The influence of Airbnb on hotel occupancy in Mexico: a Big Data Analysis (2007-2018)

La influencia de Airbnb en la ocupación hotelera de México: un análisis del Big Data (2007-2018)

Mario Gómez Aguirre[1]
Narciso Salvador Tinoco Guerrero[2]
Luis Manuel Tinoco Guerrero[3]

Recibido: 20 de marzo de 2021 Aceptado: 18 de junio de 2021
DOI: https://doi.org/10.33110/cimexus160101

Abstract

The main objective of this paper is to analyze the influence that the usage of the Airbnb's platform has had on hotel occupancy in Mexico during 2007-2018 period. The Hotel Classification System is considered to know if there are differences in this influence, according to hotels' category. To obtain the information from Airbnb, an application was created that extracted the public information of each lodging published on the website. Results were estimated by using the panel data econometric methodology, showing that the only negative impact the usage of Airbnb has on hotel occupancy is in 4-star hotels, and that an increase in the price of Airbnb's lodgings produces a rise in hotel occupancy. In other hotel categories there is no negative effect. An implication is that the usage of platforms like the one studied can be moderately regulated in Mexico.

Keywords: Sharing Economies, Airbnb, Big Data, Econometric Analysis, Panel Data.

Classification Jel: L81, L83, M21, O33.

1 Economic and Business Research Institute at Universidad Michoacana de San Nicolás de Hidalgo. E-mail: mgomez@umich.mx orcid id: 0000-0002-1178-108X

2 Economic and Business Research Institute at Universidad Michoacana de San Nicolás de Hidalgo. E-mail: narciso.tinoco@umich.mx orcid id: 0000-0003-1209-1184

3 Physical-Mathematical Sciences School at Universidad Michoacana de San Nicolás de Hidalgo. E-mail: luismanuel@thesoftwareartisans.com orcid id: 0000-0003-3028-641X

Resumen

El objetivo principal de este trabajo es analizar la influencia que ha tenido el uso de la plataforma Airbnb en la ocupación hotelera en México durante el período 2007-2018. Se considera que el Sistema de Clasificación de Hoteles sabe si existen diferencias en esta influencia, según la categoría de los hoteles. Para obtener la información de Airbnb, se creó una aplicación que extraía la información pública de cada alojamiento publicado en el sitio web. Los resultados se estimaron utilizando la metodología econométrica de datos de panel, que muestra que el único impacto negativo que tiene el uso de Airbnb en la ocupación hotelera es en los hoteles de 4 estrellas, y que un aumento en el precio de los alojamientos de Airbnb produce un aumento en la ocupación hotelera. En otras categorías de hoteles no hay efecto negativo. Una implicación es que el uso de plataformas como la estudiada puede estar moderadamente regulada en México.

Palabras clave: Airbnb, Big Data, Análisis econométrico, Datos de panel.

Clasificación Jel: L81, L83, M21, O33.

Introduction

Sharing Economies (SE) are a new business model that has been booming in recent years using digital platforms or mobile applications from which the owners of a resource encounter people who are interested in having access to it without the need to owning it. Although, ownership of goods remains of great importance to consumers in developing countries, research shows, that in most saturated markets, consumers tend to realize the disadvantage of ownership (Terporten, Bialdyga & Planing, 2012). SE by themselves could suppose an increase in the costs of the services, however new digital platforms act and change the basic conditions for the substitution since it is easy to have the right place or the suitable means of transport, contact, and contract to get access to services (Henten & Windekilde, 2015).

It is important to emphasize that information technologies, although they contribute to the development of the SE, also favor the progress of business models that are not based on sharing, for this reason it is necessary to make a distinction between both types of models. The main difference, explained by Sundararajan (2013) is that in the first case, the companies use their own assets to offer their services, while in the second case, the digital platforms use the properties of third parties to carry out their transactions. In the situation of hotel industry, there are websites that are hotels' property, which helps people easily book an accommodation online, however, they cannot be considered SE since the places to stay still belong to the big hotel chains. Airbnb

takes advantage of lodgings that other people have and want to share at low cost with users that are looking for a place to stay.

Specifically, on Airbnb, people can find a place to stay temporarily or semi-temporarily, usually in vacant rooms in another person's house, apartments, or complete houses. This implies that many tourists now have the option to leave out the traditional lodging services and choose a different space according to their needs.

According to the World Travel & Tourism Council (WTTC) In Mexico, during 2019 the total contribution of tourism to economic activity was 195,700 million dollars (USD) which represents 15.5% of the total economy, contributing nearly 7,232,900 direct and indirect jobs (WTTC, 2020). Due to this, the next question is asked: What has been the influence that the usage of Airbnb has had on hotel occupancy in Mexico? The results of this work can be used to define policies for the regulation of this kind of platforms and the design of strategies by hotels to keep customers and continue to compete in the accommodation market. So far there are no empirical studies for the case of Mexico that shows evidence of the relationship between the variables. This paper is structured as follows: after the introduction the background is presented, which contains the research hypotheses. The following section explains the source of the data and the methodology, then the results, finishing with the conclusions and references.

