Price Competition and Overtourism: A Stochastic Approach Leveraging Complex Data



*Published Version:*
Angelini, G., Costa, M., Guizzardi, A. (2025). Price Competition and Overtourism: A Stochastic Approach Leveraging Complex Data. Berlin : Springer Nature [10.1007/978-3-031-64350-7_38].

*Availability:*
This version is available at: https://hdl.handle.net/11585/973232 since: 2025-03-07

*Published:*
DOI: http://doi.org/10.1007/978-3-031-64350-7_38

# Price Competition and Overtourism: A Stochastic Approach Leveraging Complex Data

Giovanni Angelini [1[0000-0003-3000-9885]], Michele Costa[1[0000-0003-3045-7979]], and Andrea Guizzardi[2[0000-0001-8674-7604]]

[1] Department of Economics, University of Bologna, Piazza Scaravilli 2, 40126 Bologna, Italy, g.angelini@unibo.it, michele.costa@unibo.it

[2] Department of Statistical Sciences, University of Bologna, Via Belle Arti 41, 40126 Bologna, Italy, andrea.guizzardi@unibo.it

**Abstract.**
This study explores the pricing tactics and strategic decision-making processes within the hospitality industry. The availability of complex data from online travel agencies has transformed pricing into a dynamic process similar to stock market dynamics by providing continuous information on competitors' pricing decisions. The study highlights the importance of a stochastic approach to capturing price trajectories. By utilizing a Structural Vector Autoregressive (SVAR) model, we analyze how the interplay between demand seasonality, hotel features, and customers' willingness to pay affects price competition throughout the booking window. The suggested multivariate stochastic approach aligns better with modern dynamic pricing algorithms based on stochastic demand functions. The findings reveal the dominance of seasonality-based pricing strategies set by revenue managers in Venice, highlighting the potential for more sophisticated pricing policies that leverage factors such as room availability, quality, and rates fences within the booking window. However, only few hotels consider these factors to modulate price competition, and mostly in the last minute. This result underscores the importance of investing in revenue management algorithms and skilled professionals who can accurately forecast the pick-up curve and optimize pricing strategies, even in destinations facing challenges related to overtourism.

**Keywords:** complex data, multivariate stochastic approach, advance booking.

## 1 Price competition, seasonality and booking time

Pricing tactics, and strategic decision-making are more and more driven by complex data thanks to the growing availability of information from modern Property Management Systems (PMS) managing, in real time, rates, inventories and customers' fidelization and satisfaction. Consequently, firms publish on Online Travel Agencies "ask" prices akin to stock markets, providing those who seek to study competition dynamics with a continuous flow of information about decision-makers' regarding the expected pick-up curve in the booking window. This complex framework must be studied with statistical methodologies able to consider the interplay of demand seasonality, hotels' features and customers' willingness to pay, across the booking window.

Seasonality, including calendar effects, significantly drives price competition, especially in metropolitan areas and highly popular destinations where events, holidays, or weekends can significantly affect daily demand. Lower-quality hotels often offer deeper discounts during low seasons while high-priced hotels are generally less sensitive to demand fluctuations, indicating asymmetric seasonal effects based on quality differentiation [8]. However, during high-peak weeks (i.e., during fairs, periodic or mega-events), lower-scale hotels can capture customers willing to pay higher prices at the last minute due to room shortage [4].

Existing literature often treats the advance booking determinant linearly and deterministically, implicitly assuming a constant pick-up rate. However, a more accurate description of price trajectories across advance bookings (that not neglects the active intervention of yield management strategies) can be achieved by adopting a stochastic approach [5] that captures the dependence between prices set at different booking periods. This approach aligns better with modern dynamic pricing algorithms and allows consideration of price fairness issues throughout the booking period. As market competition intensifies, retailers may explore more complex selling strategies during the booking window also leveraging customers' willingness to pay for options like a premium for sustainability actions or free cancellation and date changes [7].

We consider price discrimination patterns as a multivariate stochastic process, estimating a Structural Vector Autoregressive (SVAR) model. From a theoretical perspective [2], room rates follow a stochastic dynamic process driven by hoteliers' expectations regarding stochastic demand at different arrival days. Expanding on this, we propose that the stochastic causal chain among rates should be identified also considering the advance booking period, as expectations for the occupancy rate on the arrival day, primarily change based on incoming bookings and cancellations. Therefore, our focus is on modeling the sequence of prices posted online for different arrival days and advance booking periods, considering other relevant exogenous variables affecting product differentiation (e.g., rate fences or rooms features).

