Price, Size and Density

Paul Kilgarriff

Abstract

The current discussions and analysis around housing in Ireland lack a detailed analysis of the relationship between house size and prices in Ireland. Related to house size is the density of an area. In this paper I contribute a measure of price per square metre (ppsm) for housing. Using web-scraped data a kriging methodology is used to interpolate a ppsm at the Small Area level for Dublin and its commuter area. Results show that the highest ppsm is located around the city centre. Houses however in these areas are smaller compared to areas south of the city where ppsm is remains high as well as house size. Exploring affordability shows that households can trade-off distance to the central business district (CBD) in exchange for a lower ppsm and higher house size. Properties >90m^2 remain unaffordable even for households in the 75th percentile of disposable incomes. Using radial analysis along with scaling to control for city size, comparisons in population density and building heights are made between Dublin, Vienna, Copenhagen and Paris. Controlling for city size, the analysis highlights the low density nature of Dublin and limited quantity of living space when compared to the other cities. Solutions are required to increase the supply of living space around the CBD in an attempt to improve affordability.

Introduction

Choosing where to live is one of the biggest decisions faced by individual and households. This decision is influenced by a several factors including: personal situation, income level, location of employment and level of wealth. Single person households theoretically require less living space compared to a married couple with children. A household's level of income will determine the size of mortgage and hence their price range and ultimately location. The Central Bank of Ireland mortgage rules limit households to 3.5 times gross income[1]. There are also limits related to the loan-to-value percentage, households are required to have a deposit ranging from 5-30% depending upon the type of purchaser (first-time buyer, second time buyer etc.). These limits only apply to size of the mortgage and do not give any recommendations on minimum living area per person or maximum size of dwelling.

The price of housing can be separated into dwelling and location factors that focus on characteristics and quality of the housing unit and location factors related to the neighbourhood and broader environment (Wilkinson, 1973). Hedonic regression models are used to represent house prices as a function of a bundle of these attributes and characteristics (Roback, 1982; Rosen, 1974). These attributes will consist of the characteristics of the house (size m^2, number of bathrooms, materials of construction, energy efficiency of the house) and the environment in which it is located (location of the house, distance to city centre, characteristics of the neighbourhood). Determining the exact location of the property can significantly improve the explanatory power of the model (Hill, 2013). It is possible to decompose the value of a house based on the characteristics of the house and environmental amenities.

[1] https://www.centralbank.ie/consumer-hub/explainers/what-are-the-mortgage-measures

Another factor influencing residential decision choice is commuting time and employment centres. Living in an urban area allows the resident to benefit from agglomeration economies such as local markets, a large labour supply and greater level of firms. Commuting costs consist of the monetary costs of travel (car running costs) and time costs (value of time spent travelling). An individual has a maximum commuting distance based on their income, distance to employment and housing costs. Using a monocentric city model is a simplification sued to examine these trade-offs. In the model, a city is defined as having employment centrally located at the city centre or central business district (CBD), to which individuals commute (Alonso et al., 1964; Mills, 1972; Muth, 1969).

In its simplest form, the Alonso-Mills-Muth (AMM) model consists of a featureless plain with employment concentrated in the central business district (CBD) and surrounded by a large residential area. All houses are the same size and shape and all households have the same utility. Income is also identical and the only difference between households is their decision on where to live. Households spend all of their income, divided between housing costs, commuting costs and consumption on another composite good. As commuting costs increase with distance, housing costs must decrease with distance. The result is a negative rent gradient (Coulson; 1991). In equilibrium no alternative location exists which would increase utility (Coulson and Engle; 1987).

AMM in its simplest form will have a rent gradient equal to the marginal cost of commuting. An increase in the marginal cost of commuting, would increase the slope and hence the rent gradient (Coulson and Engle; 1987). As commuting costs increase, there is an incentive to live closer to the CBD. Similarly, if commuting costs decrease, the slope of the rent gradient would decrease and the fringe distance increase. As an individual move away from the CBD, increasing commuting costs are counterbalanced with lower housing costs (Henderson; 1985). Previous research found the housing market and population density in several European cities to have a strong monocentric structure with the highest values located at the city centre with the gradient decreasing with distance to the city centre (Glumac et al., 2019; Helbich et al., 2014; Lemoy and Caruso, 2018; Manzoli and Mocetti, 2019).

The simple monocentric model assumption of one standard house size can be relaxed. Typically, households will live in smaller housing units closer to the CBD or trade-off higher commuting costs with lower housing prices and larger housing units at greater distances to the CBD (Mieszkowski and Mills; 1993). Both land costs and structural density decrease as distance to the CBD increases (Brueckner et al.; 1987). As land close to the CBD is more expensive compared to the periphery, developers will maximise rents by building more densely closer to the CBD. This concept has implications for population density, the level of urbanisation and house size. As houses closer to the CBD are smaller compared to the periphery and structural density higher, it follows that a $1km^2$ plot at the CBD will contain more housing units and therefore more people compared to the periphery (Henderson; 1985). Population density will therefore decrease as distance to the CBD increases. Falling transport costs make suburban and periphery locations more attractive with cheaper, larger housing on offer (Brueckner; 2000). This will have a feedback effect on house prices in the suburbs and periphery.

Commuting costs increase with distance to the CBD, whereas house prices decrease with distance to the CBD (Wheaton, 1977). For a city, this translates to the city core having higher house prices relative to the periphery and the city core having smaller properties relative to the periphery.

Individual's trade-off house price, house size and commuting costs in a way, which maximises their utility subject to their income (Fujita, 1989). The monetary costs are the same for rich and poor but there is an opportunity cost related to the time cost of commuting which is higher for the rich as they have higher wages (Glaeser et al., 2001). In an Irish context, an increase in rents of 10% around the CBD is associated with a 0.6 minute increase in the national average one-way commuting time (Ahrens and Lyons, 2021). The supports the fact that Dublin is a monocentric city and that affordability around the CBD is leading to outward migration.

Housing size, location and density are all important factors in determining house prices. Related to house price or 'rents' is the concept of housing affordability stress (HAS). Housing can be classified as not affordable when housing costs as a proportion of household income (gross or disposable income) exceed a specified threshold (Corrigan et al., 2019). Corrigan et al., (2019) found using a 30/40 affordability benchmark provided more useful results compared to using strict cut-offs. The 30/40 rule classifies housing as unaffordable if a households housing costs are in excess of 30% of household income and the household belongs to the bottom 40% of the income distribution (Baker et al., 2015). However, gross or disposable household income does not account for household size, equivalised income is preferred (OECD, 2014). Household size will determine the size of the house and ultimately the price. A trade-off exists between having data on household income deciles but at an aggregated local authority level or having data on median equivalised household income at an Electoral Division (ED) level but no income distribution information. Using ED level income assumes the market of potential buyers for a property in an ED is limited to residents of that ED. It is therefore more appropriate to use the local authority level income data and test for a range of scenarios. I acknowledge the limitation in using a non-equivalised income measure however in this case a range of scenarios are tested which cover a range of house sizes. Using a weighted average income for the Dublin FUA, the market of potential buyers includes all households within the region (Dublin, Meath, Kildare and Wicklow). The 30/40 affordability benchmark is used to assess the maximum size of property that meets an affordability threshold of 35% of income.

In larger cities buildings at the city centre tend to be taller and hence density higher (Henderson, 1985). At you approach the city centre the land to capital ratio decreases as the cost of land increases. It can therefore be difficult to compare population density between cities of different sizes. One method to control for city size is the use of scaling (Batty, 2008; Bettencourt, 2013). Lemoy and Caruso (2018) previously discovered a homothetic scaling law to transform both artificial land use and density. The density of a city was found to scale with city size measured by its total population in a homothetic manner. This is the standard relationship between the area and the side length of a surface in three dimensions (square or disc for instance). We note that homothetic or isometric scaling uses a fixed factor for all parts of the considered system, in comparison to allometry which uses different rates of growth (Thompson; 1917; Huxley; 1932) for different parts of the system. Density decreases with distance to the city centre at an increasing rate for small cities. At the same time density is also higher at the centre in larger cities. To control for the city size effects, both the horizontal (distance) and vertical (density) axes are rescaled with respect to the cities population (Lemoy and Caruso, 2018).

The overall objective of this paper is to explore the relationship between house prices, house size and density. This is in response to typical metrics such as average house price. Using Dublin FUA as a case study, I firstly calculate the price per square metre (€/m^2) at the Small Area (SA) level. I

apply €/m^2 to the test various scenarios to calculate housing cost for each SA in the Dublin FUA. The average disposable income at the 30th, 40th and 50th percentiles are used to examine affordability. The final component of the analysis will examine density. Using scaling, the density profile of Dublin is compared to several European cities. Net density is also measured using total residential floor space and population per area to estimate the persons per km^2 of floor space. Residential floor space is measured using the EU Copernicus Urban Atlas Land Use and building heights data. This is the first study in Ireland to combine, property value data, income data, building heights and detailed land use data in the same model.

Policy Context

The majority of published reports and data on house prices in Ireland use median price of a residential property. Along with the RPPI (CSO, 2017), the CSO also produce a median price per Eircode routing key, this is the middle rank of all properties sold in that Eircode. No distinction is made in relation to housing characteristics such as size. The Daft.ie Irish House Price Report (Lyons, 2021) lists more detail on size with prices broken down by postal district in Dublin and number of bedrooms. In the Irish property market there is a lack price information which is both spatial and attribute rich. One solution to this issue is the calculation of a price per square metre (ppsm) of housing at a small area level that is the standard in the majority of European countries. Publishing a ppsm indicator provides a missing link between house prices, house characteristics and density.

The National Planning Framework (NPF) (NPF, 2017) mentions "higher density housing" and compact urban growth with National Policy Objective 35 specifically targeted at infill development increased building height and re-use of existing buildings. The NPF however does not explicitly define high-density thresholds or dwelling size. The Sustainable Residential Development in Urban Areas (DoEHLG, 2009) published under section 28 of the Planning and Development Act 2000 (as amended) defines medium density in urban areas as 35-50 dwellings per hectare (dph) and >50 dph in central urban areas or areas served by high frequency public transport. It also recommends utilising the plot ratio (ratio of total floor area of the building to the area of the site) to account for dwelling size. Combining both dph and plot ratio can however place limits on the maximum height of a building. One approach not considered is to benchmark Dublin's density (inhabitants per km^2) against other European cities.

Methodology

I used web-scraping to gather house-listing data from the property website Daft.ie. The original scarping yielded 15,570 listings, after filtering those with no price (price on application or ATM listings) and houses with no size in m^2 (listings with a size over 500^2 were dropped), this left 7,629 listings with a price and a house size in m^2. The data contains an address string but no specific co-ordinates. I used the Google Maps API geo-coding product to geocode the address string. The Google Maps API works in the same way as the Google Maps product on your phone or laptop. Each address string is searched using Google Maps and a latitude and longitude in the coordinate reference system (CRS) WGS84 (EPSG: 4326) is returned. Addresses in cities that contain a street name or house number are geocoded more accurately compared to rural areas where several houses have the same address. As Small Areas are typically larger in rural areas, I am confident Google Maps will return a co-ordinate within the correct Small Area or at least to a neighbouring Small

Area. The latitude and longitude co-ordinates are projected into the Irish Transverse Mercator CRS (EPSG:2157). Projecting co-ordinates enables us to measure distances in metres and kilometres. After geocoding all listings, I attribute each property to the nearest Small Area. Neighbouring characteristics from the census are attributed to the property such as the unemployment rate, household income and tertiary education rate.

Data

Income and Delineation

I am using local authority level data for this analysis as the potential buyer of a property in an ED is not limited to households living within that ED or neighbouring EDs. Instead, the market of potential buyers includes all households within the FUA. EU-OECD definition of a functional urban area is the most widely used definitions of a city in Europe. The definition of a Functional Urban Area/city region uses commuting thresholds. All local units where at least 15% of their employed residents are working in a city, these local units are considered part of the commuting zone or FUA. Enclaves are included and exclaves are excluded, i.e. if a local unit is surrounded by units considered part of the commuting zone, that unit is also included as part of the commuting zone. Non-contiguous local units are excluded (Dijkstra et al., 2019; Dijkstra and Poelman, 2014). It is reasonable to assume the potential buyers of a property within the FUA, includes all households currently living within that FUA.

