The Impact of Sharing Economy: New Evidence from Airbnb in Taiwan*

Fang-Chang Kuo[†], Hsin-Hsi Shih[‡]

April 6, 2021

Abstract

We investigates the impact of Airbnb on incumbent hotels' revenue in Taiwan. Combining listing information scraped from Airbnb website with a panel of hotel revenues, we propose a novel set of instrumental variables to identify the casual impact. Relative to OLS estimate, IV-2SLS estimates indicate larger negative effects from Airbnb. Furthermore, smaller and lower quality hotels are heavily affected by Airbnb listings.

Keywords: Airbnb, hotel, competition, instrumental variable

JEL Codes: L13, L83

*We are grateful for valuable comments received from discussants and participants at National Taipei University, Industrial Economics Conference, and TEA 2020 Annual Conference. All remaining errors are our own.

[†]National Chung Cheng University, Department of Economics, No.168, Sec.1, University Rd., Minhsiung, Chiayi 62102, Taiwan; Phone: +886-5-2720411 ext. 34121. Fax: +886-5-2720816. Email: *fckuo@ccu.edu.tw*; https://fangchangkuo.weebly.com/

[‡]Cathay Life Insurance. Email: *c1109123@gmail.com*

1 Introduction

Founded in 2008, Airbnb has grown rapidly in the past decade. As of 2020, Airbnb has hosted more than 750 million guests with around 7 million listings in 100 thousand different cities. It has become one of the most well-know examples in the sharing economy. The success brings various implications to lodging industry. Potential impacts on the conventional hotels are relevant to both regulators and firms. However, previous studies have shown mixed results about the impact of Airbnb on incumbent hotels' performances. Some find that Airbnb is a formidable competitor to the lodging industry (Zervas et al., 2017; Dogru et al., 2019). Others find insignificant or no effects from Airbnb (Blal et al., 2018; Chang and Sokol, 2021). The main challenge for identifying a casual link lies in the fact that entry of Airbnb listings is not exogenous. Specifically, Airbnb hosts can flexibly adjust their listings in response to potential future or present demand shocks. The common trend assumption for difference-in-difference (DD) research design is not valid in this setting. Hotels in treatment group will not follow the same trend when Airbnb's entry is absent since a positive demand shock, which attracts more listings from Airbnb, could increase treated hotels' performances at the same time.

In this paper, we investigates the impact of Airbnb on incumbent hotels' revenue in Taiwan. This study uses a unique data-set containing various hotel performance metrics from the entire population of legitimate hotels in Taiwan, and Airbnb listing data scraped from Airbnb website. Different from the previous literature, we propose a novel set of instrumental variables (IV) and use two-stage least square approach to address the endogeneity issue. The instrumental variables measures the total number of vacant houses and housing inventories. The logic behind these two variables follows the mechanism of sharing economy, featuring the utilization of idle capacity, as more unused housing capacity could drive the supply of Airbnb but not correlate with specific demand shocks in a market.

The empirical estimates show noticeable difference between estimates in OLS and IV-2SLS estimation, indicating bias caused by potential endogeneity. IV-2SLS estiamte indicates larger negative effect of Airbnb listings on hotel revenues. The effect is both statistically and economically significant. In addition, hotels of lower tiers and smaller capacities are heavily affected by Airbnb.

Our empirical results add to a growing literature on various impacts of sharing economy (Einav et al., 2016). Focusing on Airbnb, our empirical results are consistent with Zervas et al. (2017) and Dogru et al. (2019). However, this paper distinguishes from previous studies on Airbnb and hotel performance in a way that we are able to find instrumental to address endogeneity issue. Our estimated impacts are larger in contrast to previous studies. Besides, we find that Airbnb imposes stronger business stealing effects on low quality and small hotels.

2 Data

In this study, we construct a unique data-set containing monthly performance panel of Taiwanese hotels, Airbnb listing, and vacant houses and housing inventories. Three main data-sets collected from various sources, and are merged according to county and month.

