Econometric Game 2021

Team 23

April 8, 2021

1 || Introduction

In this paper, we analyze the effect of the number of Airbnb listings on house prices in Amsterdam. Amsterdam is the top tourist destination in the Netherlands, drawing millions of tourists every year. Many of these tourists use Airbnb to find their stay in Amsterdam. However, there is a housing shortage in Amsterdam, putting pressure on policy makers to pass legislation to ensure housing in Amsterdam is used for residents, not tourists.

Proponents of stricter legislation for the use of homes for short term rentals on Airbnb make the argument that more homes on Airbnb means fewer homes are sold to residents. Economic theory then suggests that the inelasticity of housing supply in the short run would cause an increase in the prices of houses. This argument is supported by the fact that the development of the number of Airbnb listings and house prices in Amsterdam have followed very similar trajectories since the first Airbnb listing appeared in Amsterdam in 2008. Since 2008, the number of listings has grown exponentially. This is demonstrated in Figure I, which shows that there were only a few listings in 2008 and 2009, but in 2013 the amount increased dramatically and they appeared in all parts of the city. Finally, the number of listings grew even further by 2018. Furthermore, in Figure II, we demonstrate the dynamics of the housing prices. We observe that the prices slightly rose after 2008 right after the introduction of Airbnb (first dotted red line). After 2013, they continued rising exponentially, which may be associated with Amsterdam City Council policy of making short-term rental easier for the local residents (second dotted red line).[1] At first sight, it is not difficult to see why some would argue that an increase in Airbnb listings had caused an increase in house prices in Amsterdam, since the correlation is positive.

However, opponents of Airbnb could argue that it is also possible that increased short term rentals via Airbnb have a negative effect on prices. For instance, it is possible that increased tourism is a nuisance to residents, hence driving them out of neighbourhoods where Airbnb is prevalent. Mathematically, perhaps without realizing it, opponents of stricter legislation for Airbnb rentals are arguing that the relationship in Figure II might be due to confounding variables affecting both the housing prices as the number of Airbnb listings.

Hence, it is possible to economically argue both a positive and negative relationship between Airbnb listings and house prices. This paper provides an empirical investigation of the nature of the relationship between the two variables in question.

[1] https://www.airbnbcitizen.com/wp-content/uploads/2016/12/National_PublicPolicyTool-ChestReport-v3.pdf

Similar relationships have been investigated in previous literature. A recent study by Barron et al. (2021) found that a 1% increases in Airbnb listings lead to a 0.026% growth in house prices in the United States. In a paper by Koster et al. (2018), it was shown that Airbnb had a large effect on house prices in areas attractive to tourists. Garcia-López et al. (2020) found in their investigation that for the average neighbourhood in terms of Airbnb activity, the house prices increased by 5.3%. However, not much research has been done on Amsterdam housing market which is the main focus of this paper.

The relationship of interest in this paper is the effect of the number of Airbnb listings on housing prices. We start the analysis by estimating a neighborhood and time fixed-effects model that shows that after controlling for amenities, there is a strong positive relationship between Airbnb activity and house prices. Additionally, we observe the effect is even stronger after the city of Amsterdam implemented more lenient legislation in 2013. Furthermore, we find no strong evidence of a non-linear relationship between Airbnb activity and housing prices.

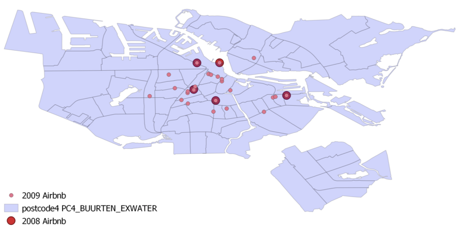
There are a number of ways the relationship between Airbnb activity could be obscured by endogeneity. First, there might be a reverse effect of housing prices on Airbnb listings, causing simultaneity bias. Second, there are possible variables that affect both housing prices and numbers of Airbnb listings. For instance, population growth or GDP growth could positively affect both house price and listings. To tackle these sources of endogeneity we use an instrumental variables approach. The instrumental variable we use is an interaction (product) between a measure of how 'touristy' a postal code is and a measure of the interest in Airbnb. Firstly, as a measure of how 'touristy' a postal code is, we use the density of monuments (such as parks, musea or statues) in that postcode. The validity of the instrument is based on the assumption that having monuments nearby matters to tourists, thus effecting Airbnb listings, but does not matter to local residents, hence not affecting house prices which is reasonable. Secondly, we use Google trends to measure the interest in Airbnb. Clearly, higher interest in Airbnb through Google searches is a valid proxy for the amount of Airbnb listings made. And due to the fact that the changes in Google searches will most likely be due to changes in popularity and not changes in housing prices, Google searches for Airbnb is a valid instrument.

