Entropy of Tourism: the unseen side of tourism accommodation

Miguel Ángel Ruiz Reina[1]

[1]University of Málaga (Spain). Economics Applied (Statistics and Econometrics) Department

*Faculty of Economics and Business

PhD Program in Economics and Business

Edificio de Servicios Múltiples (Office B201)

Address: s/n, Plaza del Ejido, 29013 Málaga

Abstract.

The measurement of random situations is a relevant fact within the field of Econometrics. In the usual practice of the field of Econometrics assumptions are made about the statistical distributions of the data. In this paper, a concept from Physics is introduced, specifically the use of Entropy as an explanatory factor in the decision-making of tourist accommodation in apartments. The emergence of a concept called Sharing Economy has made the housing market to evolve. Assuming a direct competition against traditional accommodation such as hotels. For this, the Entropy time series have been calculated and defined for the periods between January 2005 and August 2018, to forecast and compare them with a time horizon of 12 months with the most usual predictive models. To carry out the forecast comparisons, we work with the new Matrix U1 Theil which allows quantifying the gain of the use of the Entropy models. The main locations in which the study has been applied are Spain, Catalonia and Barcelona. The theoretical implications and economic consequences are relevant for stakeholders in decision-making. The study details the relations of exchange in the decision making of accommodations in tourist apartments and details the relationship of exchange in situations of uncertainty with a high explanatory capacity of the models.

Keywords: Forecasting, Tourism Demand, Behavioural Economics, Entropy, Randomness.

1. Introduction

Decision-making by economic agents can be influenced by psychological, cognitive, emotional, economic or cultural factors. The modelling of decision-making is a complex problem due to the high randomness of socials agents. Randomization before the social science revolution is a fact as in fields such as Economics, Finances or Tourism. It is common to have a lack of knowledge of the behaviour of economic agents previously.

The modelling of the randomness that occurs in the social sciences, particularly in the field of Economics, allows defining the behaviour of economic agents and markets implications. This modelling allows stakeholders to adjust market offers, reduce costs, make consumption estimates and ultimately, make decisions in efficiency. The sharing economy is a revealing that fact that consumer preferences have changed since the global economic crisis of 2008 (Ramos et al, 2019)

Regarding to this paper, we will define a concept from Physics and that will allow us to define uncertainty decisions in different time of scenarios (Golan, 2008). For this, we will define the concept of Information Theory named Shannon Entropy (Shannon, 1948). This concept will allow us to order information, and reduce uncertainty about the behaviour of economic agents. The application of econometric models with the use of Entropy implies a gain in the explanatory and predictive capacity of the forecasting models.

According to the definition we will make, Entropy is related to uncertainty and situations with high Entropy values will be associated with less information provided by economic agents. Situations with low Entropy will be associated with high information provided by economic agents (Delgado-Bonal & Marshak, 2019).

Regarding the review of the literature, we will focus on how economic agents make their tourist accommodation decisions. Specifically, we are interested in modelling the behaviour of the market in the decision-making of tourist accommodation in apartments versus hotel accommodation in Spain, Catalonia and Barcelona. According to the extensive literature reviewed, since 2008 a concept called Sharing Economy has emerged and in regards to tourist accommodation, the best known is the Airbnb application (Bakker & Twining-Ward, 2018). The prices of the apartments (Portolan, 2013), the quality (Raya & Sastre, 2011), the benefits of accommodation in apartments or even the effect of images on the internet have been studied (Ert et al, 2016).

But, as far as the authors' knowledge reaches, no uncertainty measures have been provided in the lodging decision-making. The innovation of this paper is the application of Shannon Entropy as a measure of uncertainty and this measure will be compared with traditional predictive models such as SARIMA (Box et al, 2008). As a dimensionless measure of the predictive capacity, we will use the novel Matrix U1 Theil proposed by Ruiz-Reina (2019).