Table 1 shows how the income deciles vary by LA. Although DLR has the highest incomes at the upper end of the distribution, Fingal and South Dublin both have higher incomes in the bottom decile compared to DLR. Taking the population of households within each LA, I calculate a weighted average for each decile across the LAs.

Table 1: CSO Income Deciles by Local Authority

	Dublin city	DLR	Fingal	Kildare	Meath	South Dublin	Wicklow
0.1	14,855	17,854	22,378	18,167	17,044	19,240	15,000
0.2	23,341	30,672	32,435	28,271	27,073	28,382	24,054
0.3	31,868	43,591	42,388	38,216	36,782	37,452	33,100
0.4	40,437	56,610	52,237	48,000	46,171	46,451	42,138
0.5	50,869	71,206	63,239	58,835	56,399	56,747	52,152
0.6	63,606	88,242	75,801	71,378	68,283	68,493	64,127
0.7	80,375	112,141	92,179	87,532	83,229	83,635	79,605
0.8	105,275	148,517	116,264	110,049	104,502	105,123	102,405
0.9	138,305	197,372	148,057	138,931	132,101	132,956	132,525

Table 2 shows the weighted average 35th-45th deciles of income for Dublin FUA. This study considers only the bottom 40% of households when measuring affordability. The 40th decile ranges from €40,437- €56,610, this shows there is considerable heterogeneity in income across the region. I will also test the 25th, 50th and 75th percentiles to assess affordability.

Table 2: Weighted Average by Income Decile

Decile	Weighted Average	Dublin City	DLR	Fingal	South Dublin
35	41,611	36,153	50,100	47,312	41,951
36	42,570	37,010	51,402	48,297	42,851
37	43,528	37,867	52,704	49,282	43,751
38	44,487	38,724	54,006	50,267	44,651
39	45,446	39,580	55,308	51,252	45,551
40	46,405	40,437	56,610	52,237	46,451
41	47,503	41,481	58,070	53,337	47,480
42	48,601	42,524	59,529	54,437	48,510
43	49,699	43,567	60,989	55,537	49,540
44	50,797	44,610	62,449	56,638	50,569
45	51,895	45,653	63,908	57,738	51,599

Property Price Data

Property price listing data comes from Daft.ie. Listings were web-scraped on the 30th May 2020. Web scraping is performed using the R-Selenium package in R studio along with Docker and a virtual network controller (VNC). Docker creates a virtual network on which the web browser operates. The browser is then automated and controlled within R-studio. The required text on the web page is specified, copied and downloaded. The browser automatically proceeds to the next set of results without user input until there are no more property listings in the user specified area remaining. The web-scraping procedure 'grabs' information in relation to house sales listings including the asking price, the size of the property (m2), number of rooms and/or bedrooms, property type (apartments, detached house etc.) and the location/address of the property. To convert the data into a usable spatial format, each listing is geocoded. The address string of the property is searched using the Google Maps API, which presents an approximation of the co-ordinates. Matching the property to a neighbourhood is sufficient for our analysis and we are not interested in matching to specific properties.

Another issue in relation to the web-scraping procedure is overburdening the website servers. In order to perform the web scraping in a manner that will not cause disruption to the website being scraped and the other users, a sleep command is performed before the procedure moves onto another web page. This is to avoid putting a bandwidth load on the property listings website.

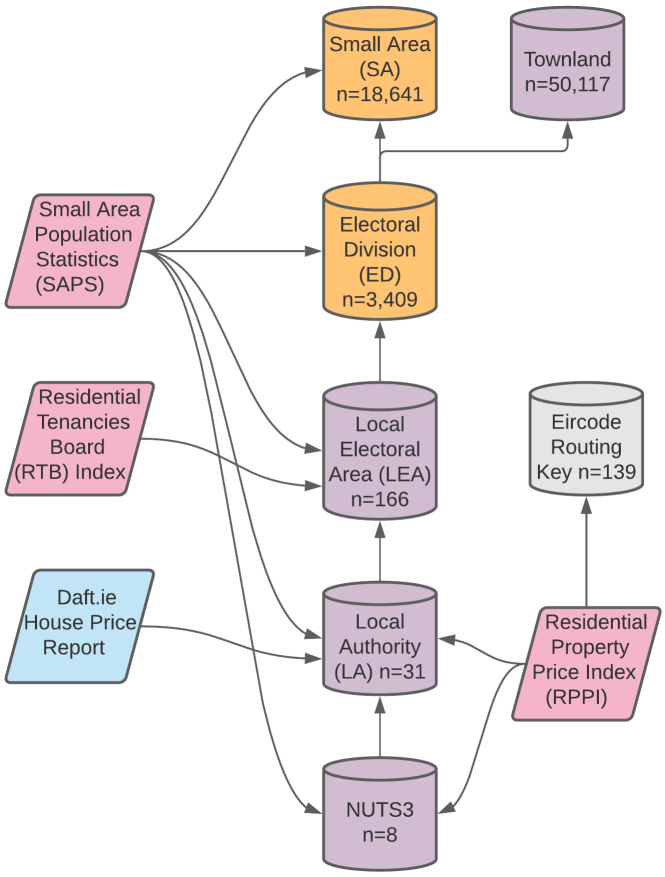
Unit of Analysis

The size and arrangement of areal spatial units used will have an impact on the analysis, this is termed the Modifiable Areal Unit Problem (MAUP) (Gehlke and Biehl, 1934; Openshaw, 1984). Figure 1 illustrates how the various geographies of Ireland are nested. The smallest division in the Small Area Population Statistics (SAPS) of the Census are the Small Areas (SA) of which there are 18,641. The SA's were created in response to growing variation in populations between EDs. In 2011, population within EDs ranged from 55 to 24,405. 2% of the population of Ireland was contained in three EDs (Blanchardstown-Blakestown, Lucan-Esker & Navan Rural). Overall less than 1% of EDs, accounted for 11% of the overall population of Ireland (Charlton, 2010). SAs were created treating the ED as sacrosanct (SAs were sub-divisions of an ED) and ensuring they were consistently small with a minimum size of 65 households. The Geodirectory (postal

addresses), townlands, street centrelines and boundaries (watercourses, major roads, railways) were also used to create the final SA dataset of 18,641. Using SAs over EDs, can help to minimise the effects of MAUP (Charlton, 2010).

The published data from the RPPI is available at an aggregated level; the Eircode Routing key is the lowest level of disaggregation for which the data is available publicly. Figure 1 shows that the Eircode Routing Key geography is a silo and not connected to the others as it cuts across and does not overlap seamlessly with any one geography. Similarly, the Daft.ie House Price Report is at a Local Authority level and Dublin postal districts. The lowest level of disaggregation for the RTB index is at the Local Electoral Area (LEA). The lack of available property price data at a spatially disaggregated level is the main justification for spatial interpolating property prices at the SA level.

Figure 1: Hierarchy of Spatial Units and Demographic and Property Related data. Pink represents data available from the CSO, orange are the two smallest units at which SAPS is available.



Once the web-scraped property listing information is linked to a SA, all of the census data is available for that property. Table 3 shows the breakdown of listings by Local Authority (LA) within the Dublin FUA. Unsurprisingly the Dublin LA's contain the majority of property listings, 32% of all listings. To improve the efficiency of the interpolation, we take a subset and model price per m^2 for Dublin Functional Urban Area (Area). This includes the counties of Dublin,

Kildare, Wicklow and Meath. To improve the interpolation at the edge of the FUA, all listings within a 50km buffer of the FUA are included in the final subset of listings.

Table 3: Daft.ie listings by Local Authority

	All listings	Dublin FUA
LA	# listings	# listings
DLR	503	473
Dublin City	1,143	1,076
Fingal	437	427
Kildare	224	215
Meath	218	209
South Dublin	345	344
Wicklow	277	259
Carlow	50	42
Cavan	96	27
Kilkenny	139	6
Laois	69	57
Louth	118	61
Offaly	89	24
Westmeath	89	30
Wexford	274	15
Other	3,558	0
Total	7,629	3,265

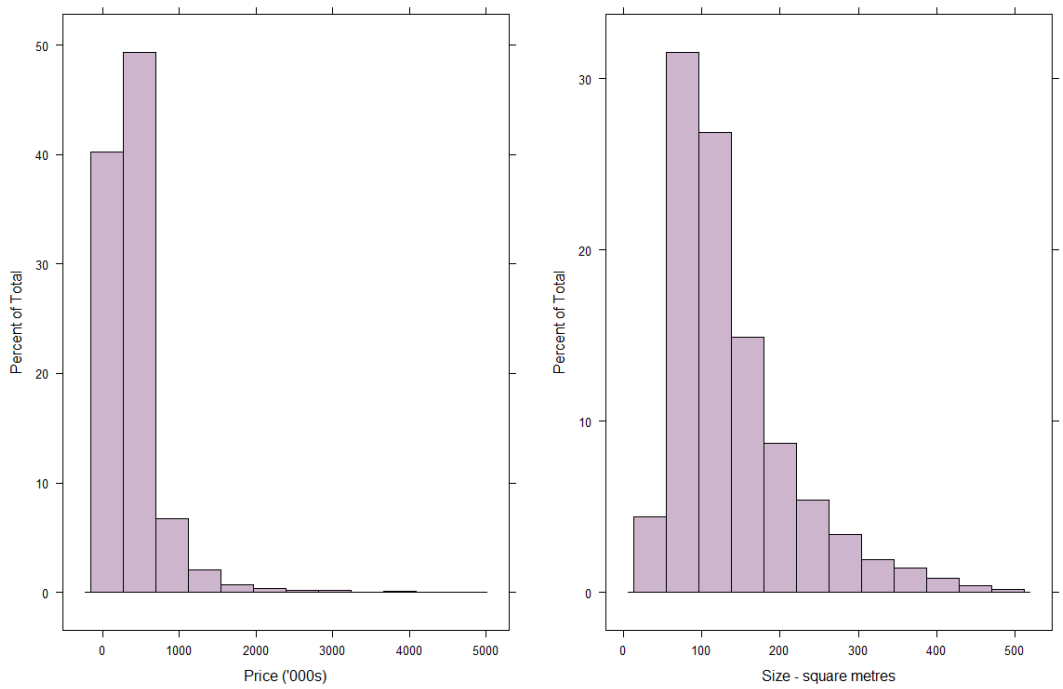
Table 4 shows the summary statistics of the Daft.ie listings within the Dublin FUA. The maximum values indicates the data contains outliers that may prove problematic, as they are not representative of the majority of property listings. A property with 18 bedrooms might be a former hotel or country manor and not your typical property.

Table 4: Property listing summary statistics within Dublin FUA

	List Price (€)	No. Beds	No. Baths	Size m2
Min.	29,000	1	1	31
1st Q.	200,000	3	1	84
Median	300,000	3	2	116
Mean	393,263	3.386	2.269	137.4
3rd Q.	450,000	4	3	168
Max.	5,250,000	18	15	494
				N=7,629

Figure 2 shows the majority of listings are priced below €1,000,000 and are less than 200m^2 in size. There are more outlier listings in relation to price with several properties costing between €1-5 million. The data is truncated by dropping the top and bottom five percentiles from the data in relation to price and size. This reduces the influence outlier properties have on the final interpolated rental value.

