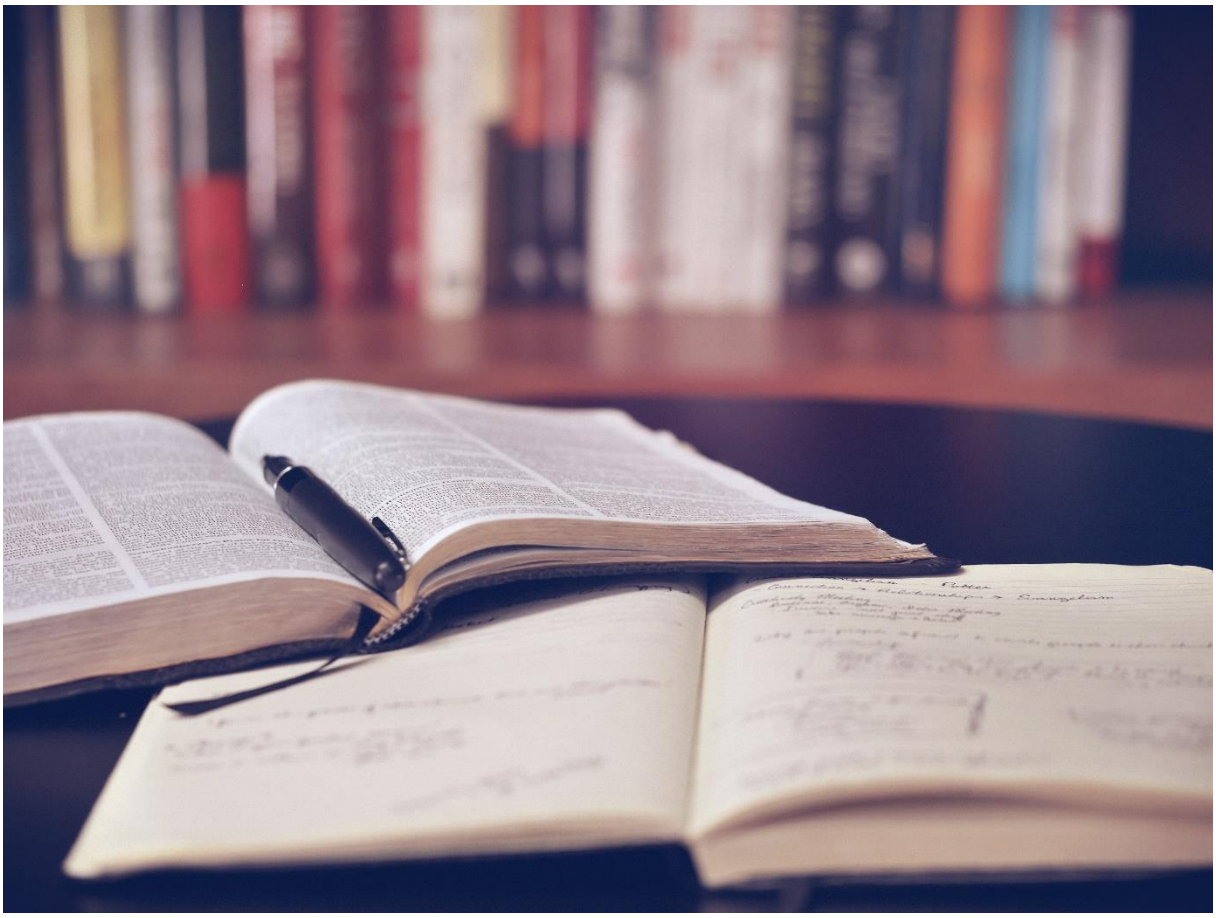
UMR CNRS 6240 LISA

Working Paper TerRA n°06

08-06-2017



Insiders vs outsiders in the hotel sector: is it worth entering an official classification system?

Olivier Beaumais, Sauveur Giannoni

Insiders vs outsiders in the hotel sector: is it worth entering an official classification system?

Olivier Beaumais*†, Sauveur Giannoni†

June 2, 2017

Abstract

We investigate whether entering an official hotel classification system is as lucrative as suggested in the tourism management literature. Indeed, in countries in which the official hotel classification system is voluntary, a substantial fraction of hotels choose not to enter the system, and are outsiders. Considering that being classified (being insider) as a predictor of the rate structure may raise an endogeneity issue, we apply the recursive semi-ordered probit model to control for endogeneity and appropriately assess the effect of being classified on price rates.

Using a sample of 357 hotels of Corsica, we show that, in contrast to previous research, classification does not provide any rate premium. We also fully derive conditional probabilities and partial effects on differences in conditional probabilities within the recursive semi-ordered probit model.

Keywords: hotel rating, rate premium, endogeneity, recursive semi-ordered probit.

JEL codes: L15, L83, C35.

*Corresponding author: olivier.beaumais@hotmail.fr

†UMR CNRS 6240 LISA; Università di Corsica - Pasquale Paoli; Campus Mariani, BP 52, 20250 Corte, France.

1 Introduction

As to hotel classification systems, a wide variety of situations can be found worldwide. In some countries, *e.g.* in Finland, no classification system is in use and hotels are ranked through customers' reviews posted on websites. In other countries, only an informal classification system exists, provided by private organizations, which get recognition from customers. Private organizations indeed appear to have played a key-role in the emergence of classification systems. It is well-known that in France, the tyre company Michelin published the first edition of a guide for French motorists in 1900 - about 2400 cars were in circulation at the time - including lists of hotels, car mechanics and petrol stations (Vine, 1981). Likewise, with the advent of motoring, automobile clubs (*e.g.,* Royal Automobile Club in England, American Automobile Association in the US) provided their members with rankings of accommodations, including hotels. Over time, these rankings have gained reputation far beyond the initial beneficiaries and are a reference for many customers. More recently, the need for standardized procedures of classification has led many countries to propose official hotel classification systems, defined by law. In turn, official classification systems can be statutory (*e.g.,* Belgium, Denmark, Estonia, Italy, etc.) or voluntary (*e.g.,* Czech Republic, France, Germany, etc.). Nowadays, feedback customers' reviews, informal and official systems actually coexist, and when an official voluntary hotel classification system is in use, hoteliers can choose whether to enter it.

Among the benefits that hoteliers expect from classification are increased rates and margins (UNWTO, 2015). Providing information to customers acts as a marketing tool in a highly competitive market, where differentiation from competitors can be a matter of survival. To illustrate the importance of classification systems, Israeli (2002) points out that, despite the Israeli Ministry of Tourism having abandoned the national star rating system in 1995, due to costs of inspection and enforcement which were found to be too high, hotels nevertheless continued using and advertising the star rating they had previously been awarded. According to Israeli (2002), the star rating may be considered as an asset that gives rise to a significant rate premium. Such a rate premium, related to entering an official hotel classification system, is commonly reported in the tourism management literature (see Abrate et al., 2011).

