

## The effect of the Dubai Metro on the value of residential and commercial properties

Sara I. Mohammad

Imperial College London  
sara.mohammad07@imperial.ac.uk

Daniel J. Graham

Imperial College London  
d.j.graham@imperial.ac.uk

Patricia C. Melo

The James Hutton Institute  
Patricia.Melo@hutton.ac.uk

**Abstract:** This paper analyzes the impact of the newly operated Dubai Metro on the sale transaction value of dwellings and commercial properties. The effect is estimated for properties within different catchment zones of a metro station using difference-in-differences and hedonic pricing methods on both repeated cross-sectional data and pseudo panel data. Our estimates show a positive effect of the metro on sale values of both residential and commercial properties, although the effect is stronger for commercial properties. The models also reveal that the effect of the metro on the value of dwellings and commercial properties is largest within 701 to 900 meters of a metro station and is about 13 percent and 76 percent, respectively.

### Article history:

Received: March 26, 2014  
Received in revised form: March 8, 2015  
Accepted: July 18, 2015  
Available online: October 6, 2015

## 1 Introduction

A large number of researchers have examined the effect of rail systems on property values and the range of estimates varies substantially across studies (Mohammad et al. 2013). The majority of studies suggests that proximity to rail stations enhances property values (e.g., Laakso 1992; Pan and Zhang 2008; Voith 1991; Weinberger 2001), some indicate a negative impact at certain locations mainly due to negative environmental externalities (e.g., Bolling, Ihlanfeldt, and Bowes 1998; Cervero 2003; Du and Mulley 2006), and a few show no noticeable effect (Gatzlaff and Smith 1993).

The empirical literature has generally adopted hedonic pricing (HP) methods to estimate the relationship between rail and property values. Although HP models can control for unobserved heterogeneity across properties, they cannot identify a causal relationship between rail access and property value, only indicative results. For example, stations may be found at high-valued areas like a commercial hub zone, and HP models typically do not account for the impact of these attributes on the reported values (Billings 2011; Gibbons and Machin 2005).

As an improvement to the HP models, a number of studies have recently started using an innovation-based model, the difference-in-differences (DID) estimator (e.g., Agostini and Palmucci 2008;

---

Copyright 2015 Sara I. Mohammad, Daniel J. Graham & Patricia C. Melo

<http://dx.doi.org/10.5198/jtl.u.2017.750>

ISSN: 1938-7849 | Licensed under the [Creative Commons Attribution – Noncommercial License 3.0](https://creativecommons.org/licenses/by-nc/3.0/)

Billings 2011; Dewees 1976; Dubé, Thériault, and Des Rosiers 2013; Gibbons and Machin 2005; McDonald and Osuji 1995; McMillen and McDonald 2004). Although the application of this method differs across studies, it generally relies on comparing prices before and after the treatment for the properties that experienced the treatment (treated) compared to those that did not (control).

Some studies have also compared results obtained from DID to those from conventional HP models. In a study on the effect of new rail stations on dwelling prices in London, Gibbons and Machin (2005) found that HP models produced statistically larger effects than DID models. On the other hand, Agostini and Palmucci (2008) found that the DID models produced higher estimates of the effect of Santiago Metro on property values, compared to HP models. The difference in the magnitude of estimates between the DID and HP models in the two studies may be related to the data structure; while Agostini and Palmucci (2008) used repeated cross-sectional data, Gibbons and Machin (2005) used pseudo panel data.

This paper contributes to existing research in the following ways. It provides the first estimates of the effect of the Dubai Metro on residential and commercial property values using both HP and DID methods. Dubai Metro is the first metro system in the rapidly growing Middle East region, and hence this study is not only beneficial for Dubai but also for other similar cities in the Middle East. We estimate the effect of the metro on property values using a set of consecutive 500-meter distance bands, as well as smaller buffer zones around a station to examine how the magnitude of the effect changes with increased distance from the station.

In the absence of genuine panel data, we construct pseudo panel data from repeated cross-sectional data to control for bias in the results by comparing the same groups of properties over time. Therefore, the second contribution of this research is that for the first time in the literature, we estimate the effect of the metro using both repeated cross-sectional and pseudo panel data from the same datasets and draw conclusions on the most appropriate data structure. Third, we have created a new database that combines residential and commercial property data, with transport data and neighborhood and accessibility data, which can be used by other researchers.

Our findings suggest that the new metro has, in general, positively affected the sale value of residential and commercial properties, although more substantially for the latter. The nature and magnitude of the effect differs by type of property. The DID model suggests that the effect on the sale value of residential properties is significant and positive (7.8 percent) within 1 kilometer of a station, whereas the effect for commercial properties is significant and positive (41 percent) up to 1.5 kilometers. The results from the preferred HP model also indicate that the effect of the metro is significant and positive for residential (1.2 percent) and commercial properties (14 percent) within 1.5 kilometers of a station, but the magnitude effect is considerably lower. We also find that the peak of the impact of the metro on the value of residential and commercial properties occurs within 701 to 900 meters of a station and is equal to 13 percent and 76 percent, respectively.

This paper is structured as follows. Section 2 discusses the sources of bias in the empirical models of the effect of passenger rail systems on property values and the methods that can be used to address them. Section 3 describes the transport system and the property data for Dubai, while Section 4 explains the empirical models. The results are presented and discussed in Section 5. Section 6 concludes with the main research findings.

## **2 Sources of bias in the estimation of the effect of transport on property values**

Reverse causality, omitted variables, and unobserved heterogeneity are possible sources of bias that can occur in studies on the effect of a transport system on the value of properties. Reverse causality can occur

if, for example, properties in close proximity to a metro station are located in high-valued neighborhoods for reasons other than access to the metro, or these neighborhoods may have affected the choice of location of the metro. Only a limited number of studies has attempted to address the issue of reverse causality by using a DID method (e.g., Agostini and Palmucci 2008; Ahlfeldt 2013; Billings 2011; Dubé, Thériault, and Des Rosiers 2013; Gibbons and Machin 2005). However, we were not able to find any study that tested the hypothesis of reverse causality, i.e., that there is an association between the value of properties and the choice of metro location. Studies using the DID method distinguish between “treated” properties and “control” properties. Treated properties are defined as those located within a relatively close distance from a station, generally not more than 2 kilometers (e.g., Agostini and Palmucci 2008; Ahlfeldt 2013; Billings 2011; Gibbons and Machin 2005), while the remaining properties are defined as control. Alternatively, some researchers define the group of treated and control properties using a propensity score matching method (PSM) (Billings 2011; Concas 2012).

Another source of bias refers to omitted variables, or confounding. A number of researchers have controlled for the effect of factors that might be confounded with the transport system. These variables typically refer to accessibility to employment centers (e.g., Billings 2011; Dewees 1976; Du and Mully 2007; Dubé, Thériault, and Des Rosiers 2013; Gibbons and Machin 2005; Kim and Zhang 2005; Martinez and Viegas 2009; Wu 2012) and accessibility to major shopping centers (e.g., McDonald and Osuji 1995; Vichiensan and Miyamoto 2010). Failure to include these variables leads to potential inconsistent and biased estimates of the effect of the transport system on property values.

A more elaborate example is as follows. In estimating the effect of proximity to a light rail (LRT) station on the value of residential and commercial properties, Billings (2011) found significant differences in the characteristics of the treated and control communities for reasons other than access to LRT stations (e.g., income and crime levels). The study adopts a DID model and a distance-based HP model and argues that the results are unbiased by controlling for neighborhood attributes and using repeated cross-sectional data. Another example is Bowes and Ihlanfeldt (2001), which uses a HP model to study the effect of proximity to a rail station on the value of residential properties. Neighborhood attributes such as employment density, median income, or crime levels are included in the model and are found to have a significant effect on property values.

The third issue is unobserved heterogeneity across properties. Access to panel data can help correct for time-invariant unobserved heterogeneity, but a genuine panel data is rarely available. One way of overcoming this limitation is by constructing a pseudo panel dataset (Cameron and Trivedi 2005; Colorado 1997; Deaton 1985). A pseudo panel dataset is formed by grouping individual observations in the repeated cross-sectional data into cohorts. The mean value of the individual observations in a cohort is the value of the pseudo panel data. Therefore, a value for each cohort is available for several periods of time even if the individuals in that cohort are not the same over time. A very limited number of studies used pseudo panel data to estimate the effect of rail on property values (e.g. Ahlfeldt 2013; Gibbons and Machin 2005). The majority of studies have used repeated cross-sectional data (e.g., Agostini and Palmucci 2008; Billings 2011; Dubé, Thériault, and Des Rosiers 2013; Koster, Ommeren, and Rietveld 2010; McMillen and McDonald 2004; Martinez and Viegas 2009; Weinberger 2001). To our best knowledge, no study has used both repeated cross-sectional data and pseudo panel data in the same analysis.

Both the data structure and the estimation method play important roles in controlling for the different sources of bias. The majority of researchers control for unobserved heterogeneity and omitted variable bias using repeated cross-sectional data, and a few use panel data or pseudo panel data. The next section describes the transport system and the property data, while the empirical methods are presented in Section 4.