Figure 2: Histogram of Property Price and Size

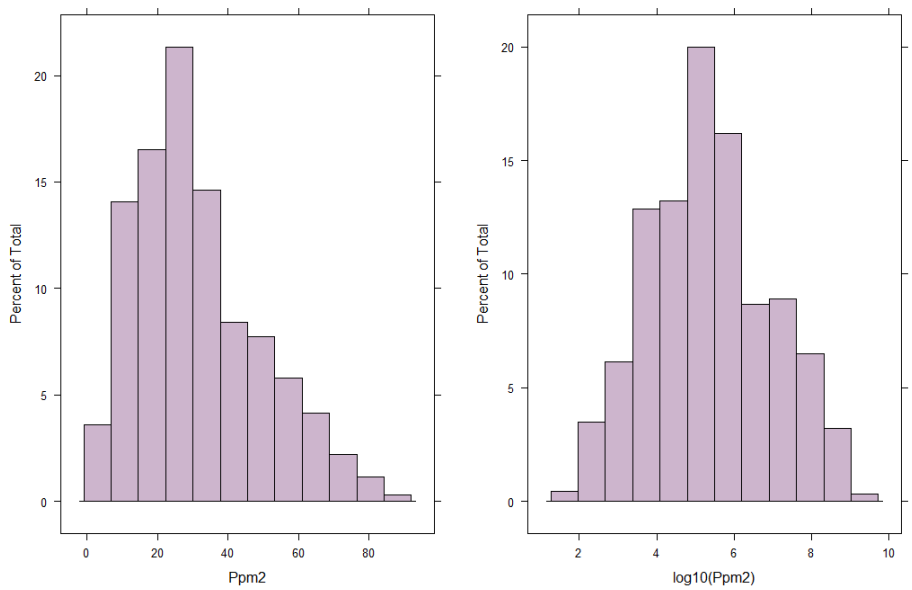


After this process I now have a dataset of 7,000 property listings containing information on price, size (m2), number of bedrooms and bathrooms. Each listings is geocoded and linked to a SA. Several demographic and economic indicators are also calculated at the SA level including, employment rate, tertiary education rate, old age dependency ratio, crime rate, housing tenure, household size, rooms per household and modal share. In addition several distance related measures are calculated, distance to hospital, university, garda station, coastline, lake, rail/tram station and motorway junction.

Kriging

A measure of ppsm is calculated using ordinary kriging (Cressie, 1990). Kriging operates on best linear unbiased prediction (BLUP) (Goldberger, 1962) and accounts for spatially correlated data. Kriging is preffered over other interpolation methods including nearest neighbour interpolation, inverse distance weighting, pycnophylactic interpolation (Anselin and Lozano-Gracia, 2008). Kriging assumes the variance-covariance matrix can be estimated as a function of distance only. When applied, kriging creates a smooth interpolation surface between data points. The variance-covariance matrix is estimated by computing a variogram (Pace et al., 1998). The pair-wise squared differences among all errors, are plotted against the distance between the pair points (Bailey and Gatrell, 1995). There is a limit to the distance at which the value of one point is related to the value of another point (Hoshino and Kuriyama, 2010). As distance increases the covariance converges towards zero. Beyond this range distance, points will have zero impact on the points inside the range. In kriging a greater weight is applied to points which are closer in distance to the dependant (Dubin et al., 1999) following Tobler's first law of geography (Tobler, 1970). Kriging was previously used to estimate real estate values in Ireland (Kilgarriff et al., 2018). I have the price per metre squared for a range of sample points from Daft.ie and the web-scraping process.

Using Kriging I interpolate the price per metre squared at points where I do not have price information. Before fitting the model, ppsm is log transformed (log10) to normalise the distribution.



To examine the robustness of the kriging methodology and the dataset, the data is split in two. The model is calibrated using the training data and kriging is performed to interpolate ppsm across the study area (Dublin FUA). The difference is calculated between the predicted value and the observed value at the locations not sampled.

Table 5: Absolute difference between listing ppsm (observed) and predicted ppsm (kriging)

	Difference (€)
Min.	-4,937
1st Qu.	-543
Median	-48
Mean	-145
3rd Qu.	376
Max.	3,284

Figure 3 and table 5 show the variance in the absolute difference between the predicted ppsm value from kriging and the observed listing ppsm. Overall, the majority of differences fall within +/- €500 ppsm. For an average size house of 120m^2 that translates to a difference of +/- €60,000. The dataset of property listings used has a mean of €405,000; this difference represents an error of +/- 15%. Using a larger dataset of house sales would result in greater accuracy of the interpolated ppsm. Using the CSO RPPI data is one avenue that will be explored. Differences between the interpolated price for an area and a listing may also be due to specific characteristics of the property. Overall the process illustrates that kriging provides us with an accurate interpolated value. Large differences between the observed and predicted are explained by the high level of heterogeneity in property characteristics.

Figure 3: Absolute difference between listing ppsm (observed) and predicted ppsm (kriging)

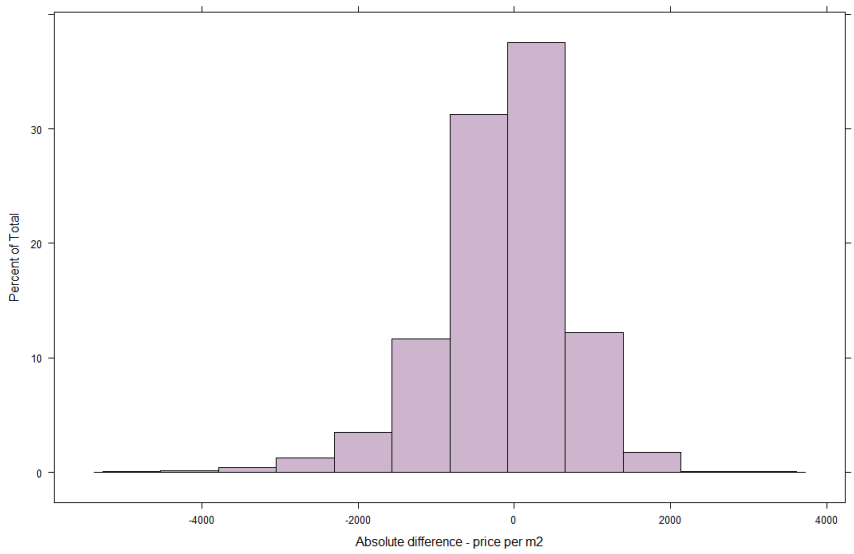


Figure 4: Daft.ie Property Listings Value (€)

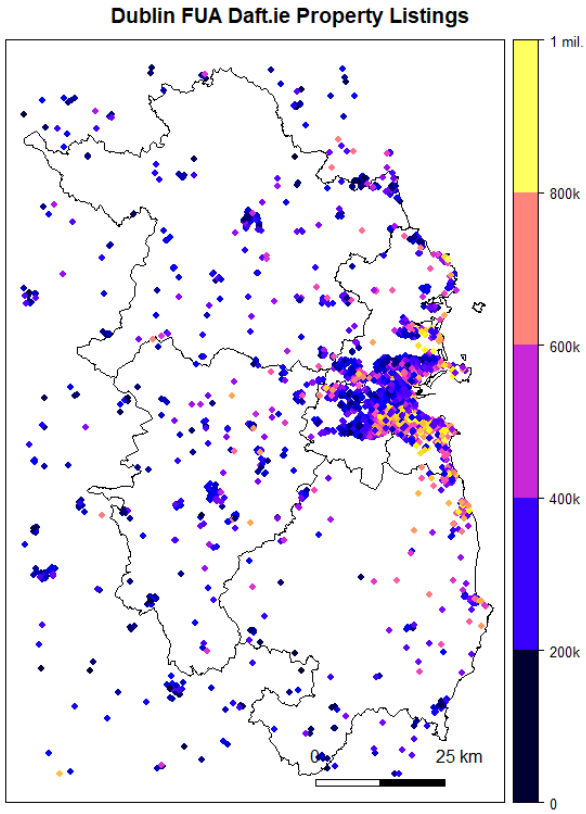
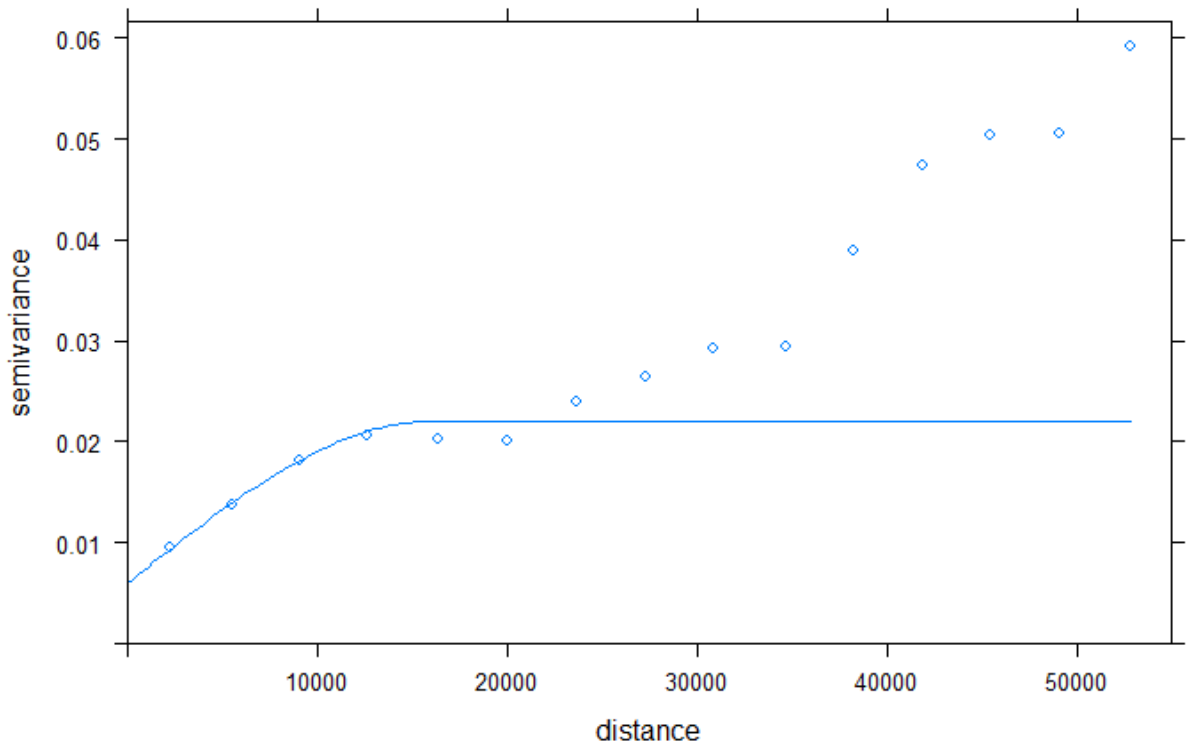


Figure 4 shows the location of the listings within the Dublin FUA. As we can see the majority of listings are in Dublin city local authority. There is less coverage in the Wicklow mountains and some rural areas of Meath. One advantage of using kriging is the error variance. This is an

estimation of the level of uncertainty and is typically higher where we have less sampling points. In this case the rural areas. Given the high frequency of points around Dublin CBD, I can be more confident around the interpolation estimate for these areas.

Figure 5: Fitted variogram



Kriging requires a model which is fitted to the data. The variogram in figure 5 is fitted to the data to describe the spatial continuity of the data. Semi-variance (γ), half the square difference between two points. γ should be smaller for points which are closer to each other. The nugget (y-intercept 0.006) represents small-scale variability of the data. The range is the distance at which the variogram levels off. Beyond this distance of ~15km pairs of points are not spatially correlated with each other. The sill (0.014) is the maximum variability between pairs of points (Cameron and Hunter, 2002).

Figure 6: Kriging - error variance

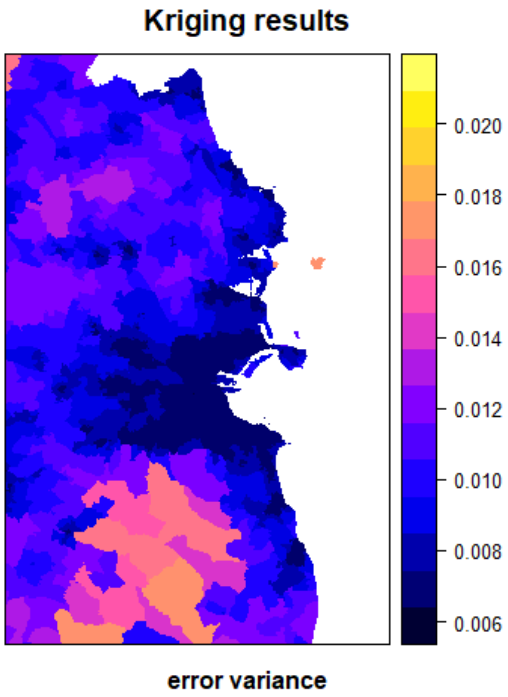


Figure 6 shows the error variance of the kriging predicted estimates. The error values are lowest in the city centre where there the highest frequency of sampling points are located. Remote areas particularly the Dublin mountains area (south) have the highest error variance. One of the advantages of kriging is the production of error values that indicate where we have the highest level of confidence in the results.