Background

According to Jiménez (2008), since the beginning of the 1970s, certain points were made in relation to the economic and political link of transnational companies' growth, including hotel companies. There is a close and bidirectional relationship between the reason to be, of multinational companies and the nature of globalization, so that multinational companies have adapted, in some cases, to the changes of globalization, and in other cases, they have introduced and led the transformations of it (Torres, 2011). In tourism, transnational companies participate, among others, in the hotel sector. According to Propin, López and Sánchez (2004), a hotel consortium is the affiliation of owners of independent hotel units that voluntarily associate to seek benefits derived from the access to more important resources such as financing or propaganda (Martorell, 2002).

Generally, hotel chains have their own website where tourists can consult information and make reservations to contract accommodation within them, however, there are third party websites that can facilitate access to this information such as Airbnb, Booking, Marriott, Hotels.com, Trivago and others.

Airbnb is a platform that has been studied empirically from different points of view: there are studies that analyses the reasons why people use Airbnb (Guttentag, 2016), reasons for discrimination within the platform

(Edelmar & Luca, 2014), importance and impact of the rating system (Ert, Fleischer & Magen, 2015; Zervas, Proserpio & Byers, 2017; ke, 2017), the effect of Airbnb on the price of housing and rents (Barron, Kung & Proserpio, 2018), Airbnb comparisons with traditional reservation and accommodation services (Gyódi, 2017), Airbnb's impact on hotel industry (Zervas, Proserpio & Byers 2017; Coyle & Yu-Cheong, 2016; Fissha & Shrestha, 2017) and some recommend the regulation and taxation of this type of service (Nguyen, 2014; Barron, Kung & Proserpio, 2018).

The main reasons why people choose to use and stay at Airbnb's places are mainly that these accommodations are lower priced, have amenities that hotels generally do not provide and are in different locations, so guests can choose a space close to a place of interest (Guttentag, 2016). Gyódi (2017) made a comparative analysis between the services offered by Booking.com and those of Airbnb in Warsaw, discovering, that in general, prices are more accessible in the spaces provided by Airbnb than those listed in Booking.com, due to this, the first hypothesis of this investigation is:

Hypothesis 1: The prices of Airbnb's lodging services have had a negative effect on hotel occupancy in Mexico.

It can be thought that the price of the Airbnb's lodging services is related to the amenities it offers, its location, and its ratings in the reputation system. However, in general the ratings obtained on Airbnb are much higher than those obtained in other platforms such as TripAdvisor (Zervas, Proserpio & Byers, 2017). What most affects the price is the realism and quality of the photos where the facilities can be appreciated, and the final decision of guests, comparing similar accommodations, can also be affected by personal photos of the hosts (Ert, Fleischer & Magen, 2015).

Airbnb's usage can be expected to have a positive effect on the price of housing in towns and cities because owners have the option to use the capacities of their properties to obtain additional income, which at the same time causes the price of rents to increase as they go from having long-term to short-term rents (Barron, Kung & Proserpio, 2018).

According to Fang (2015), the entry of sharing economies benefits the tourism industry since it generates new jobs, more tourists travel, incentivized by the lower cost, and stay more nights, however, as low-category hotels are being replaced by Airbnb (Zervas, Proserpio & Byers, 2017) the effect diminishes as the size of sharing economies increases and low-category hotel employees may lose their jobs.

The effect that usage of Airbnb has in hotel industry has been investigated in different times and cities of the world: Coyle and Yu-Cheong (2016) analyzed the effect that the platform has on the hotel industry of fourteen European cities, showing that the usage of Airbnb is not as negative as can be thought because there is only a negative effect on low-category hotels since Airbnb could replace its services.

Similarly, in Helsinki the website is not considered to be a direct competition for hotels, because the type of accommodation is very different (Fissha & Shrestha, 2017). In Texas most affected hotels by the rapid growth of the application, are as well, those that have low categories because Airbnb can compete on prices with them, in addition that lodgings published in the site may offer different features than low-category hotels cannot (Zervas, Proserpio & Byers, 2017); due to this, the following research hypotheses are:

Hypothesis 2: The usage of Airbnb has not had a negative effect on the occupancy rate of high-category hotels in Mexico.

Hypothesis 3: The usage of Airbnb has had a negative effect on the occupancy rate of low-category hotels in Mexico.

Data and Methodology

To verify the research hypotheses in this study, it was necessary to obtain the number of accommodations published on Airbnb's website at different years (2007-2018) as well as their prices and the hotel occupancy rate by category. A balanced panel was analyzed where the individuals were each of the states belonging to Mexico, giving a total of 32 individuals since it was not possible to obtain the information about the state of Tlaxcala. The information on hotel occupancy was obtained directly from the statistical compendiums of the Ministry of Tourism of Mexico (SECTUR, 2019). To obtain information on the number of Airbnb's lodgings published on the website an own algorithm was developed, simulating a web browser application with a graphical interface using Java language and JavaFX. JavaScript queries were used to obtain the Hypertext Markup Language (HTML) code. After this, a second Java application was created, that processed the HTML codes to obtain the Uniform Resource Locator (URL) links of each of the accommodations found in the searches carried out. Once URL links were obtained, a JavaFX application with a web browser was developed to access the web page of each accommodation and thus be able to access their HTML code. To process the files individually and store them, a script with the method of string processing called "Shift - Or" was used. Once the information was obtained, it was converted to a basic text format that could be used programmatically for analysis and review.