For the study of this paper, previous data from official Spanish sources (INE[1]) on Apartments Overnight Stays (AOS) and Hotel Tourist Accommodation (HTA) for the areas of Spain, Catalonia and Barcelona have been collected. The sample period studied corresponds between January 2005 and August 2018. For the forecasting evaluation of tourist accommodation, an out-sample period of 12 months has been selected.

The paper is organized as follows: section 2, describes the methodological framework of econometric modelling for AOS demand; in section 3, data analysis and empirical results are obtained for Spain, Catalonia and Barcelona; in section 4, it is possible to find the main

---

[1] INE: Instituto Nacional de Estadística (España). www.ine.es

conclusions and economics implications of this study. Finally, there are two sections for the bibliographical references used and appendix section, respectively.

2. Methodology

In this section we will focus on the definition of Shannon Entropy and the application within an econometric model. To make comparisons according to the Matrix U1 Theil methodology we will use SARIMA models. This methodological framework will allow us to define situations of uncertainty in the tourist accommodation market and in the section of conclusions the main economic relations obtained from this analysis will be discussed.

2.1 Cobb-Douglas model with uncertainty factor

Assuming that all types of accommodation provide the same utility and that the information is not asymmetries of information in the tourists´ decision-making process, we will begin by defining the concept of Shannon Entropy for a temporary period:

$$H(X_t) = -\sum_{j=1}^{n} p(x_{jt}) \log_2 p(x_{jt}) \qquad (1)$$

Thus, we can define situations of greater uncertainty for probability values close to 0.5 (Figure 1), and for two events (probability of HTA and probability of AOS) it would be represented as follows:

Figure 1: Maximum Entropy represented for two system of information. j=1,2.
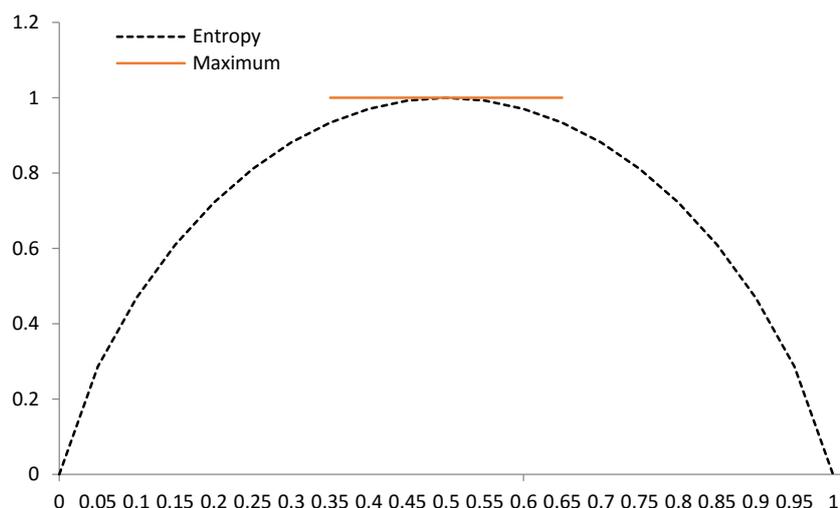
Source: Own ellaboration.

The temporal arrangement will facilitate the time series of uncertainty. Values of Entropy close to one will indicate a lot of uncertainty and those near to zero will indicate little uncertainty. According to Entropies' definition, situations with greater certainty means that economic agents define their decisions perfectly.

For the definition of the econometric uncertainty model, we will define a demand for apartment accommodation type Cobb-Douglas (Wooldridge, 2012). The Cobb–Douglas form is remarkably robust in a vast variety of applications and we are defining as follows (Brown, 2008):

$$y_t = \beta_1 x_t^{\beta_2} H_t^{\alpha_1} e^{u_t} \tag{2}$$

Where y_t represents the endogenous variable (AOS); x_t represents the explanatory variable (HTA); H_t represents the Entropy factor or uncertainty measure in decision-making process; u_t represents exogenous and uncontrollable effects in the modelling; $\beta_1, \beta_2, \alpha_1$ are the parameters that should be estimated.