In 2008, the Bureau of Tourism in Taiwan requested all legitimate hotels to submit monthly reports of operating performances including room revenues, sales, number of employees, total number of room available, and age. The data spans the period between January 2009 and June 2016. Revenue per available (RevPAR) is the dependent variable measuring hotel performance.[1] We also supplement hotel data with a panel of consumer ratings from major online review platforms, including TripAdvisor, Agoda, Expedia, and Bookings.com. Following Lewis and Zervas (2016), online ratings are aggregated across different rating

[1] RevPAR is the most commonly-used performance measure in the lodging industry. With room revenues and sales in each month, we can calculate average daily rate (ADR) for each room-night sold. Occupancy rate is obtained by dividing sales with total number of room available in each month. Revenue per available is the product of average daily rate and occupancy rate.

websites and serve as a proxy for time-varying hotel-specific quality measure.[2]

Listing information is scraped from Airbnb website during Nov. 2019 and Dec. 2019. In each listing, we extract the registration date of Airbnb host to proxy the entry date of the listing. Using the time information over all available listings in all counties, we construct the cumulative number of listings in the past. This strategy is also adopted by Zervas et al. (2017) to measure Airbnb supply in each county.[3]

Two instrumental variables are used in the IV-2SLS estimation to identify the causal impact of Airbnb listings. Vacant house is defined as residential property with power usage lower than 60 kilowatt hours per month. Housing inventory is the number of newly-built houses, within five years, to be sold in a city or county. The intuition for choosing vacant house and housing inventory as IVs will be explained in the next section.[4]

3 Empirical Model

Our data consists of individual firm level performance information and market level Airbnb listing information. The panel structure of our data-set allows for the possibility of controlling various levels of fixed-effects. Following Zervas et al. (2017), we use similar specification and estimate the following regression equation:

$$\log(\text{RevPAR}_{jkt}) = \beta \cdot \log(\text{Airbnb listing}_{kt}) + \boldsymbol{X}^{'}_{jkt}\boldsymbol{\gamma} + \nu_j + \tau_t + \text{County}_k \times \text{Month}_t + \epsilon_{jkt} \quad (1)$$

where j indexes hotels, k indexes county, and t indexes year-month. Our main dependent variable, RevPAR, is a common measure of hotel performance in the lodging industry. $\boldsymbol{X}_{jt}$

[2] Some platform use five point scale while others use ten point scale. We use ten point scale and convert five point scale into ten by multiplying rating scores by two.

[3] In Online Appendix A, we provide a table of cumulative counts of Airbnb listings in each year for all counties.

[4] Both variables are collected from Platform of Real Estate Information from Ministry of the Interior in Taiwan. For more details, see Online Appendix.

are controls for hotel characteristics, including age, number of rooms, number of employees, and online rating scores. We also includes total number of hotel rooms supplied in a county to control for changes in total hotel room supplied. Airbnb listing$_{kt}$ is cumulative number of listings in a county. Parameter of interest is β, which measures the impact of Airbnb's listing on hotel revenues. τ_j is hotel fixed effect, which absorbs any time-invariant firm-specific unobserved factor and ν_t is time fixed effect, which captures overall macroeconomic conditions or demand shocks common to all firms in every market. Seasonality in specific market is included by adding county-and-month fixed effects. Finally, ϵ_{jkt} is an error term.

This econometric model is widely-used by previous studies as it is essentially a generalized difference-in-difference regression model. However, the identifying assumption in a DD framework may not be satisfied in this case since Airbnb hosts can flexibly adjust their room supply in response to future demand shocks, which are unobserved by researchers. Parallel trend assumption is violated as hotels compete with more Airbnb listings would benefit from positive demand shocks such as music festival, or major sports events. The revenue trends would not have been the same in the absence of increase in Airbnb listings.