The results of the instrumental variable approach are similar to the approach without instrumental variables. The effects of Airbnb density on housing prices is even stronger in the former case. We thus conclude that increases in Airbnb density lead to higher housing prices in Amsterdam.

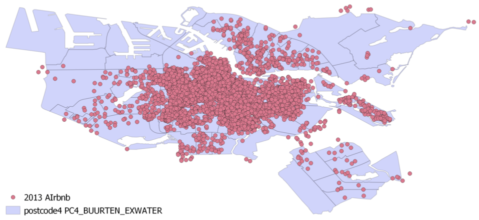
The structure of the paper is as follows. In part 2 we introduce the data and describe the transformations required for the subsequent analysis performed in part 3. Part 4 describes and interprets the results. Finally, part 5 is the conclusion.

2 || Data

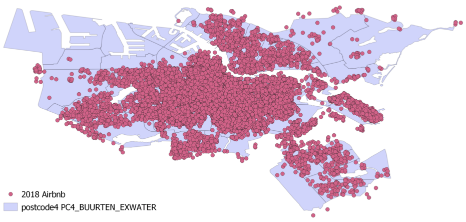
The data comes from the Dutch Association of Realtors (NVM) and contains microdata on housing prices in Amsterdam for the period 2000-2018. Specifically, it contains transaction housing prices as well as housing characteristics on an address level. We consider housing prices as a dependent variable. The main independent variable of interest is the Airbnb activity in



(A) 2008 and 2009



(B) 2013



(C) 2018

Figure I

Airbnb listings at different points in time

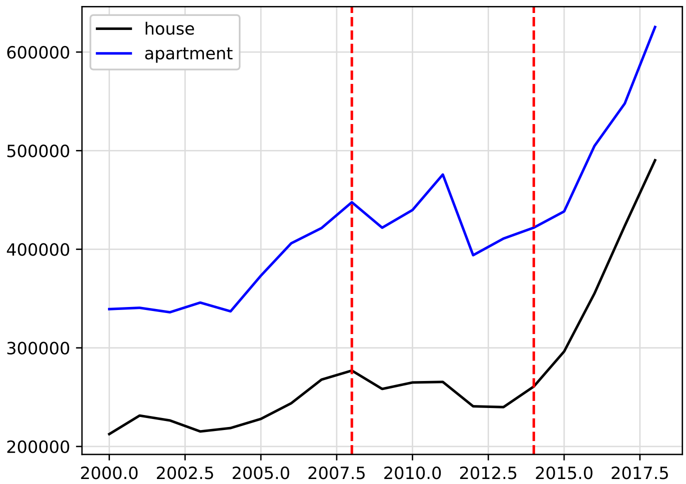


Figure II

Housing prices dynamics for different types of housing

Amsterdam. There are several ways to measure this activity. First, we focus on the number of the Airbnb listings within a radius of 250 metres, since the more listings there are the more Airbnb can potentially affect the house prices (if there is any effect present). Next, we consider the distance to the nearest Airbnb listing as an Airbnb activity measure since the prices of the houses that are located closer to the Airbnb places might be affected more.

2.1 || Control variables

The housing prices are usually modeled using hedonic regression which models housing prices as being a composite of several housing characteristics. It is based on the consumer demand theory and was originally proposed by Rosen, 1974. Therefore, in the main regressions we control for housing characteristics that are usually related to the housing prices. Particularly, we control for the size, year of construction, whether there is parking and garden, type, the quality of house and monumental status. Moreover, we include buyerpaysorfree variable as a covariate since it can affect the housing price.

2.2 || Instrumental variables

To resolve the endogeneity issue we construct an instrumental variable, which is an interaction of two components: monuments per square kilometre in a postal code district (four digits) and the interest over time in the google search term "Airbnb Amsterdam".