Yet, in countries in which the official hotel classification system is voluntary, a substantial fraction of hotels choose not to enter the system, and are outsiders. In France the share of unclassified hotels amounts to 25% (INSEE, 2015a), while in the UK this fraction is even higher, 29% according to Hotel Data Limited[1].

Thus, is the return from entering an official classification system as important as suggested in the literature? The ambition of the present paper is to investigate this question, focusing on the existence of a rate premium for insiders, classified hotels, compared to outsiders, unclassified ones. To that end, we adopt mainly an empirical approach, relying on a sample of 357 hotels of Corsica, while proposing a novel methodology to appropriately analyze the data.

Corsica is a small French island in the Mediterranean sea with 320,000 inhabitants. This region is of special interest for this study for at least two reasons. First, Corsica is one of the most popular tourist destinations in France. According to official data, INSEE (2015a), 35 million night stays are registered each year while the total amount of tourism spendings is 2.5 billion euros a year, one third of the regional GDP. Second, competition in the hospitality sector in Corsica is fierce with a listed supply of around 400 hotels, more than 200 guesthouses, 142 campsites (www.visit-corsica.com) and a significant unofficial supply.

In our search of a rate premium, the probability for a hotel to offer a given rate category is modeled using an ordered probit model. Being part of the official classification system,

[1]http://www.hoteldatauk.com.

being an insider, is captured through an independent binary variable (insider *vs* outsider, classified *vs* unclassified). As this variable raises an endogeneity issue, we use an appropriate method, the recursive semi-ordered probit model, the scope of which goes beyond the present research question, by allowing both to correct for endogeneity and to compute relevant quantities and partial effects.

In the following Section, we present a short review of the literature on hotel pricing and classification. Then, the presentation of the data is followed by our methodological solution to the endogeneity issue raised by the classification variable, along with the derivation of the partial effects within the framework of the recursive semi-ordered probit model. Empirical results stemming from the recursive semi-ordered probit model are compared to the result from a naive ordered probit model, before the last Section concludes.

2 Hotel pricing and classification in the literature

The analysis of price determinants is a key topic of the literature related to the hospitality industry within economics and management[2]. Some authors introduce original methods such as behavioral process method or conjoint analysis in order to address the issue of pricing. Danziger et al. (2006) use a behavioral process measure to investigate the contribution of strategic assets in determining customer perceptions of hotel room price in the Israeli hospitality industry. 114 participants (MBA students and hotel employees) are asked to estimate the market price of a single occupancy hotel room in the cities of Eilat and Tel Aviv after acquiring information on competing hotels. The information available for the competing hotels are price, brand name, star rating, number of rooms, number of restaurants, location and pool size. To insure the quality of the experiment, participants receive a monetary reward that is increasing with the accuracy of the estimation. The

[2] The reader will find interesting additional references in Abrate et al. (2011) and Heo and Hyun (2015).

main result shows that consumers select relatively few information items, with price and star rating information being the most frequently selected.

Conjoint analysis is used in a few studies and notably in Goldberg et al. (1984). The latter shows how the categorical conjoint model provides an efficient way to estimate utilities for large numbers of attribute levels while still retaining individual differences. Yet, no clear practical result is obtained in this study.

Recently Masiero et al. (2015) carried out a choice experiment in order to investigate the willingness to pay for hotel room attributes in a medium sized luxury hotel in Hong Kong. In a first stage, a meeting with hotel managers was arranged in order to identify the key attributes and levels to be used in the stated choice experiment so that the study could be beneficial to the manager. Seven attributes with different levels were identified: price per room per night, view, floor, access to hotel club, free mini bar, guest smartphone, cancellation policy. After a pilot survey aiming at testing the accuracy of the experiment, the final survey took place between march and may 2014. The results reveal that business travelers are less price sensitive than leisure travelers while first-time visitors to the city put much more weight on some hotel room attributes such as free beer and wine in the mini-bar. This method is proven to yield interesting results for hotel managers but still is of seldom use in the hotel literature.

Conversely, a vast strand of literature relies on the very popular hedonic price method (Rosen, 1974). This method appears popular due to its relative simplicity, its tractability and the possibility for managers to identify the key attributes to extract some rate premiums. Results obtained in a large part of these papers stress the role of hotels rating in the existence of a rate premium.

The seminal work by Carvell and Herrin (1990) aims at identifying the key determinants of hotel prices in the San Francisco area in order to suggest profit enhancing policies to hotel managers. In this study, an improvement in the rating of a particular hotel raises the price per night of about $15.

Wu (1999) studies the impact on pricing of franchising strategy and finds that franchised motels in the USA extract a rate premium compared to non franchised ones. Furthermore, he shows that a higher rating has a positive significant impact on the rate per night (around $3).

Accordingly, Israeli (2002) is interested in the role of star rating and corporate affiliation on room prices in the hospitality sector in Israel. In that case, an improved star rating is associated with a rate premium between $42.7 and $66.3 per night depending on the season. Espinet et al. (2003) investigate the effect on prices of the attributes of a sample of holiday hotels in Spain between 1991 and 1998. They emphasize that 4-star hotels charge prices that are 50% higher than 3-star hotels. The premium amounts to 64% compared to 1-star hotels.

Following Bull (1994) a still increasing number of papers focus on how certain locational and site-specific attributes affect room prices and may create some location specific premiums. Rigall-I-Torrent et al. (2011) establish interesting results related to the effect of the proximity of a beach on rates. Using a sample of 4,934 room prices of coastal Catalunia they show that in a hotel in front of a beach, the price is up to 17% higher. White and Mulligan (2002) study 584 budget hotels and motels in four southwestern US states and find that along with chain membership locational attributes are significant predictors of the rates.Yet they remark that an important geographic issue is still not considered, the spatial interdependency of room rates. Zhang et al. (2011) fill this gap using geographically weighted regression in a sample of 228 Beijing hotels. Evidence of spatial autocorrelation is found. Furthermore, once again the importance of star rating on the level of rates is confirmed even after controlling for spatial autocorrelation.

Balaguer and Pernías (2013) investigate the relationship between prices and the density of competitors in the area of Madrid. The results indicate that a greater density of competitors implies a lower average and less dispersion of local prices.

To sum up, according to the literature, the key determinants of hotel prices seem to be star rating, location and the intensity of competition. Implicitly, these three main factors reflect the role of quality and differentiation on hotel rates. Several recent empirical papers directly address the impact of differentiation and quality signals on prices. Bacerra et al. (2013) study the effects of vertical and horizontal differentiation on pricing policy in a large sample of Spanish hotels. Vertically differentiated hotels, with more stars, offer smaller discounts over listed prices and charge higher prices. Similarly, chain hotels, horizontally differentiated, also charge higher prices and provide smaller discounts. Specifically, when hotels are differentiated by the number of stars, the degree of local competition moderates the effect of differentiation on pricing policy.

Lee (2015) studies the relationship between quality differentiation and price competition in a sample of more than 4,250 Texas hotels. Using a modified two-stage least squares in order to test for spatial price competition, he shows that hotels compete with more distant neighbors of similar quality than those who are quality-differentiated.