To address this issue, we combine fixed-effects model in Equation 1 with two instrumental variables, vacant houses and housing inventories. We choose these two variables as IVs since they measure unused housing resources in each county. The number of cumulative Airbnb listings could be positively driven by spare housing capacities in a county. The two IVs are likely to be uncorrelated with unobserved demand shocks since they are determined several years ago in the residential housing market. Equilibrium in residential housing market should be orthogonal to demand shocks from tourists. In addition, the surplus in vacant houses or newly-built houses cannot adjust for county-month specific demand shocks in time. Therefore, we believe they are valid instruments satisfying both relevance condition and exclusion restriction.

4 Empirical Results

Table 1 reports empirical results for Equation 1 with OLS and IV-2SLS estimation respectively. In the first column in Table 1, in which OLS estimates are presents, we find a negative but insignificant relationship between Airbnb listings and hotel revenues. However, the estimate is close to zero and not economically significant as 1% increase in Airbnb listings is associated with only 0.006% decrease in RevPar. The OLS estimate is potentially driven by omitted variable bias. As unobserved demand shocks could be positively correlated with Airbnb listings and RevPar, the direction of bias is likely to be upward, resulting in a less negative estimate of Airbnb listings.

Consistent to our intuition of bias direction, after addressing the endogeneity issue, IV estimates of column 3 is much more negative and precisely estimated. On average, 1% increase in cumulative Airbnb listings is causing around 0.19% decrease in RevPar.[5] Our estimated impact of Airbnb listings is larger when compared to results in Zervas et al. (2017). However, despite the fact we implement IV approach in estimation, majority of hotels in Taiwan are in general small hotels in lower tiers.[6] Airbnb could be close substitute for small low quality hotels since these hotel only provides basic services without various amenities.

We further separate our sample based on quality levels and capacities to investigate heterogeneous effects of Airbnb listings. Estimation results are presented in Table A4. Column 1 uses hotel with lower than 3 stars and Column 2 focus on hotel with 3 or more stars. The difference between estimates in Column 1 and 2 is stark. Estimate of -0.237 is larger in magnitude comparing to -0.191 in Table 1. Hotels with lower qualities are heavily affected by Airbnb while hotels of higher quality are less affected with insignificant estimate. For

[5] The decrease in RevPar is mainly driven by lower occupancy rates. See Online Appendix Table A3.

[6] According to Hollenbeck (2018) and Zervas et al. (2017), mean capacity of Texas hotels is 86. In contrast, mean capacity of Taiwanese hotels is 33. We present a distribution of hotel capacities in Online Appendix Figure A1.

Table 1: Effect of Cumulative Airbnb Listings on Hotel Revenue

	(1) OLS	(2) IV-2SLS 1st	(3) IV-2SLS 2nd
log Cum. Airbnb listings	-0.006		-0.191***
	(0.005)		(0.069)
Hotel Age	-0.119***	-0.018	-0.123***
	(0.029)	(0.028)	(0.029)
log No. of Rooms	-0.678***	-0.014	-0.680***
	(0.044)	(0.043)	(0.045)
log No. of Employees	0.285***	0.014	0.288***
	(0.024)	(0.016)	(0.024)
Is Reviewed	-0.101	-0.045	-0.110
	(0.125)	(0.153)	(0.127)
Average Rating	0.044**	0.013	0.046***
	(0.017)	(0.021)	(0.017)
log Hotel Room Supply	0.047	-0.381***	-0.043
	(0.032)	(0.076)	(0.047)
log Vacant Houses		0.748**	
		(0.309)	
log Housing Inventories		0.163***	
		(0.027)	
Hotel FEs	Yes	Yes	Yes
Year-Month FEs	Yes	Yes	Yes
County-Month FEs	Yes	Yes	Yes
N	201,052	201,052	201,052

Note: Dependent variable in column 1 and 3 is logarithm of revenue per available room. Column 2 is the first-stage regression results of IV-2SLS. Endogenous variable is log cumulative Airbnb listings. Associated F-statistic for excluded instruments is 19.64. The Hansen overidentification test statistic is 0.081, which has a p-value 0.776. Standard errors in parentheses are robust and clustered at firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Column 3 and 4, we run separate IV-2SLS for hotels below and above median capacity, 33 rooms. The estimates of Airbnb listings indicates that smaller hotels are heavily affected by increase in Airbnb listings.