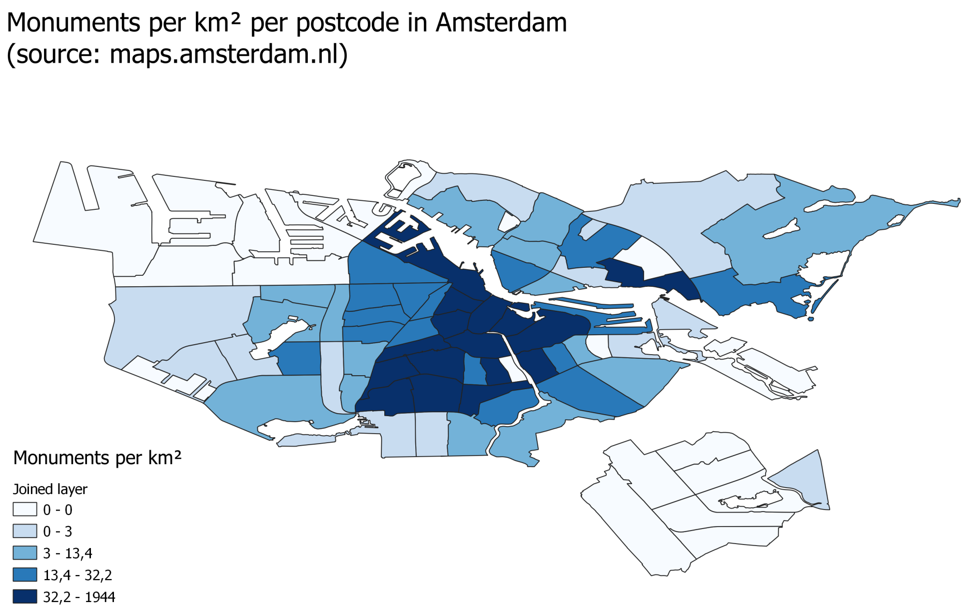


Figure III

Monuments per square kilometer per four-digit postcode

The first component, the number of monuments per square kilometre in a postal code district

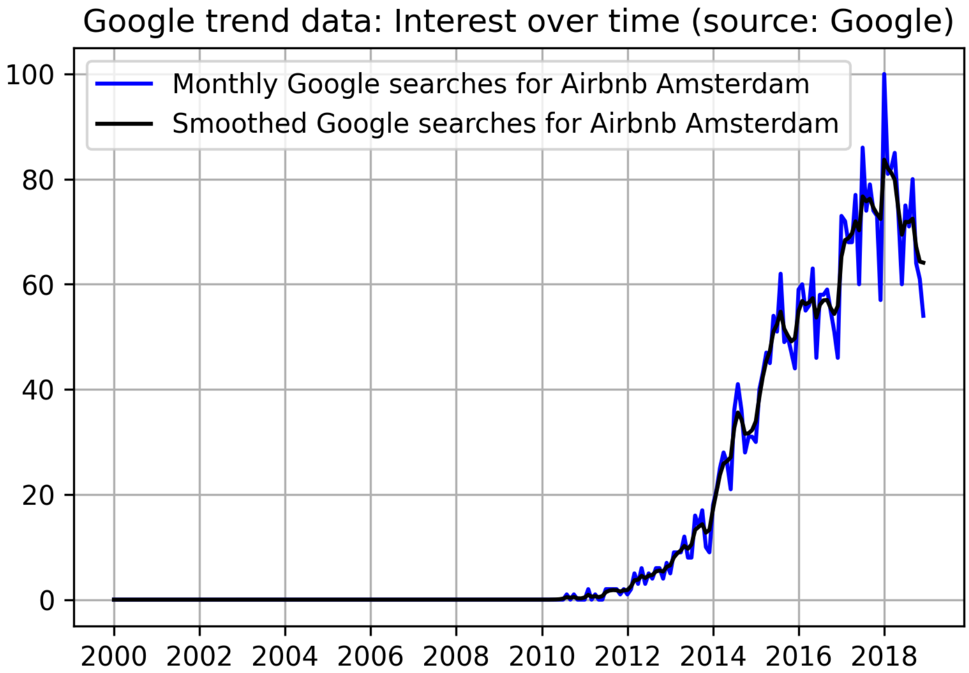


FIGURE IV

Interest over time in "Airbnb Amsterdam: raw data and smoothed estimate"

is publicly available on the Amsterdam municipality website[2]. These include monuments like statues, museums, and parks. We use this component as it is varies across the different postal codes in Amsterdam, and it is not correlated with the house prices. Using the software QGIS we generated the density of the monuments per $\mathbf{km^2}$ per postal code district (Figure III).