In the context of the present paper, an article by Abrate et al. (2011) is of particular relevance. The authors study the impact of quality signals, specifically star rating and subscription to a quality assurance program, on a dataset of 145 hotels in Turin, Italy. The model developed is unusual since it associates a classic hedonic price equation with two probit equations in order to account for the number of stars and the subscription to quality assurance program. They find that reputation based quality signals, star ratings and quality assurance programs, are significant predictors of prices. Indeed, hotels in which quality is assured benefit from a specific price premium suggesting that some limitations exist in the traditional star rating system.

From the methodological point of view, this model lies somewhere between the classic hedonic method of current use in the literature and the original recursive semi-ordered bivariate probit approach that is developed in the present paper. The topic of our paper is related to the question of the role of quality signals on the level of hotel prices but has never been considered in the existing literature. Our point is to stress that two types of hotels exist. We can refer to the first type of hotels as insiders, hotels which enter the official star rating system and to the second type as outsiders, hotels which do not enter the official star rating system. Since the proportion of outsiders is significant, one must question the rationale of being an outsider. To put it differently, is the financial incentive to be an insider high enough to encourage hoteliers to enter an official classification system?

3 The data

The new classification system enacted by law in France in December 2009 offers us the opportunity to assess in a proper way the incentive to enter an official classification system. This voluntary system introduces two new major features: hotels are now rated from 1 to 5 stars instead of 1 to 4 stars in the previous system; and the rating is granted for 5 years instead of 25 years in the previous system.

According to *Atout France*[3], the French authority in charge of classification, the logic of the new classification system relies on the modernity and the qualitative nature of the criteria. A distinctive feature of the system is the large number of criteria considered. The total number amounts to 246. It is possible to gather these criteria into three categories: facilities, customer service and accessibility and sustainable development.

The data used in this study have been obtained from the Corsican Tourism Agency

[3] www.atout-france.fr.

website[4]. In December 2013, around 400 hotels were indexed on this website. The reference year is 2012. Due to missing data, the initial sample consists of 369 hotels, 67 unrated outsiders and 302 insiders rated from 1 to 5 stars. According to this sample the proportion of outsiders lies above 18% in Corsica. The final sample contains 357 observations after 1-star and 5-star hotels have been suppressed due to their very small number.

For each hotel, 38 attributes are available for analysis including for example the presence of a swimming pool in the hotel or the possibility of exchanging currency in house. All the available attributes are listed in Table 1 in accordance with the logic of the classification system.

Insider *versus* outsider hotels are captured through a dummy variable, $CLASS$, which takes on the value of 1 for classified hotels (insiders) and 0 for the unclassified ones (outsiders).

Data on rates consist of two prices for each hotel, an off-peak rate and a peak rate[5]. Both prices are voluntarily given by hotel managers to the Corsican Tourism Agency in order to be published on the advertising page of the hotel on the agency website. The off-peak rates are those charged at the start of the tourism season in Corsica (between March and May) whereas the peak rates are charged in mid-August. This kind of data is similar to brochure of prices used in several previous studies (Rigall-I-Torrent et al., 2011, Espinet et al., 2012).

It may appear as less accurate than daily rates accounting for discounts associated with the use of pricing techniques such as yield management. Nonetheless, in the context of this paper, prices given directly by hotel managers are of particular interest since it expresses the maximum rate that they are willing to charge and therefore reflects the highest potential rate premium derived from classification. The off-peak rate and the peak rate are ordinal

[4] www.visit-corsica.com.

[5] Rates are for a standard double room for double occupancy without breakfast.

Table 1: List of available attributes of hotels.

Equipments	Services to customers	Accessibility and others
Patio	Booking of services	Disabled access
Parking	Currency exchange	Beach
Park	Airport shuttle	Pets allowed
TV lobby	Laundry service	
Seminar room		
Lobby		
Bar		
Restaurant		
Public Phone		
Swimming pool		
Garden		
Boules pitch		
Spa		
Sauna		
Moto garage		
Cable TV		
Library		
Lift		
Double glazing		
Internet access		
Private terrace		
Air-conditioner in room		
Safe in room		
Mini-Bar		
Air-dryer		
Sauna in room		
Spa in room		
Phone in room		
TV in room		
Wifi		
Room service		

variables broken down, respectively, into six categories and five categories. Reflecting differentiated competitive conditions during off-peak and peak seasons, none of the hotels in the sample charges peak rates within one of the categories ([€50-€80], see Table 2). Given that off-peak rates exhibit more variability, as a response to more intense competition, we chose to focus our analysis on off-peak rates, keeping in mind that peak rates and off-peak rates are highly dependent (Spearman's rank correlation coefficient highly significant, equal to 0.74).

In each of the rate categories, an expensive hotel of an inferior star category (or an hotel without star rating) is likely to coexist with a relatively cheap hotel of an upper star category.

The variable $ROOMS$, the number of rooms in the hotel, is used to control for the effect of the hotel size on rates. The dataset also contains other categorical variables in order to control for location ($REGION$) and for the opening period of the hotel ($OPEN$).

$OPEN$ has four levels and hotels that are open throughout the year are chosen as the reference level. The variable $REGION$ is made of nine micro-regions defined according to the list of the Corsican Tourism Agency: Southern; Valinco; Ajaccio; West Corsica; Bastia (which is the reference micro-region); Center; East Coast; Castagniccia.

The discussion of the literature shows that some authors found evidence of the impact of the density of competitors on rates (Balaguer and Pernías, 2013). Although this issue is not at the core of the present paper, we believe that it has some relevance and has to be controlled for. Accordingly, the dummy variable DUM_NUM is defined. It takes on the value of 1 if the number of hotels in the town is above the sample mean and the value of 0 otherwise[6].To complement this dummy variable, we include an interaction variable

[6]The mean number of hotels is 16.86.

Table 2: Description of categorical variables used in the model.

Variable	Variable description	Frequency
OPR	Off-peak rate	
	Below 50€	21%
	€50 $\leq OPR <$€80	53.5%
	€80 $\leq OPR <$ €110	14.3%
	€110 $\leq OPR <$€140	5%
	€140 $\leq OPR <$€170	2.8%
	Above €170	3.4%
PR	Peak rate	
	Below €50	31.09%
	€80 $\leq OPR <$€110	26.33%
	€110 $\leq OPR <$€140	15.41%
	€140 $\leq OPR <$€170	9.80%
	Above €170	17.37%
CLASS	Hotel has a star rating (yes=1; no=0)	81.2%
REGION	Region of Corsica in which the hotel is located	
	Southern	16.5%
	Valinco	8.7%
	Ajaccio	12.6%
	West Corsica	12.9%
	Balagna	17.4%
	Bastia (reference)	20.2%
	Center	6.7%
	East Coast	3.6%
	Castagniccia	1.4%
OPEN	Lenght of opening of the hotel	
	Less than 4 months	2.2%
	4 to 6 months	13.4%
	6 to 9 months	46.8%
	The all year (reference)	37.5%
POOL	Pool in the hotel	38.1%
BEACH	Beach in the neighborhood of the hotel	9.5%
SPA	Spa in room	3.6%
CABLE	Cable TV in room	51.8%
MINIBAR	Mini-bar in room	26.3%
BAR	Bar in the hotel	80.7%
RESTAURANT	Restaurant in the hotel	56.3%
AIR	Air-conditioner in the room	64.1%
INTERNET	Internet access in the hotel	59.9%
DUM_NUM	The number of hotels in the location is above the mean	47.1%
DUM_CLASS	Classified hotel with DUM_NUM=1	41.5%
ROOMS	Number of rooms in a hotel	3.79
N	Number of hotels in the sample	357

DUM_CLASS which takes on the value of 1 if an hotel has a star rating and is established in a town with a number of hotels above the mean and the value of 0 otherwise. The underlying idea is to capture some density effects on the slope of the $CLASS$ (insiders *versus* outsiders) variable.