Table 2: Heterogeneous Effects for Different Qualities and Capacities

	(1) Low Quality	(2) High Quality	(3) Small Capacity	(4) Large Capacity
log Cum. Airbnb listings	-0.237**	-0.054	-0.288***	-0.131
	(0.093)	(0.065)	(0.109)	(0.102)
Hotel Age	-0.119***	-0.137	-0.083*	-0.137***
	(0.030)	(0.105)	(0.042)	(0.044)
log No. of Rooms	-0.690***	-0.522***	-0.811***	-0.521***
	(0.048)	(0.138)	(0.054)	(0.085)
log No. of Employees	0.298***	0.199***	0.278***	0.277***
	(0.025)	(0.072)	(0.034)	(0.034)
Is Reviewed	-0.252	0.094	-0.511*	-0.052
	(0.154)	(0.193)	(0.302)	(0.140)
Average Rating	0.069***	0.009	0.110**	0.033*
	(0.022)	(0.024)	(0.043)	(0.019)
log Hotel Room Supply	-0.059	0.042	-0.042	-0.059
	(0.057)	(0.077)	(0.074)	(0.066)
Hotel FEs	Yes	Yes	Yes	Yes
Year-Month FEs	Yes	Yes	Yes	Yes
County-Month FEs	Yes	Yes	Yes	Yes
N	174636	26410	88816	112184
F-statistic	13.38	14.54	9.78	8.93
Overidentification test-stat.	0.004	0.984	0.384	1.934

Note: Column 1 limits to 1-star or 2-star hotels or hotels without any star rating, and Column 2 focuses on hotels of 3-star or more. Dependent variable for all columns is logarithm of revenue per available room. Column 3 and 4 use median capacity, 33 rooms, to define hotel with small capacity and larger hotels. Endogenous variable is log cumulative Airbnb listings. First-stage results are presented in Online Appendix Table A4. Standard errors in parentheses are robust and clustered at firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5 Conclusion

Using novel instruments to address endogeneity of unobserved demand shocks, we find larger effects of Airbnb listings on hotel revenues. The discrepancy between OLS and IV-2SLS estimates indicates potential endogeneity issue with unobserved demand shocks. Our findings also show that small lower quality hotels face stronger competition from Airbnb listings.

Reference

Blal, I., Singal, M., and Templin, J. (2018). Airbnb's effect on hotel sales growth. *International Journal of Hospitality Management*, 73:85–92.

Chang, H.-H. and Sokol, D. D. (2021). How incumbents respond to competition from innovative disruptors in the sharing economy-the impact of airbnb on hotel performance. *Strategic Management Journal*.

Dogru, T., Mody, M., and Suess, C. (2019). Adding evidence to the debate: Quantifying airbnb's disruptive impact on ten key hotel markets. *Tourism Management*, 72:27–38.

Einav, L., Farronato, C., and Levin, J. (2016). Peer-to-peer markets. *Annual Review of Economics*, 8:615–635.

Hollenbeck, B. (2018). Online reputation mechanisms and the decreasing value of chain affiliation. *Journal of Marketing Research*, 55(5):636–654.

Lewis, G. and Zervas, G. (2016). The welfare impact of consumer reviews: A case study of the hotel industry. *Working Paper*.

Zervas, G., Proserpio, D., and Byers, J. W. (2017). The rise of the sharing economy: Estimating the impact of airbnb on the hotel industry. *Journal of marketing research*, 54(5):687–705.