Second, we also consider the interest over time index in the google search "Airbnb Amsterdam"[3]. This index represents search interest relative to the highest point, which is set to 100. This index is depicted in Figure (IV). As can be seen, the raw data is rather spiky and a bit erratic as a signal. This is potentially because the moment of searching for an Airbnb accommodation may not be the same as the moment of going to Amsterdam. To estimate a clearer signal, we use an Kalman exponential smoother in a state space model (Durbin and Koopman, 2012):

$$google_t = \mu_t + \epsilon_t, \qquad \epsilon_t \sim \mathcal{NID}(0,\, \sigma_\epsilon^2) \tag{1}$$

$$\mu_{t+1} = \mu_t + \eta_t, \qquad \eta_t \sim \mathcal{NID}(0,\, \sigma_\eta^2) \tag{2}$$

$$\mu_1 \sim \mathcal{N}(0, \kappa), \qquad \kappa \to \infty \tag{3}$$

We then use an interaction of these two components to obtain an instrumental variable. Since we work with the logarithm of the Airbnb activity we consider the logarithm of the Google trend. We discuss this below.

[2]https://maps.amsterdam.nl/open_geodata/?k=122

[3]https://trends.google.com/trends/explore?date=all&q=airbnb%20amsterdam

2.3 || Data transformations

We take the logarithm of the housing prices as it is usually done in hedonic regressions. Moreover, to make large observations less influential we also work with the logarithm of the Airbnb measures and hence also with the logarithm of the Google Trend search. Furthermore, we transform the variable *quality* into four categories: "bad", "poor", "good" and "excellent". Initially, the variable quality is between 0 and 20, then "bad" quality corresponds to the 1 percentile, "poor" to the 1-10 percentile, the majority corresponds to "good" quality (10-90 percentile) and top 10 correspond to the "excellent" group. Additionally, since tourists rent apartments more often than houses we expect that Airbnb had a different effect on the apartment and houses prices houses. However, we do not expect the effect to be different for different types of houses. Therefore, we transform the variable type into accommodation and house type (0 and 1).

3 || Model

3.1 || Baseline model

Our baseline specification is the following:

$$log(Y_{it}^r) = \alpha_r + \tau_{year} + \beta log(Airbnb_{it}^r) + \gamma_i X_{it}^r + \epsilon_{it}^r, \tag{4}$$

where Y_{it}^r transaction price of house i which is located in region r, α_r are region-specific fixed effects that account for time-invariant region characteristics, τ_{year} are yearly time fixed effects, and X_{it}^r are housing characteristics. $Airbnb_{i,t}$ is the measure of Airbnb activity which is measured either using distances or density. We name this Model (1). The regions are defined based on $pc4$ zip code level and the standard errors are clustered at the region level.

As there could be some nonlinear effects of the Airbnb activity on the house prices we consider the baseline model with additional nonlinear terms: $log(Airbnb_{it}^r)^2$ and $log(Density_{it}^r) \times log(\frac{1}{Distance_{it}^r})$ as regressors.

Furthermore, in 2014 the Amsterdam City Council made it more accessible for the residents to share their private apartments. Therefore, we expect after the 2014 the effect of Airbnb to be even larger. For this we consider a modified version of the baseline specification with an additional term $\tilde{\beta} log(Airbnb_{it}^r) \times D_{t>2014}$, where D is equal to 1 after and including 2014 year.

3.2 || Instrumental variables

The causal graph for the effect of Airbnb listings on the housing price is given below. We wish to investigate the effect of the number of Airbnb listings on the housing prices. However, there might be numerous causal paths making the analysis of this effect problematic.