### 3 The transport innovation and property data

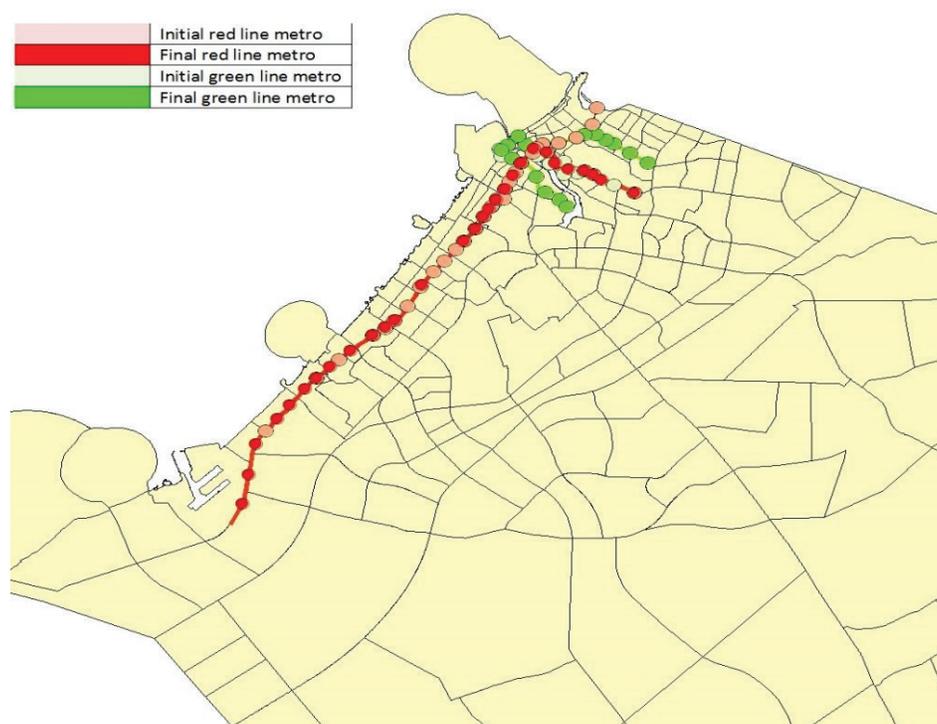
#### 3.1 The transport innovation

Dubai has been growing rapidly. The population increased from 827,000 in 2000 to over 2.1 million in 2012. This has led to a strong increase in the demand for transport services. To cater to high travel demand in a car-oriented environment, the Road and Transport Authority (RTA) in Dubai has implemented measures to promote public transport. These include the enhancement of the public transport network of buses and marine services and the building of the first metro in the Middle East (Figure 1).

The initial study on the construction of a metro system in Dubai was conducted by Dubai Municipality in 1992, and the study of the potential location for the metro was conducted in 2003 (Dubai Municipality 2005). The public announcement of the metro was made in July 2005, and construction started nine months later. The first 10 stations of the red line opened in September 2009, while the rest of the stations of the red line started operating in May 2010. The green line opened in late 2011. Currently, Dubai Metro is the longest driverless metro in the world. The red line is 52 kilometers long with 29 stations, and the green line is 22 kilometers long with 18 stations. The metro carried around 36 percent of the total public transport passengers in Dubai in 2011 (Dubai Statistics Center 2011).

The metro was proposed to support the increasing travel demand to the major existing and future employment centers, shopping malls, and Dubai Airport. The potential demand was calculated using a transport model for the year 2010 and included details on the distribution of land uses, shopping densities and employment densities (Dubai Municipality 2003). The major part of the red line route runs along one of the main road corridors in Dubai (known as Sheikh Zayed Road), along which are located a large employment center and five of the main shopping malls. It connects residential properties located close to either end of the line with employment areas and major shopping malls located in the middle of the line. The green line serves mainly the old central business district (CBD) area. The alignment of the metro was chosen to serve the areas with the highest estimated and projected trips in Dubai and with available right of way.

The initial and final routes of the metro are very similar (see Figure 1), with the exception of a few changes. The initial route consisted of a (red) line running from the west to the east of Dubai, whereas the final route diverts southeast and includes a second (green) line. In addition, more stations were added to the green line, especially in the old CBD area, as well as in the southwestern part of the Dubai Creek, which serves a planned employment center. The final red line southeast of Dubai Creek was initially part of the green line.



**Figure 1:** Initial and final routes of the Dubai Metro

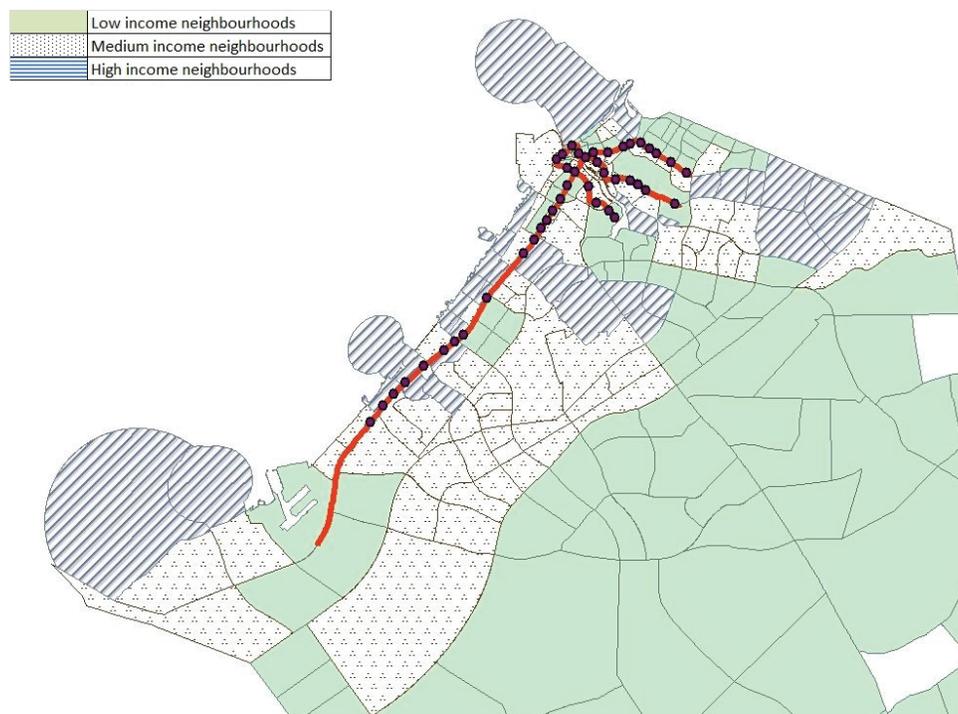
Source: Dubai Municipality 2003

To check for potential reverse causality between property values and the location of the metro stations, we consider whether the route may have been selected to serve neighborhoods with higher income levels. Unfortunately, there are no data available for actual income, only an indication of income level according to three levels: low, medium, and high income. Figure 2 shows the classification of communities into high-, medium-, and low-income communities<sup>1</sup>. We can see that the metro serves communities with different income levels and there is no obvious pattern between higher income and the location of stations. As a result, we argue that the choice of metro location does not depend on the income level (and thus property value) of a given area, and hence reverse causality is not likely to be an issue.

Nevertheless, the choice of station location is based on the spatial distribution of employment and shopping densities. In fact, as mentioned earlier, one of the reasons for building the metro system was to serve existing and future major employment centers, existing and future major shopping malls, and the Dubai Airport. Therefore, it is important to consider these variables in the empirical analysis. We provide a summary of the characteristics of the communities served by the metro (i.e., treated) and those not served (i.e., control)<sup>2</sup> in Table 1. The table provides summary statistics for population, employment, and shopping densities for all treated and control communities in Dubai and for the communities in our dataset.

<sup>1</sup> The classification of income levels is based on the rental indices for residential properties in a community, which was chosen in association with RERA.

<sup>2</sup> The definition of treated and control groups in the context of this study is provided in more detail in Section 3.2



**Figure 2:** The location of Dubai metro and community income level

Source: Authors

While the mean population density for all treated and control communities is different, it is similar for the sample of treated and control communities in our datasets. Nonetheless, the mean employment and shopping densities for treated communities is considerably higher than for control communities for both the population of communities in Dubai and the sample of communities in our datasets. This substantial difference between the treated and control communities confirms that the metro was chosen based on the location of employment and shopping destinations.

In addition, we assume that the new metro does not lead to redistribution of major employment and shopping destinations during the time period of our study (2007-2011). This is a reasonable assumption, since both the land development and land use distribution in Dubai had been fixed for that period at the time of the announcement of the metro in 2005. Nevertheless, employment densities have changed across Dubai due to an increase in the supply of offices; therefore, we need to control for employment density in our models.

**Table 1:** Summary of community characteristics

Variable	Treated communities		Control communities	
	Mean	Standard deviation	Mean	Standard deviation
All communities in the urban area of Dubai				
Population density*	16,600	25,500	3700	6200
Employment density*	8100	13200	1300	3200
Shopping density**	1543	4344	937	1975
Communities in our datasets				
Population density*	2300	2600	1600	1200
Employment density*	1300	1200	100	200
Shopping density**	1138	1598	578	1079

\* These are measured as the number of population/employment per square kilometer.

\*\* This is measured as the number of shopping commuters to a community in a peak hour.

### 3.2 Description of the property data

The effect of the Dubai Metro on property values is tested using two datasets provided by the Dubai Real Estate Regulatory Authority (RERA) for sale transaction records of residential and commercial properties. RERA provides individual records on property data from 2007 to 2011. Due to confidentiality reasons, we do not provide the raw data and do not report the actual property values, only the mean and the standard deviation (table 2), as well as the results from the regression analyses (Section 5). Table 2 shows that the mean values per unit area of residential and commercial properties are larger for the treated groups than for the control groups. In addition, the value of the standard deviation is also larger for the treated groups, which indicates there is more variation in the value of treated properties than in the value of control properties.

**Table 2:** Mean and standard deviation of the value of residential and commercial properties

Sale value per unit area	Residential properties		Commercial properties	
	Treated	Control	Treated	Control
Dirhams per square meter				
Mean	11,253	7418	10,192	7280
Standard deviation	19,250	3881	5572	2636

Similar to previous empirical work, treated properties are defined as those located within a given distance of a station, while control properties are those located at a greater distance (e.g., Agostini and Palmucci 2008; Ahlfeldt 2013; Billings 2011; Gibbons and Machin 2005). For example, Agostini and Palmucci (2008) divided properties to those within 1 kilometer (treated) and those beyond (control). Billings (2011) defined treated properties using two catchment zones, within 0.5 mile (0.8 kilometer) and within 1 mile (1.6 kilometers) of a station, whereas Bowes and Ihlanfeldt (2001) classified properties using a set of contiguous catchment zones between 0.25 miles (0.4 kilometers) and 3 miles (4.8 kilometers) from a station.

