

# The Data Open

## Impact of Hosting Olympics on Real Estate Prices with Difference-in-difference Model

Eric Frankel, Mengxiao Li, Cangcheng Tang, Shiyu Liu

July 19, 2020

### 1 Executive Summary

We analyzed real estate market data in order to better understand the impact of the Olympic Games on the economy of host cities, taking the 2012 London Olympics and the 2000 Sydney Olympics as case studies. We chose these two cities because they represent different types of Olympics hosts. London was already an established global city at the time of its hosting, with the goal of juicing economic growth in disadvantaged boroughs, while Sydney was less well-known relative to London, hoping to emerge on the global stage through an Olympic bid.

Using a difference-in-difference approach, we determined that hosting the Olympics contributed 4% to the residential real estate market in the London boroughs that hosted the events, which corresponds to an increase of about £1.5bn (\$1.88bn) in residential property value in those areas.

In Sydney, hosting the Olympics added 5.92% to residential property values. This increase in the areas that hosted Olympic events alone represents about 1.81bn AUD (\$1.27bn), which doesn't include the impact on the real estate market of the greater Sydney Metropolitan Area.

These increases in residential property prices covered about 18% of the cost of the Sydney Olympics, and about 16.5% of the cost of the London Olympics. This suggests that host cities should focus on planning for the long-term legacy of Olympic venues in order to benefit from these indirect economic impacts.

### 2 Background

Hosting the Olympic games is often seen as a way to bring economic growth to the host city and country, reasoning that is often used to justify the rising cost of hosting the Olympics, which has exceeded \$10bn since the 2004 Olympics in Athens. Much of this investment goes towards improving existing infrastructure and building new facilities, like stadiums and hotels, to accommodate the massive influx of tourists and athletes into the host city. These improvements are made under the assumption that they will lead to a lasting, positive legacy for the host. However, as vacant stadiums, or "white elephants," dot the landscape of many past hosts, questions remain about how effective hosting the Olympics is at delivering long term prosperity.

Real estate prices provide a unique means to assess the indirect economic impact of the games, especially over the long-term, as elevated real estate prices years after the games are a signal that the host area continues to be desirable. Thus, we have chosen to investigate the following question:

**How does hosting the Olympics impact the real estate market in the host city?**

We answer this question through the use of detailed land registry data, in combination with aggregate statistics, applying a difference-in-differences model to assess how hosting the Olympics affect real estate prices in host neighborhoods compared to the rest of the host city.

## 3 Data

### 3.1 England Real Estate Pricing Data

England House Price Index data were retrieved from UK government statistics database. From the price index data, we were able to retrieve the time series of average England real estate price data across different regions in UK. The different regions include 33 boroughs of the greater London area, North East, North West, South East, South West, East of England, Yorks & the Humber, East Midland, and West Midlands. The time series data were used in our exploratory data analysis and earlier stage of modeling. We further classify the regions as River Zone, Central Zone, and others, where River Zone is the core hosting area for Olympic games and the Central Zone is the non-core venues for the Olympic.

### 3.2 Sydney Real Estate Pricing Data

Sydney land registry data from 1990-2010 were collected from the Valuer General of New South Wales, which provides a weekly report of real estate transactions in the province. These data were combined with the quarterly, housing-specific CPI for the Sydney Metropolitan Area from the Australian Bureau of Statistics in order to produce the inflation-adjusted price per square meter for every real estate transaction during the period in constant 4Q1999 AUD. These data were further classified as host or non-host, based on the postal codes that hosted Olympic venues, and commercial or residential based on the zoning code for the transaction.

Using these data we can compare the effect of the Olympics on real estate prices in host vs. non-host areas. Further, we can contrast the impacts on the commercial real estate market vs. those on the residential real estate market.

## 4 Methodology

### 4.1 General Method

We utilize a difference-in-difference approach via linear regression and its semi-logarithmic functional form to estimate real estate prices. The model takes the value of property as its target variable, regressing on the attributes (such as a house's neighborhood environment). Since normality is an important assumption for statistical models, the semi-logarithmic functional form aims to transform the non-normal price distribution to a normal shape. For each of our datasets, we implement the Box and Cox transformation to the target variable, which represents the value of the real estate. In this case, the target variable is inflation-adjusted price per square meter, which is transformed using the following function:

$$price^{(\theta)} = \begin{cases} \frac{p^\theta - 1}{\theta} & \text{for } \theta \neq 0 \\ \ln p & \text{for } \theta = 0 \end{cases}$$

Figure 1 is the QQ-plot for England Housing Price. The un-transformed distribution is on the left, which is skewed right. The transformed distribution appears on the right, and it follows the normal distribution.