It is necessary to note in the theoretical specification for problem optimization of the demand for tourist accommodation. We will assume that the rent and prices are given, focusing only on the exchange relationship between accommodation in AOS and HTA.

With the linearization of the function by applying natural logarithms we can estimate the demand for tourist apartments with the uncertainty factor. The method for estimating the model parameters is ARMA Maximum Likelihood (OPG - BHHH) available in the version of the econometric software Eviews 10.

2.2   Forecasting accuracy

As a dimensionless measure of forecasting accuracy, we will use the Matrix U1 Theil proposed by Ruiz-Reina (2019). The matrix allows us to know which model predicts closest to the real values of the variable and quantifies the gain of one model against another one in terms of the best predictor model.

Values close to zero indicate that the model in the row has a lower Theil inequality index (Theil, 1958) than the one in the column. Values equal to one will have the same predictive capacity and values greater than one will indicate that the value found in the row has a better forecasting capacity than that found in the column.

3. Data and Empirical Results

For the development of the empirical results section, data from official sources of the National Statistical Office of Spain (INE) have been collected. The two variables (AOS and HTA) under study are for the period from January 2005 to August 2019. According to Spanish official data, these types of accommodation represent approximately 90% of the demand for tourist accommodation in Spain for that period.

As can be seen in the Figure 2 both series show a similar systematic behaviour characterized by similar trends and seasonal behaviours.

Figure 2: Hotel (HTA) and Apartment (AOS) demand in Spain. Jan. 2005- August. 2019.
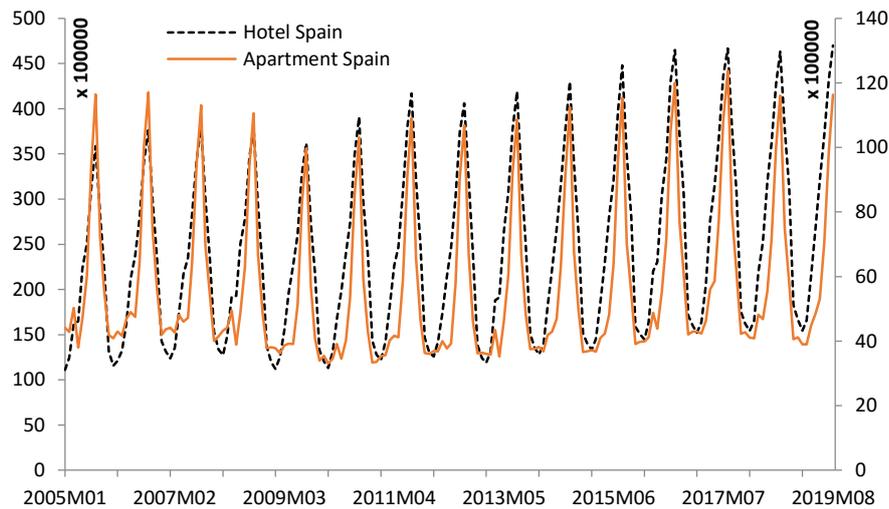
Source: Own ellaboration.

As for the descriptive statistics, it should be noted that the maximums of the three series analysed (Spain, Catalonia and Barcelona) are for the periods of August of all the years. The minimums occur approximately between the periods of November to March of each year. Regarding the coefficient of variation (CV) we can say that the series show values between 0.39 and 0.53, with the exception of the series of tourist apartments in Catalonia showing a value of 0.94. These results in a representative decrease of the average, mainly due to the seasonal peaks detected.

Table 1: Descriptive statistics of Hotel (HTA) and Apartment (AOS). Jan. 2005- August. 2019.