In this study, we consider three catchment areas—0.5 kilometer, 1 kilometer, and 1.5 kilometers—for the DID and HP models. For example, defining treated properties at 0.5 kilometer means that all properties located at a distance equal to or less than 0.5 kilometer from a station are considered treated, while those located at a greater distance are considered control. Hence the results for each model using a particular catchment area (i.e., 0.5, 1, or 1.5 kilometers) represent the effect of the metro on the value of properties located within that catchment area, compared to all other properties located outside that catchment area.

To estimate the effect of the metro on property values, we obtained data pre- and post-opening of the metro in 2009. Depending on the availability of property data, researchers estimating the effect of a transport system on property values have used data either from pre-announcement time to post-operations or from after-announcement (i.e., during construction) to post-operations. For example, Ahlfeldt (2013) and Gibbons and Machin (2005) used property data during construction (and define it as pre-treatment) and post-operations (as post-treatment). On the other hand, Agostini and Palmucci (2008) used property data from pre-announcement to during construction, whereas Billings (2011) and Concas (2012) used property data from pre-announcement to post-operations.

In the ideal situation, one would obtain data from pre-announcement of the transport system until a few years after operations started. In the case of Dubai, we were only able to use data from 2007 (i.e., during construction) until 2011 (within 2 years of the metro operations). Therefore, we can only capture the short-term effect of the metro. In addition, because our data do not cover the pre-announcement period, we expect that the impact of the Dubai Metro estimated in this study may be undervalued. In

summary, we can only estimate the marginal effect of the operations phase of the metro, and we are not able to capture initial land market reactions to the announcement of the metro. Other studies that also use data from after-announcement include Ahlfeldt (2013) and Gibbons and Machin (2005).

Although the red line started operations in September 2009, this year is chosen to be a pre-treatment year for the following reasons. First, only 10 of 29 stations operated in that year and it happens that all the treated properties in our sample are located within 1.5 kilometers of the stations that operated in the year 2010. In addition, all control properties in our datasets are located at a distance greater than 1.5 kilometers from any opened or planned metro stations. This happens to be the case in Dubai as the available transactions and listings records are for properties located in the relatively newly developed communities in Dubai, where properties experienced the most transaction activities, and the majority of the metro stations that operated in the year 2009 were located in the older and less-active communities. Therefore, 2009 is a pre-treatment year for all the properties in our datasets.

Second, the metro in Dubai is the first of its kind in the whole Middle East region, and given that it operates in a car-oriented city, the perception of the residents toward the metro within the first few months of operation and the possible impact of the metro on the value of properties are likely less obvious compared to cities in which railways are already available and the residents have experience and a perception of the system benefits. In addition, other researchers also have identified the opening year as a pre-treatment year (e.g., Concas 2012; Gibbons and Machin 2005). In addition, the datasets did not identify the sale month of the properties, and hence we are not able to distinguish the treatment time per month. With regard to the possibility that properties may have received double treatment by being close to more than one station, we do not consider the amount of treatment, only whether the property received a treatment or not (i.e., whether it is close to at least one station).

The effect of the metro on property values is tested up to a maximum distance of 1.5 kilometers of a metro station. This threshold is chosen for various reasons. First, it is based on what appears to be an acceptable distance of users to a metro station, according to a survey conducted by RTA, which revealed that accessibility via walking is up to 1.2 kilometers, and kiss-and-ride, park-and-ride and/or feeder buses for distances beyond 1 kilometer (Roads and Transport Authority 2011). Second, a large number of previous studies have used similar thresholds, between a 1- to 2-kilometer distance from a rail station (e.g., Agostini and Palmucci 2008; Du and Mulley 2006; Gibbons and Machin 2005).

To account for additional factors that can affect property values (other than property characteristics, property location and transaction year), we enrich our original dataset by adding contextual variables that need to be included in the empirical models: distance to main amenities (schools, hospitals and shopping malls); employment density; shopping density; as well as measures of transport accessibility (i.e., distance to the nearest highway, distance to the nearest metro station, and the number of metro stations within a given catchment area). All these variables are matched manually to each property record. Table 3 provides some basic descriptive statistics for these variables.

**Table 3:** Descriptive statistics of contextual variables

Variable	Residential properties				Commercial properties			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Property area (sq m)	142.6	96.9	16	1155	94.8	62.4	14	1114
Distance to nearest school (m)	1992.2	1133.9	100	3990	1992.2	1133.9	100	3990
Distance to nearest hospital (m)	3883.4	1274.0	1513	5725	3716.2	1256.0	789.8	5604.3
Distance to nearest shopping mall (m)	902.2	581.3	71	2567	624.4	603.7	75.0	2494.2
log employment density	5.6	1.9	3.7	7.2	4.7	1.5	3.7	8.3
log shopping density	5.4	2.2	1.9	9.1	7.1	2.4	1.9	9.1
Distance nearest station (m)	2767.2	3667.1	149	12,000	4996.5	4372.4	187	13,750
Number of stations within 1.5 km	0.9	1.2	0	3	0.6	1.1	0	3
Distance nearest highway (m)	3901.6	2474.8	844	13,940	2898.0	2254.6	1000	7538.1

### 3.3 Creating pseudo panel data

In the absence of a genuine panel data, pseudo panel data may be created to control for temporal variation within the “same” groups. We consider different criteria to create pseudo panel data from repeated cross-sectional data. Previous studies have created pseudo panel data using either one or a number of time-invariant measures. The main challenge is to balance the sample size (i.e., number of cohorts) with the cohort size (i.e., number of records within a cohort). There is no right grouping criterion; however, the aims are to group individual records to homogenous cohorts, assuming that observations within a group share almost the same unobserved factors, ensure heterogeneity across groups, and retain the most optimal sample size.

In almost all cases, the unobserved factors of a given cohort in a pseudo panel data can change over time because individual records within a cohort can change. Researchers argue that if the cohort size is large enough, the unobserved factors of the constructed cohort can be approximated to the actual unobserved factors of that cohort (Deaton 1985; Tsai et al. 2013). Although Verbeek and Nijman (1992) show that a minimum of 100 observations per cohort provide a sufficient number of records for the pseudo panel data to reduce estimation bias, one can also argue that the sufficient number of observations per cohort may depend on the percentage of observations per cohort to the total number of observations in that group. However, existing literature does not suggest any percentage threshold that can be used as guidance.

We have considered different criteria to create pseudo panel data. Although all the grouping choices are time-invariant, they differ in the level of aggregation. Grouping records to the plot of land in which they are sited provides a reasonable level of aggregation without losing many records. This grouping criterion controls for the effect of the geographical location of property and allows for the unobserved factors related to the building and the plot of land to be captured. However, this may still lead to considerable differences in property characteristics within a cohort, especially if the number of properties

within a given parcel or building is large. As a result, within-group variation is likely to be large and close to between-group variation.

To reduce the degree of heterogeneity within cohorts, a second option is used that groups records to both the plot of land and (a measure of) property size. This accounts for the effect of the unobserved characteristics of properties that can arise due to its size. Properties are grouped by area for every 150 square meters in size except for those less than 50 square meters as they correspond mainly to studio apartments and hence are expected to have unique characteristics.

To validate the grouping criteria, we compared between- and within-cohort standard deviations of the variables included in our models for the two options considered. The results confirmed that the second grouping option provides larger between- than within-cohort variation. The average cohort in the residential and commercial datasets contains 17 observations. Hereafter the constructed cohorts are referred to as pseudo panel groups. The two data structures, repeated cross-sectional data and pseudo panel data, are summarized in Table 4.

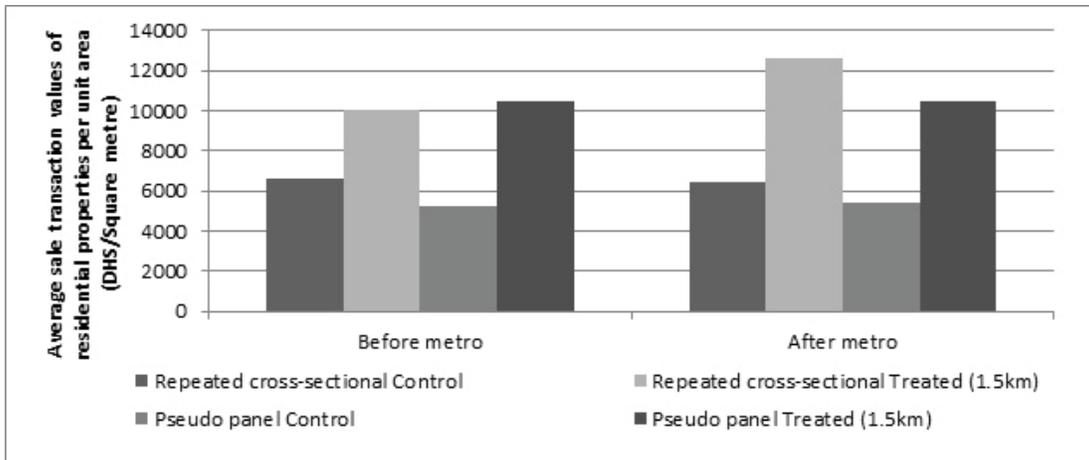
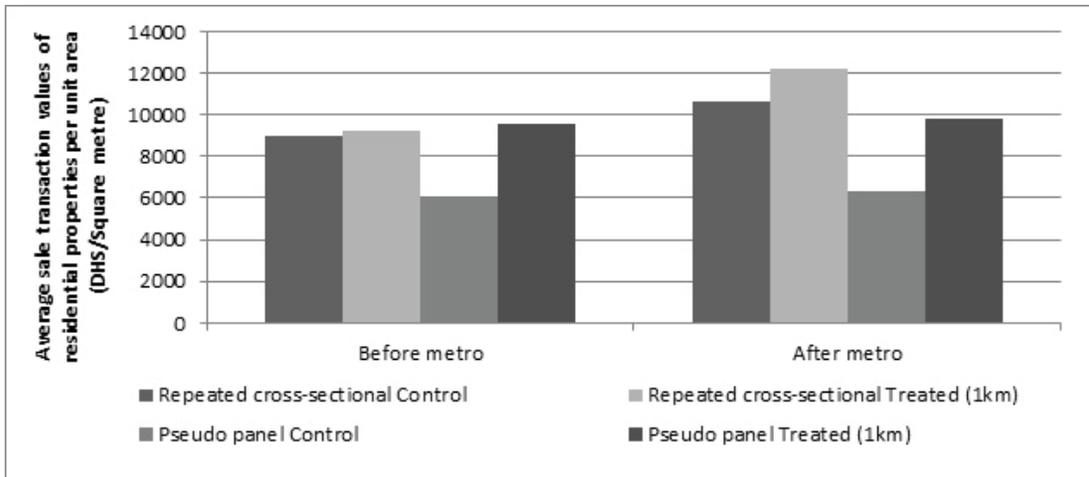
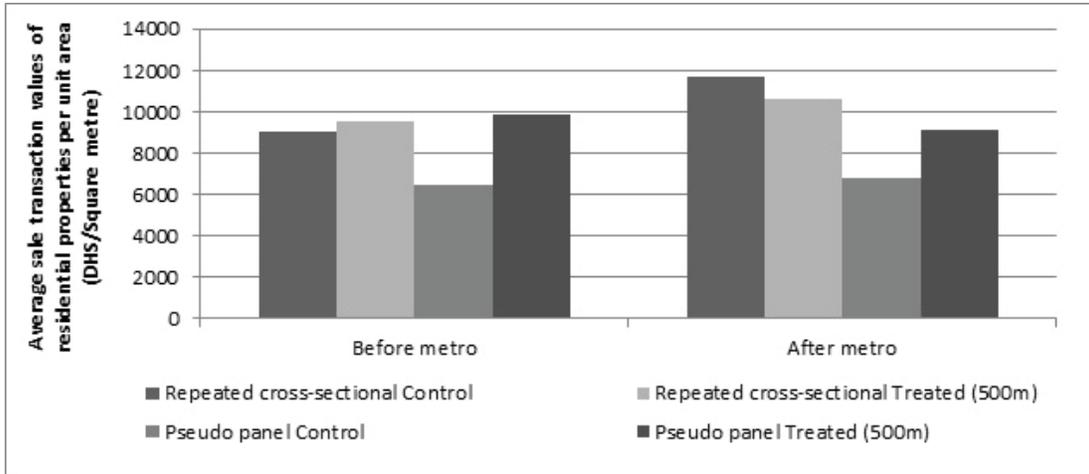
**Table 4:** Summary statistics of the data structures