In general, *Big Data* process proposed by Labrinidis and Jagadish (2012) was followed. The use of *Big Data* in scientific research has generated controversy, Kitchin (2014) proposes two potential ways in which research could be conducted during the following years. The first is empiricism, where "the data speaks for itself" and where the theory does not matter. The second is the science based on data, where the scientific method is modified by combining abduction, induction, and deduction approaches. However, in social sciences, Kitchin (2014) ensures that the *Big Data* analysis will only improve the set of available information, allowing researchers to use new approaches and te-

chniques without completely replacing traditional small data studies. Due to ethical considerations that have been proposed for the use of *Big Data* (Boyd & Crawford, 2011; Manyika, et al., 2011), in this investigation only the public records of the Airbnb's lodgings were extracted, subsequently eliminating the name of each of them, to maintain anonymity.

Following the Big Data process (Labrinidis & Jagadish, 2012), the steps of Acquisition and Recording; Extraction, cleaning, and Annotation; and Integration, Aggregation and Representation were conducted by using the set of programs and applications described before. Once the data on the Airbnb's website and the hotel occupancy in the states by category were obtained, the Modeling and Analysis step was conducted, where research variables were tested, for which the Panel Data methodology was used.

Panel Data warn observations for different individuals at different moments in time (Baltagi, 2001), so two dimensions should be considered in this paper. Because a panel data model includes both, a cross section and a longitudinal dimension (Pesaran, 2004), it is important to perform cross-section dependence test and unit root tests to check if the variables are stationary, and to obtain their order of integration. If a variable is stationary at levels, it is considered to have an integration order I (0), if it is at first differences it could be said that its integration order is I (1) and if it is at second differences, then it is considered to have an integration order I (2) (Gujarati, 2010). In this paper, the unit root test carried out, is the one suggested by Pesaran (2007).

To test if the variables are cointegrated, the cointegration test used were those suggested by Kao (1999) and Pedroni (2001). Then, the short- and long-term causal relationships were obtained through the usage of the vector error correction model (VECM).

To check if there are dependency relationships between the variables of the model in which no cointegration was found, the general model applied is as follows:

$$Y_{it} = \beta_0 + X_{1it}\beta_1 + X_{2it}\beta_2 + \varepsilon + u_{it}$$

which is the same as:

$$Y_{it} = \beta_0 + X_{1it}\beta_1 + X_{2it}\beta_2 + \omega_{it}$$

Where represents the hotel occupancy rate, is the usage of Airbnb, represented by the number of accommodations published on the website, is the price of Airbnb's lodgings and $\omega_{it} = \varepsilon_i + u_{it}$.

The Granger's causality (1969) was analyzed assuming that the lack of predictability of a variable corresponds to Granger non-causality. A variable is not Granger causal when it cannot be predicted in a better way and with the

information that is available of , than what can be done with the absence of in the prediction model.

To prove the Granger causality, it is necessary that the variables are stationary, previously applying unit root test. VECM was applied to obtain Granger's causality in short and long term.

Once the Modeling and Analysis of the data has been carried out, the next step is the interpretation, which can be noticed both, in the results of the present investigation and in its conclusions.

Results

In this section the results obtained from the tests carried out are shown. First the results of Pesaran (2004) test for cross-sectional dependence (CD) are exposed, then the results of Cross-sectional Im, Pesaran, and Shin (CIPS) panel unit root test (Pesaran, 2007) are presented. The results of cointegration tests, long-term coefficients and causality tests are grouped according to the category of the hotels, starting with a group that includes all the hotels, followed by the 5-star hotels, then the four, three and two-stars hotels, finish with the 1-star hotels results.

Table 1.
Cross-sectional dependence test results.

Variable	AL	P	NHO	5-star HO	4-star HO	3-star HO	2-star HO	1-star HO
CD statistic	71t.51923***	73.86474***	73.86474***	71.51923***	71.51923***	71.51923***	71.51923***	71.51923***

Note: *** denotes the rejection of the null hypothesis at the 1% level.
Source: Own elaboration based on test results.

The results of the cross-section dependence test are presented in Table 1. The null hypothesis of non-dependence is rejected for all variables at a level of significance of 1%. Therefore, there is a transversal dependency in all variables, data of each state are correlated with each other. According to that, unit root tests that generate consistent results in the presence of cross-section dependence must be applied, for which the CIPS unit root test was used in this study.

Table 2.
Cross-sectional CIPS unit root test's results.

Variable	Parameter	CIPS
AL	Constant and trend	-0.313
P	Constant and trend	-2.121
NHO	Constant and trend	-1.754
5-star HO	Constant and trend	-2.026
4-star HO	Constant and trend	-1.817
3-star HO	Constant and trend	-1.813
2-star HO	Constant and trend	-1.586
1-star HO	Constant and trend	-1.761
First differences		
ΔAL	Constant	-2.598**
ΔP	Constant	-2.825***
ΔNHO	Constant	-2.833***
Δ5-star HO	Constant	-2.483**
Δ4-star HO	Constant	-2.479**
Δ3-star HO	Constant	-2.715***
Δ2-star HO	Constant	-2.945***
Δ1-star HO	Constant	-2.723***

Note: *** and ** denote the rejection of the null hypothesis at the 1%, and 5% and levels.
Source: Own elaboration based on test results.