Various adjustments were made prior to this transformation, including the inflation adjustment and controlling for rare events. We control for the impact of the Global Financial Crisis of 2008 by multiplying price by the financial crisis multiplier, as the financial crisis had a substantial negative impact on real estate prices. Note that this adjustment is made only for London as the time period used for the Sydney data does not include the financial crisis. We created three dummy variables in order to quantify the impact of hosting the Olympic Games on real estate prices. The first represents whether the data are before or after the announcement of which boroughs would host Olympic events. The second is a geographical dummy variable that classifies an area as host or non-host. Finally, we have an interaction dummy variable, which is equal to unity for host areas after the announcement and zero otherwise. This interaction term helps us understand the difference-in-difference for the change in prices of host boroughs after the announcement relative to the change in prices of non-host boroughs. Using this approach should control for other factors that impact real estate prices in the Greater London Area, and allow us to understand the impact of the

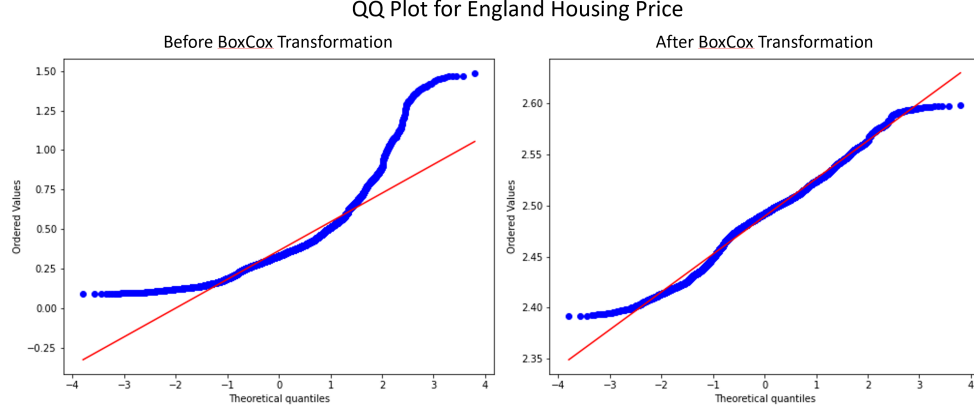


Figure 1: The left QQ-plot shows a strong right skew, the right QQ-plot shows the distribution follows the  $N(0, 1)$ .

Olympics themselves. Our general model for this problem is:

$$\log p_{j,t} = c + \delta_1 \text{TimeStamp} + \delta_2 \text{GeographicalFactor} + \delta_3 \text{TimeStamp} * \text{GeographicalFactor} + \sum_{i=1}^n \beta_i x_i + \epsilon_j$$

We slightly modified this model as appropriate for each host city, and computed the coefficients using Ordinary Least Squares regression.

The percentage change brought by each coefficient is calculated by comparing to  $c$ , the constant, this is because  $c$  represents the average in non-host areas before any announcement for the Olympics plans. Since we were using a Box and Cox transformation to preprocess our data, we need to first transform them back for comparison, specifically, the formula we used is as follows:

$$\sqrt[\theta]{(\delta + c) \cdot \theta + 1} - \sqrt[\theta]{c \cdot \theta + 1}$$

## 4.2 Olympic Country Level Impact: England

To investigate how hosting the Olympic Games influence the overall real estate price across the country, we first did a difference-in-difference model for different regions in England using the England Average Real Estate dataset. We use July 1st, 2007 as the break point for the *TimeStamp* dummy variable because this was about the time when the host boroughs were announced to the public. This factor is equal to unity if the price is recorded after the announcement and zero if the price is recorded before the time point. We considered the time period from 2001 to 2016 to ensure there was an equal amount of years before and after the announcement. The geographical dummy variable is equal to unity for regions that had a high density of Olympic venues, and zero for other regions. Specifically, Greater London, Eastern England and South East England are grouped as the Olympic regions. Such regions are shown in dark blue in Figure 2. The specific regression model for this dataset is:

$$\text{price}_j^{(\theta)} = c + \delta_1 \text{Timestamp} + \delta_2 \text{HostRegions} + \delta_3 \text{Timestamp} \cdot \text{HostRegions} + \epsilon_j$$

## 4.3 City Level Impact

In addition to the impact on large regions of the host countries, we wanted to assess how prices in the specific host areas changed relative to other areas of the host city. We collected data from London and Sydney in order to analyze the impacts on an intra-city level.

Again, we used two dummy variables, one for analyzing differences before and after the announcement for the Olympics municipal plan, the other for analyzing the effect of the district/borough. The interaction of the two dummy variables is the difference-in-difference term, which evaluates the Olympics effect on host districts' housing price.