<b>Data type</b>	<b>Residential properties</b>	<b>Commercial properties</b>
Panel years	2007-2009 (before metro); 2010 – 2011 (after metro)	
Coefficient of variation of property value	1.5 (repeated cross-sectional data) 0.76 (pseudo panel data)	1.12 (repeated cross-sectional data) 1.49 (pseudo panel data)
Number of observations in the repeated cross-sectional data*		
Total number	39,308	3419
Properties within 0.5 km of a metro station	7911 (20%)	708 (21%)
Properties beyond 0.5 km of a metro station	31,397 (80%)	2711 (79%)
Properties within 1 km of a metro station	21,326 (54%)	1237 (36%)
Properties beyond 1 km of a metro station	17,982 (46%)	2182 (64%)
Properties within 1.5 km of a metro station	28,959 (74%)	1510 (44%)
Properties beyond 1.5 km of a metro station	10,349 (26%)	1909 (56%)
Number of observations in the pseudo panel data*		
Total number	3344	336
Cohorts within 0.5 km of a metro station	223 (7%)	75 (22%)
Cohorts beyond 0.5 km of a metro station	3121 (93%)	261 (78%)
Cohorts within 1 km of a metro station	573 (17%)	171 (51%)
Cohorts beyond 1 km of a metro station	2771 (83%)	165 (49%)
Cohorts within 1.5 km of a metro station	891 (27%)	280 (83%)
Cohorts beyond 1.5 km of a metro station	2453 (73%)	56 (17%)

\* The cells contain the number of properties and the figure in brackets is the percentage of observations to the total sample size.

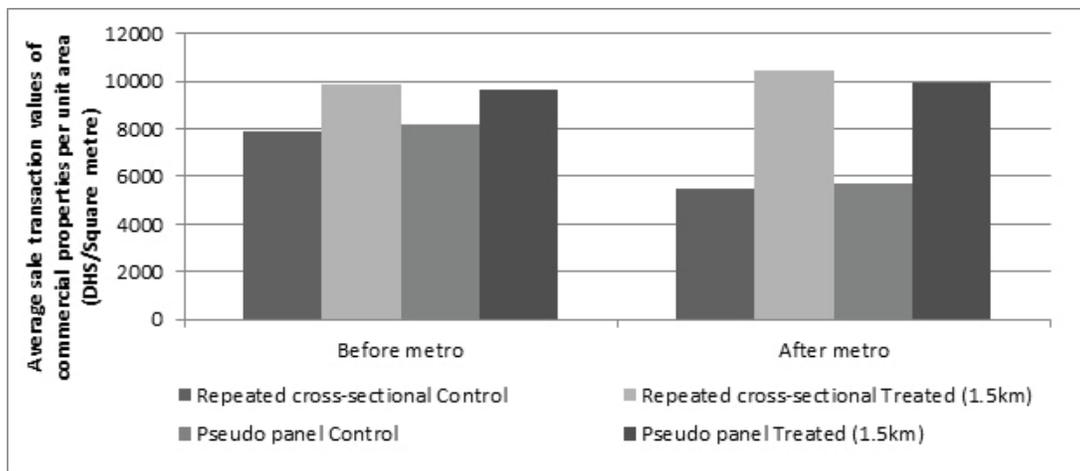
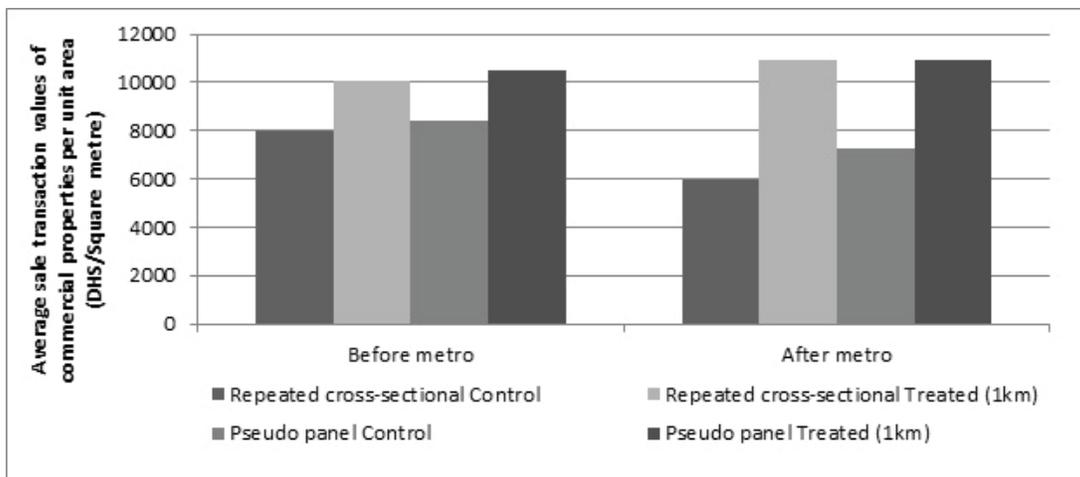
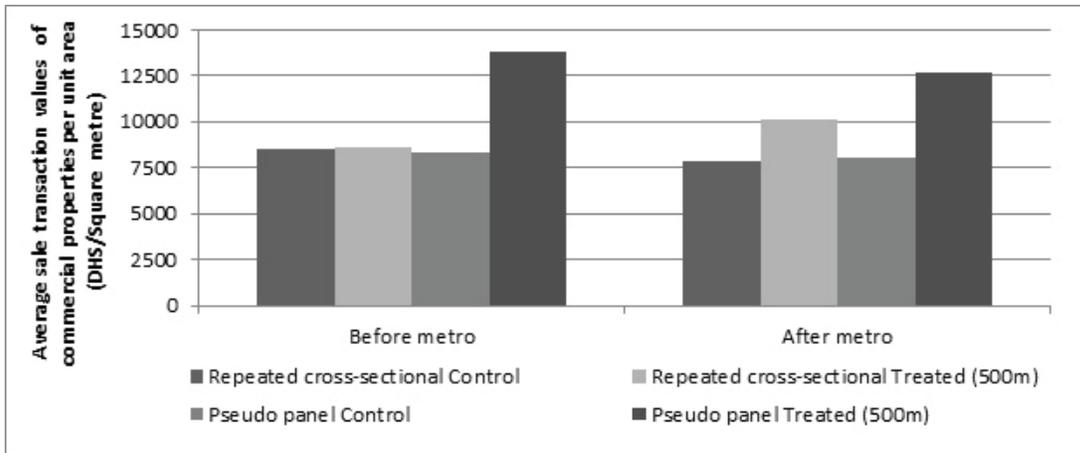
We compare the average property values per unit area between the pseudo panel data and the repeated cross-sectional data in Figures 3 and 4. In general, we find that the values are similar for the catchment area defined at 1.5 kilometers, followed by the catchment areas based on 1 kilometer and 0.5 kilometer. This is because with larger distance bands the number of properties in each group increases, which in turn results in more comparable average values between the two data structures.

We acknowledge that, by creating pseudo panel groups, some information is lost. The difference in average property values between the repeated cross-sectional data and the pseudo panel data indicates that measurement error occurring in the pseudo panel data may bias estimations. We expect that the larger the difference, the more dissimilar the results generated using the two data structures.



**Figure 3:** Average sale transaction values of residential properties per unit area before and after the metro for repeated cross-sectional data and pseudo panel data for different catchment zones

Source: Authors



**Figure 4:** Average sale transaction values of commercial properties per unit area before and after the metro for repeated cross-sectional data and pseudo panel data for different catchment zones

Source: Authors

## 4 Empirical methodology

In this section we describe the models used to estimate the impact of the Dubai Metro on property values. The first is a DID model that examines the effect of proximity to the metro. The second and third are both HP models of the effect of accessibility to one or more metro stations, but they differ in the way they specify accessibility.