As our variable of interest (off-peak rates) is an ordered categorical variable, the simple ordered probit model could appear as a natural candidate for assessing the effect of the $CLASS$ variable on the off-peak rates, thus for assessing the incentive to apply for classification. However, the $CLASS$ variable, when considered as a determinant of price rates, probably raises an endogeneity issue: the decision to enter a rating scheme and the choice of a price rate are likely to be dependent and, most importantly, influenced by the same unobservable confounders. Indeed, the cost of the classification procedure, in itself, is low: about €650 for 1 to 3-star hotels and about €2,600 for 4 to 5-star hotels; thus, this cost can not be considered as a determinant of the decision to enter the classification system, nor a determinant of the price rates. Most of the costs related to a classification decision actually are investment costs, which are unobservable, but are potentially strong determinants of both the decision to enter the rating scheme and to choose a rate structure. Handling the associated endogeneity issue requires appropriate econometric methods, which we now describe in detail.

4 Handling the endogeneity issue

4.1 The recursive semi-ordered probit model

Indeed, dealing with endogeneity in non linear models is still challenging, especially when stemming from discrete (binary, multinomial, ordered categorical, count data) regressors. Usual instrumental variables (IV) approaches are likely to produce inconsistent estimates, and a current practice to handle endogeneity in non linear models consists in applying maximum likelihood (see Geraci et al., 2014; Wooldridge, 2014) to a set of equations which includes the basic equation for the independent variable under consideration and equations for the potentially endogenous independent variables. For example, when the outcome is a binary variable and the potentially endogenous variable is also a binary, consistent estimates can be derived from the estimation of a recursive bivariate probit model (see Greene, 2011). When the outcome is an ordered categorical and the potentially endogenous variable is a binary, the recursive semi-ordered probit model applies, under the assumption of normally distributed unobserved errors.

The semi-ordered probit is a special case of the bivariate ordered probit model (Greene and Hensher, 2010), that rests on the assumption that two latent variables y^*_{1i} and y^*_{2i} are determined by the following system:

$$y^*_{1i} = x'_{1i}\beta_1 + \varepsilon_{1i}$$
$$y^*_{2i} = x'_{2i}\beta_2 + \varepsilon_{2i}$$

where β_1 and β_2 are vectors of unknown parameters, x_{1i} and x_{2i} are vectors of covariates, and ε_{1i} and ε_{2i} have standard bivariate normal distribution:

$$E[\varepsilon_{1i}|x_{1i}, x_{2i}] = E[\varepsilon_{2i}|x_{1i}, x_{2i}] = 0$$
$$Var[\varepsilon_{1i}|x_{1i}, x_{2i}] = Var[\varepsilon_{2i}|x_{1i}, x_{2i}] = 1$$
$$Cov[\varepsilon_{1i}, \varepsilon_{2i}|x_{1i}, x_{2i}] = \rho.$$

We observe two ordered categorical variables y_{1i} and y_{2i} such that:

$$
y_{1i} = \begin{cases} 0 \text{ if } \mu_{-1} < y^*_{1i} \leq \mu_0 \\ 1 \text{ if } \mu_0 < y^*_{1i} \leq \mu_1 \\ . \\ . \\ . \\ J \text{ if } \mu_{J-1} < y^*_{1i} \leq \mu_J \end{cases} \qquad y_{2i} = \begin{cases} 0 \text{ if } \delta_{-1} < y^*_{2i} \leq \delta_0 \\ 1 \text{ if } \delta_0 < y^*_{2i} \leq \delta_1 \\ . \\ . \\ . \\ K \text{ if } \delta_{K-1} < y^*_{2i} \leq \delta_K \end{cases}
$$

The unknown thresholds satisfy the condition that $\mu_0 < \mu_1 < \ldots < \mu_{J-1}$ and $\delta_0 < \delta_1 < \ldots < \delta_{K-1}$. As usual, we assume $\mu_{-1} = \delta_{-1} = -\infty$ and $\mu_J = \delta_K = +\infty$.

If y_{1i} is binary ($J = 2$, $y_{1i} = 0, 1$), then the model is called the semi-ordered bivariate probit model (Greene and Hensher, 2010). The model is recursive when y_{1i}, the observed binary realization[7] of the latent variable y^*_{1i}, appears on the right-hand side of the second equation of the system under consideration:

$$
y^*_{1i} = x'_{1i}\beta_1 + \varepsilon_{1i}
$$
$$
y^*_{2i} = x'_{2i}\beta_2 + \gamma y_{1i} + \varepsilon_{2i}
$$

where γ is an unknown parameter. Testing the recursivity of the model is done by testing $\gamma = 0$, whereas testing the endogeneity of y_{1i} in the second equation amounts to testing the hypothesis that $\rho = 0$ using a likelihood ratio or Wald test. The model can be estimated by full information maximum likelihood[8].

In our case y^*_{1i} is the latent variable associated with the decision to apply or not for classification, and y^*_{2i} is the latent variable associated with the hotel rate scheme, so that y_{1i} actually is the $CLASS$ variable, and y_{2i} is the off-peak rates variable.

When relevant, various partial effects can be computed (J. Mullahy, 2011): partial effects on joint probabilities and/or partial effects on conditional probabilities. Given the

[7]When the latent realization appears on the right-hand side of the second equation, the model is called the simultaneous bivariate probit model. The simultaneous bivariate probit model is presented in detail in Sajaia (2008).

[8]Note that Donat and Marra (2015) propose a semi-parametric estimation of the model.

recursive structure of the model, some computations are quite involved.