Figure 7: Kriging interpolated price per square metre of property (purchase)

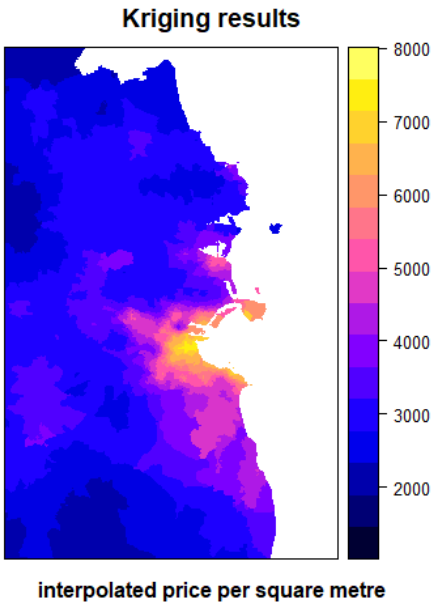


Figure 7 shows the interpolated price per square metre (ppsm) from the kriging model. Ppsm is highest around Dublin central business district (CBD), this is an expected outcome and one that corresponds to urban economic theory and Alonso-Mills-Muth moncentric city model. House prices having a decreasing gradient with distance to the CBD.

One last step in the kriging process is to cross-validate the model. The N-fold cross validation method partitions the data into N parts. For all observations in a part, predictions are made using the remaining N-1 parts. This step is repeated for each of the N parts. From the CV we calculate the mean error and the correlation between the observed and the predicted. From the CV I obtain a correlation coefficient of 0.88 between the observed and the predicted (an ideal value is 1). A mean error of -0.000199 is obtained with a preferred value being zero. From these two measures I am satisfied that the method is robust and the variation given the sample of data points is satisfactory. I will now proceed to examine the affordability of housing under different scenarios and the relationship between price per square metre in an area and density.

Urban Atlas

Land use and building height data comes from the EU Copernicus Urban Atlas (EC, 2016) which is available at a 5 metre resolution. The 2012 version of the Urban Atlas is used to complement the Building Heights data that is also available for 2012 and is at a 20 metre resolution. In the Urban Atlas the boundary of each city corresponds to its Functional Urban Area (FUA) consists of an urban core and areas within a commuting threshold (Dijkstra et al., 2019).

The population for each city is calculated using the 2011 EU Eurostat GEOSTAT 1km^2 population grid (Eurostat, 2020). Population is downscaled by weighting the Urban Atlas data based on the land use code. Categories 1 – 5 are given the highest weighting and defined as residential (Continuous urban fabric (S.L. : > 80%), discontinuous dense urban fabric (S.L. : 50% - 80%), discontinuous medium density urban fabric (S.L. : 30% - 50%) , discontinuous low density urban fabric (S.L. : 10% - 30%) and discontinuous very low density urban fabric (S.L. : < 10%). These categories all have a soil sealing level greater than 10%. Weighting cells makes it possible to downscale the population from a 1km cell into the 5m subdivision cells.

The Dublin General Post Office is used at the point to represent Dublin CBD. Other studies have used the city hall or historic city hall to represent the CBD. In Ireland the post office is the most appropriate location and focal point to use to represent the CBD. For other cities analysed the historic city hall is used as the CBD of the FUA. As every city has a city hall this is a method which enables us to perform the radial analysis in a systemic and consistent way (Walker; 2018; Wilson; 2012).

Radial Analysis & Scaling

Taking a monocentric approach allows me to convert spatially detailed data into a two dimensional curve. Drawing concentric rings of equal distance around the CBD. I measure several aspects, population (count of inhabitants within each ring), area/surface of the ring (land area, total area) and average house price (mean ppsm within the ring). A ring is the difference in area between a disc of radius r and a disc of radius r-1.

This analysis uses the homothetic scaling law devised by Lemoy and Caruso (2018). The artificial land use and population density of a city was found to scale with city size measured by its total population in a homothetic manner. More precisely, the total artificial area of a city is proportional to its total population, and the radius of the city scales with the square root of its total population whereas population density scales with the cube root of population. This is the standard relationship between the area and the side length of a surface in two dimensions (square or disc for instance). This homothetic or isometric scaling uses a fixed factor for the entire considered system. In contrast allometry uses different rates of growth (Thompson; 1917; Huxley; 1932) for different parts of the system.

Lemoy and Caruso (2018) found that the radial population density profiles p(r) of different cities are quite similar if the distance r to the city center is rescaled to a distance r' given by:

$$r' = r \times \sqrt[3]{\frac{N_{Paris}}{N}} = r \times k$$

where N is the population of the city being analysed and N_{Paris} is the population of the largest city in the dataset, used as a reference. k is the rescaling factor. For London and Paris, k≈1. Both r (distance) and p (density) are rescaled using the scaling factor k. See Kilgarriff et al. (2020) for a detailed working of the scaling methodology. Unlike artificial land use where only the horizontal axis (distance) is rescaled, in the case of population density both the horizontal and vertical axis are rescaled. The vertical axis is rescaled to reflect the finding that larger cities tend to have taller buildings around the CBD compared to smaller cities.

In addition to Dublin, Vienna is analysed as it is considered to have a desirable housing market, Copenhagen is analysed as it is of similar size to Dublin and also a capital city and finally Paris as it is a mega-city and is the largest European city. Scaling makes it possible to compare cities independent of city size, in the case of these cities a distance of 10km in Paris is similar to 6km in Vienna, 5.4km in Copenhagen and 5.3km in Dublin. That is, at a distance of 10km in Paris and 5.3km in Dublin, we should expect to see similar density and urban structures.

Results:

The results are broken down into three sections. The first section examines the interpolated price per square metre and average house size. The second section examines affordability of housing under several scenarios and the final section examines density in Dublin with a comparison of European cities.

House price per square metre

Figure 8: Price per square metre (jenks breaks)

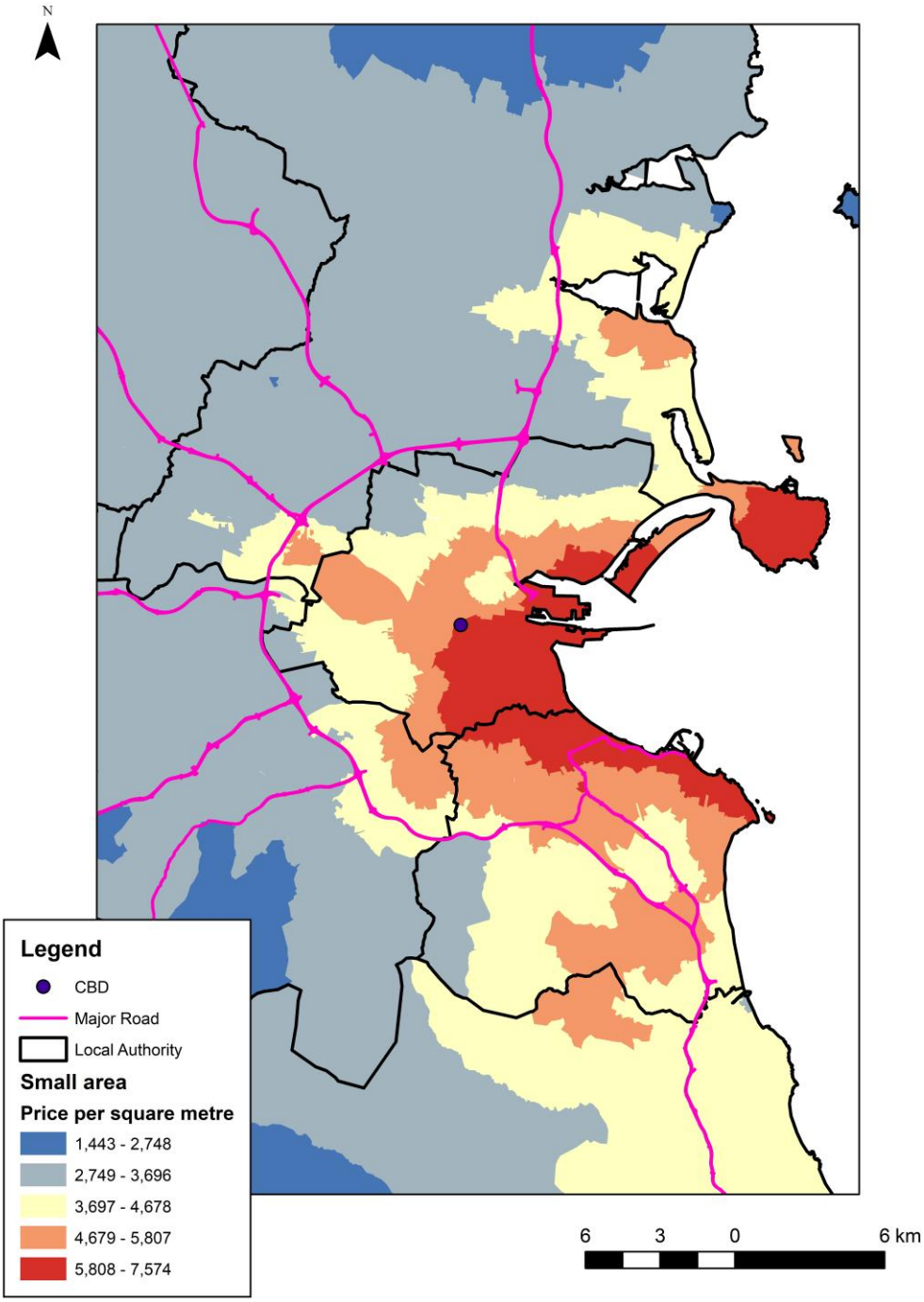
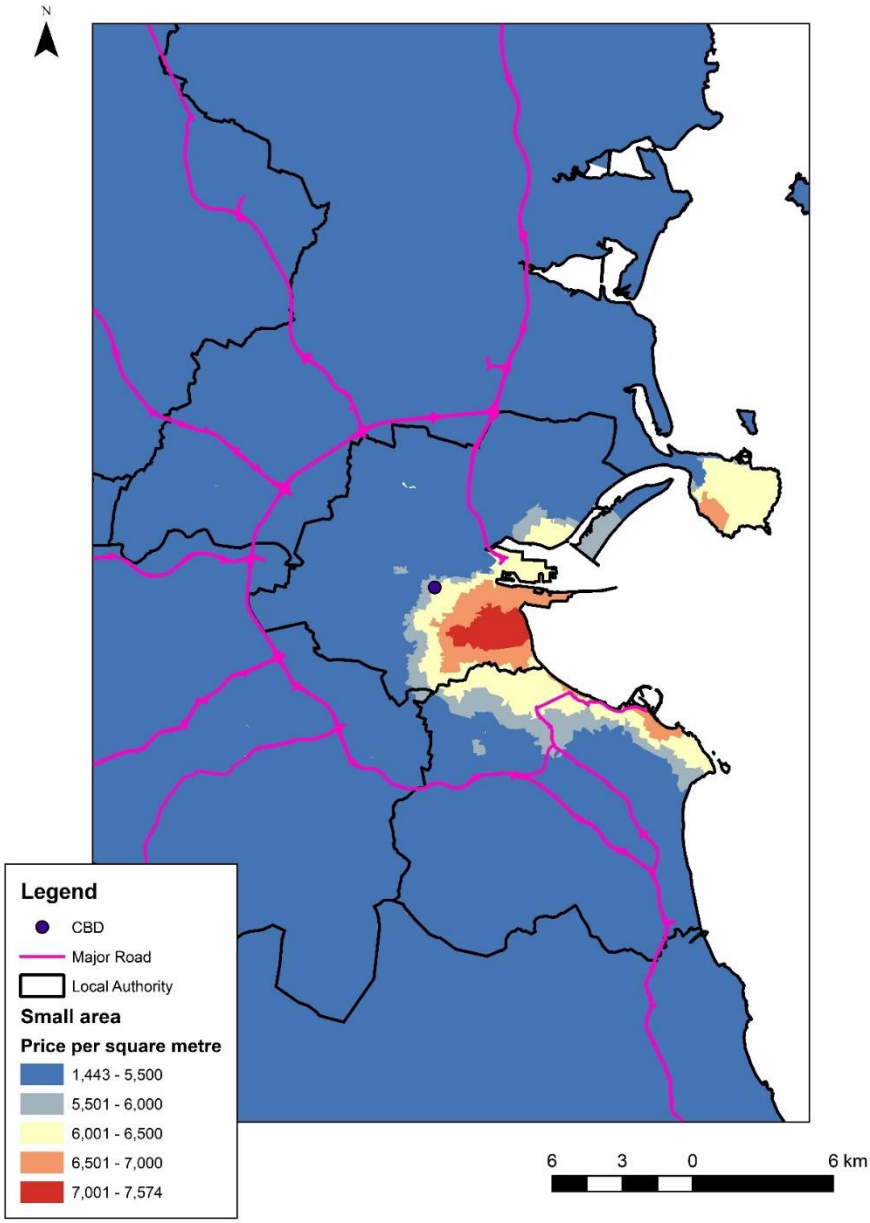