Table 2 shows the results of the CIPS unit root test, where no variable is stationary at the levels, however, when transforming them to their first differences, they are all stationary, so they are integrated in order 1 at different levels of significance. To prove the presence of a long-term relationship between variables, cointegration tests of Kao (1999) and Pedroni (2001) were made.

The estimation of the long-term coefficients and the causality tests were then carried out. The results are grouped according to the category of hotels.

National Hotel Occupancy

Table 3.
Results of Kao's cointegration test for national hotel occupancy.

Test	t-statistic
ADF	-24.1895***

Note: *** denotes the rejection of the null hypothesis at the 1% level.
Source: Own elaboration based on test results.

Table 4.
Results of Pedroni cointegration test for national hotel occupancy.

Alternative hypothesis: common AR coefficients.

	Statistic	Weighted statistic
Panel v-Statistic	0.561262	0.561262
Panel rho-Statistic	0.100224	0.100224
Panel PP-Statistic	-4.266879***	-4.266879***
Panel ADF-Statistic	-1.398500***	-1.398500***

Alternative hypothesis: individual AR coefficients.

	Statistic
Group rho-Statistic	2.820268
Group PP-Statistic	-3.343061***
Group ADF-Statistic	-1.548049***

Note: *** denotes the rejection of the null hypothesis at the 1% level.
Source: Own elaboration based on test results.

In Table 3 and Table 4 the results of cointegration tests are shown for the model that considers all hotels in the country, regardless of their category. In general, it can be concluded from the two tests that variables are cointegrated.

Table 5.
Estimation of long-term coefficients for national hotel occupancy.

Variable	FMOLS Coefficients
Airbnb's lodgings	0.000211***
Price	0.000279***

Note: *** denotes statistical significance at 1% level.
Source: Own elaboration based on test results.

In Table 5 it can be noted that both Airbnb's lodgings and their price have positive and significant coefficients, which indicates that they move in the same direction as the hotel occupancy, that is, when Airbnb's lodgings or the price increase, also hotel occupancy increases.

Table 6.
Causality test results for national hotel occupancy.

Dependent Variable	Short Run			Long Run
	ΔNHO	ΔAL	ΔP	ε
ΔNHO	-	-0.000253***	-0.00135***	-0.65546***
ΔAL	4.2725**	-	-0.024022*	-140.0012
ΔP	1.9813***	-1. 9403***	-	-4.1453***

Note: ***, ** and * denote statistical significance at 1%, 5% and 10% levels, respectively. Source: Own elaboration based on test results.

From Table 6 it can be concluded that there is a unidirectional short-term causal relationship that goes from Airbnb's Lodgings to national hotel occupancy, which considering the long-term coefficient indicates that the more Airbnb's lodgings there are in a city, the hotel occupancy will be higher. In the same way there is a short-term causal relationship that goes from the price of Airbnb's accommodations to hotel occupancy, which indicates that the higher the prices of Airbnb's lodgings, tourists will prefer to stay in hotels. Long term, both explanatory variables influence the dependent variable. In addition, Airbnb's lodgings and hotel occupancy influence the price of Airbnb's accommodation.

5-star Hotel Occupancy

Table 7.
Results of Kao's cointegration test for 5-star hotel occupancy.

Test	t-statistic
ADF	-20.2031***

Note: *** denotes the rejection of the null hypothesis at the 1% level.
Source: Own elaboration based on test results.

Table 8.
Results of Pedroni's cointegration test for 5-star hotel occupancy.

Alternative hypothesis: common AR coefficients.		
	Statistic	**Weighted statistic**
Panel v-Statistic	1.493039*	1.493039*
Panel rho-Statistic	1.702231	1.702231
Panel PP-Statistic	-6.174498***	-6.174498***
Panel ADF-Statistic	-1.519283***	-1.519283***
Alternative hypothesis: individual AR coefficients.		
	Statistic	
Group rho-Statistic	4.514808	
Group PP-Statistic	-5.725577***	
Group ADF-Statistic	-1.698901***	

Note: *** and * denote the rejection of the null hypothesis at the 1% and 10% levels, respectively. Source: Own elaboration based on test results.

According to Table 7, it can be concluded that for the model that includes 5-star hotels, the variables show a long-term relationship. The same happens with the Pedroni test, the results are observed in Table 8, where most of the results imply the rejection of the null hypothesis of non-cointegration at different levels of significance.

Table 9.
Estimation of long-term coefficients for 5-star hotel occupancy.

Variable	FMOLS Coefficients
Airbnb's lodgings	3.35E-05
Price	0.002841***

Note: *** denotes statistical significance at 1% level.
Source: Own elaboration based on test results.

Results shown in Table 9 implies that Airbnb's lodgings coefficient is positive but not significant, which means as expected, that does not have a relationship with hotel occupancy. The price does have a statistically significant coefficient.

Table 10.
Causality test results for 5-star hotel occupancy.