### 4.1 Difference-in-differences models

The DID method compares the change in the value of the treated group with the change in value of the control group before and after the operations of the metro. Treated properties are those located within 0.5 kilometer, 1 kilometer, and 1.5 kilometer of a station, whereas control properties are those located beyond these catchment areas.

Equations 1 and 2 show the DID models for the repeated cross-sectional data and the pseudo panel data, respectively. The index  $j$  denotes the building to which a property belongs to in equation 1 and the pseudo panel group in equation 2. The term  $\mu_j$  controls for time-invariant unobserved factors related to the building (equation 1) and the pseudo panel group (equation 2). It is worth mentioning that, for pseudo panel data, the group effect (D) is eliminated to avoid model over-specification as the same group repeats over time and the fixed-effect is on the group level. The term  $X$  refers to property related characteristics, while  $L$  refers to neighborhood attributes such as distance to nearest amenity (e.g., highway), and  $C$  refers to location attributes such as employment density. Table 3 provides a summary of these variables. More information on the basic DID model is presented in Appendix A.

$$\ln y_{i(j)t} = \alpha + \alpha_T T + \gamma_D D + \beta(T.D) + \alpha_x X_{i(j)t} + \alpha_L L_{jt} + \alpha_C C_{jt} + \mu_j + \varepsilon_{i(j)t} \quad (1)$$

$$\ln y_{jt} = \alpha + \alpha_T T + \beta(T.D) + \alpha_x X_{jt} + \alpha_L L_{jt} + \alpha_C C_{jt} + \mu_j + \varepsilon_{jt} \quad (2)$$

### 4.2 Hedonic pricing models

The hedonic pricing models estimate the effect of accessibility to metro stations within defined catchment areas. We consider two versions, formalized in equations (3)-(4) and equation (5), respectively.

The first version considers the effect of proximity to a metro station, measured by the variable  $N_j$ , that is, the number of metro stations (0, 1, 2, 3, etc.) within a given distance (i.e. 0.5 kilometer, 1 kilometer, and 1.5 kilometers) from unit  $j$ . Similarly to the DID models, index  $j$  denotes the building in which a property is sited in for the repeated cross-sectional data (equation 3) and the cohorts for the pseudo panel data (equation 4). We also account for temporal changes through a year trend ( $t$ ) and include contextual variables to control for property and neighborhood characteristics (included in the terms  $X$  and  $L$ , respectively). In addition, to reduce the potential for bias due to unobserved heterogeneity and confounding in the treatment effect, we further include location-specific variables ( $C$ ) for employment and shopping densities. Table 3 provides a summary of these variables.

$$\ln y_{i,j} = \alpha + \alpha_t t + \beta N_j + \alpha_x X_{i(j)t} + \alpha_L L_{jt} + \alpha_C C_{jt} + \mu_j + \varepsilon_{i(j)t} \quad (3)$$

$$\ln y_{jt} = \alpha + \alpha_t t + \beta N_j + \alpha_x X_{jt} + \alpha_L L_{jt} + \alpha_C C_{jt} + \mu_j + \varepsilon_{jt} \quad (4)$$

Similar to previous literature (Billings 2011; Bowes and Ihlanfeldt 2001; Duncan 2008; Efthymiou and Antoniou 2013; Pan and Zhang 2008; Weinberger 2001), the DID and HP models above estimate

the effect of the metro separately for different catchment zones (i.e., 0.5 kilometer, 1 kilometer, and 1.5 kilometers). Ideally, we would have used smaller distance bands; however, this was not possible for the analyses using pseudo panel data due to the resulting limited number of observations in each distance band.

Nonetheless, we can estimate an additional HP model using smaller distance bands for the repeated cross-sectional data. This allows us to provide a finer description of the spatial impact of the metro on the value of residential and commercial properties. We set a threshold of a minimum of 30 observations per distance band, which results in a set of consecutive 200-meter distance bands, starting at 300 meters. The model is described in equation (5). The only difference compared to equation (3) is that we now use a set of dummy variables  $d$  to identify the series of consecutive 200-meter distance bands.

$$\ln y_{i(j)t} = \alpha + \alpha_t t + \sum_k \beta_k d_{jk} + \alpha_x X_{i(j)t} + \alpha_L L_{jt} + \alpha_C C_{jt} + \mu_j + \varepsilon_{i(j)t} \quad (5)$$

## 5 Results and discussion

The results obtained from the DID model and the two versions of the HP models are reported in Tables 5, 6 and 7, respectively, for dwellings and commercial properties. We only show the findings for the effect of the metro on property value, the main purpose of this study. Tables 5 and 6 present the effect of the metro estimated separately for the three catchment zones (i.e., 0.5 kilometer, 1 kilometer, 1.5 kilometers), while Table 7 presents the effect of the metro on property values over a range of contiguous 200-meter distance bands starting at 300 meters (i.e., 300 meters, 0.5 km, 0.7 km, ... 1.5 km).

The majority of previous studies adopted a random-effects (RE) estimator to examine the effect of a transport system on property values. However, in the presence of omitted variable bias, this estimator will lead to inconsistent results. As an alternative, the fixed-effects (FE) estimator provides consistent results, but the results generated may be inefficient when within-group variation is small. Some researchers argue, and demonstrate, that when within-group variation is small, the widely used Hausman test to decide between estimator type (FE vs. RE) is problematic, as it relies on both the between- and within-variations to be large enough (Clark and Linzer 2013; Hahn, Ham, and Moon 2011; Plümper and Troeger 2007).

Clark and Linzer (2013) used Monte Carlo Simulation to decide on the estimator that produced consistent results in different circumstances, basing their decision on the least Root Mean Square Error (RMSE). In cases like ours (i.e., low within-variation, many and relatively small cohorts and low correlation between unobserved factors and regressors<sup>3</sup>), they showed that the RE estimator produced more reliable results than the FE estimator. Similarly, Tsai et al. (2013) also used Monte Carlo Simulation on pseudo panel data and found that, in cases where between-group variation is larger than within-group variation, the RE models outperformed FE models based on RMSE. This was the case even in the presence of moderate correlation between the explanatory variables and the unobserved factors.

The discussion above suggests that one should decide on the more suitable estimator depending on the degree of within- and between-group variation. The property records in our dataset exhibit low within-group variation, hence we select the RE estimator as preferable to the FE estimator. We discuss the results for the preferred RE estimator but also report the results for the FE models in Appendix B.

### 5.1 Difference-in-differences models

The results of the DID models are presented in Table 5 and discussed below, using equation 1 for the repeated cross-sectional data and equation 2 for the pseudo panel data. The models for residential and

<sup>3</sup> We have checked for the correlation between unobserved factors and regressors by conducting a correlated random effects model. The results show low correlations.

commercial properties explain between 21 and 35 percent and 50 and 62 percent of the variation in property values, respectively. Starting with the sale value of residential properties, the level of statistical significance is relatively similar between repeated cross-sectional data and pseudo panel data, except for the 1-kilometer catchment zone.

We estimate a significant positive impact of the Dubai Metro on residential property values, except for properties very close to the metro (i.e., within 0.5 kilometer) for which the effect is negative. The effect of the metro on the value of residential properties within 0.5 kilometer is estimated to be -9 percent and -17.7 percent for the repeated cross-sectional and pseudo panel data, respectively, compared to properties located farther away. A possible explanation is that properties located very close to metro stations may be adversely affected by increased noise levels (see, Diaz 1999; Du and Mulley 2006). For example, Du and Mulley (2006) estimated the effect of proximity to light rail stations and found a negative effect (between 5 and 40 percent) on the value of residential properties located within 200 to 500 meters.

On the other hand, the metro uplifted the sale value of residential properties within 1 kilometer by 7.8 percent and 4.3 percent using repeated cross-sectional data and pseudo panel data, respectively, although the latter estimate is not statistically significant. As for the 1.5-kilometer distance band, the model fails to identify a statistically significant effect. The difference in the magnitude of the effect of the metro on property values between the two data structures is likely to result from measurement error due to averaging of observations in the pseudo panel data.

We now consider the results for commercial properties. In contrast to dwellings, we find an overall positive and significant effect across the three catchment zones. In addition, the magnitude of the impact of the metro is also much higher. The results obtained using repeated cross-sectional data suggest that property values are uplifted by 39 percent, 42 percent, and 41 percent within 0.5 kilometer, 1 kilometer, and 1.5 kilometers of a metro station, respectively. However, we find a significant effect of the metro on commercial property values only for properties located within 1 kilometer (+30 percent) and 1.5 kilometers (+46 percent) using pseudo panel data.

**Table 5:** The effect of proximity to a metro station in Dubai on sale values of residential and commercial properties (DID equations 1 and 2)

Dataset	Residential properties (RERA)	
	Repeated cross-sectional	Pseudo panel
Data structure	Repeated cross-sectional	Pseudo panel
Estimator type	RE	RE
No. Observations	39,308	3344
Catchment zone	Defining treated properties at 0.5 km of a station	
Group * time effect	-0.090*** (0.014)	-0.177*** (0.054)
Overall R <sup>2</sup>	0.3361	0.2872
Catchment zone	Defining treated properties at 1 km of a station	
Group * time effect	0.078*** (0.008)	0.043 (0.039)
Overall R <sup>2</sup>	0.3467	0.2366
Catchment zone	Defining treated properties at 1.5 km of a station	
Group * time effect	0.018 (0.013)	0.033 (0.038)
Overall R <sup>2</sup>	0.3416	0.2275

Dataset	Commercial properties (RERA)	
	Repeated cross-sectional	Pseudo panel
Data structure	Repeated cross-sectional	Pseudo panel
Estimator type	RE	RE
No. Observations	3419	336
Catchment zone	Defining treated properties at 0.5 km of a station	
Group * time effect	0.394*** (0.054)	0.166 (0.114)
Overall R <sup>2</sup>	0.5080	0.5070
Catchment zone	Defining treated properties at 1 km of a station	
Group * time effect	0.420*** (0.037)	0.303*** (0.078)
Overall R <sup>2</sup>	0.6065	0.5379
Catchment zone	Defining treated properties at 1.5 km of a station	
Group * time effect	0.405*** (0.036)	0.458*** (0.096)
Overall R <sup>2</sup>	0.5366	0.5441

Legend: \*, \*\*, \*\*\* indicate significance at 10 percent, 5 percent, and 1 percent, respectively. Values in parentheses are standard errors. Note: Although other variables as listed in Table 3 (e.g., property characteristics, neighborhood characteristics, distance to amenities) are included and estimated in the model, the results above only present the impact of the Dubai Metro, as it is the main purpose of this research. The results for the contextual variables are available from the authors upon request.