The joint probability for $y_{1i} = j$ and $y_{2i} = k$ is (denoting the standard bivariate normal cdf as $\Phi_2(\varepsilon_{1i}, \varepsilon_{2i}, \rho)$):

$$
\begin{aligned}
\Pr(y_{1i} = j, y_{2i} = k) = \quad & \Phi_2(\mu_j - x'_{1i}\beta_1, \delta_k - x'_{2i}\beta_2 - \gamma \times j, \rho) \\
- \quad & \Phi_2(\mu_{j-1} - x'_{1i}\beta_1, \delta_k - x'_{2i}\beta_2 - \gamma \times j, \rho) \\
- \quad & \Phi_2(\mu_j - x'_{1i}\beta_1, \delta_{k-1} - x'_{2i}\beta_2 - \gamma \times j, \rho) \\
+ \quad & \Phi_2(\mu_{j-1} - x'_{1i}\beta_1, \delta_{k-1} - x'_{2i}\beta_2 - \gamma \times j, \rho)
\end{aligned}
$$

Following the notations of Greene and Hensher (2010), and dropping the observation subscript for convenience:

$A_{Uj} = \mu_j - x'_1\beta_1$, $B_{Ukj} = \delta_k - (x'_2\beta_2 + \gamma \times j)$, $A_{Lj} = \mu_{j-1} - x'_1\beta_1$, $B_{Lkj} = \delta_{k-1} - (x'_2\beta_2 + \gamma \times j)$

Using the general result from the bivariate normal probability,

$\frac{\partial \Phi_2(A,B,\rho)}{\partial A} = \phi(A)\Phi(\frac{B-\rho A}{\sqrt{(1-\rho^2)}})$

where ϕ and Φ are the pdf and cdf of a univariate standard normal variable, we can first derive the partial effects of a variable in the model on the joint probability (Greene and Hensher (2010))[9]:

$$
\frac{\partial \Pr(y_{1i}=j, y_{2i}=k)}{\partial x_1} = \begin{pmatrix} \phi(A_{Uj})\Phi(\frac{B_{Ukj}-\rho A_{Uj}}{\sqrt{(1-\rho^2)}}) - \phi(A_{Lj})\Phi(\frac{B_{Ukj}-\rho A_{Lj}}{\sqrt{(1-\rho^2)}}) \\ -\phi(A_{Uj})\Phi(\frac{B_{Lkj}-\rho A_{Uj}}{\sqrt{(1-\rho^2)}}) + \phi(A_{Lj})\Phi(\frac{B_{Lkj}-\rho A_{Lj}}{\sqrt{(1-\rho^2)}}) \end{pmatrix} (-\beta_1)
$$

$$
\frac{\partial \Pr(y_{1i}=j, y_{2i}=k)}{\partial x_2} = \begin{pmatrix} \phi(B_{Ukj})\Phi(\frac{A_{Uj}-\rho B_{Ukj}}{\sqrt{(1-\rho^2)}}) - \phi(B_{Lk})\Phi(\frac{A_{Uj}-\rho B_{Lkj}}{\sqrt{(1-\rho^2)}}) \\ -\phi(B_{Ukj})\Phi(\frac{A_{Lj}-\rho B_{Ukj}}{\sqrt{(1-\rho^2)}}) + \phi(B_{Lk})\Phi(\frac{A_{Lj}-\rho B_{Lkj}}{\sqrt{(1-\rho^2)}}) \end{pmatrix} (-\beta_2)
$$

But, the conditional probabilities may be more useful than the joint probability:

[9] As noted in Greene and Hensher (2010), if any variables appear in both equations, the effects are added.

$\Pr(y_{2i} = k \mid y_{1i} = j) = \frac{\Pr(y_{1i}=j, y_{2i}=k)}{\Pr(y_{1i}=j)}$

where the marginal univariate probability is : $\Pr(y_{1i} = j) = \Phi(A_{Uj}) - \Phi(A_{Lj})$

Straightforward derivation gives:

$$\frac{\partial(\Pr(y_{1i}=j, y_{2i}=k)/\Pr(y_{1i}=j))}{\partial x_1} = \frac{\partial \Pr(y_{1i}=j, y_{2i}=k)/\partial x_1}{\Pr(y_{1i}=j)} - \Pr(y_{2i} = k \mid y_{1i} = j)\frac{(\phi(A_{Uj})-\phi(A_{Lj}))}{\Pr(y_{1i}=j)}(-\beta_1)$$

and

$$\frac{\partial(\Pr(y_{1i}=j, y_{2i}=k)/\Pr(y_{1i}=j))}{\partial x_2} = \frac{\partial \Pr(y_{1i}=j, y_{2i}=k)/\partial x_2}{\Pr(y_{1i}=j)}$$

In our model specification, y_{1i} takes on two values $(0, 1)$. Therefore, one can easily compute the difference in conditional probabilities, *i.e.*, $Dif_{ki} = \Pr(y_{2i} = k \mid y_{1i} = 1) - \Pr(y_{2i} = k \mid y_{1i} = 0)$ to assess the effect of the binary variable y_{1i} on y_{2i}. Note that Dorat and Marra (2015), after estimating a recursive bivariate ordered model to study the effect of education on alcohol consumption of individuals in the UK, actually assess this effect by computing the conditional probabilities of an individual to consume a certain quantity of alcohol given her/his observed educational achievement. In so doing, they find that individuals with higher education have a larger probability to intake alcohol above the National Health Service recommendations, than the lesser educated ones. Computing differences in conditional probabilities extends this approach, notably in allowing the assessment of whether the differences are significant, and in allowing the calculation of partial effects on the differences. Therefore,

$$Dif_{ki} = \frac{\Pr(y_{1i}=1, y_{2i}=k)}{\Pr(y_{1i}=1)} - \frac{\Pr(y_{1i}=0, y_{2i}=k)}{\Pr(y_{1i}=0)}.$$

The average difference is obtained by averaging the individual differences over all observations.

This allows to compute the following partial effects (when relevant):

$$\frac{\partial Dif_{ki}}{\partial x_1} = \left(\begin{array}{c} \frac{\partial \Pr(y_{1i}=1, y_{2i}=k)/\partial x_1}{\Pr(y_{1i}=1)} - \Pr(y_{2i} = k \mid y_{1i} = 1) \frac{(\phi(A_{U1}) - \phi(A_{L1}))}{\Pr(y_{1i}=1)} (-\beta_1) \\ -\frac{\partial \Pr(y_{1i}=0, y_{2i}=k)/\partial x_1}{\Pr(y_{1i}=0)} + \Pr(y_{2i} = k \mid y_{1i} = 0) \frac{(\phi(A_{U0}) - \phi(A_{L0}))}{\Pr(y_{1i}=0)} (-\beta_1) \end{array} \right)$$

and

$$\frac{\partial Dif_{ki}}{\partial x_2} = \left(\begin{array}{c} \frac{\partial \Pr(y_{1i}=1, y_{2i}=k)/\partial x_2}{\Pr(y_{1i}=1)} \\ -\frac{\partial \Pr(y_{1i}=0, y_{2i}=k)/\partial x_2}{\Pr(y_{1i}=0)} \end{array} \right)$$

Notice that when $\rho = 0$, the expression of the difference in conditional probabilities simplifies: in that case, the joint probabilities factor into the products of the marginals. Thus,

$$\begin{aligned} DIf_{ki} &= \frac{\Pr(y_{1i} = 1, y_{2i} = k)}{\Pr(y_{1i} = 1)} - \frac{\Pr(y_{1i} = 0, y_{2i} = k)}{\Pr(y_{1i} = 0)} \\ &= \frac{\Pr(y_{1i} = 1) \times \Pr(y_{2i} = k \mid y_{1i} = 1)}{\Pr(y_{1i} = 1)} - \frac{\Pr(y_{1i} = 0) \times \Pr(y_{2i} = k \mid y_{1i} = 0)}{\Pr(y_{1i} = 0)} \\ &= \Pr(y_{2i} = k \mid y_{1i} = 1) - \Pr(y_{2i} = k \mid y_{1i} = 0) \end{aligned}$$

The latter is the simple marginal effect of the binary variable (y_{1i}) on the ordered variable.