Dependent Variable	Short Run			Long Run
	Δ5-star	ΔAL	ΔP	ε
Δ5-star	-	-8.45E04***	-0.00179***	-0. 9901***
ΔAL	4.8653***	-	-0.42224***	35.0183
ΔP	2.1332***	-0.3312	-	-2.1282***

Note: *** denotes statistical significance at 1% level.
Source: Own elaboration based on test results.

Table 10 shows that there is a one-way short-term causality relationship that goes from Airbnb's lodgings to the occupancy of 5-star hotels, however the coefficient is not significant. In the same way there is a short-term causal relationship from the price of Airbnb's lodgings to hotel occupancy, which considering the coefficient, indicates that the higher the price of Airbnb's lodgings, in short term, tourists will choose to stay in 5-star hotels. In long term the causal relationship in similar, so the same would happen.

4-star Hotel Occupancy

Table 11.
Results of Kao's cointegration test for 4-star hotel occupancy.

Test	*t*-statistic
ADF	-30.05343***

Note: *** denotes the rejection of the null hypothesis at the 1% level.
Source: Own elaboration based on test results.

Table 12.
Results of Pedroni's cointegration test for 4-star hotel occupancy.

Alternative hypothesis: common AR coefficients.

	Statistic	Weighted statistic
Panel v-Statistic	2.057773**	2.057773**
Panel rho-Statistic	1.129661	1.129661
Panel PP-Statistic	-3.583436***	-3.583436***
Panel ADF-Statistic	-2.774653***	-2.774653***

Alternative hypothesis: individual AR coefficients.

	Statistic
Group rho-Statistic	3.909166
Group PP-Statistic	-2.489476***
Group ADF-Statistic	-3.266792***

Note: *** and ** denote the rejection of the null hypothesis at the 1% and 5% levels, respectively. Source: Own elaboration based on test results.

Results of Kao's cointegration test, which are presented in Table 11, show that in the model that includes the 4-star hotels, the variables are cointegrated. Table 12 shows results of the cointegration test of Pedroni, where it is observed that most of the tests reject the null hypothesis of non-cointegration at levels of significance of 1% and 5%.

Table 13.
Estimation of long-term coefficients for 4-star hotel occupancy.

Variable	FMOLS Coefficients
Airbnb's lodgings	-0.00350***
Price	1.00013***

Note: *** denotes statistical significance at 1% level.
Source: Own elaboration based on test results.

The estimation of the long-term coefficients in Table 13 indicates a negative and significant relationship between the occupancy rate of 4-star hotels and the amount of Airbnb's lodgings, which indicates that as one decreases, the other increases. The price's coefficient is positive and significant, so it moves in the same direction as hotel occupancy. Table 14 shows that there is a short-term causal relationship that goes from Airbnb's lodgings to the occupancy rate of 4-star hotels, which, considering the coefficient, means that the more Airbnb's lodgings there are in a city, 4-star hotels will tend to receive fewer visitors. However, due to the short-term causal relationship that goes from the price to the occupancy, it can be thought that the higher the price of Airbnb's lodgings, the tourists will prefer to stay in 4-star hotels. In the long term, causality relations run in the same direction, so the same situation is expected to happen, but also, Airbnb's lodgings and hotel occupancy influence the price of Airbnb's accommodation.

Table 14.
Causality test results for 4-star hotel occupancy.

Dependent Variable	Short Run			Long Run
	Δ4-star	ΔAL	ΔP	ε
Δ4-star	-	-6.84E-05***	-0.000128***	-0.955859***
ΔAL	7.19786	-	-0.001828	-8.36749
ΔP	8.78505	-1.491869***	-	-5.441725***

Note: *** denotes statistical significance at 1% level.
Source: Own elaboration based on test results.

3-star Hotel Occupancy

Table 15.
Results of Kao's cointegration test for 3-star hotel occupancy.

Test	t-statistic
ADF	-19.09790***

Note: *** denotes the rejection of the null hypothesis at the 1% level.
Source: Own elaboration based on test results.

Table 16.
Results of the Pedroni cointegration test for 3-star hotel occupancy.

Alternative hypothesis: common AR coefficients.

	Statistic	Weighted statistic
Panel v-Statistic	2.841065***	2.841065***
Panel rho-Statistic	0.571830	0.571830
Panel PP-Statistic	-3.891784***	-3.891784***
Panel ADF-Statistic	-6.168964***	-6.168964***

Alternative hypothesis: individual AR coefficients.

	Statistic
Group rho-Statistic	2.19084
Group PP-Statistic	-1.88721***
Group ADF-Statistic	-4.87336***

Note: *** denotes the rejection of the null hypothesis at the 1% level.
Source: Own elaboration based on test results.

In Table 15 by performing Kao's cointegration test, the rejection of the null hypothesis of non-cointegration at a level of significance of 1% is obtained. In the same way, Table 16 indicates that in Pedroni's cointegration test, the majority of null hypotheses of non-cointegration are rejected at a level of significance of 1%.

Table 17.
Estimation of long-term coefficients for 3-star hotel occupancy.

Variable	FMOLS Coefficients
Airbnb's lodgings	0.000143***
Price	0.000241***

Note: *** denotes statistical significance at 1% level.
Source: Own elaboration based on test results.

