Monograph:

Cuberes, D., Roberts, J. orcid.org/0000-0003-2883-7251 and Sechel, C. (2019) Household location in English cities. Working Paper. Sheffield Economic Research Paper Series (2019001). Department of Economics , University of Sheffield. ISSN 1749-8368

The University Of Sheffield.

Department Of Economics.

Sheffield Economic Research Paper Series.

Household Location in English Cities

David Cuberes, Jennifer Roberts and Cristina Sechel

ISSN 1749-8368

SERPS no. 2019001

January 2019

Household Location in English Cities

David Cuberes*
Clark University, US
d.cuberes@clarku.edu

Jennifer Roberts
University of Sheffield, UK
j.r.roberts@sheffield.ac.uk

Cristina Sechel
University of Sheffield, UK
c.sechel@sheffield.ac.uk

* *Corresponding author*

January 2019

Abstract
This paper is the first to test an amenity-based sorting model for cities in England. We use individual level data on urban households for the period 2011-2016, combining this with data on local amenities to explore household location under both monocentric and polycentric assumptions about city structure. On average we find that there is no systematic relationship between income and household distance to the 'city centre', once neighbourhood amenities and other household characteristics are taken into account. Household heterogeneity is important, and as well as influencing location directly, we also find interactions between the effects of household characteristics and local amenities. There are also important differences between cities in England; for example higher income households seem to live further from the city centre in Birmingham, but closer to it in Newcastle. Our results reveal some important differences to the US evidence that has dominated this literature. Migrant status is important in England, and on average migrants live much closer to the city centre than non-migrants, but race per se does not seem to influence household location. Also it appears that in England only the employed (and those above the poverty line) are influenced by the availability of public transport; which is in direct opposition to the US evidence. Overall we conclude that the standard urban land use model provides a partial explanation of how households sort by income in cities, but that the role of amenities and household heterogeneity is large and warrants more attention.

JEL classification: R20; R23. **Keywords:** cities; household location; income; amenities.

Acknowledgments
We thank Paul Cheshire, Gordon Dabinett, Christian Hilber, Alasdair Rae, Jordan Rappaport, Craig Watkins and seminar participants at the Urban Economics Association Meetings and the Federal Reserve Bank of Boston for their useful comments. We are also very grateful to Mike Harding of Lancaster University for providing the google maps travel data. The UK Household Longitudinal Study (UKHLS) is an initiative by the Economic and Social Research Council, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by the National Centre for Social Research. National Grid Reference information for UKHLS was accessed via the Secure Data Service. Neither the original data creators, depositors or funders bear responsibility for the further analysis or interpretation of the data presented in this study.

1. Introduction

The standard urban model of Alonso (1964), Mills (1967), and Muth (1969) (AMM hereafter) predicts that, as one moves away from the centre of a city, housing prices should decrease. This must be the case in equilibrium in order to generate a compensating differential for the associated commuting costs of living in the suburbs. Under certain conditions, the model also predicts that household incomes should increase the further households are located from the city centre. The introduction of amenities into this model complicates this pattern and it is plausible that richer households will live closer to the centre than poorer ones because of differing preferences for amenities. The model presented in Brueckner et al. (1999), BTZ hereafter, precisely addresses this point: how the predictions of the AMM model are affected by the inclusion of amenities.

Our main goal is to test the amenity-based model of BTZ using recent data for England. We use data from a large representative survey of households to study household location in the eight largest English cities (excluding London) over the period 2011 to 2016.[1] We consider household distance from the central business district (CBD), and a number of other possible 'town centres', as well as their proximity to a range of 'amenities', such as public transport, retail outlets and property crime. Our micro level data allows us to account for household heterogeneity, including characteristics of the head of the household (such as gender, age, level of education and migrant status), and household level variables (such as income, access to a car and home ownership status).

