Intraweek and weekend.

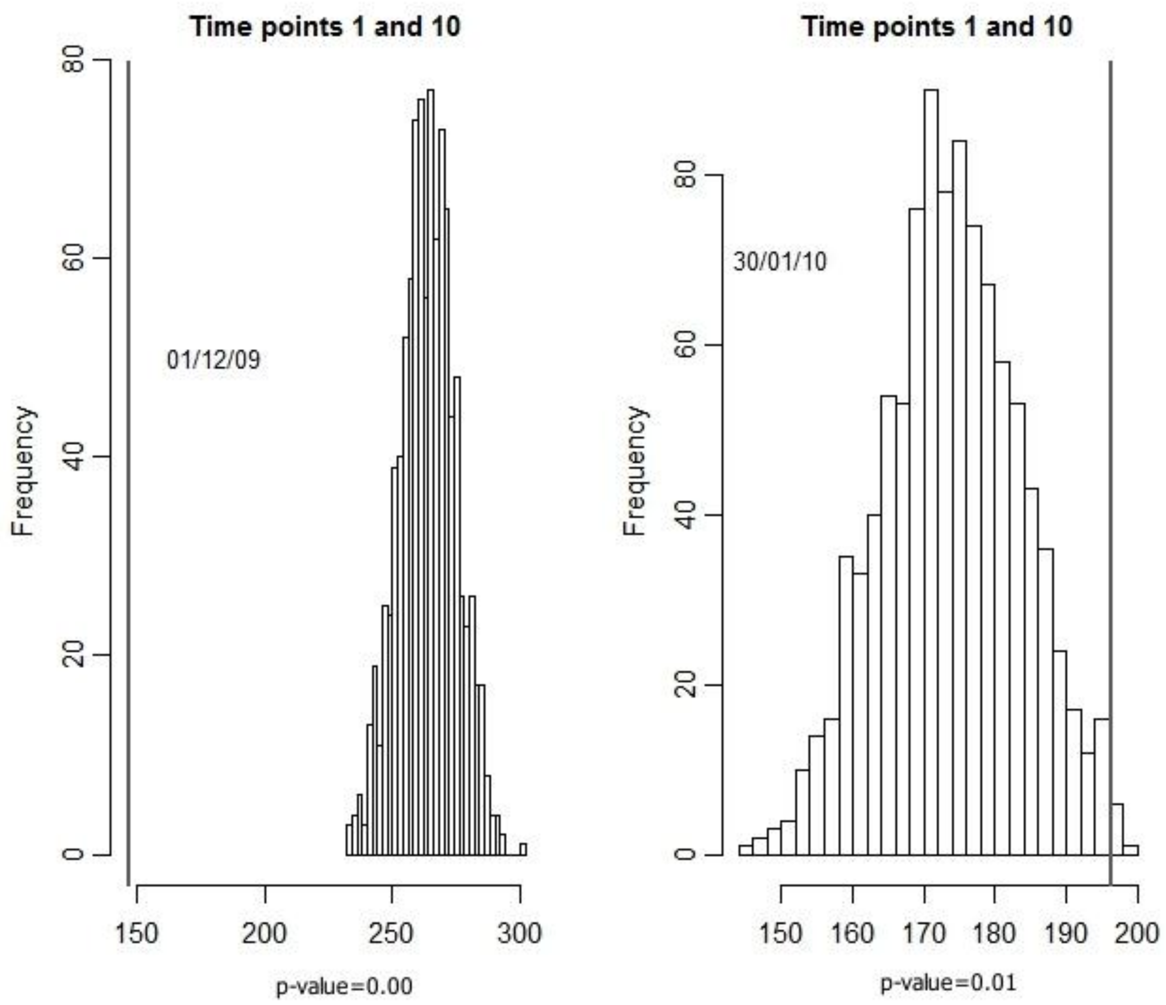


Figure 2. Example of box plot of prices concerning the first check-in date, Madrid.