Figure 2: Dark blue parts are the Olympics hosting boroughs in England, and the light green parts are the other boroughs. The grey parts are out of England

#### 4.3.1 Sydney

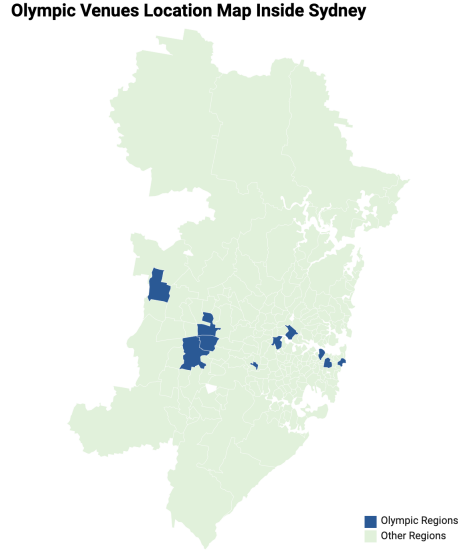


Figure 3: The Olympics hosting districts are highlighted in dark blue, other districts are in light green

In this section, we want to explore how residential real estate prices in host districts were affected by the Olympics. We set the time breakpoint as 1995, since that was the time when the New South Wales government announced its planning for the Olympics. All the eight hosting districts are in Sydney, shown in Figure 3.

The regression model can be written as:

$$pirce_j^{(\theta)} = c + \delta_1 Timestamp + \delta_2 HostDistrict + \delta_3 Timestamp \cdot HostDistrict + \epsilon_j$$

### Olympic Venues Location Map Inside London

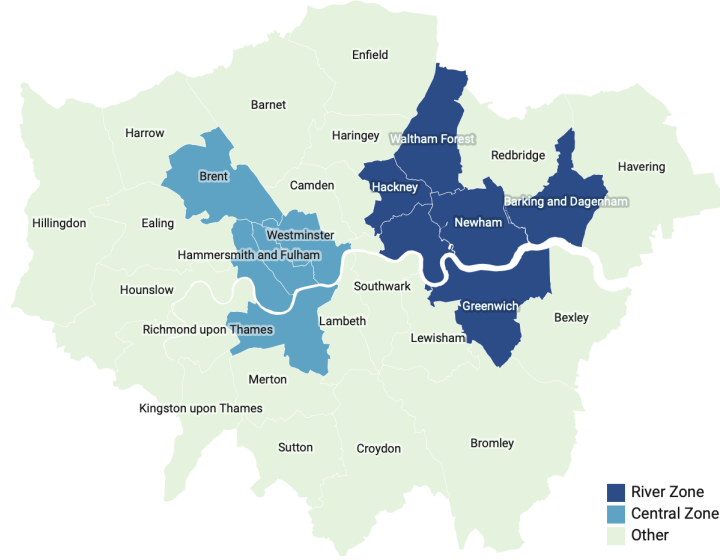


Figure 4: The dark blue boroughs are the River Zone, the light blue boroughs are the Central Zone

#### 4.3.2 London

The London boroughs were separated into three categories, River Zone, Central Zone, and Other, based on how much investment went into each region. This breakdown is highlighted in Figure 4. The mean residential real estate values of the three zones over time is shown in Figure 5. Similar to the timestamp factor in country level impact, July 1st, 2007 was used to classify data as before or after the announcement. The regression model can be written as:

$$price_j^{(\theta)} = c + \delta_1 Timestamp + \delta_2 RiverZone + \delta_3 CentralZone + \delta_4 Timestamp \cdot RiverZone + \delta_5 Timestamp \cdot CentralZone + \epsilon_j$$