The estimation of long-term coefficients, shown in Table 17, indicates that there is a positive and significant relationship between the number of Airbnb's lodgings, its price, and the occupancy rate of the 3-star hotels. Due to this it can be concluded that the three variables move in the same direction.

Table 18.
Causality test results for 3-star hotel occupancy.

Dependent Variable	Short Run			Long Run
	Δ3-star	ΔAL	ΔP	ε
Δ3-star	-	-0.000216***	-0.000156***	-0.278280***
ΔAL	9.3720***	-	-0.026984**	-5.261006***
ΔP	-2.91482**	-2.270018***	-	-8.167243***

Note: *** and ** denote statistical significance at 1% and 5% levels, respectively. Own elaboration based on test results.

Analyzing the causality shown in Table 18, there is short-term relationship that goes from Airbnb's lodgings to hotel occupancy, and a bidirectional relationship between the occupancy rate of 3-star hotels and the price of Airbnb's lodgings. According to this, it can be thought, that in the short term, the higher the price of Airbnb's lodgings, tourists will prefer to stay in hotels, however, the more difficult it is for visitors, due to the high occupancy rate, to find accommodation in a hotel, the price of Airbnb's accommodations will tend to increase. Log-terms, a causal relationship can be observed from Airbnb's lodgings and its price towards hotel occupancy, also price and hotel occupancy to Airbnb's lodgings, and Airbnb's lodgings and hotel occupancy to the price.

2-star Hotel Occupancy

Table 19.
Results of Kao's cointegration test for 2-star hotel occupancy.

Test	t-statistic
ADF	-39.57291***

Note: *** denotes the rejection of the null hypothesis at the 1% level.
Source: Own elaboration based on test results.

Table 20.
Results of Pedroni's cointegration test for 2-star hotel occupancy.

Alternative hypothesis: common AR coefficients.

	Statistic	Weighted statistic
Panel v-Statistic	-3,365,648	-3,365,648
Panel rho-Statistic	2.172219	2.172219
Panel PP-Statistic	-9.055679***	-9.055679***
Panel ADF-Statistic	-3.113358***	-3.113358***

Alternative hypothesis: individual AR coefficients.

	Statistic
Group rho-Statistic	3.221842
Group PP-Statistic	-9.324022***
Group ADF-Statistic	-3.689816***

Note: *** denotes the rejection of the null hypothesis at the 1% level.
Source: Own elaboration based on test results.

In Table 19 can be observed that the null hypothesis of non-cointegration is rejected at a level of significance of 1%. In the same way, in Table 20 can be seen that, for the Pedroni's cointegration test, in most cases the null hypothesis of non-cointegration is rejected. Due to this it is concluded that the three variables are cointegrated in the model that includes the 2-star hotels.

Table 21.
Estimation of long-term coefficients for 2-star hotel occupancy.

Variable	FMOLS Coefficients
Airbnb's lodgings	0.000294***
Price	0.000218***

Note: *** denotes statistical significance at 1% level.
Source: Own elaboration based on test results.

From Table 21, considering the long-term coefficients, it can be concluded that both, Airbnb's lodgings, and their price, have positive and statistically significant relationship with the occupancy of 2-star hotels. Due to this, it can be affirmed that three variables move in the same direction.

Table 22.
Causality test results for 2-star hotel occupancy.

Dependent Variable	Short Run			Long Run
	Δ2-star	ΔAL	ΔP	ε
Δ2-star	-	-0.000371***	-0.000151***	-1.12254***
ΔAL	8.12412***	-	-0.123698***	-9.32361
ΔP	3.13989***	-0.112416***	-	-7.19163***

Note: *** denotes statistical significance at 1% level.
Source: Own elaboration based on test results.

According to Table 22, in short term the direction of the causal relationships goes from Airbnb's lodgings and their price towards the occupancy of 2-star hotels, so the higher the price of them, tourists will prefer to spend the night in 2-star hotels. In long-term, the causal relationships go in the same direction, so the same situation is expected to happen.

1-star Hotel Occupancy

Table 23.
Results of Kao's cointegration test for 1-star hotel occupancy.

Test	t-statistic
ADF	-7.576121***

Note: *** denotes the rejection of the null hypothesis at the 1% level.
Source: Own elaboration based on test results.

Table 24.
Results of Pedroni's cointegration test for 1-star hotel occupancy.

Alternative hypothesis: common AR coefficients.

	Statistic	Weighted statistic
Panel v-Statistic	-1,541901*	-1,541901*
Panel rho-Statistic	-0.097185	-0.097185
Panel PP-Statistic	-4.185421***	-4.185421***
Panel ADF-Statistic	-0.054132	-0.054132

Alternative hypothesis: individual AR coefficients.

	Statistic
Group rho-Statistic	0.099183
Group PP-Statistic	-5.54812***
Group ADF-Statistic	1.00173

Note: *** and * denote the rejection of the null hypothesis at the 1% and 10% levels, respectively. Source: Own elaboration based on test results.

Table 23 shows that in the model that includes 1-star hotels, the non-cointegration hypothesis is rejected at a level of significance of 1%, in Kao's test. However, Table 24 shows that it cannot be said that the variables are cointegrated, so the next step is to calculate the dependency relationship between the variables and Granger's causality.