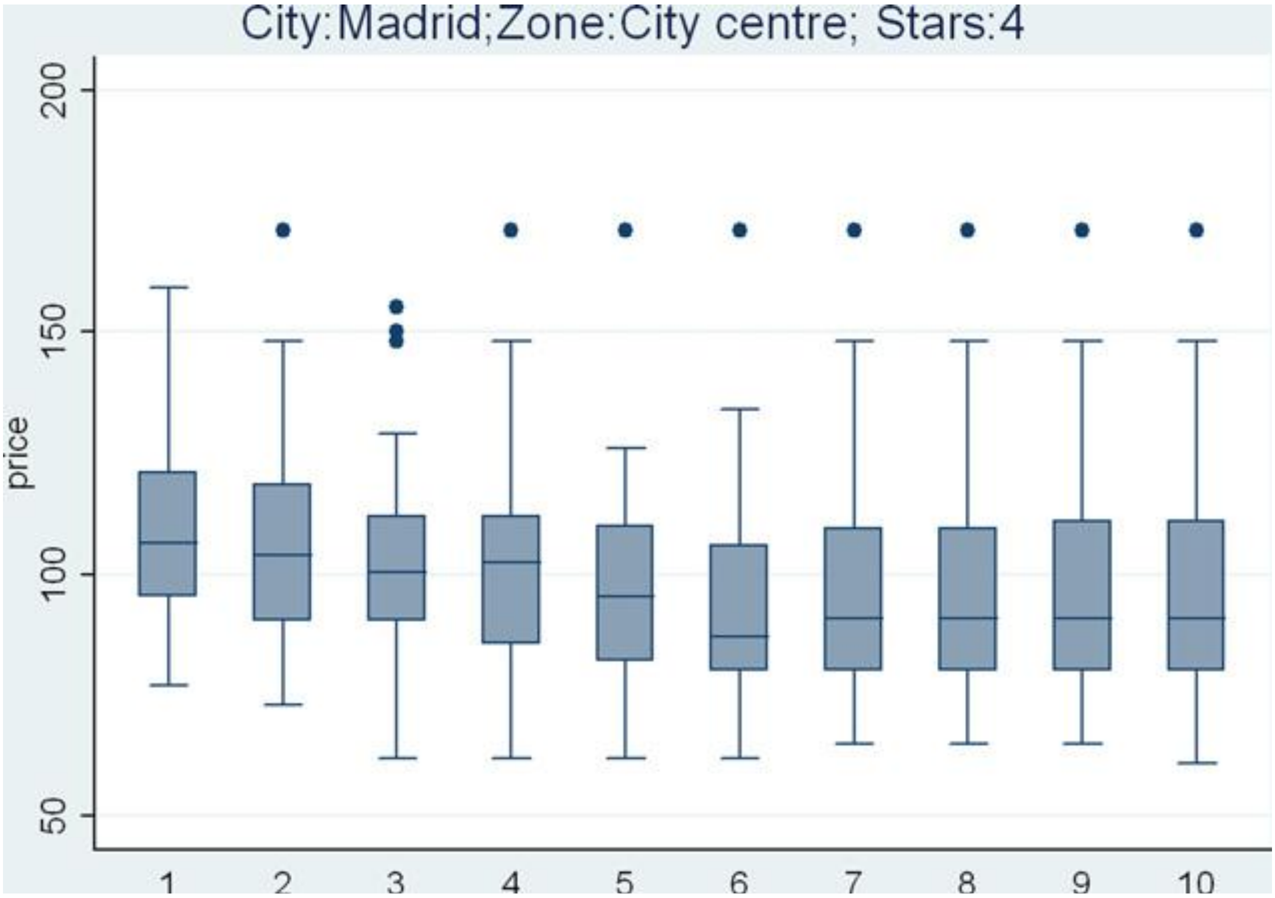


Figure 3. Example of box plot of prices concerning the first check-in date, Amsterdam.

Figure 4. Example of box plot of prices concerning the intraweek check-in date, London.

Figure 5. Example of box plot of prices concerning the weekend check-in date, Prague.