## 5.2 Hedonic pricing models

We start with the model for equations (3) and (4), shown in Table 6. The models explain between 34 percent and 38 percent of the variation in residential property values and between 59 percent and 63 percent of the variation in commercial property values.

Similar to the DID models, the effect of metro on sale values for residential properties using the first version of the HP models tends to be negative for properties very close to the metro but positive for larger distance bands. The model suggests that the value of properties within 1 kilometer and 1.5 kilometers increased by about 2 percent and 1.2 percent, respectively, using repeated cross-sectional data. However, the results for the models using pseudo panel data are not significant. The difference in statistical significance between the repeated cross-sectional data and the pseudo panel data is due to the

difference in the average values between the individual records and the constructed cohorts, as shown in Figure 3.

As for the effect of the metro on the value of commercial properties, we find an overall positive and significant impact. The model for repeated cross-sectional data suggests an increase in values of about 36 percent, 34 percent, and 14 percent for properties within 0.5 kilometer, 1 kilometer, and 1.5 kilometers of a station, respectively. The model using pseudo panel data also indicates a positive and significant impact on commercial property values. The results show enhancements in value of 23 percent, 25 percent, and 18 percent for properties within 0.5 kilometer, 1 kilometer, and 1.5 kilometers, respectively.

The estimates obtained from repeated cross-sectional data and pseudo panel data are more comparable for commercial properties than for residential properties. A possible explanation is the larger consistency of average values between the two data structures in the commercial property dataset compared to the residential property dataset (see Figures 3 and 4). Similarly, we find that the coefficient of variation of property values in the repeated cross-sectional data versus pseudo panel data is more comparable for commercial properties (Table 4).

The results of the second version of the HP model (equation 5) using consecutive 200-meter distance bands between 300 meters and 1.5 kilometers (i.e., 300 meters, 0.5 kilometer, 0.7 kilometer, ... 1.5 kilometers) are presented in Table 7 and Figure 5. We obtain a concave shape for the relationship between property value and the distance to the metro. Similar to the previous models, the impact of the metro is substantially larger for commercial properties than for residential properties. Moreover, there appears to be some discontinuity in the significance of the metro effect for residential properties, particularly for distances up to 300 meters, between 501 and 700 meters, and greater than 1.1 kilometers.

As with the DID models, there is a negative impact (-6 percent) for properties located very close to the metro, that is, between 301 and 500 meters of one or more metro stations. The peak of the positive impact of the metro on sale values occurs for properties located between 701 meters to 900 meters (+13 percent) of a metro station. This is followed by a 10 percent increase in sale values for properties located between 901 meters and 1.1 kilometers, after which there is no significant impact. In contrast, the effect of the metro on commercial property values is significant across all distance bands and ranges between 25 percent and 76 percent; it reaches its peak for properties located between 701 and 900 meters (+76 percent) of a metro station.

**Table 6:** The effect of proximity to a number of metro stations in Dubai on sale values of residential and commercial properties (HP equations 3 and 4)

Datasets	Residential properties	
	Repeated cross-sectional	Pseudo panel
Data structure	Repeated cross-sectional	Pseudo panel
Estimator type	RE	RE
No. Observations	39,308	3344
Catchment zone	Defining treated properties at 0.5 km of a station	
Accessibility to a number of metro stations	-0.086*** (0.013)	-0.031 (0.051)
Overall R <sup>2</sup>	0.3766	0.3411
Catchment zone	Defining treated properties at 1 km of a station	
Accessibility to a number of metro stations	0.019* (0.010)	0.004 (0.034)
Overall R <sup>2</sup>	0.3805	0.3413
Catchment zone	Defining treated properties at 1.5 km of a station	
Accessibility to a number of metro stations	0.012** (0.005)	0.002 (0.015)
Overall R <sup>2</sup>	0.3818	0.3415

**Table 6:** *Continued*

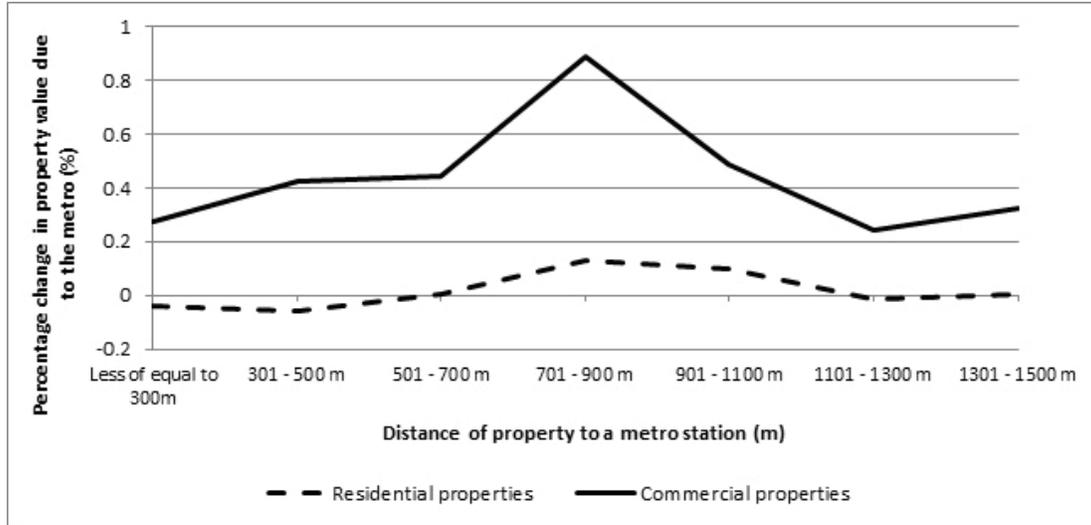
Datasets	Commercial properties	
	Repeated cross-sectional	Pseudo panel
Data structure	Repeated cross-sectional	Pseudo panel
Estimator type	RE	RE
No. Observations	3419	336
Catchment zone	Defining treated properties at 0.5 km of a station	
Accessibility to a number of metro stations	0.356*** (0.052)	0.234** (0.108)
Overall R <sup>2</sup>	0.612	0.6018
Catchment zone	Defining treated properties at 1 km of a station	
Accessibility to a number of metro stations	0.336*** (0.029)	0.245*** (0.058)
Overall R <sup>2</sup>	0.614	0.6188
Catchment zone	Defining treated properties at 1.5 km of a station	
Accessibility to a number of metro stations	0.138*** (0.014)	0.175*** (0.037)
Overall R <sup>2</sup>	0.5868	0.626

Legend: \*, \*\*, \*\*\* indicate significance at 10 percent, 5 percent, and 1 percent respectively. Values in parentheses are standard errors. Note: Although other variables as listed in Table 3 (e.g., property characteristics, neighborhood characteristics, distance to amenities) are included and estimated in the model, the results above only present the impact of the Dubai Metro, as it is the main purpose of this research. The results for the contextual variables are available from the authors upon request.

**Table 7:** The effect of the proximity to a metro station at smaller catchment areas on the sale value of residential and commercial properties (HP equation 5)

Dataset	Residential properties	Commercial properties
Data structure	Repeated cross-sectional	Repeated cross-sectional
No. Observations	39,308	3419
Distance to a metro station		
Less or equal to 300m	-0.042 (0.025)	0.313*** (0.061)
301 - 500 m	-0.059*** (0.020)	0.482*** (0.072)
501 - 700 m	0.005 (0.020)	0.438*** (0.072)
701 - 900 m	0.129*** (0.019)	0.757*** (0.075)
901 - 1100 m	0.099*** (0.018)	0.390*** (0.091)
1101 - 1300 m	- 0.012 (0.016)	0.253*** (0.084)
1301 - 1500 m	0.0003 (0.019)	0.323*** (0.123)
overall R <sup>2</sup>	0.3341	0.5664

Legend: \*, \*\*, \*\*\* indicate significance at 10 percent, 5 percent, and 1 percent, respectively. Values in parentheses are standard errors. Note: Although other variables as listed in Table 3 (e.g., property characteristics, neighborhood characteristics, distance to amenities) are included and estimated in the model, the results above only present the impact of the Dubai Metro, as it is the main purpose of this research. The results for the contextual variables are available from the authors upon request.



**Figure 5:** The effect of the Dubai Metro on the value of residential and commercial properties at different distances

Source: Authors

### 5.3 Comparison with existing evidence

Here we discuss the preferred estimates and compare them with existing evidence. The DID model is thought to be better at capturing the causal effect of the Dubai Metro on property values than the HP model. In addition, since we control for unobserved heterogeneity in our models using repeated cross-sectional data, we argue that the models based on repeated cross-sectional data are likely to produce more reliable results compared to those based on pseudo panel data. First, the number of observations in the repeated cross-sectional dataset is much larger than in the pseudo panel dataset. The sample size in the pseudo panel dataset is less than 10 percent of the sample size of the repeated cross-sectional dataset for both types of properties (see Table 4). The larger the sample size, the more likely the accuracy of the results. Second, since the average cohort size is relatively limited, this can lead to measurement error in the average values of the pseudo panel groups and hence bias results. Agostini and Palmucci (2008) and Koster, Ommeren, and Rietveld (2010) also argue that individual repeated cross-sectional data corrected for unobserved heterogeneity are preferable. Therefore, our preferred results are those obtained from the DID models using repeated cross-sectional data.

Comparing our results for the car-oriented city of Dubai with those obtained for other cities with low levels of public transport, we find similar estimates for cities in the United States (e.g., Agostini and Palmucci 2008; Billings 2011; Bowes and Ihlanfeldt 2001; Cervero 2003; Duncan 2008). Although the measure of transport accessibility used in our study may differ from the measures used in previous work, we can still find a basis for comparison.