Table 25.
Results for 1-star hotel occupancy.

	Coefficient	Standard Error	T stat
Airbnb's lodgings	0.001286***	3.37E-05	8.98153
Price	2.28E-05	2.20E-05	0.843249
Square R	0.75592	Adjusted Square R	0.610212

Note: *** and ** denote statistical significance at 1% and 5% levels, respectively. Source: Own elaboration based on test results.

Hausman test confirms that random effects model should be used to perform the dependency test and Table 25 shows that Airbnb's lodgings is a statistically significant variable that has a positive relationship with the occupancy rate of 1-star hotels. However, the price is not significant, this may be because, generally Airbnb's lodgings, have a range of services and amenities that usually 1-star hotels do not have.

Table 26.
Results of Granger's causality test for 1-star hotel occupancy.

Null Hypothesis	F-statistic
Airbnb's lodgings does not Granger Cause Occupancy	175.182***
Occupancy does not Granger Cause Airbnb's lodgings	18.1691***
Price does not Granger Cause Occupancy	19.7412***
Occupancy does not Granger Cause Price	4.34267

Note: *** denotes Null Hypothesis rejection at 0.01.
Source: Own elaboration based on test results.

Looking at Table 26, it can be concluded that there is a bidirectional causal relationship between the occupancy rate of 1-star hotels with Airbnb's lodgings, this could mean that the more occupied 1-star hotels are, the greater the amount of Airbnb's lodgings, and the same situation happens in the opposite direction.

Discussion and conclusion

The main objective of this paper is to analyze the influence that usage of Airbnb's platform has had on the hotel occupancy in Mexico during the 2007-2018 period. On this subject, previous work has been done for different tourist destinations: European cities (Coyle & Yu-Cheong, 2016), Helsinki (Fissha & Shrestha, 2017), and Texas (Zervas, Proserpio & Byers, 2017), among others, concluding that in general, Airbnb does not have a negative effect on the hotel industry, except for low-category hotels.

To test research hypothesis, the Big Data process proposed by Labrinidis and Jagadish (2012) was followed, using for the Modeling and Analysis section, econometric tests of cross-sectional dependence, unit root, cointegration, dependence and causality, that are specifically indicated for each of the cases.

Results indicate that the first research hypothesis es rejected since there is no evidence that Airbnb's prices had had a negative effect on hotel occupancy in Mexico and there is a causal relationship that goes from Airbnb's price towards hotel occupancy, with a positive coefficient, so the higher the price of Airbnb's lodgings, tourists will prefer to stay in hotels of different categories. This is consistent with Kaplan and Nadler (2015) stating that Airbnb can be a cheaper option to visitors when hotels have high prices.

The second hypothesis states that the usage of Airbnb has not had a negative effect on the occupancy rate of high-category hotels in Mexico, which cannot be rejected since it was found only in 4-star hotels. In literature revision, platforms like Airbnb are called to be a real threat to traditional accommodation companies such as hotels (Sigala, 2017), and according to this paper's results, the proliferation of Airbnb's lodgings has a short-term and long-term causal relationship with hotel occupancy, however in most cases studied in this paper, this relationship has positive coefficients, which implies that in general, hotel occupancy in Mexico is not negatively affected by Airbnb's lodgings. The only negative effect found between the usage of Airbnb and hotel occupancy occurs in 4-star hotels, where a reduction in hotel occupancy is related to the increase in Airbnb's lodgings, however, the higher the price of Airbnb's lodgings, the hotels will have more demand, so an implication of the results is that the usage of platforms such as Airbnb can be moderately regulated in Mexico and 4-star hotels' managers need to think about reducing their prices, so they can compete with Airbnb's lodgings and not having their occupancy rate decreased.

The third hypothesis indicates that the usage of Airbnb has had a negative effect on the occupancy rate of low-category hotels in Mexico, which according to the results is rejected since for 1-star hotels there is a positive and bidirectional causal relationship between their occupancy and Airbnb's lodgings, but not with the price, which is not significant in the model. This

may be because Airbnb's lodgings tend to offer better services and amenities than the hotels belonging to that category.

Future lines of research may focus on studying the specific reasons why tourists traveling to Mexico may prefer to use Airbnb versus staying in hotels.

References

Baltagi, B.H. (2001). *Econometric Analysis of Panel Data,* Second Edition. Wiley, New York.

Barron, K., Kung, E., & Proserpio. (2018). The Sharing Economy and Housing Affordability: Evidence from Airbnb. Available at: http://dx.doi.org/10.2139/ssrn.3006832

Boyd, D., & Crawford, K. (2011). *Six Provocations for Big Data. A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society*. Available at: https://ssrn.com/abstract=1926431

Coyle, D., & Yu-Cheong, T. (2016*). Understanding Airbnb in Fourteen European Cities*, working paper. Toulouse School of Economics. Available at: https://www.tse-fr.eu/sites/default/files/TSE/documents/ChaireJJL/PolicyPape
rs/2016_30_12_pp_understanding_airbnb_in_14_european_cities_coyle_yeung_v.3.1.pdf

Edelmar, B., & Luca, M. (2014). *Digital Discrimination: The Case of Aribnb.com*. Harvard Business School NOM United Working Paper No. 14-054.