Figure 8 shows the interpolating price per square metre estimated using kriging. The highest levels of ppsm is located south-east of the CBD with pockets along the coastline and in Howth. As you move away from the CBD prices start to fall, with prices beyond the M50 typically in the bottom

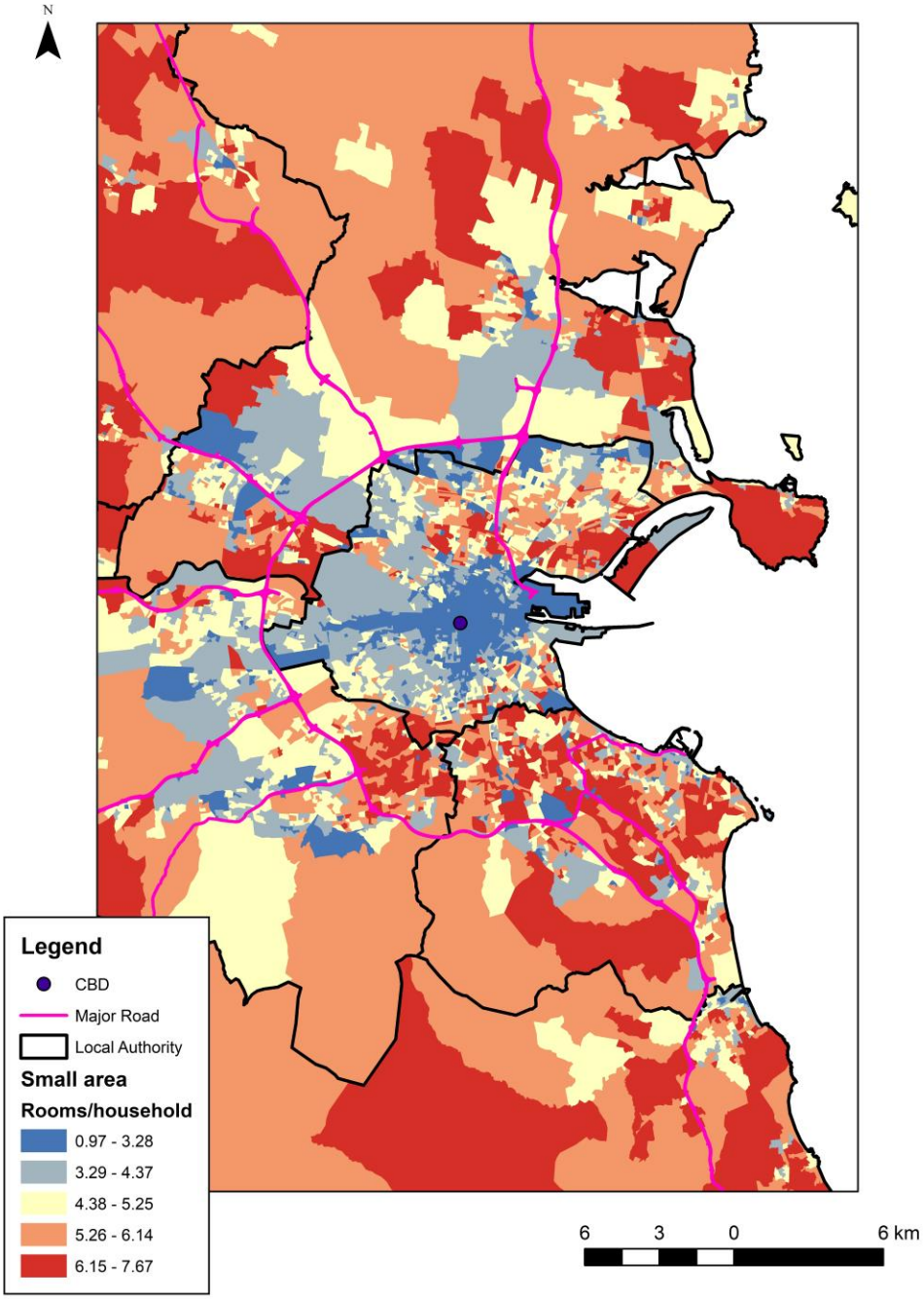
bands. Dublin is a monocentric city, characterised by higher prices close to the CBD and a decreasing price gradient.

Figure 9: Price per square metre (narrow bands)



In figure 9 the bands are narrowed at the upper limit to understand where the highest prices are located. The highest prices are located in and around the Ballsbridge area of Dublin with prices decreasing as you move away from this point. Interestingly this is not the area with the highest listed house prices. These are located south of Ballsbridge in DLR local authority where houses are larger in size and ppsm remains close to the maximum value.

Figure 10: Rooms per household Census 2016 (small area)



From figure 10 we can see that although Dublin city LA has the highest ppsm, house size in this area is smaller in comparison to DLR LA. Households in DLR have almost twice the number of rooms per household on average compared to Dublin city LA. One consequence of this is although ppsm is lower in DLR, given house size is larger compared to other areas of Dublin, this makes the majority of properties unaffordable. This highlights the need for a broad range of properties to

satisfy households at different stages on the life-cycle, e.g. studio apartments for students and young workers, 1-2 bed properties for young families, 3-4 bed properties for middle aged couples with children and 1-2 bed apartments for the elderly. From the figure 10 we would expect the centre to have a higher density, with density increasing with distance to the CBD. In the next section I examine both the house price gradient and density gradients for Dublin.

Housing Affordability

Mortgage payments are amortised and are calculated based on a 25 year 3% fixed-term mortgage with the requirement of a 20% deposit. Affordability is then estimated using the 30th, 40th and 50th percentile weighted average household disposable income of €36,818, €46,405 and €57,385 respectively. House size breaks of 45 m^2, 74 m^2, 90 m^2, 140 m^2 and 190 m^2 are used. These house size breaks are based on the DoHPLG "Sustainable Urban Housing: Design Standards for New Apartments" (DoHPLG, 2018). 45 m^2 represents the minimum size for a one-bedroom unit, 74 m^2 the minimum size for a two-bedroom unit and 90 m^2 the minimum size for a three-bedroom unit. Increments of 50 m^2 are added to the minimum three-bedroom size to represent four and five bedroom properties.

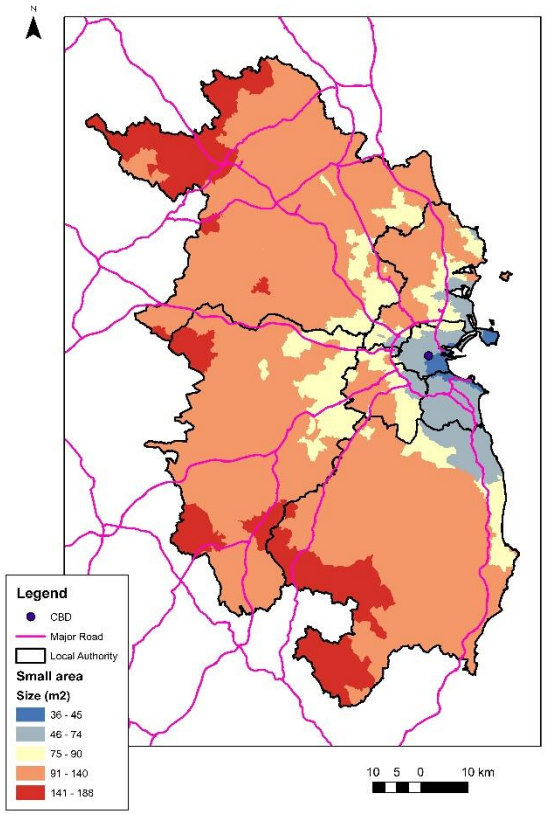
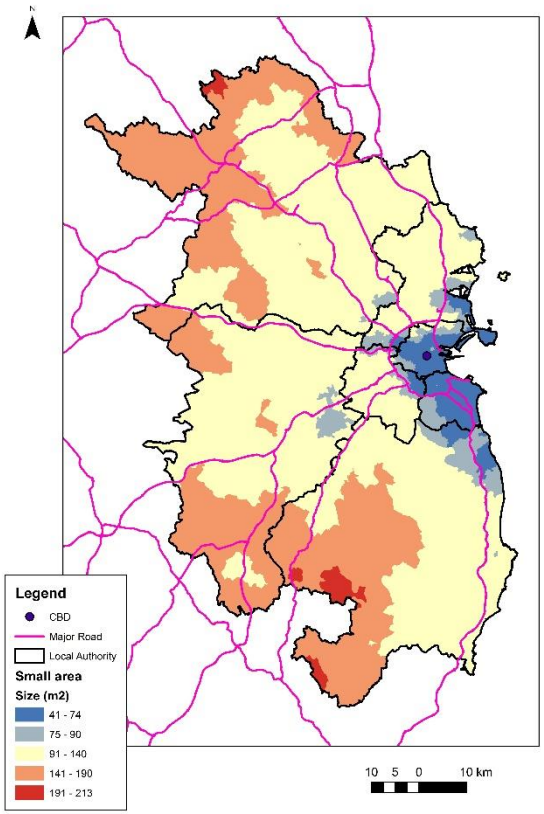
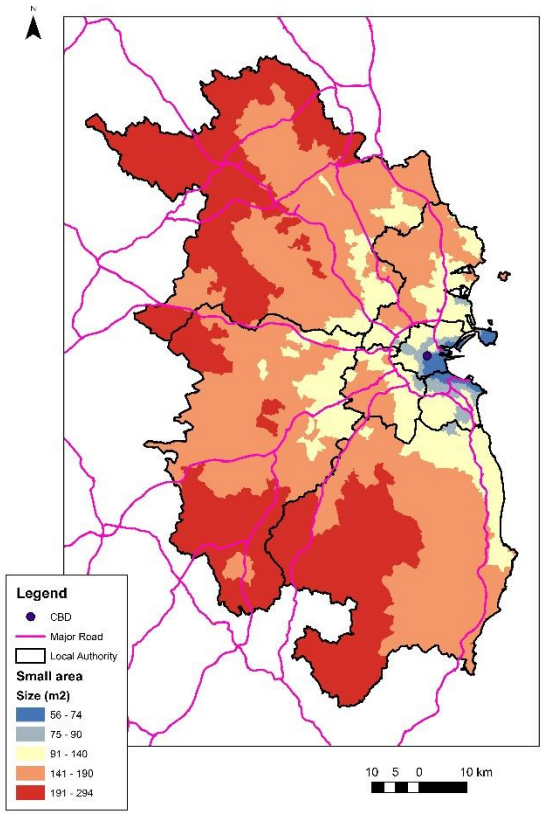
Figure 11 shows the results of an analysis that examines the maximum house size, which is affordable in a given area given a household's disposable income. Any property where the monthly mortgage payment would exceed 35% of income is deemed unaffordable. Throughout all scenarios any property larger than 90 m^2 in the city core (CBD) and south of the city urban is unaffordable. The maps highlight the attraction for households in moving to the suburbs and periphery where households can afford properties in excess of 90 m^2 in the majority of areas.