In particular for residential properties, we find that the effect is concave, where it is negative (-9 percent) on the value of properties located very close to a railway station and positive (7.8 percent) within 1 kilometer of the metro. Other studies also find a concave effect, where the impact is either negative or insignificant and peaks some distance away (e.g., Billings, 2011; Seo, Golub, and Kuby 2014). For example, Billings (2011) for Charlotte, North Carolina, found that the effect of the LRT on the value of single family homes was not significant within 0.8 kilometer while it was positive in the range of 4 percent to 13 percent for all properties located within 1.6 kilometers. Bowes and Ihlanfeldt (2001) also found that the effect of proximity to a rail in Atlanta ranges from -19 percent to 2.4 percent. Agostini

and Palmucci (2008) estimated the effect at one catchment area for dwellings within 1 kilometer of a metro station, and found that the value increased during the construction stage by 4.3 percent.

A very limited number of studies estimated the effect of a railway on the value of commercial properties. To our best knowledge, only one study has used a DID model to estimate the effect of rail on commercial property values (Billings 2011) but did not find any significant estimates. Therefore, we compare our results to studies using HP models. The preferred model estimates a maximum effect of just above 40 percent on the value of commercial properties in Dubai. This estimate is between the range of estimate values obtained in previous studies. While Cervero (2003) found that proximity to a light rail station increased the value of commercial properties within 0.5 miles (just above 800 meters) by 72 percent, Weinberger (2001) estimated that the increase in the value of commercial properties located within 0.5 miles of a light rail station was in the range of 7 percent to 10 percent.

#### **5.4 Policy implications**

Some researchers argue that capturing part of the increase in property value due to a railway investment can be justified to fund at least part of the transport service (Dickens 1992; Smith and Gihring 2006). There are three main fiscal mechanisms for land value capture: taxes or charges, partnership deals, and endowment schemes (see Enoch, Potter, and Ison 2005; Lari et al. 2009; Martínez and Viegas 2012; Salon and Shewmake 2012). The first is the most common and is applied through taxes or levies on properties that benefit from accessibility to a rail station; examples include direct land levee or tax increment financing (TIF). The second mechanism revolves around partnering with developers or landowners to develop areas around the railway. In the third scheme, land is granted to the transport provider to either resell after it gains value due to the railway or invest in and then sell. In these two cases, the profit of land value increase or the return on investment from the development is used to fund the railway. While the first mechanism allows for a continuous contribution of land value capture, the second and third are generally a one-off or an irregular plan toward funding transport services.

Dubai does not apply any taxation on properties, land, and income, similar to all of the cities in the Gulf Corporation Council in the Middle East region. It is, therefore, a unique case in this regard and all transport systems are fully funded by the government. Should the government decide to capture at least a part of the value increase due to the metro, the positive results generated from the different models estimated in this study support this decision for commercial properties located up to 1.5 kilometers from a station and for residential properties located between 0.5 kilometer to 1.1 kilometers. Nonetheless, as the rate and the impact radius vary across type of property, the amount captured can differ as well.

There are perhaps other considerations that the government could ponder before endorsing a given land value capture policy, such as the status of Dubai's economy, the competitiveness of the city worldwide, and the wider plans for the city's growth (Salon and Shewmake 2012). In addition, the legal framework of property acquisition has to be modified to consider the rights of the transport provider to capture at least a part of the increases in property values due to investments in rail projects (Martínez and Viegas 2012; Smith, Gihring, and Litman 2010).

## **6 Conclusions**

This paper estimated the effect of the Dubai Metro on residential and commercial property values using both DID and HP models and two types of data structure: repeated cross-sectional data and pseudo panel data. Our findings indicate an overall positive effect of the Dubai Metro on the sale value of residential and commercial properties, although the nature and magnitude of the effect differs by type of property and distance band. Similar to previous studies, we find that the effect is negative for residen-

tial properties located very close to the metro (i.e., less than 0.5 kilometer) and that it reaches its peak between 701 to 900 meters of a station. Our findings also suggest that the effect of the metro is considerably higher for commercial properties than for residential properties. Moreover, unlike residential properties, the effect is positive and high for properties located very close to the metro. This is expected since the impact of negative externalities associated with noise and pollution from the transport system are less of an issue for businesses than for households.

The increase in the value of residential and commercial properties due to the metro has potential policy implications for the funding of future public transport proposals through a land value capture mechanism. Should Dubai choose to capture at least a part of the increase in property values due to the metro to fund future metro services, it could use the estimates produced by this study for different catchment zones. In addition, the results obtained in this analysis may also be informative for other cities in the Middle East region also planning to construct their own urban rail systems.

The main limitations of this study relate to data availability. One such issue is the lack of data for pre-announcement years, implying that the estimates produced in this study can only capture the effect of the operations phase of the metro compared to the construction phase, and hence are likely to be undervalued<sup>4</sup>.

## **7 Acknowledgements**

The authors express their gratitude to Dubai Real Estate Regulatory Authority (RERA) for providing property data and special thanks to the Roads and Transport Authority (RTA) in Dubai for providing transport data.

<sup>4</sup> We thank one anonymous reviewer for making this point clear.

## References

- Agostini, C. A., and G. A. Palmucci. 2008. The anticipated capitalization effect of a new metro line on housing prices. *Fiscal Studies* 29: 233–256.
- Ahlfeldt, G. M. 2013. If we build, will they pay? Predicting property price effects of transport innovations. *Environment and Planning A* 45: 1977–1994.
- Bertrand, M., E. Dufló, and S. Mullainathan. 2004. How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics* 119: 249–275.
- Billings, S. B. 2011. Estimating the value of a new transit option. *Regional Science and Urban Economics* 41: 525–536.
- Bollinger, C. R., K. R. Ihlanfeldt, and D. R. Bowes. 1998. Spatial variation in office rents within the Atlanta Region. *Urban Studies* 35: 1097–1118.
- Bowes, D. R., and K. R. Ihlanfeldt. 2001. Identifying the impacts of rail transit stations on residential property values. *Journal of Urban Economics* 50: 1–25.
- Cameron, A. C., and P.K. Trivedi. 2005. *Microeconometrics: Methods and Applications*. New York: Cambridge University Press.
- Cervero, R. 2003. *Effects of Light and Commuter Rail Transit on Land Prices: Experiences in San Diego County*. Berkeley, CA: University of California.
- Clark, T. S., and D. A. Linzer. 2013. Should I use fixed or random effects? <http://userwww.service.emory.edu/~tclark7/randomeffects.pdf>.
- Collado, M. D. 1997. Estimating dynamic models from time series of independent cross-sections. *Journal of Econometrics* 82: 37–62.
- Concas, S. 2012. *Accessibility and Housing Price Resilience: Evidence from Limited-Access Roadways*. Tampa, FL: Center for Urban Transportation Research, College of Engineering, University of South Florida.
- Deaton, A. 1985. Panel data from time series of cross-sections. *Journal of Econometrics* 30: 109–126.
- Deweese, D. N. 1976. The effect of a subway on residential property values in Toronto. *Journal of Urban Economics* 3: 357–369.
- Diaz, R. B. Impacts of rail transit on property values. 1999. Commuter Rail/Rapid Transit Conference, American Public Transit Association, Toronto.
- Dickens, I. 1992. Transport investment, economic development and strategic planning: The example of light rail transit. *Planning Practice and Research* 7: 9–12.
- Du, H., and C. Mulley. 2006. Relationship between transport accessibility and land value. Local model approach with geographically weighted regression. *Transportation Research Record* 1977: 197–205.
- Du, H., and C. Mulley. 2007. *Transport Accessibility and Land Values: A Case Study of Tyne and Wear*. RICS Research paper series 7(3).
- Dubai Municipality. 2003. A study on alternative systems for public transport in Dubai: PS002. In Dubai Municipality.
- Dubai Municipality. 2005. Project PS007 preliminary engineering stage for Dubai rail project. In Dubai Municipality.
- Dubai Statistics Center. 2011. Dubai in figures. <http://dsc.gov.ae/EN/Publications/Pages/Publication-List.aspx?PublicationId=2&Year=2011>.
- Dubé, J., M. Thériault, and F. Des Rosiers. 2013. Commuter rail accessibility and house values: The case of the Montreal South Shore, Canada, 1992–2009. *Transportation Research Part A: Policy and Practice* 54: 49–66.
- Duncan, M. 2008. Comparing rail transit capitalization benefits for single-family and condominium