We make three main contributions to the literature. First, we test for the effect that a large set of amenities has on household location and on the income-distance gradient under both monocentric and polycentric assumptions about city structure. Second, we explore how a number of other important household characteristics influence location. The importance of household heterogeneity has long been acknowledged in the literature (Wheaton, 1977; Anas, 1990; Epple and Platt, 1998), and it has been explicitly taken into account in recent economic geography models (see Redding and Rossi-Hansberg (2017) for a review). However, most empirical studies of the effect of income on location rely on area-level analysis, and hence abstract from household heterogeneity. Third, the existing literature on urban income gradients and amenity-based sorting is very US-dominated. We are the first to systematically study this phenomenon in English cities. As well as exploring how the patterns we identify vary across the eight English 'core cities', we also compare our findings with those from US studies. In Section 2 we present the BTZ model, and in Section 3 we review the relevant

[1] These are the 8 English 'core cities' (DCLG, 2011). London is excluded, as is often the case in the UK urban literature (e.g. Rae, 2013) because of its size. With a population of over 8 million in 2011, London is by far the largest UK city; the total population for the other cities in our sample is just over 4 million.

background literature. Data are described in Section 4 and the econometric method, empirical results and robustness checks are discussed in Section 5. Finally, Section 6 concludes the paper.

2. Theoretical framework

To guide our empirical work we rely on BTZ, who extend the AMM model by introducing amenities. This framework provides a way to test how amenities and household characteristics may affect sorting by income. Let x be distance to the CBD, and $a(x)$ be the level of amenities at that distance. We assume that household preferences are given by $U(e, q, a)$ where e is consumption of a numeraire non-housing good, and q is consumption of housing. The utility function is increasing in the three arguments. Commuting cost (per kilometre) is given by t and income by y; thus household disposable income at distance x is y-tx. Denoting p the price per unit of housing, the household's budget constraint is

$$\boldsymbol{e + pq = y - tx} \tag{1}$$

so the utility function can be written as $u(y - tx - pq, q, a)$. Maximization with respect to q, taking p as given, leads to the first-order condition

$$\boldsymbol{u^q = pu^e} \tag{2}$$

where superscripts denote partial derivatives. Letting $\bar{u}$ be the exogenous utility level, in a spatial equilibrium p must vary with distance so that every household has the same utility, thus $max_q u(y - tx - pq, q, a) = \bar{u}$. The slope of the bid-price function $p(x)$ can then be found by differentiating (2) with respect to x and using the first-order condition:

$$\boldsymbol{p'(x) = -\frac{t}{q(x)} + \frac{v^a\,[y-tx, p(x), a(x)]}{q(x)} a'(x)} \tag{3}$$

where $v[y - tx, p(x), a(x)]$ is the consumer's indirect utility function and $v^a = \frac{u^a}{u^e}$ represents the marginal valuation of amenities after optimal adjustment of housing consumption. The AMM model abstracts from amenities and so $v^a = 0$, which implies that the price gradient is a function only of the commuting cost and housing consumption.

Assume that there are two types of household: the poor, with income y_0, and the rich, with income y_1 ($y_0 < y_1$). Preferences can vary by income group so that both the utility function and the indirect utility function have sub-indexes 0 or 1. Since the rich have a higher opportunity cost of time, it is assumed that $t_0 < t_1$; thus there are two price gradients, one for the poor, $p_0'(x)$, and one for the rich, $p_1'(x)$. As is customary in standard urban models, land is occupied by the group who bid the highest for housing. This implies that the boundary between the area where the poor and the rich live, $\hat{x}$, satisfies $p_0(\hat{x}) = p_1(\hat{x})$. If $p_1'(\hat{x}) > p_0'(\hat{x})$, i.e. the slope of the price gradient for the rich is steeper than for the poor, then the rich live near the CBD and the poor live in the suburbs. This simple analysis shows that the slope of the price gradients of the two groups depend on whether amenities are more abundant near the CBD ($a'(x) < 0$) or far away ($a'(x) > 0$), and also on how much amenities are valued by each group (how large v_1^a is relative to v_0^a). Using this framework, BTZ argue that, since high income households are willing to pay more to live close to attractive amenities, city centres (like Paris) with high concentration of these amenities attract higher income residents.

It is worth stressing here that central to urban spatial models is the concept of either a single 'city centre' or CBD (as in the AMM and BTZ models) or multiple town centres or business districts (as in the polycentric models that originated with Fujita and Ogawa (1982)). The assumption of monocentricity has been rejected in many empirical studies (for example Giuliano and Small (1991) and Ahlfeldt and Wendland (2013)) thus in our empirical work we consider the possibility of both monocentricity and polycentricity, by specifying both a single CBD and some alternative 'areas of town centre activity' for each of our cities (see Section 4).