Hotel demand	Mean	Median	Min	Max	Stand. Desv.	CV
Spain	24,315,574	22,859,238	11,088,055	46,998,612	9,852,892	0.41
Catalonia	4,061,645	3,880,689	1,222,648	8,994,168	2,165,811	0.53
Barcelona	2,156,634	2,060,275	8,614,56	4,049,501	844,824	0.39
Apartment demand	Mean	Median	Min	Max	Stand. Desv.	CV
Spain	5,706,667	4,688,535	3,302,242	12,373,783	2,403,836	0.42
Catalonia	518,299	292,694	112,574	1,835,570	485,564	0.94
Barcelona	43,603	39,124	14,670	86,734	19,414	0.45

Source: Own ellaboration.

With regard to Figure 3, Entropy shows certain characteristics: on the one hand, there are increases in Entropy when tourist demand peaks are greater. This has a direct relationship with the increase in accommodation in tourist apartments due to the seasonality detected (it will be demonstrated with econometric models); on the other hand, since 2005 Entropy has shown a slight decreasing trend.

Figure 3: Entropy in Spain since Jan. 2005 to August. 2019.
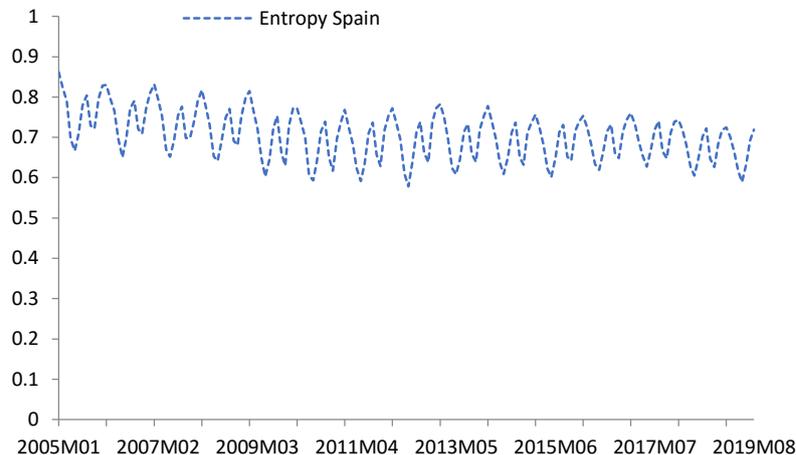
Source: Own ellaboration.

The results of the log-linearized models are those presented in the following Table 2. The econometric models estimated with the use of Entropy as an explanatory factor have three common characteristics for the three areas analysed (Spain, Catalonia and Barcelona): first, the R-squared are close to one; second, the elasticities between the demand for accommodation in HTA and AOS are approximately unitary; third, Entropy is positively related to the demand dependent variables of tourist apartments. In the case of Spain's total, it has the highest value, that is, an increase of 1% in Entropy implies an increase of 2.16% in demand in tourist apartments (in the case of Catalonia its elasticity of Entropy is 1.55 and in Barcelona it is 1.28).

In addition, from an econometric point of view, we can guarantee the statistical inference and the efficiency of the estimated parameters with the contrast that in the residuals there is no serial correlation with at least a 95 percent confidence level (Ljung & Box, 1978). Where $\hat{e}_t \sim SARMA$ that guarantee that the residues are white noise.

Table 2: Econometrics models with Entropy as explanatory variables[2]. Sample period Jan. 2005- Aug. 2018.

	log(Spain)	log(Cat)	log(Barc)
Variable	Coeffi.	Coeffi.	Coeffi.
C	-0.45***	-0.95***	-1.36***
LOG(hotel)	0.98***	0.99***	0.99***
LOG(entropy)	2.16***	1.55***	1.28***
Residual	SARMA $(0,1)(1,1)_{12}$	SARMA $(2,0)(1,0)_{12}$	SARMA $(2,0)(1,0)_{12}$
R-Squared	0.9998	0.9999	0.9999
p-value Q-Stat (up to 12)	0.173	0.391	0.236

Source: Own ellaboration.