Figure 11 highlights the trade-off between price per square metre (ppsm) and house size in m^2. The figure illustrates why there is a migration of households into the suburbs, periphery and beyond. Households are trading off access to the CBD in order to benefit from a lower price per square metre and larger dwelling size. Given the higher levels of working from home in the future, there is an increasing demand for more space to accommodate a home office. Initial simulations suggest a movement in all activity towards the city edge (Batty, 2021).

Current mortgage lending rules may have the unintended consequence of overcrowding in the urban core. Households can trade-off housing size against price in order to qualify for a mortgage, purchase and afford a property. This property however may be small given the household size and living space requirements. Mortgage lending rules however make no recommendation around minimum property size or minimum m^2 per inhabitant.

One solution is to build more densely around the CBD, thus increasing the supply of living space. An important consideration is the provision of shared space/common areas/green space. Building densely on a given site area, can free up adjacent areas for shared space while at the same time maintaining a high population density.

Figure 11: Maximum size affordable property given income percentile

30th Income percentile – maximum affordable (35% of income) house size given ppsm in SA	40th Income percentile – maximum affordable (35% of income) house size given ppsm in SA	50th Income percentile – maximum affordable (35% of income) house size given ppsm in SA
		

Radial Profiles

Radial analysis is a powerful method to summarise complex data in an easy to understand format. In this section I summarise the results of the radial analysis focusing on the rent (house price) gradient for Dublin and density profiles for Dublin comparing to several European cities.

Figure 12: Rent (price) gradient of house prices for Dublin

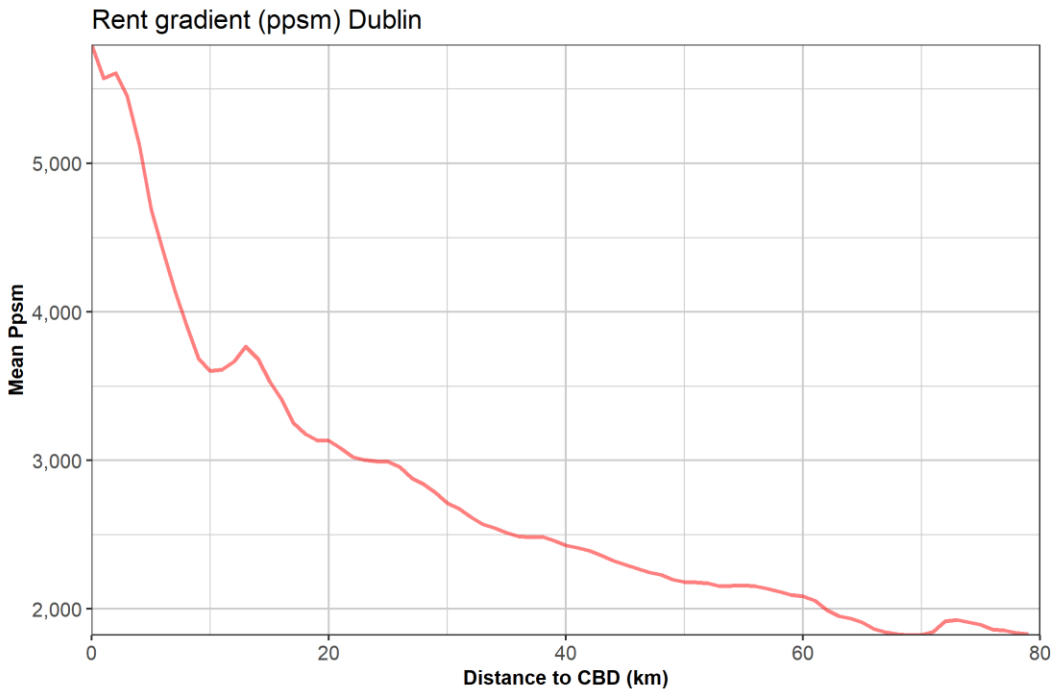
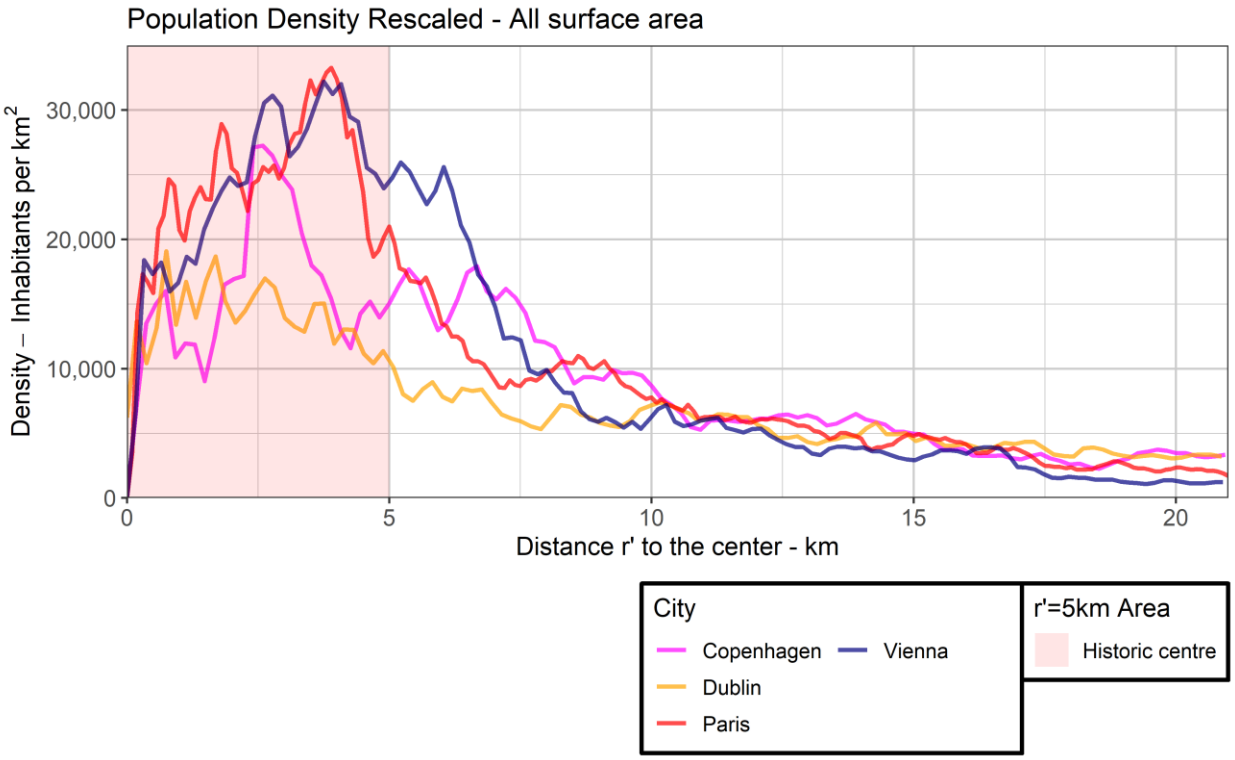


Figure 12 shows the 'rent' gradient for Dublin (curve of ppsm with respect to distance to the CBD). As we can see, the gradient for Dublin has a strong monocentric structure with the highest prices located in the city with a price gradient that decreases with distance to the centre. At 20km from the centre of Dublin, ppsm has halved in value. These findings however are not unique to just Dublin with house prices in many European cities having a gradient which decreases with distance to the city centre (Glumac et al., 2019; Helbich et al., 2014; Manzoli and Mocetti, 2019). Interestingly there is a smaller peak at 12km to the CBD. This corresponds to the suburban areas of Dalkey, Killiney and Lucan where prices remain high.

Figure 13: Rescaled population density for European cities



Rescaling on both horizontal and vertical axis, I control for city size. The maximum distance of r'=20km is equivalent to an actual distance of 10km non-rescaled for Dublin. In figure 12 we saw that the ppsm was ~3,500 at 10km from CBD. Figure 13 shows the rescaled population density for Dublin along with the European capital cites of Vienna, Copenhagen and Paris. The populations are: Dublin FUA = 1.76m, Copenhagen = 1.83m, Vienna = 2.68m and Paris = 11.7m. In figure 13 density is calculated:

Population density = Population of Ring / Total land area of ring

Figure 13 shows Dublin has a low population density in the historic centre area compared to other European cities. Beyond the rescaled distance r'=10km all cities have a similar population density. The majority of heterogeneity in density occurs within the historic centre area. The measure of population density is calculated using the entire surface area of the ring. Using only land area covered by residential structures a measure of net density is calculated.