3. Related Literature

Traditionally, urban economics has focused mostly on the benefits of agglomerations in terms of productivity. However, as Glaeser et al. (2001) argue, another, perhaps equally or more important reason why people agglomerate is to enjoy consumption amenities. Many recent studies have explored the empirical relevance of this type of amenity (see Duranton and Puga (2014) for a comprehensive review). Building on the seminal models of Rosen (1974) and Roback (1982) urban economists have constructed quality of life indices as a way to measure the direct utility contribution of local consumption amenities (see Gyourko et al. (1999) for a review, as well as more recent work by Gabriel and Rosenthal (2004); Rappaport (2008); and Albouy (2016)). Most of this literature has focused on differences in quality of life between cities rather than differences across neighbourhoods within cities (e.g. Albouy, 2008; Carlino and Saiz, 2008). Perhaps because of this focus, most papers study how these amenities affect the city's population growth rate (Duranton and Puga (2014). Less work has considered how amenities within a given city matter for residential sorting. Exceptions include Rosenthal and Ross (2015) and Couture (2015) who consider amenities across neighbourhoods and Handbury (2013) who studies variation across different groups of individuals. Evidence on these issues from outside of the US is still very scarce.

Most of the existing literature on consumption amenities tends to explore these one at a time, (rather than considering a broad set simultaneously) with the aim of analysing how much these amenities are valued, rather than to explain how they determine the location of rich and poor households. Some of the amenities studied include sports arenas (Ahlfeldt and Maennig, 2009, 2010a, 2010b), architecture (Ahlfeldt and Holman, 2016), historical buildings (Koster et al., 2016), cultural heritage (Van Duijn and Rouwendal, 2013), urban renewal (Ahlfeldt, 2011), ocean views (Rappaport and Sachs, 2003), climate (Cheshire and Magrini, 2006; Rappaport, 2007), and natural amenities (Ng, 2008).

To our knowledge, a multi-city study of the effect of income and amenities on household location has only been carried out in few instances.[2] Glaeser et al. (2008) consider the role of public transportation in US cities. Koster and Rouwendal (2017) look at the impact of historic amenities in Dutch cities.[3] In a paper closely related to ours, Rosenthal and Ross (2015) study income sorting in US cities and analyse the role of several characteristics including public transit, public services, and the age of the housing stock. We consider all of these, as well as a number of others studied in the literature, and, in an extension to existing work we combine individual, household and neighbourhood level data, which enables us to explore the effects of household heterogeneity.

4. Data and Descriptive Statistics

We use individual level data from the five most recent waves of the UK Household Longitudinal Study (UKHLS) (University of Essex, 2016), and we combine this at the neighbourhood level with information on local amenities derived from a number of sources. The UKHLS is a panel survey of around 40,000 households. We use waves 3 to 7, which cover the period 2011 to 2016.[4] All adults in each household are interviewed and the data contain rich information on social and economic circumstances. We also utilise small area geographical identifiers for every household in the UKHLS; these are defined at the level of 'lower layer super output area' (LSOA). There are over 32,000 LSOAs in England, with an average population size of 1500; they are a UK equivalent to the 'neighbourhoods' described in much of the US urban economics literature.

[2] There is a different branch of the literature focusing on neighbourhood change, including many studies on gentrification (see Rosenthal and Ross (2015) for a review) and other aspects of neighbourhood change (see for example Lee and Lin (2018) who explore the influence of natural amenities on neighbourhood dynamics). A common theme in these papers is that neighbourhood change often takes several decades. Since our data covers a 5-year period, we instead focus on spatial variation at a given point in time.

[3] Madariaga et al. (2014) carry out an analysis similar to ours but they only consider the metropolitan area of Barcelona, and they consider the 'reverse' regression of income on distance.

[4] We exclude waves 1 and 2 because compatible data on some of our amenities is not available for those years.