Once the training period of the models was used between January 2005 to August 2018, the predictions for the locations in Spain, Catalonia and Barcelona have been made (Figure 4). The time horizon used is h = 12, through the Matrix U1 Theil (Table 3) we can guarantee that in all cases the predictions have been clearly higher in the Entropy models than in the SARIMA models. The graphs of the predictions for the areas of Catalonia and Barcelona can be seen in the appendix section (Figures A1 and A2).

Figure 4: Forecasting methods of AOS (Spain). Out-sample period Sept. 2018- Aug. 2019.
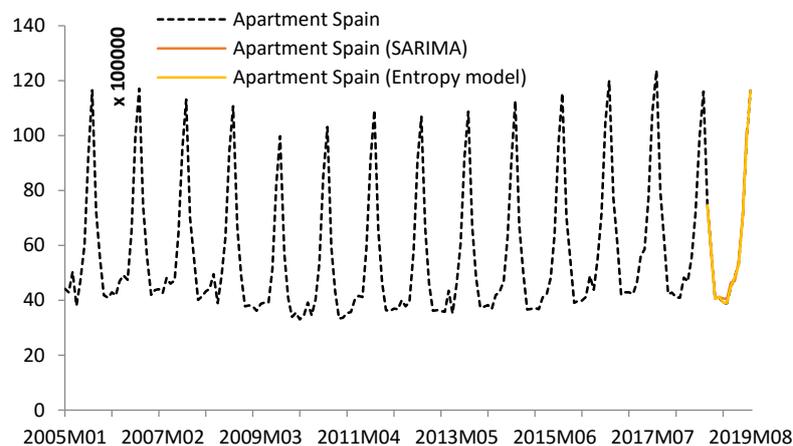
Source: Own ellaboration.

In Table 3 we can observe the gains of studies with Entropy models. In the case of Spain, the Entropy models represent 7.26 times better predictive capacity than the SARIMA model for a time horizon of 12 months (for Catalonia and Barcelona, they are 6.84 and 149.84, respectively).

---

[2] *** parameters are significant and different from 0 with the 99% of confidence.

Table 3: Forecasting Evaluation: Matrix U1 Theil for Spain, Catalonia and Barcelona. Horizon h=12.

Spain h= 12	SARIMA $(0,1,1)(0,1,1)_{12}$	Entropy models
SARIMA $(0,1,1)(0,1,1)_{12}$	1,00	7,26
Entropy models	0,14	1,00
Catalonia h= 12	SARIMA $(1,0,0)(0,1,1)_{12}$	Entropy models
SARIMA $(1,0,0)(0,1,1)_{12}$	1,00	6,84
Entropy models	0,15	1,00
Barcelona h= 12	SARIMA $(0,1,1)(0,1,1)_{12}$	Entropy models
SARIMA $(0,1,1)(0,1,1)_{12}$	1,00	149,84
Entropy models	0,01	1,00

Source: Own ellaboration.

4. Conclusions and Economics implications

The proliferation of new technologies, and the greater accessibility to the information of the economic agents in the decision-making are modifying the tourist markets. Traditionally, statistical models involve a distribution of data, which mathematically implies limitations due to their theoretical restrictions. With the application of Shannon Entropy, these statistical assumptions are omitted, introducing the concept of order of information systems. In our work, we have related two information systems in tourism demand (AOS and HTA).

The concept of Entropy and its application in econometric models has been proposed while paying special attention to the interpretation in decision-making. The measurement of uncertainty situations through Entropy is a contribution to economic, social and institutional relations. This measure of uncertainty in the tourism market is what we can name "the unseen side of tourism accommodation", since it is an initially unobservable situation but with great information for those interested in the tourism market. For this, predictive causal models have been modelled for a training period from January 2005 to August 2018 and in a out-sample period of the 12 months.