Online Appendix

A Additional Tables and Figures

Table A1: Summary Statistics

	Mean	S.D.	P25	P50	P75	Freq.	Unit
Hotel Performance							
Price ($)	63.95	59.59	34.77	53.12	77.33	Monthly	Property
Occupancy rate (%)	46.07	24.43	25.32	43.66	65.22	Monthly	Property
RevPar ($)	31.58	35.89	9.32	22.74	43.09	Monthly	Property
Characteristics							
No. of rooms	52.03	62.51	20.00	33.00	59.00	Monthly	Property
No. of employees	29.17	76.79	5.00	11.00	21.00	Monthly	Property
Age	15.37	13.41	4.00	11.00	24.00	Monthly	Property
Online ratings	7.61	0.84	7.11	7.69	8.20	Monthly	Property
County capacity (k)	359.9	259.1	293.2	151.7	536.4	Monthly	County
Airbnb							
Cumulative listing	213.65	558.06	0.00	8.00	178.5	Monthly	County
IVs							
Vacant house (k)	39.17	35.50	16.05	24.457	65.21	Quarterly	County
House inventory (k)	1.47	2.18	2.03	5.31	18.32	Quarterly	County

Notes: This table presents summary statistics of my final dataset. In regression analysis, we restrict sample to observations with at least 50 sales per month to ensure that hotels are active.

Table A2: Cumulative Counts of Airbnb Listings in Different Counties

Couty/City	2009	2010	2011	2012	2013	2014	2015	2016
Yilan County	0	0	23	53	338	850	1653	2504
Hualien County	0	0	1	26	276	717	1464	2306
Kinmen County	0	0	0	0	15	24	71	180
Nantou County	0	0	0	30	169	313	524	1073
Pingtung County	0	0	0	49	250	600	1184	2106
Miaoli County	0	0	4	30	238	526	993	1679
Taoyuan City	0	0	8	8	62	231	404	722
Kaohsiung City	1	2	14	65	241	618	1107	1683
Keelung County	0	0	0	0	3	35	81	150
Lienchiang County	0	0	0	0	0	0	4	16
Yunlin County	0	0	2	13	87	129	267	445
New Taipei City	0	1	6	41	270	713	1227	1811
Hsinchu City	0	0	0	0	9	12	25	38
Hsinchu County	0	0	0	0	67	193	337	512
Chiayi City	0	0	0	9	35	50	83	144
Chiayi County	0	0	3	41	118	261	562	969
Changhua County	0	1	3	11	145	229	450	676
Taichung City	0	2	6	32	225	505	933	1529
Taipei City	30	37	130	393	1308	2848	5009	7326
Taitung County	0	0	0	12	61	133	321	611
Tainan City	1	1	12	82	310	714	1287	1974
Penghu County	0	0	0	2	114	156	292	564

Notes: The numbers represent the cumulative counts of Airbnb listings in different county/city at the end of every year.

Table A3: Regression table

	(1) log Price	(2) Occ. Rate
log Cum. Airbnb listings	0.054	-0.088***
	(0.044)	(0.026)
Age	-0.060***	-0.018**
	(0.018)	(0.008)
log No. of Rooms	-0.155***	-0.189***
	(0.034)	(0.016)
log No. of Employees	0.073***	0.076***
	(0.011)	(0.006)
Is Reviewed	0.054	-0.094**
	(0.082)	(0.039)
Average Rating	0.001	0.022***
	(0.011)	(0.005)
log Hotel Room Supply	0.000	-0.018
	(0.029)	(0.017)
Hotel FEs	Yes	Yes
Year-Month FEs	Yes	Yes
County-Month FEs	Yes	Yes
Observations	201052	201052

Note: Table A3 shows the separate effects on ADR and OCC. First stage estimation result is the same as Table 1 Column 2 in the main text since we use same set of instruments and endogenous variable. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A4: First-Stage Regressions of Table 2