4.2 Empirical results

The results of the recursive semi-ordered probit model[10] are given in Table 3. y_{1i} is the $CLASS$ variable, and y_{2i} is the off-peak rates variable, hereafter denoted as OPR.

Note that, as shown by Wilde (2000) in the case of the recursive bivariate probit model, the identification of the model is achieved even if exactly the same exogenous vector of regressors appears in both equations, *i.e.*, even if $x_{1i} = x_{2i}$ (see also Roodman (2011), p. 180). However, using an extensive Monte Carlo experiment, Monfardini and Radice (2008) also demonstrate that the availability of extra regressors (here in the vector of regressors

[10] The model is estimated using the Stata user-written command cmp (Roodman, 2011). A specific program was written for the calculation of the marginal effects, which can be adapted to other contexts by any Stata user with some programing skills. The Stata code is available from the authors upon request.

x_{2i}) helps to preserve the validity of the exogeneity tests in the presence of misspecification. Taking this practical guideline into account, we chose to include additional regressors in the $CLASS$ equation, *i.e.*, we include regressor in the $CLASS$ equation which are not in the OPR equation.

First, it should be noted that testing the exogeneity of the $CLASS$ variable in that recursive semi-ordered probit model amounts to testing whether $\rho = Cov[\varepsilon_{1i}, \varepsilon_{2i}|x_{1i}, x_{2i}]$ is significant. ρ is indeed found to be significant, at the 5 percent level ($\widehat{\rho} = -0.57$, $p - value = 0.016$), which provides clear evidence on unobserved heterogeneity affecting conjointly the $CLASS$ variable and the OPR variable; stated differently, being classified, as a predictor of the off-peak rates, is clearly an endogenous variable. Keeping that result in mind, we can now turn to an in-depth analysis of the econometric outcome.

Table 3 shows that the probability of being a classified hotel ($CLASS$) is positively influenced by four variables: namely, the number of rooms, the presence of a bar in the hotel, the availability of Internet access in the room and air conditioned in the room. These results are in line with the literature. Abrate et al. (2011) found that the star rating depends on the number of rooms, the availability of an air conditioner in room and other services such as Internet access.

From Table 3, being classified does in turn influence significantly and positively the probability of higher rates being proposed during the low season. The off-peak rates equation offers further interesting perspectives on the determinants of hotel rates in Corsica. Not surprisingly, a strong regional effect arises: compared to the reference region (Bastia), hotels from the Ajaccio region and the southern region of Corsica, which are known as attractive areas, tend to belong to the most expensive category of hotels. More precisely, according to recent data (INSEE, 2015b), the Ajaccio region alone accounts for 41% of total night stays. It suggests that the rate premium associated with Ajaccio is the result

Table 3: The recursive semi-ordered probit model.

	Coef.	SE	z	p-value
CLASS equation				
BAR	0.405	0.191	2.11	0.034
AIR	0.674	0.170	3.97	0.000
INTERNET	0.575	0.169	3.40	0.001
ROOMS	0.023	0.006	3.72	0.000
OPR rate equation				
CLASS	1.844***	0.326	5.66	0.000
Southern	0.784***	0.198	3.94	0.000
Valinco	-0.125	0.247	-0.51	0.612
Ajaccio	0.605**	0.217	2.78	0.005
West Corsica	-0.313	0.213	-1.47	0.141
Balagna	0.143	0.192	0.74	0.457
Center	0.026	0.275	0.10	0.922
East Coast	-0.546	0.349	-1.57	0.117
Castagniccia	-0.337	0.544	-0.62	0.536
Less than 4 months	-0.250	0.420	-0.60	0.551
4-6 months	0.698***	0.200	3.49	0.000
6-9 months	0.572***	0.138	4.13	0.000
POOL (yes/not)	0.277*	0.140	1.98	0.048
BEACH (yes/not)	0.761***	0.210	3.62	0.000
SPA (yes/not)	1.076***	0.321	3.35	0.001
CABLE (yes/not)	0.424**	0.145	2.91	0.004
MINIBAR (yes/not)	0.725***	0.153	4.73	0.000
DUM_NUM	0.332	0.309	1.07	0.283
DUM_CLASS	-0.682*	0.333	-2.04	0.041

* p-value ≤ 0.05

** p-value ≤ 0.01

*** p-value ≤ 0.001

of a demand effect. Services like cable TV, minibar, swimming pool, spa in the room and the proximity of a beach are also associated to the most expensive hotels. Likewise, establishments that are not open throughout the year seem to exhibit higher rates[11].

More importantly, the significance and the sign of the DUM_NUM and DUM_CLASS coefficients deserve highlighting. At first sight, DUM_NUM, which captures a density effect (coded one when the number of hotels in town is above the sample mean and 0 otherwise) could be seen as non significant. But DUM_CLASS is an interaction variable between DUM_NUM and $CLASS$. As the effect of DUM_NUM, if any, actually passes through two variables, the appropriate test to assess the effect of DUM_NUM is a joint hypothesis test of whether the coefficients of DUM_NUM and DUM_CLASS are simultaneously zero. The result ($\chi^2(2) = 6.91$) shows that the hotel density variable does influence positively the rate. But, DUM_CLASS is significant and negative, which means that if a hotel is located in an area where the number of hotels is above the mean, the probability of exhibiting a high rate is higher (DUM_NUM effect) and in the mean time the classification premium is lower (DUM_CLASS effect). The rationale behind this result is that, if the hotel is located in a high density area (with a number of hotels above the sample mean), it means that this area is attractive for tourists, and therefore classification is not necessary in order to charge high rates. The attractiveness of the area is sufficient enough to justify high rates.

Furthermore, estimating a recursive semi-ordered probit model allows us to compute joint probabilities, conditional probabilities and differences in conditional probabilities. Denoting the joint probability as $p_{ij} = P\left(OPR = i, CLASS = j\right)$ and $pcond_{ij} = P\left(OPR = i \mid CLASS = j\right)$ for $i = 0, \ldots, 5$ and $j = 0, 1$. Table 4 and 5 report joint probabilities and conditional probabilities for the first rate category (less than €50/night) and the second rate category

[11] Note that the coefficient associated with hotels which are open less than four months is not significant since this category comprises only eight cases.