Our analysis covers the eight English 'core cities' (Birmingham, Bristol, Manchester, Leeds, Liverpool, Newcastle, Nottingham, and Sheffield). We construct these city areas in a very similar way to the Primary Urban Areas (PUAs) defined in work done for the UK Department for Communities and Local Government: DCLG (2011), with the additional caveat that we also match these areas to the households in UKHLS. To do this we match LSOAs to postcodes and keep those UKHLS households located in LSOAs that belong to our 8 cities based on the postcode area code prefix (for example B for Birmingham, LS for Leeds).[5] To retain a focus on cities and their surrounding urban areas we exclude households who live in LSOAs that are defined as 'rural'.[6]

As DCLG (2011) admit when explaining the construction of PUAs, *"Probably the single most significant source of challenges faced was the core concept for the research, that of urban areas. Urban area boundaries change over time and are not readily matched to administrative areas."* (p.99). Thus, as DCLG (2011) also state, the process of defining these areas always involves an element of compromise. We believe that our method provides a useful approximation of those UKHLS households who live in the urban areas surrounding the 8 cities. We also exclude households who move home during the period (429 households, 10.5% of our sample). Following the logic of spatial equilibrium non-movers are assumed to be, on average, in equilibrium; they have made the trade-offs between the various aspects of their utility function to maximise utility at that location. Thus, these households represent the long-term equilibrium of the urban income gradient in our cities.

The UKHLS data are constructed at the individual level; with all individuals in each household being interviewed. We construct a dataset at the household level by using household level variables, such as household income and housing tenure, and individual level variables for the head of household, such as age, education and labour market status. Since one of the key aims of the paper is to analyse the link between household income and location, and one of the main determinants of income is employment status, we limit our analysis to those households where the household head is of standard working age i.e. between 18 and 65 years-old.

4.1 The dependent variable

Full details of all variable definitions and data sources are given in Appendix Table A1. Our dependent variable in the empirical model (described in Section 5) is distance to the CBD (and other 'town centres'). This is defined in a number of different ways to provide a robustness check on our

[5] A postcode is made up of four components (area, district, sector and unit).

[6] The urban/rural classification available at: www.ons.gov.uk/ons/guide-method/geography/products/area-classifications/index.html

results. To calculate distance from the main CBD in each of our cities (assuming monocentricity), we first need to locate this area; there is no consensus on how this should be done. Bowden (1971, p. 121) describes attempts to define the CBD as a 'perplexing task', and in a more recent paper Cheshire et al. (2018) state "We often talk about 'Town Centres', but defining their location and extent is surprisingly difficult. Their boundaries are hard to pin down and intrinsically fuzzy." (p.255). Brown (1987) argues that the CBD is often defined subjectively to include the principal shopping streets.[7] We employ two alternative definitions that are consistent with previous applications. Firstly, we use the location of the city's main railway station. Nathan et al. (2005) use a similar definition in their study of city centre living.[8] Secondly, we use the main Marks & Spencer (M&S) retail store; in line with real estate valuation approaches which view this as a key retail location which increases footfall to adjacent stores (see Schiller, 2001).

In further analysis we relax the monocentricity assumption and consider multiple Areas of Town Centre Activity (ATCA) defined in work carried out for the UK Office of the Deputy Prime Minister (ODPM, 2004, 2005). ATCA locations are based on the quantity and diversity of employment and the density of land in commercial use.[9] Household distance to the CBD or ATCA is estimated using the centre of the LSOA in which the household lives. We consider the distance of each household to their nearest ATCA and also the mean distance to any ATCA in their city. Given that the centre of the LSOA is an approximation, we also check the robustness of our results using data on grid reference, which gives the exact location for each household to a 1-metre resolution.[10] Since LSOAs are very small geographic units the results are very similar and none of our conclusions are altered. Further, while the use of linear distance to the CBD or ATCA derives from the standard urban model, it may not accurately reflect actual travel distance because of uneven spatial patterns in transport networks (Schuetz et al., 2018). In order to explore this we define a number of alternative dependent variables using information on travel distance and duration derived from Google Maps. For each possible journey from household location to CBD or ATCA we measure both travel distance and duration by bus and by car.

[7] Kantor et al. (2014) arbitrarily choose the city centre of New York to be Times Square and that of Los Angeles to be Pershing Square. Other papers use city halls (Asabere and Huffman, 1991; Atack and Margo, 1998; Schuetz et al., 2018), or alternative definitions based on market potential or travel-to-work areas (see, for instance, Ahlfeldt et al., 2017)

[8] The Economist has also argued *"Cities now measure their appeal by their stations. Businesses cluster around them: at King's Cross, a once-grimy part of north London, a postcode has been created for all the new buildings around the station, which was redeveloped in 2013. John Lewis, an upmarket department store, will open in the mall above New Street (which is indeed called "Grand Central") along with 60 other shops."* www.economist.com/node/21597904

[9] The ATCA data also define subset of 'retail cores' for each city, produced in a similar way but focusing purely on retail activity. In each of our cities the main railway station and M&S coincide with the 'retail core', thus adding weight to the use of these to define the main CBD.