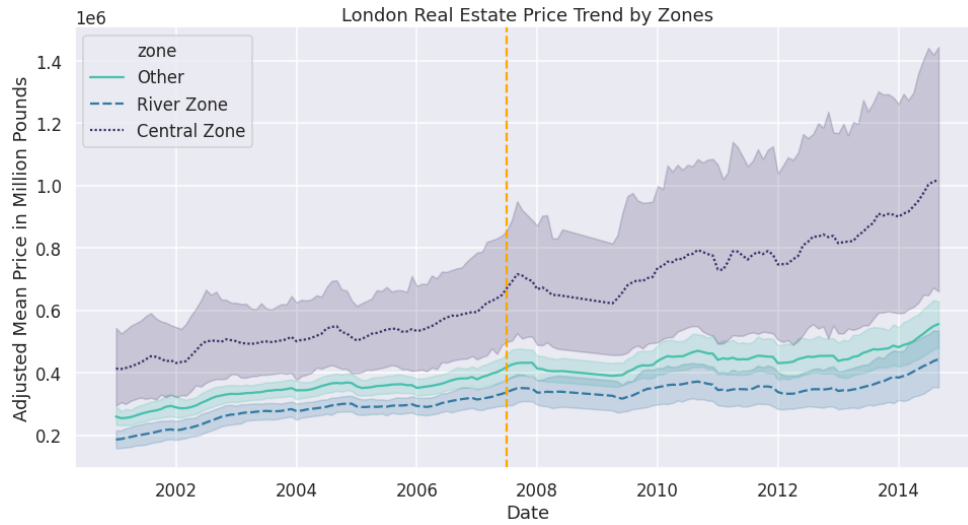


Figure 5: Each line represents a Zone in London. Central Zone's housing prices are the highest, while River Zone's are the lowest.

## 5 Results

### 5.1 Olympic Country Level Impact: England

Table 1: Coefficients for the England Model

	Coefficient	t-value	$p >  t $	Coefficient Transformed	Percentage Change by One Unit Increase
Timestamp: after announcement	7.237	23.213	0.000	110707.790	7.34%
Olympic Region	9.257	24.243	0.000	70017.810	4.64%
Time Stamp: after * Olympic Region	1.668	3.088	0.000	10765.980	0.71%
Constant	67.675	306.990	0.000	1508555.000	

The result of the England model is shown in Table 1. Based on this regression, it is clear that time had a significant impact on real estate prices, with real estate prices throughout the UK being about 7% higher post announcement. We also find that the Olympic Regions had prices about 4.6% higher than non-Olympic regions. Finally, the interaction term had a modest impact, showing that Olympic regions post-announcement were about 0.7% higher than non-Olympic regions post-announcement. Importantly, all coefficients are significant at the 95% confidence level.

### 5.2 City Level Impact

#### 5.2.1 Sydney

Table 2: Coefficients for the Sydney Model

	Coefficient	t-value	$p >  t $	Coefficient Transformed	Percentage Change by One Unit Increase
Timestamp: after announcement	0.357	71.830	0.000	68.565	27.61%
Olympic Region	0.561	22.519	0.000	115.159	46.37%
Time Stamp: after * Olympic Area	0.084	2.825	0.005	14.705	5.92%
Constant	6.688	1580.004	0.000	248.354	

The result of the residential model is shown in Table 2. All coefficients are significant at 95% confidence level. Looking at the percentage change, we can see that all variables have a positive effect on the residential housing price. Specifically, for the hosts, the increase brought by the announcement is 33.53%, for non-hosts, the increase by the same event was only 27.61%. The interaction of the two dummy variables represents the difference-on-difference effect. That is, the difference of means between the two time periods of the hosts, less the difference of means between the two time periods of the non-hosts, which is 5.92%.

#### 5.2.2 London

The result of the River model is shown in Table 3. All coefficients are significant at 90% confidence interval. Looking at the percentage change, we can see that being the River Zone lowers the housing price, but with the Olympics, its housing price actually increases. Specifically, for the River Zone (main host districts), the

Table 3: Coefficients for the London River Model

	Coefficient	t-value	$p >  t $	Coefficient Transformed	Percentage Change by One Unit Increase
Timestamp: after announcement	0.000	30.690	0.000	111303.471	32.21%
River Zone	0.000	-22.066	0.000	-84062.051	-24.33%
Time Stamp: after * River Zone	0.000	1.982	0.047	13828.585	4.00%
Constant	1.570	897000.000	0.000	345575.817	

increase brought by the announcement is 36.21%, for non-River zones, the increase was only 32.21%. The interaction of the two dummy variables represents the difference-in-difference effect. That is, the difference of means between the two time periods of the hosts, less the difference of means between the two time periods of the non-hosts is 4.00%.

Table 4: Coefficients for the London River and Central Model

	Coefficient	t-value	$p >  t $	Coefficient Transformed	Percentage Change by One Unit Increase
Timestamp: after announcement	0.000	31.044	0.000	98420.227	30.36%
River Zone	0.000	-19.185	0.000	-62690.136	-19.34%
Central Zone	0.000	26.164	0.000	150058.010	46.29%
Time Stamp: after* River Zone	0.000	2.290	0.022	13014.459	4.01%
Time Stamp: after* Central Zone	0.000	0.492	0.623	2932.350	0.90%
Constant	1.570	913000.000	0.000	324203.902	

Then we combined the River and the Central zone, its result is shown in Table 4. Only the coefficient for interaction of timestamp and Central Zone is not significant. Looking at the percentage change, we can see that only the River Zone would lower the housing price. Specifically, for the River Zone, the increase brought by the announcement is 34.37%, for the Other Zones, the increase brought by the same event was only 30.36%. The two interactions of host boroughs and timestamp represent the difference-in-difference effect. For the River Zone, its difference of means between the two time periods, less the difference of means between the two time periods of the Central and Other Zone is 4.01%. But the interaction of the Central Zone and timestamp is not significant, and the 0.90% increase is quite small.