	(1) Low Quality	(2) High Quality	(3) Small Capacity	(4) Large Capacity
log Vacant Houses	0.546*	2.379***	1.129**	0.536
	(0.327)	(0.623)	(0.493)	(0.389)
log Housing Inventories	0.149***	0.242***	0.163***	0.148***
	(0.029)	(0.066)	(0.040)	(0.035)
Age	-0.014	-0.049**	-0.018	-0.031
	(0.032)	(0.019)	(0.031)	(0.038)
log No. of Rooms	-0.017	-0.020	0.105	-0.152*
	(0.046)	(0.083)	(0.074)	(0.080)
log No. of Employees	0.020	-0.031	0.032	0.010
	(0.018)	(0.041)	(0.023)	(0.022)
Is Reviewed	-0.158	0.626*	-0.478*	0.200
	(0.167)	(0.343)	(0.262)	(0.184)
Average Rating	0.026	-0.067	0.057	-0.010
	(0.023)	(0.046)	(0.037)	(0.025)
log Hotel Room Supply	-0.364***	-0.488**	-0.346***	-0.412***
	(0.080)	(0.220)	(0.106)	(0.108)
Hotel FEs	Yes	Yes	Yes	Yes
Year-Month FEs	Yes	Yes	Yes	Yes
County-Month FEs	Yes	Yes	Yes	Yes
N	174636	26410	88816	112184

Note: Table A4 shows the first-stage regressions for Table 2 in the main text.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

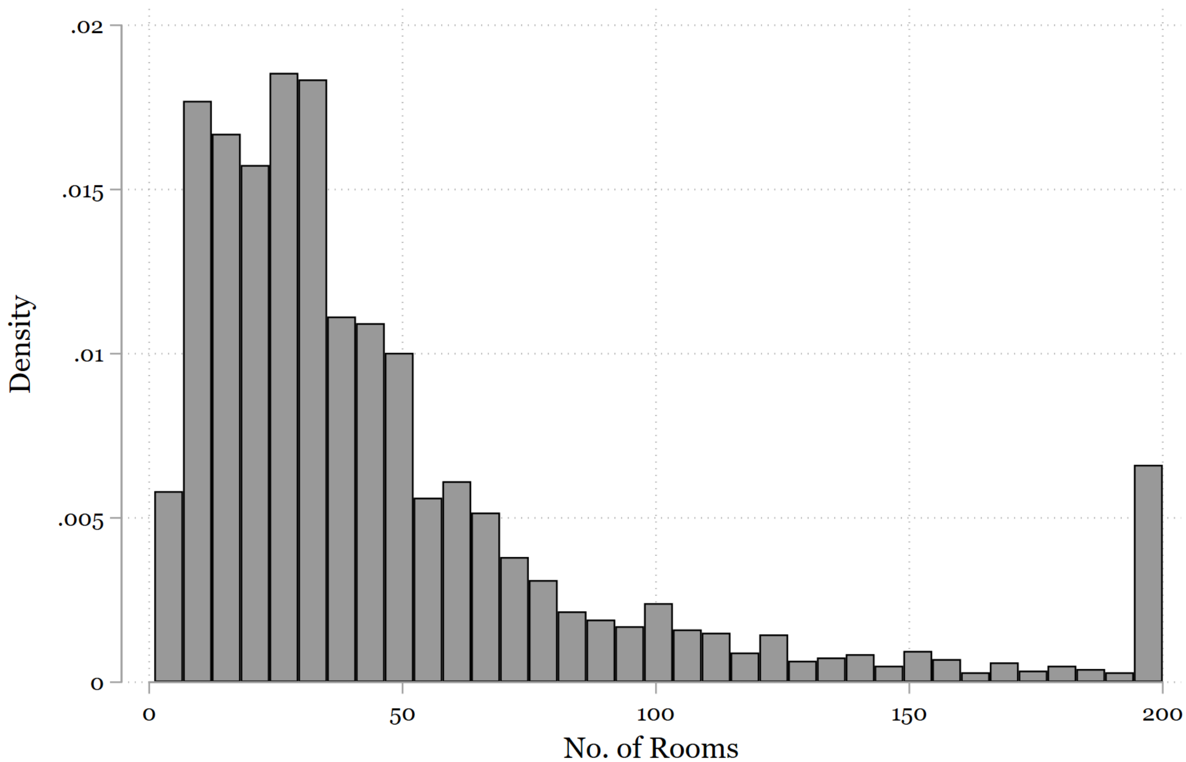
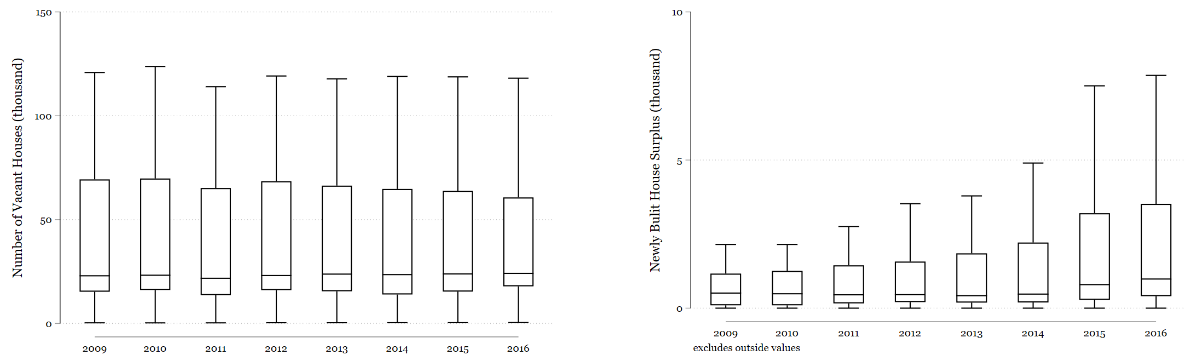