Ert, E., Fleischer, A., & Magen, N. (2015). Trust and Reputation in the Sharing Economy: The Role of Personal Photos in Airbnb. *NA - Advances in Consumer Research*, 43, p. 518-519.

Fang, B., Ye, Q., & Law, R. (2015). Effect of sharing economy on tourism industry employment. *Annals of Tourism Research*.

Fissha, H., & Shrestha, A. (2017). *Impact of Airbnb on Hotel Industry in Helsinki*, Degree Thesis. Yrkeshögskolan Arcada.

Granger, C. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37 (3), 424-438.

Gujarati, D., & Porter, D. (2010). *Basic Econometrics*. Mc Graw Hill/Irwin, Inc.

Guttentag, D. (2016). *Why tourists choose Airbnb: A motivation-based segmentation study underpinned by innovation concepts*. PhD Thesis. University of Waterloo, Canada.

Gyódi, K. (2017). Airbnb and Booking.com: Sharing Economy Competing Against Traditional Firms? Working Paper *DELab UW*, 3(2017).

Henten, A., & Windekilde, I. (2015). Transaction costs and the sharing economy. *26th European Regional Conference of International Telecommunications Society*. Spain.

Im, K. S., Pesaran, M. H., & Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, 115(1), 53-74.

Jiménez-Martínez, A. (2008). Las cadenas hoteleras en el mundo y evolución de su operación en México al inicio del siglo XXI. *Innovar*, 18 (32), 167-194.

Kao, C. (1999). Spurious Regression and Residual Based Tests for Cointegration in Panel Data. *J. Econom, 90*, p 1-44.

Kaplan, R., & Nadler, M. (2015). Airbnb: A case study in occupancy regulation and taxation. *The University of Chicago Law Review*, 82, 103.

Ke, Q. (2017). Sharing Means Renting? An Entire-market place Analysis of Airbnb. *Proceedings of the 2017 ACM on Web Science Conference*. 131-139

Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society, April-June 2014*, 1-12. DOI: 10.1177/2053951714528481

Labrinidis, A., & Jagadish, H. (2012). Challenges and opportunities with big data. *Proceedings of the VLDB Endowment*, 5(12), 2032-2033.

Levin, A., Lin, C. F., & Chu, J. (2002). Unit root tests in panel data: asymptotic and finite sample properties. *Journal of Econometrics*, 108(1), 1-24.

Maddala, G.S., & Wu, S. (1999). A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test. *Oxf. Bull. Econ. Stat*, *61*, 631-652.

Manyika. J., & Chui M. (2011). *Big Data: The next frontier for innovation, competition and opportunity*; Mckinsey Global Institute.

Martorell, O. (2002). *Cadenas Hoteleras. Análisis del Top 10*. Madrid: Ariel Turismo.

Nguyen, Q. (2014). A Study of Airbnb as a Potential Competitor of the Hotel Industry. *UNLV Theses, Dissertations, Professional Papers, and Capstones*. 2618.

Pedroni, P. (2001). Fully modified OLS for heterogeneous cointegrated panels, in Badi H. Baltagi, Thomas B. Fomby, R. Carter Hill (ed.) *Nonstationary Panels, Panel Cointegration, and Dynamic Panels (Advances in Econometrics, Volume 15)* Emerald Group Publishing Limited, pp.93-130.

Pesaran, M. (2004). *General Diagnostic Tests for Cross Section Dependence in Panels*; CESifo Working Paper Series: Munich, Germany, 1229, pp. 1-39.

Pesaran, M. (2007). A simple panel unit root test in the presence of cross-section dependence. *J. Appl. Econom*, *22*, p 265-312.

Propin, E., López, A. & Sánchez, A. (2004). Territorios preferenciales de los grupos hoteleros internacionales en América Latina y el Caribe, a principios del siglo XXI. *Investigaciones Geográficas*, 53, 122-140.

SECTUR. (2019). Compendio estadístico, Secretaría de Turismo de México. Available at: https://www.datatur.sectur.gob.mx/SitePages/ActividadHotelera.aspx

Sigala, M. (2017). Collaborative commerce in tourism: implications for research and industry. *Current issues in Tourism*, 20(4), 346-355.

Sundararajan, A. (2013). From Zipcar to the Sharing Economy. *Harvard Business Review*. Available at: https://hbr.org/2013/01/from-zipcar-to-the-sharing-eco

Terporten, M., Bialdyga, D., & Planing, P. (2012). Ver¨anderte Kundenw¨unsche als Chance zur Differenzierung, *Zuk¨unftige Entwicklungen in der Mobilit¨at. Betriebswirtschaftliche und technische Aspekte*, pp. 367-382, Springer.

Torres, D. (2011). Globalización, Empresas Multinacionales e Historia. *Pensamiento y gestión*, 30, 165-185.

WTTC. (2020). Mexico, *2020 Annual Research: Key Highlights*. World Travel & Tourism Council. Available at https://wttc.org/Research/Economic-Impact/moduleId/704/itemId/161/controller/DownloadRequest/action/QuickDownload

Zervas, G., Proserpio, D., & Byers, J. (2017). The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry. *Journal of Marketing Research: October 2017*, Vol. 54, No. 5, pp. 687-705.