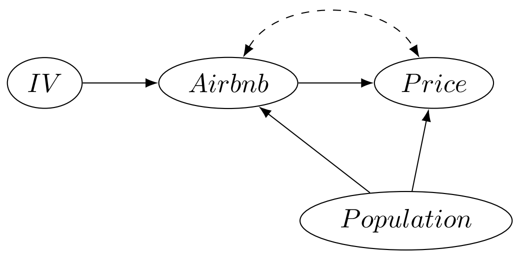
Firstly, there could also be a reverse effect between the dependent and independent variables, resulting in simultaneity. It is very likely that the house prices in a district have an effect on the number of Airbnb listings. High house prices could cause landlords to sell their property, lowering supply of the Airbnb apartments or increase in the house prices could be associated with an increase in the rental price which consequently could lead to the increase in the Airbnb activity. Additionally, there could be numerous confounding variables, such as the population of a city. An increase in the population could lead to both an increase in housing prices and an increase in rental offerings. Moreover, in general, it is hard to measure Airbnb activity since some listing can be present on the website but be not active and it is also hard to measure the distance accurately. So, we expect a measurement error to be present which leads to the endogeneity problem as well.

Given that it is very likely that there is endogeneity issue we think it is important to account for it in this study. One way to address the endogeneity issue could be to run a natural experiment to isolate the effect of the Airbnb on the housing prices by not allowing Airbnb activities in particular regions. However, we do not always have this option in real life. Therefore, another possibility would be to conduct a quasi-experiment as it was done, for example, in Koster et al., 2018.

Since we do not have the above mentioned natural and/or quasi experimental setups, we alternatively opt for the instrumental variable approach. The instrumental variable serves as a means of isolating the effect of the number of listings on house prices. To successfully accomplish this task, it is important that (i) the IV correlates with the Airbnb activity measure and (ii) the IV does not correlate with the housing price.

The number of Google searches for Airbnb in Amsterdam has grown over the measurement period, and it is plausible that it is correlated with Airbnb listings, but not with house prices, making it a relevant and valid instrument. For the density of monuments in a postal code district we make the assumption that since tourists enjoy being near the sights, the monument density will effect the Airbnb activity. Furthermore, we assume that the monument density will not affect housing prices, since local residents are less influenced by their proximity to monuments in general. The combination of both the described assumptions makes monument density a valid and relevant instrument.

To create a variable that runs over the different districts and time periods, we combine the two instruments described in the previous paragraph by multiplying them.



4 || Results

4.1 || OLS Regression

Table I reports the results of the OLS regression of the log(transaction prices). The first column is a regression on on the log(density), where we consider both zip code and year fixed effects and further control variables (which are described in the data section). In this regression, an increase of 1% in the density leads to an increase of 0.046% in house prices.

The regression in the second column additionally considers a difference in slope between the period before the change in regulation in 2014, and afterwards, by the inclusion of a dummy variable. In this regression, all fixed effects and controls are the same as in the first regression. We notice that the slope steepens quite substantially after 2014, when Amsterdam implemented more lenient regulations regarding Airbnb renting. As a consequence, the slope of the log(density) is 0.021 before 2014, and then more than doubles in size. This shows quite clearly the potential impact of legislation on the short-run impact of Airbnb on the housing market.

The third regression considers two more variables: the product of the log(density) and log(distance), as well as the log(density)^2. Here as well, the fixed effects and controls are the same as in the other regressions. The log(density) before the change in regulation in 2014 is then not significantly different from 0, whereas the log(density) after the change in regulation is. Therefore, the difference in effect from log(density) before and after the regulation change in 2014 seems to be robust.

Table I

Impact of density of Airbnb listings on house prices: OLS estimates

	log(price)		
	(1)	(2)	(3)
log(density)	.046***	.021***	.007
	(.003)	(.003)	(.007)
log(density) $\times$ regulation		.036***	.028***
		(.003)	(.006)
log(density) $\times$ log(distance)			-.001*
			(.001)
log(density)^2			.003
			(.002)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Controls	Yes	Yes	Yes
R^2	.888	.889	.889
N	108440	108440	108440

Clustered standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.2 || Instrumental Variable Regression

In order to deal with the potential endogeneity of the number of Airbnb listings within 250m of the property to explain the transaction price we performed an instrumental variable regression. In this IV analysis we instrumented the number of Airbnb listings within 250m of the property with the interaction of the logarithm of smoothed searches for Airbnb Amsterdam and number of monuments in line with Garcia-López et al. (2020) and Barron et al. (2021). The first component tracks changes in Airbnb acitivity over time, while the second component captures the proximity of the neighbourhood to the city's tourist amenities.