While VAR approaches are not new in tourism literature, they are more commonly applied to demand forecasting rather than price modelling. One notable exception is [1], who employ a panel-VAR model to describe price dynamics and competition in the Italian airline and railway markets. However, their panel-VAR methodology becomes impractical for a large cross-sectional dimension, typically the case with competing hotels in a destination, as restrictions on the coefficient matrices are hard to be imposed, leading to results with poor economic interpretation. Tensors represent a possible solution as they enable the modeling of higher-order interactions and dependencies in time series data, which we are exploring to capture the complex relationships between seasonality and advance booking across arrival time [9].

## 2 The model and data

We focus on $Y_{t-j,t}^{i}$, the logarithm of the Best Available Rates (BARs) for the *i-th* hotel, with $i = 1,2, ..., N$, available on the internet in $t - j$, for the arrival date $t$ ($j$ indicates the advance booking and $J$ is the maximum advance booking considered). According to [3], for each hotel we specify a SVAR model for the $J \times 1$ vector $Y_t^i$ which has the following representation (we drop the hotel superscript i):

$$Y_t = C + AY_{t-1} + \Gamma X_t^s + \Theta X_t^f + \eta_t, \quad \eta_t \sim N(0, \Sigma) \tag{1}$$

$$\eta_t = B\varepsilon_t, \quad \varepsilon_t \sim N(0, I_k), \quad BB' = \Sigma \tag{2}$$

where $X_t^s$ is a vector with exogenous seasonal variables and $X_t^f$ is the vector with room-dependent variables. It is important to stress how our specification is flexible with respect to different formulations of vectors $X_t^s$ and $X_t^f$. Both A and B matrices are constrained to be upper triangular. This allows us to assume that significant statistical relationships can be inferred only from larger to smaller advance bookings, coherently with the chronological ordering of the endogenous variables.

The estimation of the coefficients in model (1) (2) is performed under the assumption of Gaussian errors via constrained maximum likelihood, which is asymptotically equivalent to a feasible generalised least squares estimator [6]. In order to assess how shocks propagate both along the price trajectory and between subsequent arrival dates, we compute the Impulse Response Functions (IRFs):

$$IRF_j(h) = A^h b_j, \quad h = 0,1,\dots,H, \quad j = 0,1,\dots J \tag{3}$$

where $b_j$ is the *j-th* column of the matrix of structural coefficients *B* and *H* is the maximum horizon considered.

We focus on a set of 102 highly rated (3 stars or more) hotels located in Venice, Italy, which is a destination experiencing periodic excess demand (overtourism). Most of them (97%) are located within 2 km of Piazza San Marco (i.e., in the very center of the city). We consider 3 advance bookings, namely *j*=0,7,28, that corresponds to the prices offered in the last minute (*j*=0) and at two early bookings periods of a week (*j*=7) and a month (*j*=28). With the aid of a web-scraping software we simulate a customer searching for a room at each of the above three different advance booking in a window from January 1st, 2019 to February 29th, 2020 (425 observations) in order to exclude the covid-19 period where competition was strongly affected by the ban on (international) travel. We scrape the BAR, to ensures the highest homogeneity with respect to possible (unobservable) product differentiation practices.

The exogenous seasonal variables vector $X_t^s$ includes 6 dummy variables ($Fairs_t; Bien_t; Carn_t; Aug_t; WE_t; Hol_t$), which take value 1 in correspondence of the most visited fairs/events, the Biennale and the carnival periods, Saturdays and Sundays, August and bank holidays. As exogenous room-dependent variables $X_t^f$ we consider two dummies ($Breakf_t$ and $Cancel_t$ which takes value 1 if the breakfast is included and if the room can be cancelled with no penalties), and two variables $Items_t$ and $Price_t$, i.e., the number of features of the offered room (air conditioning, etc...) and the number of prices published by the same hotel. These variables serve as proxies for the quality offered and the stock of available rooms for day *t*, respectively.

Table 1 presents some descriptive statistics for hotel pricing. The average BARs are higher in the early booking. Revenue managers were successful in selling the top-rated rooms at higher advance booking as confirmed by the weakly increasing trend in the 10th and the 90th percentiles. Policies related to price differentiation based on breakfast offerings remain largely unchanged in the booking window and (as expected), the quote of free cancellations options strongly decrease in the last minute.