[10] UKHLS grid reference data are available via the Secure Data Service www.ukdataservice.ac.uk/get-data/how-to-access/accesssecurelab.

To summarise, we have 16 measures of location relative to the 'city centre' for each household; 4 based on linear distance, and 12 from Google Maps (or 6 based on the monocentric assumption, and 10 assuming polycentricity). These are listed in Appendix Tables A2 and A3, which provide summary statistics and a correlation matrix.[11] Table A2 reveals that households' average linear distance to the CBD is around 8.4km, but average distance to the nearest ATCA is only 4.7km. There is a large amount of variation around the mean; some households are 33km from their nearest ATCA. Actual travel distances are longer than linear distances (by a factor of around 1.4), and are very similar for both bus and car travel. However, journeys tend to be faster by car. For example it takes an average of 13 minutes to travel to the rail station by car, and 23 minutes by bus. Table A3 shows that generally these different distance and time measures are very closely correlated. In the regression analysis in Section 5 we use linear distance to the rail station as our baseline measure (and the other measures in our robustness checks). Figure 1 shows the distribution of this variable, and reveals that most of the mass of the distribution is at short distances from the city centre.[12]

4.2 Explanatory Variables

Household income is monthly equivalised real net income. Figure 2 shows the distribution of income for the pooled data; this displays a significant degree of skewness with a mean of £1570. Table 1 shows the number of households and mean income by city and wave. Household income is very stable across the period; but there is variation across cities. Bristol has the highest mean income at both the beginning and end of the period, and Birmingham the lowest. Bristol, Nottingham and Sheffield display real mean income growth. And whereas Leeds experienced very rapid employment growth prior to the Great Recession (ODPM, 2006) it has become poorer compared to the other cities in the period we observe, going from the second richest in wave 1 to the second poorest in wave 7.

Figure 3 shows mean income by linear distance quintiles from the CBD (with the corresponding fitted quadratic line) for each city. This reveals heterogeneity across cities, and, as in the US (Rosenthal and Ross, 2015), these graphs suggest that distance and household income are not always strongly correlated. Birmingham (the largest and poorest of our cities) has the clearest monotonic pattern with income increasing steeply with distance. Of the eight cities, Birmingham was estimated to be hardest hit by the Great Recession, especially in terms of job losses (DCLG, 2011). Leeds and Manchester

[11] We do not present results for M&S in any of what follows as they are virtually identical to those for the rail station.

[12] We exclude 41 households who were clear outliers, with a distance to the CBD of over 40km. They are all located around Sleaford in Lincolnshire, and have Nottingham postcodes but are not generally viewed as part of the Nottingham area because they are closer to the smaller city of Lincoln.

have a similarly steep income gradient but with slightly lower incomes further out. Nottingham and Liverpool have less steep gradients, and with more noise. Sheffield displays a clear negative gradient, which may be due to the location of two universities and a large teaching hospital in its centre; whereas in other cities these are often in the suburbs. Bristol and Newcastle (the smallest cities, and among the richest) share a relatively flat pattern. To the best of our knowledge Newcastle is unique in the UK as the site of the only large-scale twentieth century planning policy with the explicit aim of 'rebalancing' disadvantaged neighbourhoods through 'positive gentrification' (Cameron, 2003).

Using data from multiple sources, we consider a broad set of neighbourhood amenities studied in the existing literature.[13] The definitions and data sources are detailed in the Appendix Table A1. The majority of our amenity measures come from Ordnance Survey Point of Interest (PoI) data for 2011-2015. We consider the availability of each amenity within 1000m of each household using ArcGIS. The amenities include: (i) public transport access points, such as bus and tram stops and train stations; (ii) public services such as schools and hospitals; (iii) historical and cultural attractions, such as historic buildings and museums; (iv) retail services such as shops and department stores; (v) facilities for eating out such as restaurants, cafes and public houses; (vi) sport facilities such as leisure centres and gymnasiums; and (vii) outdoor recreational facilities such as commons, parks and playgrounds.