- units in San Diego, California. *Transportation Research Record* 2067: 120–130.
- Efthymiou, D., and C. Antoniou. 2013. How do transport infrastructure and policies affect house prices and rents? Evidence from Athens, Greece. *Transportation Research Part A: Policy and Practice* 52: 1–22.
- Enoch, M., S. Pottor, and P. Ison. 2005. A strategic approach to financing public transport through property values. *Public Money and Management* 25(3): 147–154.
- Gatzlaff, D. H., and M. T. Smith. 1993. The impact of the Miami Metrorail on the value of residences near station locations. *Land Economics* 69: 54–66.
- Gibbons, S., and S. Machin. 2005. Valuing rail access using transport innovations. *Journal of Urban Economics* 57: 148–169.
- Hahn, J., J. Han, and H. R. Moon. 2011. Test of random versus fixed effects with small within variation. *Economics Letters* 112: 293–297.
- Kim, J., and M. Zhang. 2005. Determining transit's impact on Seoul commercial land values: An application of spatial econometrics. *International Real Estate Review* 8: 1–26.
- Koster H. R. A., J. N. V. Ommeren, and P. Rietveld. 2010. *Estimating the Benefits of Improved Rail Access: Geographical Range and Anticipation Effects*. Amsterdam: Tinbergen Institute.
- Laakso, S. 1992. Public transport investment and residential property values in Helsinki. *Scandinavian Housing and Planning Research* 9: 217–229.
- Lari, A., D. Levinson, Z. Zhao, M. Iacono, S. Aultman, K. V. Das, J. Junge, K. Larson, and M. Scharenbroich. 2009. *Value Capture for Transportation Finance: Technical Research Report*. Minneapolis, MN: Center for Transportation Studies, University of Minnesota.
- Martínez, L. M., and J. M. Viegas. 2009. Effects of transportation accessibility on residential property values—hedonic price models in the Lisbon, Portugal, metropolitan area. *Transportation Research Record* 2115: 127–137.
- Martínez, L. M., and J. M. Viegas. 2012. The value capture potential of the Lisbon Subway. *Journal of Transport and Land Use* 5(1): 65–82.
- McDonald, J., and C. Osuji. 1995. The effect of anticipated transportation improvement on residential land values. *Regional Science and Urban Economics* 25: 261.
- McMillen, D. P., and J. McDonald. 2004. Reaction of house prices to a new rapid transit line: Chicago's midway line, 1983-1999. *Real Estate Economics* 32: 463–486.
- Mohammad, S. I., D. J. Graham, P. C. Melo, and R. J. Anderson. 2013. A meta-analysis of the impact of rail projects on land and property values. *Transportation Research Part A: Policy and Practice* 50: 158–170.
- Pan, H., and M. Zhang. 2008. Rail transit impacts on land use: Evidence from Shanghai, China. *Transportation Research Record* 2048: 16–25.
- Plümpert, T., and V. E. Troeger. 2007. Efficient estimation of time-invariant and rarely changing variables in finite sample panel analyses with unit fixed effects. *Political Analysis* 15: 124–139.
- Roads and Transport Authority. 2011. *A Survey on Metro Passengers*. Development and Corporate Dubai: Performance Department, Roads and Transport Authority.
- Salon, D., and S. Shewmake. 2012. Opportunities for value capture to fund public transport: A comprehensive review of the literature with examples from East Asia. Transportation Research Board 2012 annual meeting, January 22–26, Washington, DC.
- Seo, K., A. Golub, and M. Kuby. 2014. Combined impacts of highways and light rail transit on residential property values: A spatial hedonic price model for Phoenix, Arizona. *Journal of Transport Geography* 41: 53–62.
- Smith, J., and T. Gihring. 2006. Financing transit systems through value capture. *The American Journal*

- of Economics and Sociology* 65: 751–786.
- Smith, J. J., T. A. Gihring, and T. Litman. 2010. *Financing Transit Systems Through Value Capture—An Annotated Bibliography*. Victoria, BC, Canada: Victoria Transport Policy Institute.
- Tsai, C., W. Leong, C. Mulley, and G. Clifton. 2013. Examining estimation bias and efficiency for pseudo panel data in travel demand analysis. *Transportation Research Record* 2354: 1–8.
- Verbeek, M., and T. Nijman. 1992. Can cohort data be treated as genuine panel data. *Empirical Economics* 17: 9–23.
- Vichiensan, V., and K. Miyamoto. 2010. Influence of urban rail transit on house value: Spatial hedonic analysis in Bangkok. *Journal of the Eastern Asia Society for Transportation Studies* 8: 986–996.
- Voith, R. 1991. Transportation, sorting and house values. *Real Estate Economics* 19: 117–137.
- Weinberger, R. R. 2001. Light rail proximity: Benefit or detriment? The case of Santa Clara County, California. *Transportation Research Record* 1747: 104–113.
- Wu, W. 2012. *Does public investment spur the land market? Evidence from transport improvement in Beijing*. London: Spatial Economics Research Center.

## Appendix A: DID models

Each individual “i” in the dataset belongs to either a treated or a control group (D) in the time periods pre- and post-treatment (T). The treated group takes a value of D=1 and 0 otherwise; and the time after treatment takes a value of T=1 and 0 otherwise. We adopt the notation of  $y^D$  it such that  $y_{i0}^1$  and  $y_{i1}^1$  represent the value of an individual i, in a treated group, pre- and post-treatment, respectively; and  $y_{i0}^0$  and  $y_{i1}^0$  represent the value of an individual i, in a control group, pre- and post-treatment, respectively. The treatment effect (i.e., new metro),  $\beta$ , is

$$\beta = (\bar{y}_1^1 - \bar{y}_0^1) - (\bar{y}_1^0 - \bar{y}_0^0)$$

Where  $\bar{y}_1^1$  and  $\bar{y}_0^1$  are the mean property values for the treated group after and before the treatment, respectively;  $\bar{y}_1^0$  and  $\bar{y}_0^0$  are the mean property values for the control group after and before the treatment, respectively.

A basic DID model is

$$\ln y_{it} = \alpha + \alpha_T T + \gamma_D D + \beta(T \cdot D) + \varepsilon_{it}$$

Where:

$y_{it}$  is the value of property i at time t

T equals 1 if the treatment has occurred, 0 otherwise

D equals 1 for treated properties, 0 otherwise

$\beta$  is the DID estimator

The difference between the values of treated and control properties before and after the treatment eliminates any time effects and yields an estimation of the treatment effect. This suggests the following

$$(\ln y_{i1}^1 - \ln y_{i0}^1) - (\ln y_{i1}^0 - \ln y_{i0}^0) = \beta + (\varepsilon_{i1}^1 - \varepsilon_{i0}^1) - (\varepsilon_{i1}^0 - \varepsilon_{i0}^0)$$

The  $\beta$  is called the difference-in-differences estimator and is unbiased if

$$E[(\varepsilon_{i1}^1 - \varepsilon_{i0}^1) - (\varepsilon_{i1}^0 - \varepsilon_{i0}^0)] = 0$$

The general assumptions of a DID estimation applicable to our study are as follows (e.g., Bertrand, Dufo, and Mullainathan 2004; Cameron and Trivedi 2005). First, to allow for a correct estimation of the DID parameter, the treated and control groups do not change over the specified time period. Second, the error term in the regression model is assumed to have a zero mean and a constant variance. Finally, for unbiased estimations, we assume that the expected change in the treated and control groups would follow the same pattern had the treatment not occurred (parallel trend assumption).

## Appendix B: Results using RE and FE estimators

**Table B.1:** Main DID model results for the effect of the metro on property values using random-effects and fixed-effects estimators

Dataset	Residential properties (RERA)	
Data structure	Repeated cross-sectional	Pseudo panel
Estimator type	FE	FE
No. Observations	39,308	3344
Catchment zone	<i>Defining treated properties at 0.5 km of a station</i>	
Group * time effect	-0.106*** (0.014)	-0.179*** (0.064)
Overall $R^2$	0.3245	0.2020
Catchment zone	<i>Defining treated properties at 1 km of a station</i>	
Group * time effect	0.075*** (0.008)	-0.112** (0.053)
Overall $R^2$	0.3334	0.2054
Catchment zone	<i>Defining treated properties at 1.5 km of a station</i>	
Group * time effect	0.034*** (0.007)	-0.158*** (0.05)
Overall $R^2$	0.3675	0.1895

Dataset	Commercial properties (RERA)	
Data structure	Repeated cross-sectional	Pseudo panel
Estimator type	FE	FE
No. Observations	3419	336
Catchment zone	<i>Defining treated properties at 0.5 km of a station</i>	
Group * time effect	0.398*** (0.054)	-0.207 (0.148)
Overall $R^2$	0.1253	0.1355
Catchment zone	<i>Defining treated properties at 1 km of a station</i>	
Group * time effect	0.426*** (0.037)	0.132 (0.106)
Overall $R^2$	0.1467	0.1340
Catchment zone	<i>Defining treated properties at 1.5 km of a station</i>	
Group * time effect	0.409*** (0.036)	0.441** (0.194)
Overall $R^2$	0.1424	0.1476

Legend: \*, \*\*, \*\*\* indicate significance at 10%, 5%, and 1%, respectively. Values in parentheses are standard errors. Note: Although other variables as listed in Table 3 (e.g., property characteristics, neighborhood characteristics, distance to amenities) are included and estimated in the model, the results above only present the impact of the Dubai Metro, as it is the main purpose of this research. The results for the contextual variables are available from the authors upon request.

**Table B.2:** Main HP model results for the effect of the metro on property values using random-effects and fixed-effects estimators

Dataset	Residential properties (RERA)	
Data structure	Repeated cross-sectional	Pseudo panel
Estimator type	FE	FE
No. Observations	39,308	3344
Catchment zone	<i>Defining treated properties at 0.5 km of a station</i>	
Accessibility to a number of metro stations	- 0.096*** (0.013)	-0.066 (0.059)
$R^2$ (overall for RE and within for FE)	0.1935	0.126
Catchment zone	<i>Defining treated properties at 1 km of a station</i>	
Accessibility to a number of metro stations	0.016* (0.010)	-0.022 (0.041)
$R^2$ (overall for RE and within for FE)	0.1924	0.1256
Catchment zone	<i>Defining treated properties at 1.5 km of a station</i>	
Accessibility to a number of metro stations	0.009 (0.005)	-0.025 (0.021)
$R^2$ (overall for RE and within for FE)	0.1924	0.1261

Dataset	Commercial properties (RERA)	
Data structure	Repeated cross-sectional	Pseudo panel
Estimator type	FE	FE
No. Observations	3419	336
Catchment zone	<i>Defining treated properties at 0.5 km of a station</i>	
Accessibility to a number of metro stations	0.160*** (0.061)	-0.019 (0.132)
$R^2$ (overall for RE and within for FE)	0.1552	0.1777
Catchment zone	<i>Defining treated properties at 1 km of a station</i>	
Accessibility to a number of metro stations	0.303*** (0.041)	0.131* (0.068)
$R^2$ (overall for RE and within for FE)	0.1674	0.1910
Catchment zone	<i>Defining treated properties at 1.5 km of a station</i>	
Accessibility to a number of metro stations	0.102*** (0.022)	0.093* (0.052)
$R^2$ (overall for RE and within for FE)	0.1588	0.1893

Legend: \*, \*\*, \*\*\* indicate significance at 10 percent, 5 percent, and 1 percent, respectively. Values in parentheses are standard errors. Note: Although other variables as listed in Table 3 (e.g., property characteristics, neighborhood characteristics, distance to amenities) are included and estimated in the model, the results above only present the impact of the Dubai Metro, as it is the main purpose of this research. The results for the contextual variables are available from the authors upon request.