From an econometric point of view, we can highlight several common points:

- Entropy models have a high explanatory capacity with R-squared close to one.

- The elasticities obtained are close to one, so increases of 1 percent of hotel demand imply a 1 percent increase in the demand for ceteris paribus tourist accommodation and Entropy factor.

- In all cases, the coefficients associated with Entropy factor are statistically significant with positive sign. Therefore, given increases in uncertainty, there is an increase in demand in tourist accommodation.

- The relationship between information and probability in a binary decision-making, more information implies more probability of one of the two systems.

21

- Based on the Matrix U1 Theil, we can guarantee that the predictive capacity of the Entropy model is better than SARIMA models.

From an economic point of view, Tourism development is related to a welfare state and economic growth. There are positions in the scientific literature of no direct relationship between economic growth and Tourism. In the case of Spain, there is undoubtedly a direct relationship in recent years, proof of this is the greatest relative weight in the Gross Domestic Product. Tourism development is related to a welfare state and economic growth. Of course, the analysis of tourist accommodation is a key factor in tourism development and the development of alternative markets to the hotel is a fact (apartments). The proliferation of tourist accommodation in apartment can be highlight:

- The rise of the secondary market of tourist accommodation is verified with the appearance of these markets (Apartments).

- The collaborative economy creates direct jobs related to the maintenance and management of these types of accommodation.

- With the knowledge of market fluctuations economic agents can make adjustments to forecast costs, labour market, prices or any other type related.

- Knowledge of the client profile is essential for economies secondary to the tourist accommodation industry such as transport, restaurants, leisure or culture.

- A regulation of the sector may imply higher tax revenues for public entities.

We can conclude that this paper through a novel quantitative approach from Physical Econometrics, it is a contribution to the science of economics methodologically, allowing better adjustments in the decisions of economic agents. This type of analysis can be extended to different markets or regions to provide knowledge in efficient decision-making.

Future lines of research, in terms of location, the areas of study could be expanded; studies could be done depending on the country of origin of the tourist accommodation consumers. From a technical point of view, the Matrix U1 Theil allows comparisons to be made with as many models as desired.

References

[1] Bakker, M., & Twining-Ward, L. (2018). Tourism and the Sharing Economy : Policy & Potential of Sustainable Peer-to-Peer Accommodation (English). *World Bank Group 130054, 1*(1).

[2] Box, G. E., Jenkins, G. M., & Reinsel, G. C. (2008). *Time Series Analysis, Forecasting and Control.* United States of America: Wiley.

[3] Brown, M. (2008). *The New Palgrave Dictionary of Economics.* Springer.

[4] Delgado-Bonal, A., & Marshak, A. (2019). Approximate Entropy and Sample Entropy: A Comprehensive Tutorial. *Entropy, 21*(6), 1-37. doi:10.3390/exx010005

[5] Ert, E., Fleischer, A., & Magen, N. (2016). Trust and reputation in the sharing economy: The role of personal photos in Airbnb. *Tourism Management*, 62-73. doi: http://dx.doi.org/10.1016/j.tourman.2016.01.013

[6] Golan, A. (2008). Information and Entropy Econometrics — A Review and Synthesis. *Foundations and Trends in Econometrics, 2*, 1-145. doi:10.1561/0800000004

[7] Ljung, G., & Box, G. (1978). On a Measure of a Lack of Fit in Time Series Models. *Biometrika, 65*(2), 297-303. doi: https://doi.org/10.1093/biomet/65.2.297

[8] Portolan, A. (2013). The impacts of private accomodation attributes on tourism demand. *DIEM : Dubrovnik International Economic Meeting, 1*(1).