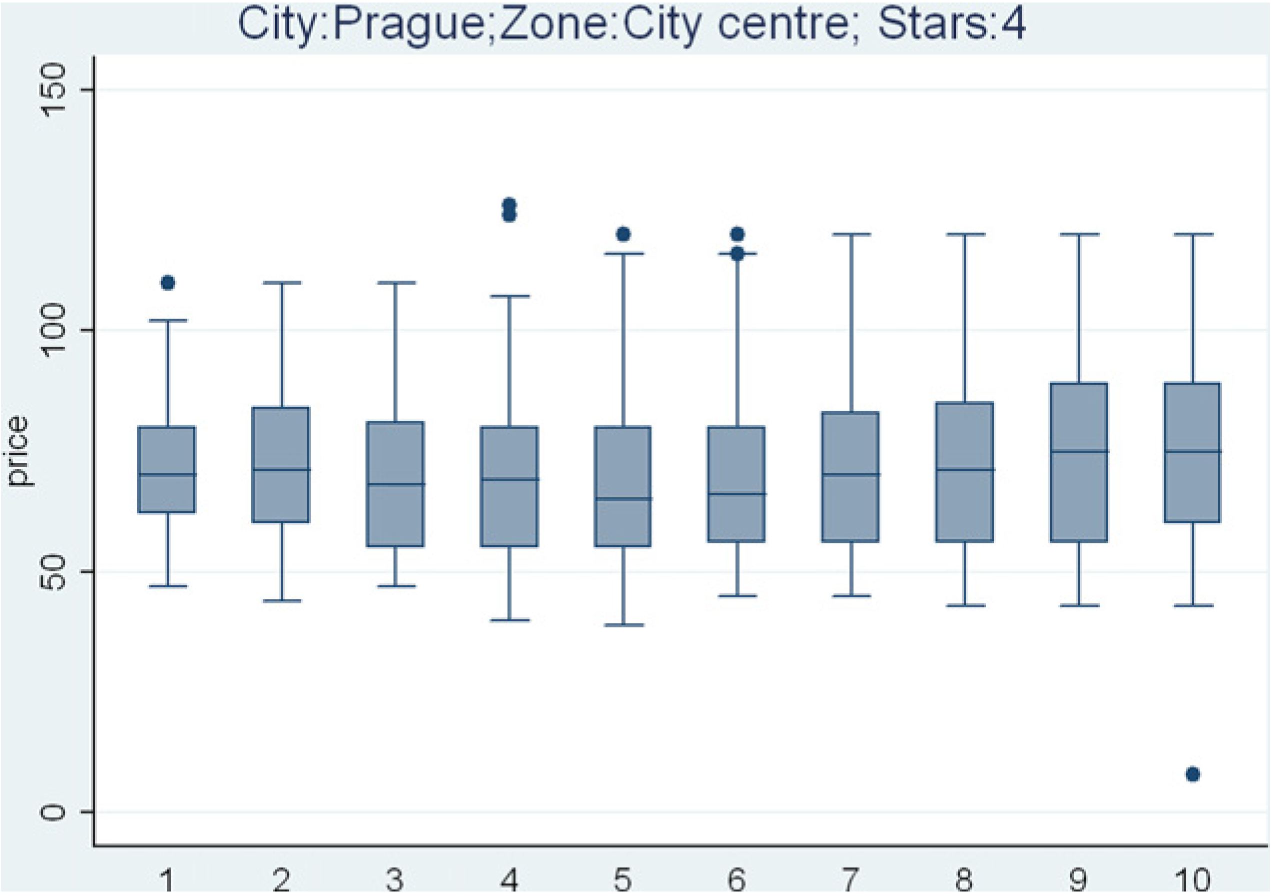


Figure 6. Proportion of price changes between consecutive observations.