Table II shows the IV regression output estimated by the 2SLS with clustered standard errors at the zip code level. In the first stage regression we find no evidence for potential weak instrument problem as the F-statistic is 16.3 > 10. From the second stage estimation we observe that the sign compared to the OLS estimates has not changed. However, the effect of an additional Airbnb listing in the near vicinity has doubled. This indicates that our OLS estimates had a strong bias towards zero, due to the endogeneity of the Airbnb listing variable. Similar to the neighbourhood and time fixed effects models estimated by OLS discussed above we included neigbourhood and time fixed effects as well as a large group of control variables capturing a large variety of housing characteristics.

Table II

Impact of density of Airbnb listing on house prices: IV estimates

	First stage log(density)	Second stage log(price)
instrument	.00001***	
	(.000)	
log(density)		.09320***
		(.012)
Zip code fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Controls	Yes	Yes
R^2	.88544	.88430
N	108440	108440

Clustered standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.3 ‖ Robustness Analysis

These results are not driven by the specific choice of the variable capturing Airbnb activity. In the appendix we show the OLS and IV estimation results in the case when we use as a measure of Airbnb activity the distance to the closest Airbnb listing. From Tables .IV and .III we observe that the same results apply using this different measure for Airbnb activity, strengthening the robustness of our core results.

5 ‖ Conclusion

This paper investigated the effect of Airbnb density on housing prices in Amsterdam. Whereas we did not find any strong non-linear effects. We found that in general, an increase of the log density of Airbnb listings led to an increase in housing prices. However, our results show that the impact is different for the period before 2014 and after, when a change of regulation was made, making it easier to list a property on Airbnb. When differentiating these periods, the increase in housing prices is less pronounced, and even insignificant in one regression model. However, when the differentiation is made, the slope of the log(density) of Airbnb listings is always significant in the period after the regulation change, and is in all models considerably larger than before the regulation change.

We considered both an OLS model, as well as a model in which an instrumental variable was added to accommodate the possible endogeneity of the variables. The instrumental variable approach are similar to the approach without instrumental variables. The effects of Airbnb density on housing prices is even stronger in the former case. We thus conclude that increases in Airbnb density lead to higher housing prices in Amsterdam, especially since the change in regulation in 2014. This may imply that legislation has a clear impact on the housing prices in the short run.

References

Barron, K., E. Kung, and D. Proserpio (2021): "The Effect of Home-Sharing on House Prices and Rents: Evidence from Airbnb", *Marketing Science*, 40(1), 23–47.

Durbin, J. and S. J. Koopman (2012): *Time series analysis by state space methods*, Oxford university press.

Franco, S. F. and C. D. Santos (2021): "The impact of Airbnb on residential property values and rents: Evidence from Portugal", *Regional Science and Urban Economics*, 88, 103667.

Garcia-López, M.-À., J. Jofre-Monseny, R. Martínez-Mazza, and M. Segú (2020): "Do short-term rental platforms affect housing markets? Evidence from Airbnb in Barcelona", *Journal of Urban Economics*, 119, 103278.

Koster, H., J. van Ommeren, and N. Volkhausen (2018): "Short-term rentals and the housing market: Quasi-experimental evidence from Airbnb in Los Angeles", CEPR Discussion Paper DP13094.

Rosen, S. (1974): "Hedonic prices and implicit markets: product differentiation in pure competition", *Journal of political economy*, 82(1), 34–55.

Appendix

Table .III

Impact of distance to nearest Airbnb listing on house prices: OLS estimates

	log(price)		
	(1)	(2)	(3)
log(distance)	-.026***	-.007**	.019***
	(.004)	(.003)	(.005)
log(distance) × regulation		-.039***	-.063***
		(.004)	(.005)
log(density) × log(distance)			-.009***
			(.001)
log(distance)^2			-.002***
			(.001)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Controls	Yes	Yes	Yes
R^2	.886	.886	.888
N	108440	108440	108440

Clustered standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table .IV

Impact of distance to nearest Airbnb listing on house prices: IV estimates

	First stage	Second stage
	log(distance)	log(price)
instrument	-.00001***	
	(.000)	
log(distance)		-.17981***
		(.027)
Zip code fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Controls	Yes	Yes
R^2	.93307	.84984
N	108440	108440

Clustered standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$