Figure 14: Rescaled net population density for European cities

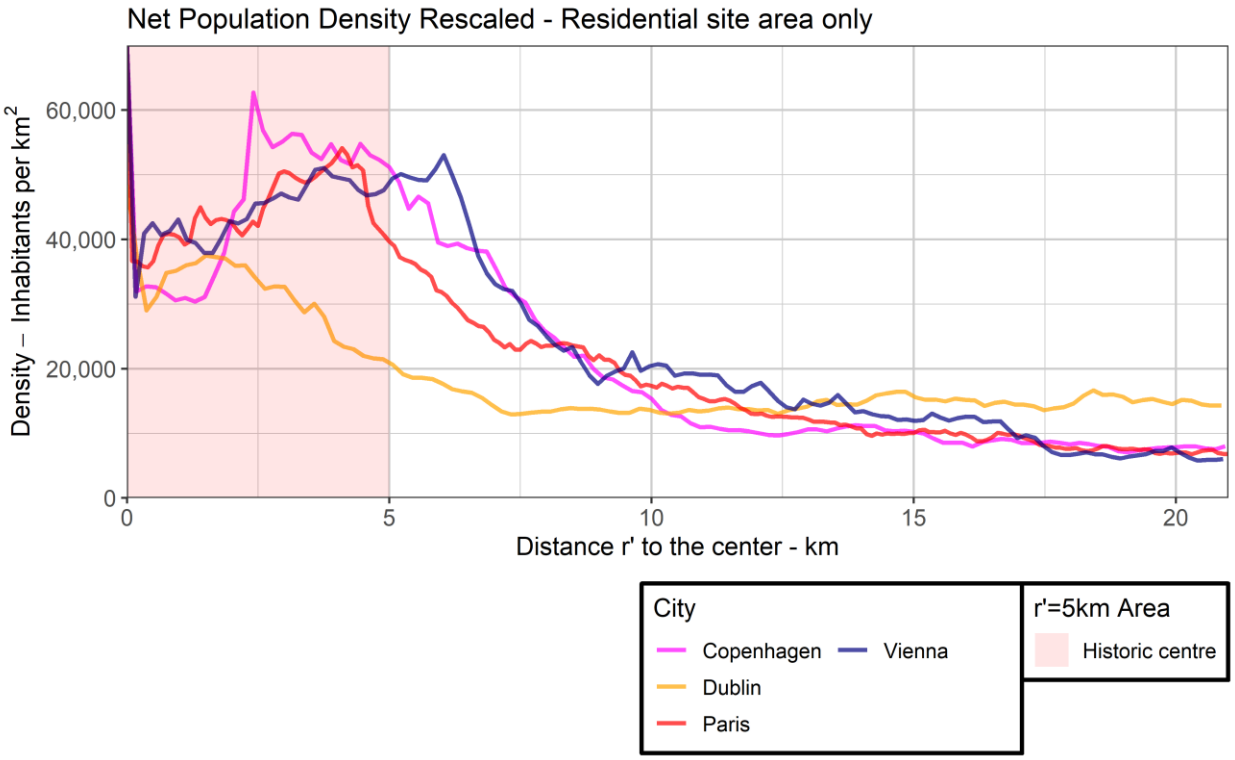
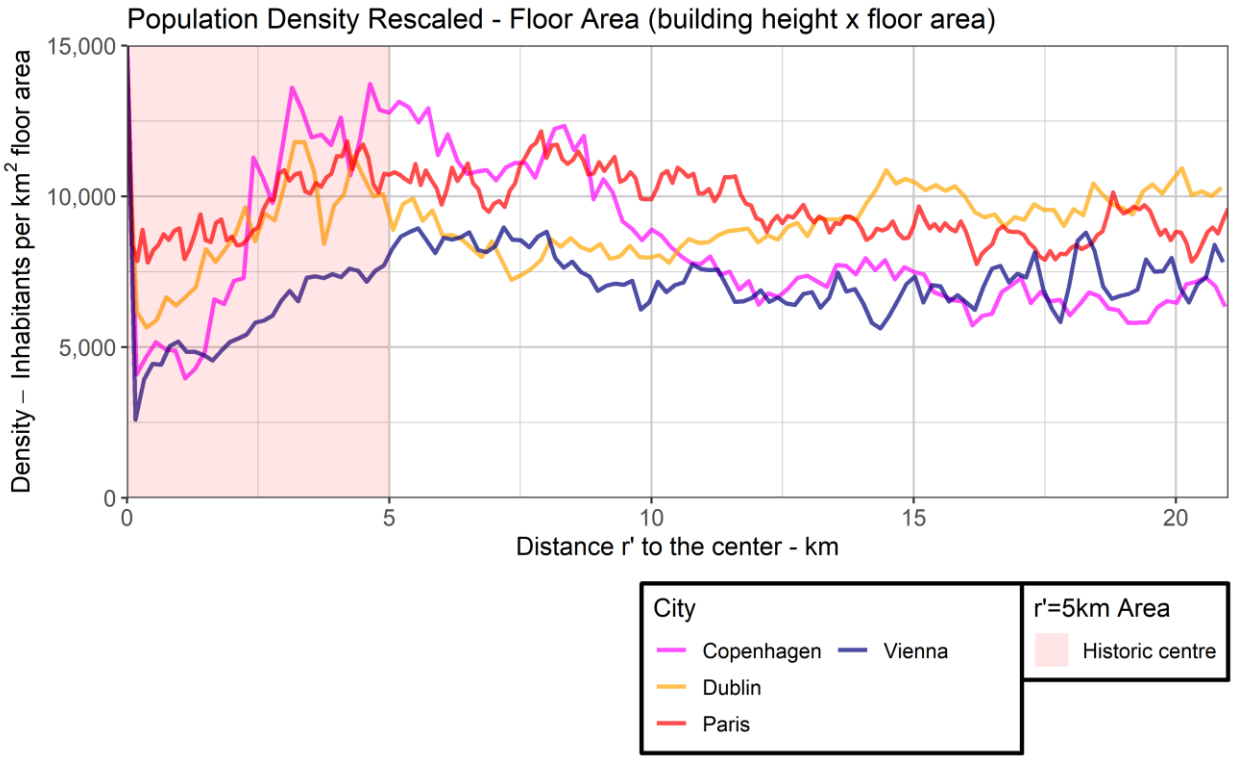


Figure 14 shows the rescaled net density for the four cities. Net density is calculated as:

Net Population density = Population of Ring / Total residential building footprint area within ring

At r'=1km all four cities have similar density levels, as we move out from the CBD net density in Dublin decreases quite rapidly whereas it increases in the other cities. Copenhagen has similar net density levels for Paris and Vienna whereas in figure 13 Copenhagen had a lower density. This suggests Copenhagen's residential living area is more densely populated compared to Paris and Vienna. In the next figure we examine this further by considering building heights and measuring total floor area and not just ground floor area.

Figure 15: Rescaled net population density (floor area and building height)

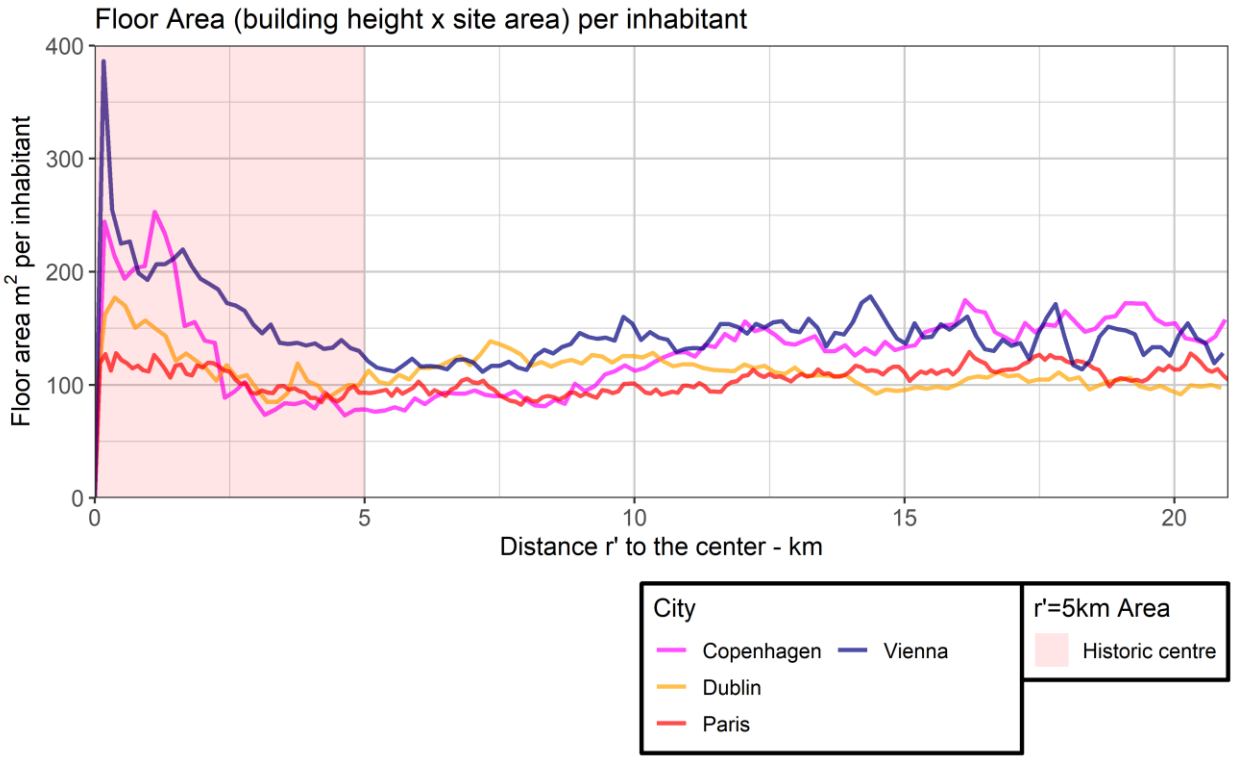


Using building heights data, total floor area is measured using the height converted to floors and site area.

Net Population density 2 = Population of Ring / Residential building footprint multiplied by number of floors of building

Figure 15 shows that density per floor area behaves in a similar way to net density. The profiles of cities are similar with the exception of Vienna which has the lowest density. This figure suggests that Copenhagen's buildings are tall, in figure 14 it had a high net density but with more added floor area the density has decreased relative to the other cities. Surprisingly Dublin has high levels of density in contrast to previous figures. This suggests that Dublin's buildings are not as tall relative to Vienna, Paris and Copenhagen leading to only a slight increase in the denominator (Density=People/Area).

Figure 16: Floor area per inhabitant for European cities



In figure 16 I measure total floor area per inhabitant. This gives an indication of the level of crowding in a city at various distances from the CBD. We can see that Vienna has the highest amount of space per person whereas Paris has the lowest. Dublin also has low levels of space around the historic centre with higher levels around 7km and low levels in the periphery. Analysing figures 13, 14, 15 and 16 together we can say that Vienna has a high population density made possible by having a high supply of living area. Paris has a high population density with space at a premium. This is not surprising given the attraction of Paris. Copenhagen has moderately high density and a moderate level of living space per inhabitant. Dublin is characterised by low density and low levels of living space per inhabitant.

Figure 17: Maximum building heights (metres) at radial distance to the CBD

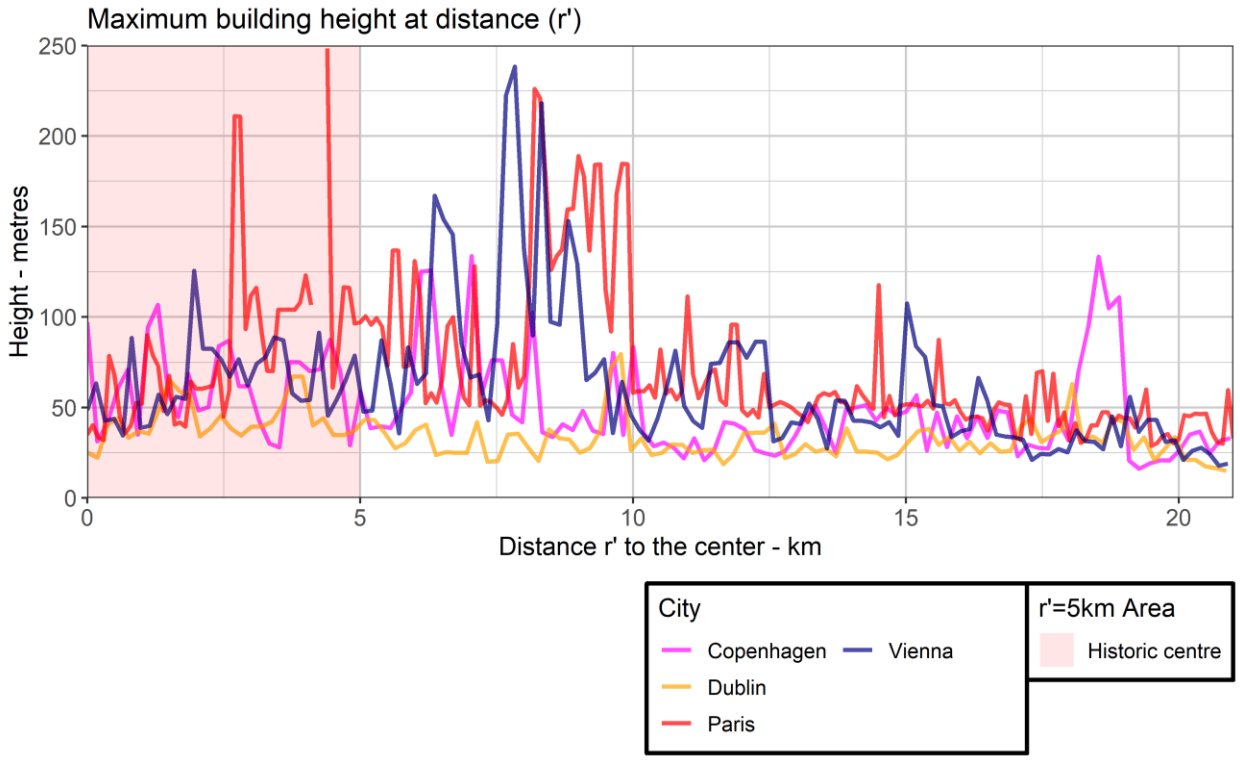


Figure 17 illustrates the differences in building heights between cities. Within the historic centre area, all four cities have similar heights. The differences however become apparent between r'=5km and r'=10km. It is clear that at these distances the cities of Vienna and Paris in particular have built taller buildings. Across the majority of distances, Dublin has the lowest height buildings. Across all distances Copenhagen's buildings are on average 13 metres taller than buildings in Dublin, over a large area this contributes to a large difference in living space. A lower supply of living space can in turn lead to a higher ppsm.