**Table 1.** Best Available Rate descriptive statistics.

| | j=0 | j=7 | j=28 |
|---|---|---|---|
| mean(BAR) | 214.7 | 234.0 | 243.9 |
| std(BAR) | 243.1 | 247.3 | 241.6 |
| 10-90perc(BAR) | 08-397 | 77-415 | 85-415 |
| %cancellation | 2.42% | 15.13% | 15.19% |
| %breakfast | 60.09% | 61.09% | 62.88% |
| Mean(#items) | 22.8 | 22.8 | 22.7 |

## 3 Results and discussion

In Table 2, we present a summary of the estimated coefficients, focusing only on those that are statistically significant (alpha < 20%) in at least one-third of the monitored hotels. As expected, all the effects on the BAR are positive. The impact of advance booking is particularly evident for the room-dependent variables $X_t^f$. Offering a cancellation policy tends to increase the price, especially for last-minute bookings. However, it is precisely for $j=0$ (last-minute) that we observe the highest variability, indicating that, in order to increase the occupancy rate, the possibility of cancellation is sometimes combined with a significant price increase. Selling a room with breakfast has a strong positive effect on the price, especially for early bookings. On the contrary, the effect of the number of features (a proxy for the offered quality) becomes clear only in the very last-minute, where less luxurious rooms are offered at discounted prices. The lack of a clear relationship between quality and price in early bookings suggests that not all hotels coordinate rate fences management with quality management. This unexpected simplification of pricing strategy could be explained by the influence of excess demand expected by the hoteliers.

In comparison to the coefficients estimated for the room-dependent variables, the estimated coefficients for $X_t^s$ show less variability across hotels and throughout the advance booking period. This suggests that seasonality plays a role in rate differentiation regardless of both the hotels' propensity to adopt dynamic pricing algorithms and their heterogeneity. It is interesting to note that the effect of all the variables measuring seasonality follows a U-shaped pattern with respect to the advance booking period, reaching a minimum at $j=7$ (one week before arrival). The only exception to this pattern is the weekend $WE_t$, where the shape is reversed. It should be highlighted that the seasonal effect associated with weekends has the highest coefficients compared to other estimates. Additionally, it is almost always significant, meaning that the majority of revenue managers consider it as a lever to differentiate prices. This is likely because is the most common seasonal pattern, which enables easy observation of its influence on the pick-up curve and facilitates price optimization through adjustments in room rates.

**Table 2.** Statistically significant (alpha < 20%) coefficients in at least one-third of the monitored hotels: mean, range, and observed frequency

| | Mean | 10-90 perc | Freq | | Mean | 10-90 perc | Freq |
|---|---|---|---|---|---|---|---|
| Fairs_0 | 0.095 | 0.031-0.164 | 66% | Hol_0 | 0.091 | 0.032-0.170 | 74% |
| Fairs_7 | 0.085 | 0.027-0.169 | 78% | Hol_7 | 0.084 | 0.029-0.139 | 75% |
| Fairs_28 | 0.090 | 0.028-0.148 | 67% | Hol_28 | 0.086 | 0.030-0.153 | 79% |
| Bienn_0 | 0.142 | 0.058-0.242 | 93% | Aug_0 | 0.084 | 0.038-0.142 | 42% |
| Bienn_7 | 0.122 | 0.044-0.214 | 88% | | | | |
| Bienn_28 | 0.143 | 0.044-0.272 | 83% | Items_0 | 0.024 | 0.003-0.051 | 45% |
| Carn_0 | 0.210 | 0.079-0.408 | 60% | Cancel_0 | 0.323 | 0.088-0.714 | 43% |
| Carn_7 | 0.201 | 0.082-0.338 | 60% | Cancel_7 | 0.222 | 0.045-0.408 | 41% |
| Carn_28 | 0.216 | 0.093-0.416 | 57% | Cancel_28 | 0.234 | 0.064-0.487 | 36% |
| WE_0 | 0.213 | 0.089-0.337 | 99% | Beakf_0 | 0.483 | 0.057-1.752 | 50% |
| WE_7 | 0.243 | 0.110-0.356 | 99% | Beakf_7 | 0.651 | 0.033-2.436 | 47% |
| WE_28 | 0.210 | 0.069-0.333 | 97% | Beakf_28 | 0.688 | 0.057-2.480 | 59% |

The analysis of the IRFs (see Fig. 1) shows that, for all the advance bookings considered, the response is weaker as *j* increases thus suggesting, as expected, that a shock has a stronger effect on closer horizons. The IRFs underscore the proficiency of hotel managers in Venice in implementing revenue management strategies. The IRFs also show persistence for 2/3 days (2,5 is exactly the average length of stay in Venice), implying the capability of hotel managers to adjust rates for the days following the shock, based on the pick-up curve that unfolds. Such adaptability reflects dynamic pricing management, allowing hoteliers to maximize revenues in response to demand fluctuations.