(€50-€80/night) for each of the nine regions represented in the model. One striking result is that the conditional probabilities (conditional on being classified or not classified) are very similar for both rate categories within each of the regions. The computation of the differences in conditional probabilities, region by region, reinforces this result, as these differences are never found to be significant at the usual levels[12].

[12]Thus computing marginal effects, given by the previous formulae, was not relevant on our data.

Table 4: Joint, conditional probabilities and differences in conditional probabilities for the first rate category.

Probability	Bastia	Southern	Valinco	Ajaccio	West Corsica	Balagna	Castagniccia	Center	East Coast
	Coef./t	Coef./t	Coef./ t	Coef./t	Coef./t	Coef./t	Coef./ t	Coef./t	Coef./t
p_{01}	0.22***	0.06**	0.26**	0.08**	0.32***	0.18***	0.33	0.21**	0.41**
	(0.05)	(0.02)	(0.08)	(0.03)	(0.07)	(0.05)	(0.19)	(0.08)	(0.13)
p_{00}	0.03*	0.007	0.04	0.01	0.05*	0.03	0.05	0.03	0.07*
	(0.02)	(0.006)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
$pcond_{01}$	0.25***	0.07**	0.30**	0.10**	0.37***	0.21***	0.38	0.25**	0.47**
	(0.06)	(0.02)	(0.09)	(0.03)	(0.09)	(0.06)	(0.22)	(0.09)	(0.14)
$pcond_{00}$	0.26*	0.06	0.32*	0.09	0.40***	0.21	0.41	0.26	0.50*
	(0.12)	(0.05)	(0.15)	(0.07)	(0.14)	(0.11)	(0.27)	(0.15)	(0.20)
$Difference$	-0.01	0.01	-0.01	0.01	-0.03	-0.003	-0.03	-0.008	-0.04
	(0.11)	(0.05)	(0.12)	(0.06)	(0.13)	(0.10)	(0.13)	(0.114)	(0.13)

* p-value ≤ 0.05
** p-value ≤ 0.01
*** p-value ≤ 0.001

Standard errors in brackets.

Table 5: Joint, conditional probabilities and differences in conditional probabilities for the second rate category.

Probability	Bastia	Southern	Valinco	Ajaccio	West Corsica	Balagna	Castagniccia	Center	East Coast
	Coef./t	Coef./t	Coef./ t	Coef./t	Coef./t	Coef./t	Coef./t	Coef./ t	Coef./t
p_{11}	0.59***	0.59***	0.56***	0.61***	0.52***	0.61***	0.51***	0.59***	0.45***
	(0.04)	(0.04)	(0.06)	(0.04)	(0.06)	(0.03)	(0.16)	(0.05)	(0.12)
p_{10}	0.09***	0.09***	0.08***	0.09***	0.07***	0.09***	0.07*	0.09***	0.06**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)
$pcond_{11}$	0.68***	0.67***	0.65***	0.70***	0.59***	0.70***	0.58***	0.68***	0.51***
	(0.04)	(0.05)	(0.07)	(0.04)	(0.07)	(0.04)	(0.18)	(0.06)	(0.13)
$pcond_{10}$	0.69***	0.74***	0.65***	0.76***	0.58***	0.73***	0.57**	0.70***	0.48**
	(0.10)	(0.06)	(0.13)	(0.05)	(0.13)	(0.08)	(0.25)	(0.11)	(0.20)
$Difference$	-0.01	-0.06	-0.001	-0.06	-0.01	-0.03	0.01	-0.02	0.03
	(0.09)	(0.05)	(0.104)	(0.03)	(0.11)	(0.08)	(0.12)	(0.10)	(0.12)

* p-value ≤ 0.05
** p-value ≤ 0.01
*** p-value ≤ 0.001

Standard errors in brackets.

Thus, it means that, when correctly assessed, classification has no impact on the probability of belonging to the most common categories in terms of rates per night. To state it in other words, the incentive for classification is almost nil as long as the hotel is not in some sense "out of the common".

At this point, one might ask whether "the game is worth the candle"[13]: what if we had not handled the endogeneity issue, and had estimated, instead of a recursive semi-ordered probit model (which is manageable, but computationally intensive), a naive ordered probit model?

4.3 Results from a naive approach

The naive approach would have consisted in estimating a simple ordered probit model, without controlling for the endogeneity of the $CLASS$ variable as a predictor of the off-peak rates. Results are reported in Table 6, corresponding marginal effects of classification on each rate category are given in Table 7. Qualitatively, the OPR equation estimates of the simple ordered probit model are very close to the estimates of the recursive semi-ordered probit model. Services like cable TV, minibar, swimming pool, spa in the room and the proximity of a beach still influence significantly and positively the probability of charging higher rates. The pattern of the regional effects is very similar to what was obtained from the recursive semi-ordered probit model, a location in the Ajaccio region or in the South of Corsica being favorable to higher rates. However, a closer inspection of the results shows a slight difference: as before, we performed a joint hypothesis test to assess the density effect; according to the result of this test ($\chi^2(2) = 5.29$) the DUM_NUM variable is not significant at the 5% level (but is at the 10% level), whereas it was clearly significant in the recursive semi-ordered probit model.

[13] Old French and... Corsican expression.

Table 6: Results of the naive non recursive probit model.

	Coef.	SE	t	p-value
OPR				
CLASS	1.008***	0.225	4.48	0.000
Southern	0.781**	0.319	2.45	0.014
Valinco	-0.136	0.329	-0.41	0.679
Ajaccio	0.648**	0.325	1.99	0.046
West Corsica	-0.356	0.319	-1.11	0.265
Balagna	0.115	0.305	0.38	0.702
Center	-0.044	0.298	0.15	0.885
East Coast	-0.601	0.422	-1.42	0.155
Castagniccia	-0.385	0.600	-0.64	0.521
Less than 4 months	-0.154	0.460	-0.33	0.738
4-6 months	0.733*	0.212	3.46	0.001
6-9 months	0.625*	0.144	4.34	0.000
POOL (yes/not)	0.356**	0.141	2.53	0.011
BEACH (yes/not)	0.816*	0.218	3.75	0.000
SPA (yes/not)	1.110*	0.336	3.30	0.001
CABLE (yes/not)	0.560*	0.143	3.91	0.000
MINIBAR (yes/not)	0.816*	0.156	5.18	0.000
DUM_NUM	0.293	0.343	0.85	0.394
DUM_CLASS	-0.624	0.367	-1.70	0.090

* p-value ≤ 0.05

** p-value ≤ 0.01

*** p-value ≤ 0.001

Table 7: Marginal effects of classification on price.