Conclusions

The aim of this research has been to highlight the importance of size in the debate around house prices, density and residential choice. Unlike many member states of the European Union, Ireland lacks a measure of price per square metre (ppsm). As a result, it is very difficult to compare areas, as house size is not homogenous across space. This paper has introduced how the kriging methodology can be used to produce a measure of ppsm at the Small Area level. The advantage of kriging over other methods is the error variance produced that indicates the level of uncertainty of the interpolated price in an area.

I showed how affordability is heavily influenced by ppsm, house size and distance to the CBD. Given recent increases in the commuting counties of Kildare, Meath and Wicklow, households are trading off access to the CBD in exchange for more living area and a lower ppsm. With greater levels of working from home and as a result less commuting, the fringe distance of Dublin will increase, that is the maximum distance that commuters consider acceptable. One of the drivers of

this outward migration is clearly the unaffordability of city centre living. Even households in the 50th income percentile can only afford a property of between 56-90 m^2.

The results of the scaling analysis clearly show that density in Dublin city centre is low by European standards. An increase in density and the supply of living space around the CBD is clearly required to increase affordability. The rent gradient curve for Dublin is extremely steep around the CBD, before decreasing at a decreasing rate. A flattening of the rent gradient curve particularly within 10km of the CBD should increase affordability and as a result increase density bringing it in line with the majority of major European cities. One of the largest determinants of the cost of providing housing in the CBD and ultimately the price is the cost of the site. The NPF has set a target of providing 40% of new housing in existing built-up areas without providing the instruments to improve land acquisition costs. As we approach the CBD, land costs increase, density increases and the capital/land ratio (given the higher land costs development on the land is more intense) should also increase. However, as we have seen, density around the CBD is low and hence the capital to land ratio is also low.

Reaching sufficient levels of density in strategic locations (around the CBD and areas with good public transport infrastructure) is required to ensure sustainable living patterns. With scarce resources, clustering developments makes the provision of services more efficient. The density gradients show that Dublin is an outlier among major European cities. Cities such as Vienna provided more living space to increase density whereas Paris lowered dwelling size requirements (minimum 9m^2). In order for Ireland to reach the same density levels, there are clearly two choices, increase building density by building higher providing more floor space or reduce size requirements. The first option is more desirable and would should provide a greater standard of living. Increasing the supply of living space should also reduce the ppsm and a result improve overall affordability particularly in close proximity to Dublin CBD.

One important component also missing from this analysis is the price per square metre for rental properties. Further research should compare the rent gradient (sales and rent prices) for Dublin with other European cities, in particular those with a higher density profile (e.g. Vienna and Paris).

References:

Ahrens A and Lyons S (2021) Do rising rents lead to longer commutes? A gravity model of commuting flows in Ireland. *Urban Studies* 58(2). SAGE Publications Sage UK: London, England: 264–279.

Anselin L and Lozano-Gracia N (2008) Errors in variables and spatial effects in hedonic house price models of ambient air quality. *Empirical economics* 34(1): 5–34.

Bailey TC and Gatrell AC (1995) *Interactive Spatial Data Analysis*. Longman Scientific & Technical Essex.

Baker E, Mason K and Bentley R (2015) Measuring housing affordability: A longitudinal approach. *Urban policy and research* 33(3). Taylor & Francis: 275–290.

Batty M (2008) The size, scale, and shape of cities. *science* 319(5864). American Association for the Advancement of Science: 769–771.

Batty M (2021) The Morphology of a Post-Pandemic City.

Bettencourt LMA (2013) The origins of scaling in cities. *science* 340(6139). American Association for the Advancement of Science: 1438–1441.

Cameron K and Hunter P (2002) Using spatial models and kriging techniques to optimize long-term ground-water monitoring networks: a case study. *Environmetrics: The official journal of the International Environmetrics Society* 13(5-6). Wiley Online Library: 629–656.

Charlton M (2010) Small Areas in the Republic of Ireland. In: 2010. Available at: https://www.yumpu.com/en/document/read/40709930/small-areas-in-the-republic-of-ireland-martin-charlton-institute-of-.

Corrigan E, Foley D, McQuinn K, et al. (2019) Exploring affordability in the Irish housing market. *The Economic and Social Review* 50(1, Spring): 119–157.

Cressie N (1990) The origins of kriging. *Mathematical geology* 22(3): 239–252.

CSO (2017) Residential Property Price Index February 2017: Background Notes. Available at: http://www.cso.ie/en/releasesandpublications/ep/p-rppi/rppifeb2017/backgroundnotes/.

Dijkstra L and Poelman H (2014) *A harmonised definition of cities and rural areas: the new degree of urbanisation, European Commission*. Regional Policy Working Papers.

Dijkstra L, Poelman H and Veneri P (2019) The EU-OECD definition of a functional urban area. OECD.

DoEHLG (2009) *Sustainable Residential Development in Urban Areas*. Dublin. Available at: https://assets.gov.ie/111252/bf943e60-30d5-4ac3-8c7a-dc4f68b3d688.pdf.

DoHPLG (2018) *Sustainable Urban Housing: Design Standards for New Apartments:*

Guidelines for Planning Authorities. Department of Housing, Planning and Local Government. Available at: https://www.opr.ie/wp-content/uploads/2019/08/2018-Design-Standards-for-New-Apartments-1.pdf.

Dubin R, Pace K and Thibodeau T (1999) Spatial autoregression techniques for real estate data. *Journal of Real Estate Literature* 7(1): 79–95.

EC (2016) *Mapping Guide v4.7 - for a European Urban Atlas*. Available at: https://land.copernicus.eu/user-corner/technical-library/urban-atlas-2012-mapping-guide-new.

Eurostat (2020) Population grids. Available at: https://ec.europa.eu/eurostat/statistics-explained/index.php/Population_grids#The_GEOSTAT_initiative (accessed 20 January 2020).

Fujita M (1989) *Urban Economic Theory: Land Use and City Size*. Cambridge university press.

Gehlke CE and Biehl K (1934) Certain effects of grouping upon the size of the correlation coefficient in census tract material. *Journal of the American Statistical Association* 29(185A). Taylor & Francis: 169–170.

Glaeser EL, Kolko J and Saiz A (2001) Consumer city. *Journal of economic geography* 1(1). Oxford University Press: 27–50.

Glumac B, Herrera-Gomez M and Licheron J (2019) A hedonic urban land price index. *Land Use Policy* 81. Elsevier: 802–812.

Goldberger AS (1962) Best linear unbiased prediction in the generalized linear regression model. *Journal of the American Statistical Association* 57(298): 369–375.

Helbich M, Brunauer W, Vaz E, et al. (2014) Spatial heterogeneity in hedonic house price models: The case of Austria. *Urban Studies* 51(2). Sage Publications Sage UK: London, England: 390–411.

Henderson JV (1985) *Economic Theory and the Cities*. Academic Press.

Hill RJ (2013) Hedonic price indexes for residential housing: A survey, evaluation and taxonomy. *Journal of economic surveys* 27(5). Wiley Online Library: 879–914.

Hoshino T and Kuriyama K (2010) Measuring the benefits of neighbourhood park amenities: Application and comparison of spatial hedonic approaches. *Environmental and Resource Economics* 45(3): 429–444.

Kilgarriff P, Charlton M, Foley R, et al. (2018) The Impact of Housing Consumption Value on the Spatial Distribution of Welfare.

Kilgarriff P, Lemoy R and Caruso G (2020) Change in Artificial Land Use over time across European Cities: A rescaled radial perspective. *arXiv preprint arXiv:2010.08401*.

Lemoy R and Caruso G (2018) Evidence for the homothetic scaling of urban forms. *Environment and Planning B: Urban Analytics and City Science*. SAGE Publications Sage UK: London, England: 2399808318810532.

Lyons R (2021) Daft.ie House Price Report: 2021 Q2. Daft.ie. Available at: https://ww1.daft.ie/report/2021-Q2-houseprice-daftreport.pdf?d_rd=1.

Manzoli E and Mocetti S (2019) The house price gradient: evidence from Italian cities. *Italian Economic Journal* 5(2). Springer: 281–305.

NPF (2017) National Planning Framework: Ireland 2040 Our Plan, Issues and Choices. Department of Housing Community and Local Government P (ed.). Dublin. Available at: http://npf.ie/wp-content/uploads/2017/02/Position-Paper-Issues-and-Choices-Ireland-2040-web.pdf.

OECD (2014) OECD: Note on Equivalence Scales. Available at: http://www.oecd.org/eco/growth/OECD-Note-EquivalenceScales.pdf.

Openshaw S (1984) The modifiable areal unit problem. In: 1984. Geo Abstracts University of East Anglia.

Pace RK, Barry R and Sirmans CF (1998) Spatial statistics and real estate. *The Journal of Real Estate Finance and Economics* 17(1): 5–13.

Roback J (1982) Wages, rents, and the quality of life. *The journal of political economy*: 1257–1278.

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

Tobler WR (1970) A computer movie simulating urban growth in the Detroit region. *Economic geography* 46(sup1): 234–240.

Wheaton WC (1977) Income and urban residence: An analysis of consumer demand for location. *The American Economic Review* 67(4). JSTOR: 620–631.

Wilkinson RK (1973) House prices and the measurement of externalities. *The Economic Journal* 83(329). JSTOR: 72–86.

Appendix:

Results using alternative method that applies Geographically Weighted Regression (GWR). Global model shows the results from the Global model, an OLS regression ignoring spatial autocorrelation.

Price = Size (m2) + number of baths + detached (dummy) + semi-detached (dummy) + apartment (dummy)

Number of beds variable dropped as was found to be highly correlated with size.

Global model

```
Call:

lm(formula = Price ~ Size_m2 + no_beds + no_baths + detached +

    semi_d + terrace + apart, data = sales1)


Residuals:

    Min      1Q  Median      3Q     Max

-556965 -103219  -17928   83764  555539


Coefficients:

             Estimate Std. Error t value Pr(>|t|)

(Intercept) 184430.0    14530.3  12.693  < 2e-16 ***

Size_m2       2363.4      115.9  20.399  < 2e-16 ***

no_beds      -7711.3     4814.2  -1.602  0.10931

no_baths     -8610.3     3915.0  -2.199  0.02793 *

detached    -33865.7    12453.1  -2.719  0.00658 **

semi_d       25305.9    11167.8   2.266  0.02352 *

terrace      29328.1    11008.7   2.664  0.00776 **

apart         2498.0    11898.5   0.210  0.83373

---

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


Residual standard error: 150700 on 3009 degrees of freedom
```

Multiple R-squared: 0.2648, Adjusted R-squared: 0.2631

F-statistic: 154.8 on 7 and 3009 DF, p-value: < 2.2e-16

Local

```
***********************************************************************
*          Results of Geographically Weighted Regression              *
***********************************************************************

*********************Model calibration information*********************
Kernel function: gaussian
Adaptive bandwidth: 16 (number of nearest neighbours)
Regression points: A seperate set of regression points is used.
Distance metric: A distance matrix is specified for this model calibration.

****************Summary of GWR coefficient estimates:******************
                Min.    1st Qu.     Median    3rd Qu.     Max.
Intercept  -12224.94   70680.98  103373.14  139741.04 340608.4
Size_m2       372.22    1905.55    2600.32    3187.50   5595.3
no_baths   -71771.60   -9067.13    2106.22   14125.02  85484.4
detached  -102669.85   32489.36   70879.65  119993.79 424291.8
semi_d    -367664.28    2489.23   34486.64   67375.55 245420.9
terrace   -147362.48  -18238.44    4142.29   26429.95 128846.7

***********************************************************************
```

Program stops at: 2021-08-29 22:23:02

GWR runs local OLS regression models for all small areas within the Dublin FUA. Local OLS is performed using small area at the centre of the kernel with a distance decay (gaussian) spatial weighted applied to neighbouring point. A variable bandwidth is utilised (selected using lowest AIC) to account for differences in the size of small areas. Figure 18 shows the results of the GWR model. The model can be considered hedonic as it controls for differences in the price attributed to house characteristics such as number of baths and house type. The figure shows the highest price per square metre is located in the city centre and south of Dublin. Next steps will explore using GWR and the RPPI dataset. Using a greater number of observations will lead to more robust coefficients.

Figure 18: GWR coefficient on the variable size (m2)