Figure A1: Distribution of Hotel Size

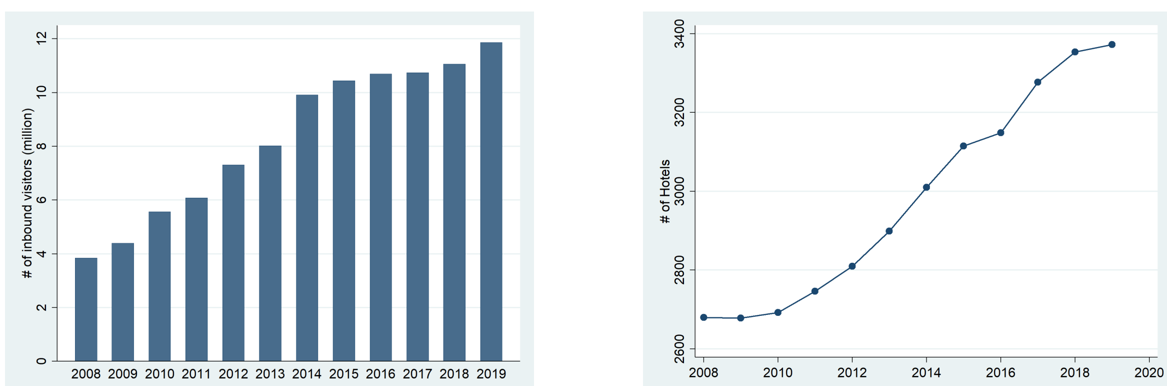
Notes: Figure A1 shows the distribution of hotel sizes. The number of room is top-coded at 200.



(a) Vacant Houses (b) Newly-built Houses

Figure A2: Unused Housing Resources

Notes: The data is publicly available on https://pip.moi.gov.tw/V3/E/SCRE0104.aspx and for detailed descriptions of the data, please see https://pip.moi.gov.tw/Eng/Default.aspx?pg=F04



(a) Total Number of Inbound Visitors (b) Total Number of Hotels

Figure A3: Increasing Numbers of Inbound Visitors and Hotels

Notes: Taiwanese hotel industry experiences rapid a growth in the past decade. The total number of inbound visitors almost tripled as shown in Figure A3a. Increasing demand also induces numerous entries in different markets across all segments. At the same time, number of hotels increase from around 2700 in 2008 to over 3200 in 2019.