[9] Ramos, I., Fitó, M., Lladós-Masllorens, J., & Liébana-Cabanillas, F. (2019). *Sharing Economy and the Impact of Collaborative Consumption.* IGI Global. doi:10.4018/978-1-5225-9928-9

[10] Raya, J., & Sastre, F. (2011). Pricing the time and location of a stay at a hotel or apartment. Tourism Economics, 17(2), 321-338. doi:10.5367/te.2011.0044

[11] Ruiz-Reina, M. (2019). Big Data: Does it really improve Forecasting techniques for Tourism Demand in Spain? *Conference: International Conference on Time Series and Forecasting. ITISE 2019. 1*, pp. 694-706. Granada (Spain): Godel Impresiones Digitales S.L. Retrieved from http://itise.ugr.es/ITISE2019_vol1.pdf

[12] Shannon, C. (1948). A Mathematical Theory of Communication. *Bell Labs Technical Journal*, 379-423. doi: https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

[13] Theil, H. (1958). *Economic Forecasts and Policy.*

[14] Wooldridge, J. (2012). *Introductory Econometrics: A modern approach.* South-Western.

BMECONF
business, management and economics

**International Conference on Applied Research
in BUSINESS, MANAGEMENT and ECONOMICS**

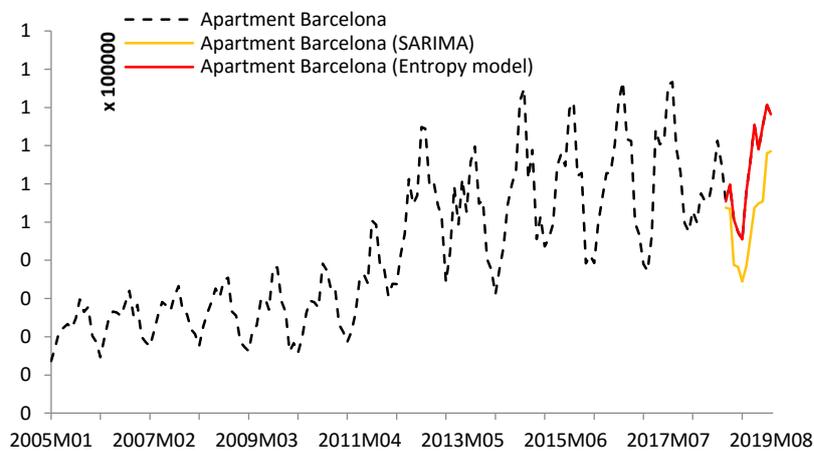**12-14 December, 2019**
Barcelona, Spain

5. Appendix

Figure A1: Forecasting methods of AOS (Catalonia). Out-sample period Sept. 2018- Aug. 2019.
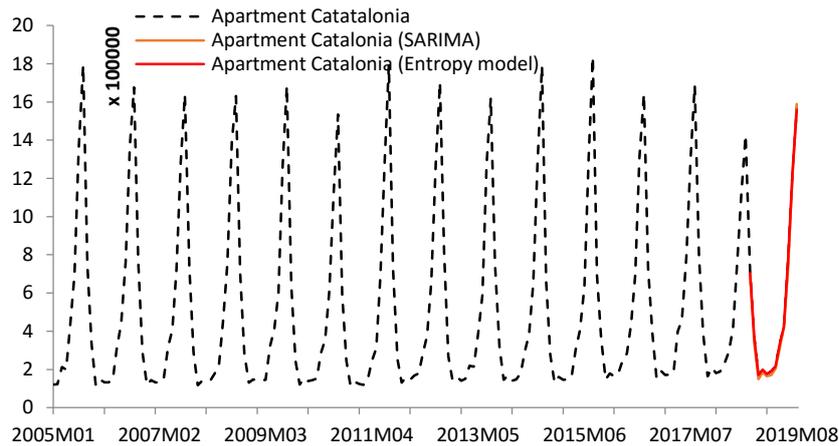
Source: Own ellaboration.

Figure A2 Forecasting methods of AOS (Catalonia). Out-sample period Sept. 2018- Aug. 2019.

Source: Own ellaboration.