	Coef.	SE	t	p-value
Marginal effect of CLASS				
$\Pr(OPR < 50€), OPR = 0$	-0.280	0.058	-4.83	0.000
$\Pr(50€ \leq OPR < 80\ €), OPR = 1$	0.129	0.042	3.11	0.002
$\Pr(80€ \leq OPR < 110\ €), OPR = 2$	0.086	0.023	3.74	0.000
$\Pr(110€ \leq OPR < 140\ €), OPR = 3$	0.032	0.011	2.93	0.003
$\Pr(140€ \leq OPR < 170\ €), OPR = 4$	0.017	0.007	2.42	0.015
$\Pr(OPR \geq 170\ €)$, $OPR = 5$	0.015	0.007	2.18	0.003

Also, in such a simple ordered model, one would have assessed the effect of being classified through the analysis of the effect of the $CLASS$ variable. The $CLASS$ variable is found to be highly significant, with a positive sign, which could lead to the conclusion that classification is a strong determinant of choosing high rates. Likewise, the analysis of the marginal effects of the $CLASS$ variable on the probability of choosing each category of the off-peak rate (see Table 7) would have shown that being classified acts negatively and strongly (-0.28 point) on the probability of charging a low price and acts positively on the probability of charging higher prices.

However, as we demonstrate above, the $CLASS$ variable is clearly endogenous, and controlling for endogeneity within the appropriate framework of the recursive semi-ordered probit model changes dramatically the picture: computing relevant conditional probabilities shows that classification actually has no effect on the probability of choosing one or the other category of off-peak rates.

5 Conclusion

Advocates of hotel classification systems often argue that classification provides advantages both to customers and to hoteliers. From official or informal classification systems, customers get valuable information regarding a product they are not able to test before they buy. When they decide to enter an official classification system, hoteliers expect to reap benefits in terms of credibility, product transparency, increased consumer satisfaction, to name but a few, and in terms of increased rate and margins (UNWTO, 2015).

Until now, the tourism management literature seemed to support the rate premium motive in the decision to enter a classification system. However, our work shows that this intuitive result has to be treated with caution. Voluntary official classification systems allow us to gather data from natural experiments, as hoteliers can choose, or not, to enter such systems. Insiders and outsiders coexist, which makes it possible to test the rate premium hypothesis. Using a sample of hotels of Corsica, a naive approach, linking rate categories to the fact of being classified, would have wrongly supported the rate premium intuition. Addressing appropriately the endogeneity issue of the classification variable (being insider or not) within the framework of a recursive semi-ordered probit model, leads to the exact opposite conclusion: on our data, classification does not provide any rate premium.

We also fully describe the calculation of conditional probabilities and partial effects on conditional probabilities within the recursive semi-ordered probit model, which, we believe, will be valuable to a large audience.

References

Abrate, G., Capriello, A. and Fraquelli, G.: 2011, When quality signals talk: Evidence from the turin hotel industry, *Tourism Management* **32**, 912–921.

Bacerra, M., Santaló, J. and Silva, R.: 2013, Being better vs. being different: Differentiation, competition, and pricing strategies in the spanish hotel industry, *Tourism Management* **34**, 71–79.

Balaguer, J. and Pernías, J. C.: 2013, Relationship between spatial agglomeration and hotel prices. evidence from business and tourism consumers, *Tourism Management* **36**, 391–400.

Bull, A. O.: 1994, Pricing a motel's location, *International Journal of Contemporary Hospitality Management* **6**, 10–15.

Carvell, S. A. and Herrin, W. E.: 1990, Pricing in the hospitality industry: An implicit market approach, *FIU Hospitality Review* **8**(2), 27–37.

Danziger, S., Israeli, A. and Bekerman, M.: 2006, The relative role of strategic assets in determining customer perceptions of hotel room price, *International Journal of Hospitality Management* **25**, 129–145.

Donnat, F. and Marrat, G.: 2015, Semi-parametric bivariate polychotomous ordinal regression, *Statistics and Computing* **27**(1), 283–299.

Espinet, J. M., Fluvià, M., Rigall-I-Torrent, R. and Saló, A.: 2012, Hotel characteristics and seasonality in prices: an analysis using spanish tour operators' brochures, *Tourism Economics* **18**(4), 749–767.

Espinet, J. M., Saez, M., Coenders, G. and Fluvià, M.: 2003, Effect on prices of the attributes of holiday hotels: a hedonic prices approach, *Tourism Economics* **9**(2).

Geraci, A., Fabbri, D. and Monfardini, C.: 2014, Testing exogeneity of multinomial regressors in count data models: does two stage residual inclusion work?, Working Paper DSE n° 921.

Goldberg, S. M., Green, P. E. and Wind, Y.: 1984, Conjoint analysis of price premiums for hotel amenities., *The Journal of Business* **57**(1), 111–132.

Greene, W. H.: 2011, *Econometric Analysis*, seventh edition edn, Prentice Hall.

Greene, W. and Hensher, D.: 2011, *Modeling Ordered Choices*, Cambridge University Press.

Heo, C. Y. and Hyun, S. S.: 2015, Do luxury room amenities affect guests' willingness to pay?, *International Journal of Hospitality Management* **46**, 161–168.

INSEE: 2015a, Dossier Corse n °2.

INSEE: 2015b, Dossier Corse n °1.

INSEE: 2015c, Esane, INSEE database.

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**, 405–424.

Lee, S. K.: 2015, Quality differentiation and conditional spatial price competition among hotels, *Tourism Management* **46**, 114–122.

Masiero, L., Heo, C. Y. and Pan, B.: 2015, Determining guests' willingness to pay for hotel room attributes with a discrete choice model, *International Journal of Hospitality Management* **49**, 117–124.

Monfardini, C. and Radice, R.: 2008, Testing exogeneity in the bivariate probit model: A monte carlo study, *Oxford Bulletin of Economics and Statistics* **70**, 270–282.

Rigall-I-Torrent, R., Fluvià, M., Ballester, R., Saló, A., Arizac, E. and Espinet, J. M.: 2011, The effects of beach characteristics and location with respect to hotel prices, *Tourism Management* **32**, 1150–1158.

Roodman, D.: 2011, Estimating fully observed recursive mixed-process models with cmp, *Stata Journal* **11**(2), 159–206.

Rosen, S.: 1974, Hedonic prices and implicit markets: Product differentiation in pure competition, *The Journal of Political Economy* **82**(1), 34–55.

UNWTO: 2015, Hotel classification systems: Recurrence of criteria in 4 and 5 stars hotels, World Tourism Organisation, Madrid.

Vine, P.: 1981, Hotel classification - art or science?, *International Journal of Tourism Management* **2**(1), 18–29.

White, P. J. and Mulligan, G. F.: 2002, Hedonic estimates of lodging rates in the four corners region, *The Professional Geographer* **54**(4), 533–543.

Wilde, J.: 2000, Identification of multiple equation probit models with endogenous dummy regressors, *Economics Letters* **69**, 309–312.

Wooldridge, J. M.: 2014, Quasi-maximum likelihood estimation and testing for nonlinear models with endogenous explanatory variables, *Journal of Econometrics* **182**, 226–234.

Wu, L.: 1999, The pricing of a brand name product: Franchising in the motel services industry, *Journal of Business Venturing* **14**(87-102).

Zhang, H., Zhanga, J., Lua, S., Chengb, S. and Zhanga, J.: 2011, Modeling hotel room price with geographically weighted regression, *International Journal of Hospitality Management* **30**.