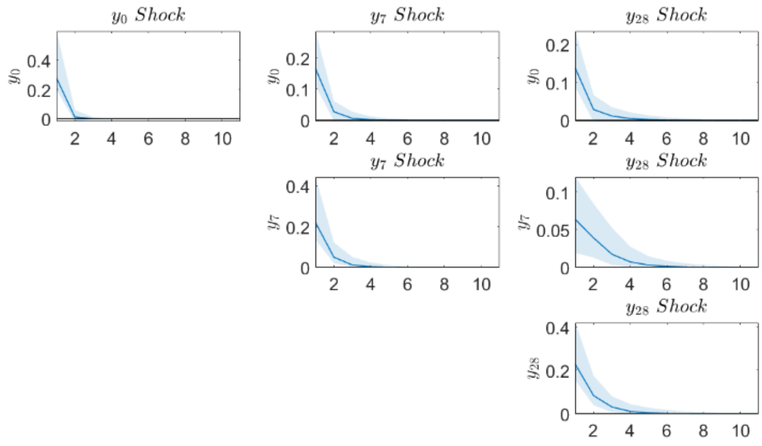


**Fig. 1.** Impulse Response Function to one standard deviation price shock.

## 4 Conclusion

In this study, we conducted a comprehensive examination of pricing strategies in a popular tourist destination like Venice. Employing a SVAR methodology, we explored the relationship between seasonality and advance booking, considering price discrimination patterns as a multivariate stochastic process. We found that the majority of hotels employ simplified pricing strategies where seasonality, particularly the arrival day, takes precedence over price discrimination based on booking time. Specifically, events like the Biennale and weekends are utilized by almost all hotels to set higher BAR prices. Prices are higher in the early booking, an unexpected result as it leaves room for speculative behaviours such as cancellation and rebooking, which is controlled by offering a low frequency of refundable BAR rates. This trend may be attributed to the dynamics of overtourism.

Furthermore, our findings indicate that the response to simulated price shocks weakens as the booking horizon extends. Only a few hotels consider factors such as room availability, quality, and rate fences to modulate price competition, and mostly in the last minute. This proves there is room to develop more complex pricing policies than those based solely on seasonality and emphasizes the importance of investing resources in revenue management algorithms and/or highly skilled professionals to accurately forecast the pick-up curve, even in destinations experiencing overtourism patterns.

## References


1. Bergantino, A. S., Capozza, C., Capurso, M.: Pricing strategies: who leads and who follows in the air and rail passenger markets in Italy. Applied Economics, 50, 4937-4953 (2018).
2. Croes, R. Semrad, K. J.: Discounting works in the hotel industry: a structural approach to understanding why. Tourism Economics 18, 769-779 (2012).
3. Guizzardi, A., Angelini, G., Pons, F.: Does advance booking matter in hedonic pricing? A new multivariate approach. International Journal of Tourism Research, 22, 277-288 (2020).
4. Guizzardi, A., Pons, F., Ranieri, E.: Advance booking and hotel price variability online: any opportunity for business customers? International Journal of Hospitality Management, 64, 85-93 (2017).
5. Lee, M.: Modeling and forecasting hotel room demand based on advance booking information. Tourism Management, 66, 62-71 (2018).
6. Lutkepohl, H.: New introduction to multiple time series analysis. Springer, Berlin (2005).
7. Masiero, L., Viglia, G., Nieto-Garcia, M.: Strategic consumer behavior in online hotel booking. Annals of Tourism Research, 83, 102947 (2020).
8. Wang, X., Sun, J., Wen, H.: Tourism seasonality, online user rating and hotel price: a quantitative approach based on the hedonic price model. International Journal of Hospitality Management 79, 140-147 (2019).
9. Wang, D., Zheng, Y., Li, G.: High-dimensional low-rank tensor autoregressive time series modeling. Journal of Econometrics, 238, 105544 (2024).