## 6 Discussion

### 6.1 Understanding the Results

Based on these results, it is clear that hosting the Olympics positively impacted real estate prices in the areas that served as venues compared to other areas of the same city. At least for Sydney and London, the infrastructure improvements and other revitalization efforts seem to have left a lasting positive legacy.

While both cities experience and increase in real estate prices, the host areas in Sydney seem to have

benefited slightly more from the hosting of the Olympics, gaining 5.92% compared with 4% for London. There are a few possible reasons for why this might be the case. First, East London, where the Games were hosted, was a primarily industrial zone prior to the Games, and had depressed price levels relative to other areas of the city. Thus, one of the goals of the Games was to revitalize this area, which was accomplished in part by the construction of Queen Elizabeth Olympic Park, which converted marshlands into one of the largest urban parks in Western Europe, and certainly increased the desirability of the surrounding land. However, it is possible that these improvements did not fully impact real estate prices during the time period analyzed, which ended in 2014. Queen Elizabeth Olympic Park reopened to the public in that same year, and residential property values may have further increased in the following years once the park was reopened. Further, we categorized the areas of Sydney as host/non-host using the postal codes for the Olympic venues. This contrasts with our method of using host boroughs in London, each of which contains multiple post codes. This may have increased the relative difference between the specific host postal codes and the non-host postal codes, as it is a more localized approach. Under the assumption that the local areas around Olympic venues receive most of the benefit from those venues, the region defined by the venues' postal code would receive the most benefit, and prices would be higher relative to the rest of the entire host district.

## 6.2 Quantifying Total Impact

To quantify the total economic impact of the increases in property values attributable to the Olympics, we collected data on the housing stock by region for both London and Sydney, along with the average home price in both cities. We then multiplied the total housing stock by the average home price to estimate the total residential market value for the host areas, and applied our percentage change to this total value. In London, the impact represents about £1.5bn (\$1.88bn), while in Sydney the total impact was approximately 1.81bn AUD (\$1.27bn). Given that hosting the London Olympics is estimated to have cost \$11.4bn, and hosting the Sydney Olympics cost about \$6.9bn, these increases represent about 16.5% and 18% of the total cost of hosting the Olympics for London and Sydney respectively.

## 6.3 Future Work

While our analysis quantifies the indirect economic impact of the hosting Olympics on the real estate prices, it does not address the implications of these changes at a societal level. For example, rising residential real estate prices are great if you own the land, but if you are a renter, rents could rise faster than your income and you may be forced to relocate. This phenomenon manifests as gentrification of the host neighborhoods, which benefits some but seriously hurts others. Future work should assess the societal impacts of these changing real estate prices, perhaps by analyzing data related to incomes and ethnic makeup of the affected areas.

Furthermore, while Sydney and London saw a positive impact on their real estate markets from the Olympics, this may not be the case for all cities. Both Sydney and London benefited from relatively strong urban planning during hosting period, which weren't always the case. Future work could address this by taking a broader look at different Olympics over time, and assessing which factors contribute the most to positive outcomes from hosting the Games. These results could further be used to assist cities considering an Olympic bid, and help the IOC decide between candidate cities.

## 7 References

1. Sydney Inflation Data - <https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/6401.0March>
2. NSW Land Sales Data - <https://valuation.property.nsw.gov.au/embed/propertySalesInformation>
3. UK House Price Index - <https://www.gov.uk/government/statistical-data-sets/uk-house-price-index-data-downloads-january-2020>
4. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.366.200&rep=rep1&type=pdf>



5. <https://www.jstor.org/stable/26150896>
6. Cost of Hosting the Olympics - [https://www.researchgate.net/figure/Costs-of-Hosting-Recent-Olympic-Games\\_tbl2\\_301940770](https://www.researchgate.net/figure/Costs-of-Hosting-Recent-Olympic-Games_tbl2_301940770)
7. Sydney Mean Home Price - [http://www.econ.mq.edu.au/\\_data/assets/pdf\\_file/0018/220581/Abelson\\_9\\_04.pdf](http://www.econ.mq.edu.au/_data/assets/pdf_file/0018/220581/Abelson_9_04.pdf)