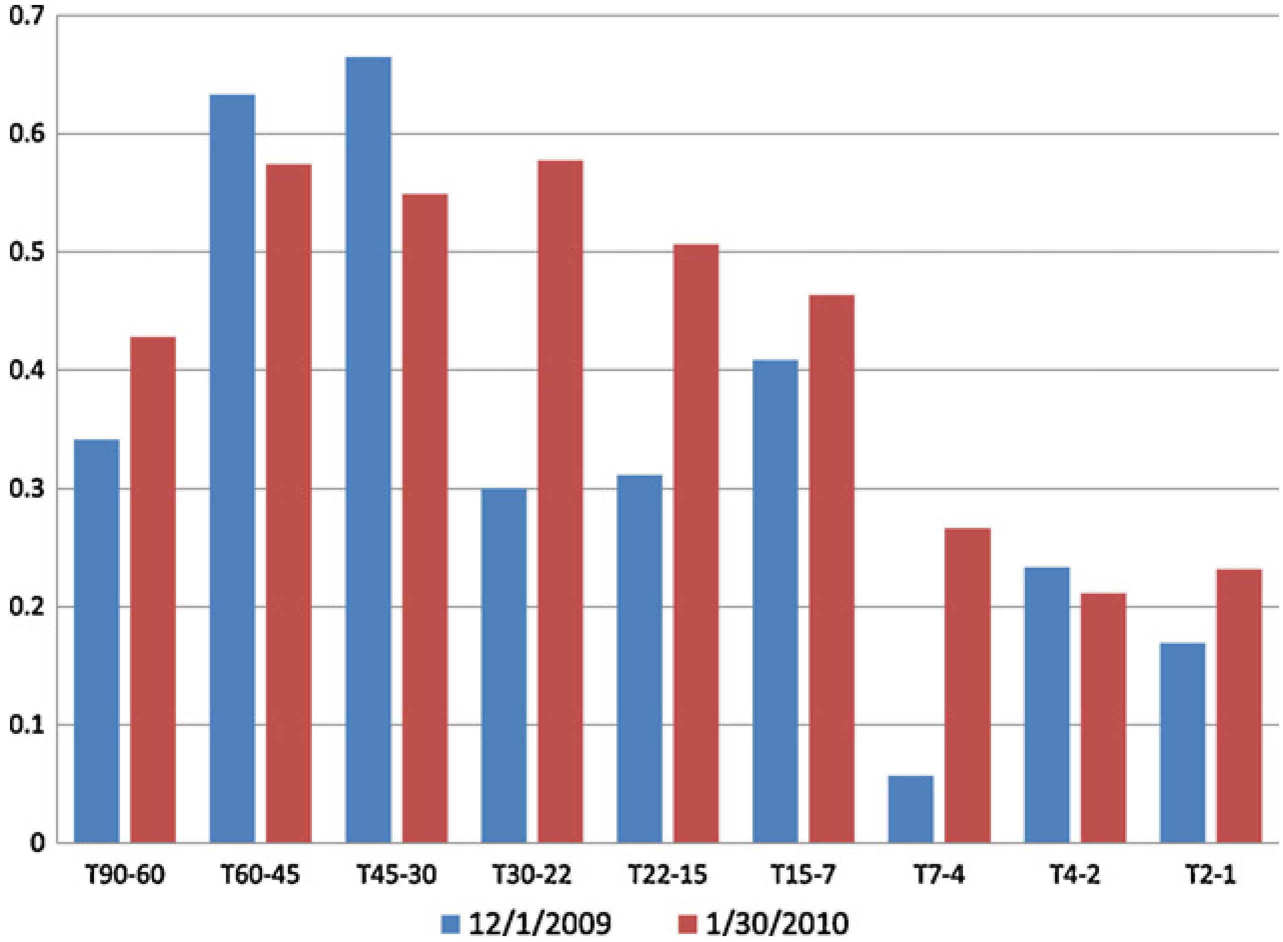
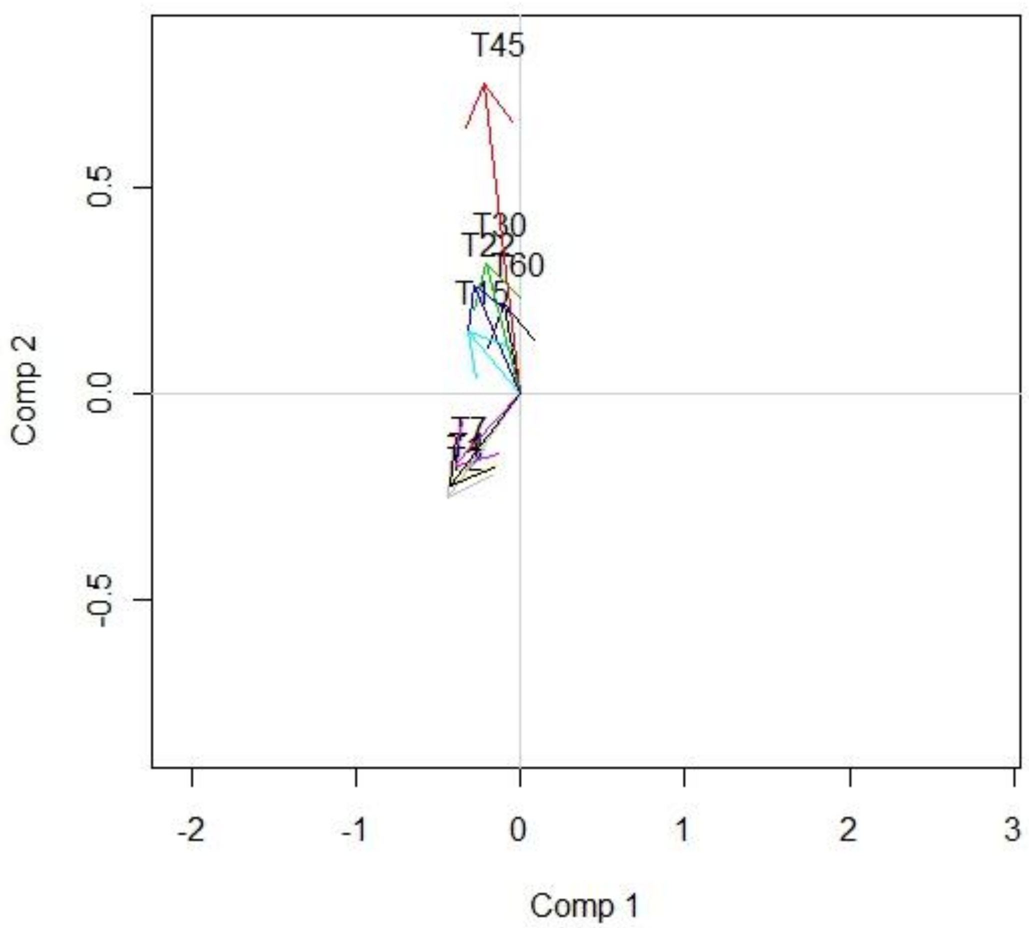


Figure 7. Principal Component Analysis of the standardized data.


T45
T30
T22
T60
T15
T7
Comp 2
Comp 1
0.5
0.0
-0.5
-2
-1
0
1
2
3