In addition to the PoI data we also include four further amenities. Firstly, the age of housing stock to capture filtering theory. Using the number of dwellings by build period for each LSOA, we construct the ratio of old dwellings (built before 1900) to the total number of dwellings within 1000m of each household. Brueckner and Rosenthal (2009) argue, in the US context, that the rich may decide to live nearer the CBD because they value new construction more than the poor. However, we might expect differences in England where older buildings are often preferred by high income households for their uniqueness and historical value.[14] Second, we include the property crime rate, defined as the ratio of the number of property crimes to the resident population (again within 1000m of the household).[15] Third, the amount of social housing within 1000m of the household. This is likely to be valued very differently according to the housing tenure of the household in question. If they rent their home from the Local Authority then the amount of social housing is a positive amenity, but owner occupiers may perceive that large social housing estates convey a negative externality (Baum-Snow and Marion

[13] Climate as an amenity has been studied in the US literature but we do not consider it here as there is much less variation in climate between cities in England.
[14] The UK Home Owners Alliance Report 2017 found that more than twice as many people would prefer an older home (47%) to a new home (21%). https://hoa.org.uk/campaigns/publications-2/homeowner-survey-2017/
[15] We also consider a broader definition including robbery, shoplifting, criminal damage and arson, and a measure of non-property crime. Varying the definition does not affect our main findings.

(2009)). Finally, as a proxy for unobserved amenities provided by the home itself we include the household's Council Tax Band; which, while it reflects home value, is also a cost. In summary we include a set of 11 amenities. We assume that eight of these are positive, one is negative (the crime rate), and two (Council Tax band and social housing) are ambiguous in their effect on household utility.

Figure 4 shows plots of amenities by distance quintiles and clear spatial patterns emerge. Public transport access (graph 1) is highest in the CBD and sharply decreases with distance. All of our other positive amenities (graphs 2-7) are concentrated in the CBD, though there are signs of increased numbers at the periphery compared to the middle quintiles. The share of old housing declines with distance, but with some increase in the furthest quintile (graph 8). Property crime is highest closest to the CBD (graph 9), and the share of social housing (graph 10) also declines steadily with distance. The average Council Tax band displays the least clear pattern since the lowest bands are in quantiles 2 and 4 with the highest in the furthest quintile.[16] These graphs make it clear that, as BTZ show, households face trade-offs when making location choices, in addition to the fundamental trade-off between housing consumption and commuting time that is central to the standard urban model. Most amenities are more prevalent closer to the CBD, including the crime rate, an important negative amenity. Heterogeneity in households' valuation of these amenities will determine location choices; as Ng (2008) shows, households who value amenities may be willing to accept longer commutes in order to be closer to these amenities.

Spatial distributions of amenities by city are shown in Figure 5. Overall these reflect the patterns illustrated in Figure 4, but there are also interesting differences across cities. There is variation in the quantity of amenities in each city, and for most the greatest variation is nearest to the CBD. Bristol, for example, has more retail and eating out services in and near the CBD than any other city, although these drop-off very steeply (graphs 4 and 5). This may be a result of Bristol having two 'city centres' in close proximity; the older historic centre stretching from the River Avon to Clifton (a wealthy central neighbourhood), and the more modern Broadmead area immediately adjacent to the main rail station. Newcastle is least well served by these two amenities close to the CBD but has a shallow u-shaped distribution over distance, so that the furthest suburbs are relatively well served.

[16] This pattern is also reflected by the average number of rooms in the respondents' homes, and their self-reported approximate current valuations for their homes; information which is available in UKHLS.

Most cities display a steady decline in public transport access with distance from the CBD, but in Bristol there is a u-shape, so that those households furthest away have relatively good public transport (graph 1). In Birmingham and Manchester public services decline steadily as we move away from the CBD, whereas in the other cities there is a u-shape (graph 2). The least amount of variation appears to be in historic and cultural amenities (graph 3). Bristol has more of these in and near the centre than any other city, reflecting its long history as a trading port (Liverpool, another old port city, has the second highest level). Nottingham dominates in terms of sports facilities (graph 6), which decline with distance for all cities. Nottingham and Bristol also seem to have more recreational spaces than most cities, in all except the furthest suburbs (graph 7). There is a lot of variation in the share of old housing near the CBD (graph 8), with the largest amounts in Liverpool and Birmingham, and the lowest in Bristol, which has seen a lot of recent city centre residential development. Most cities display a reduction in the property crime rate as we move further from the CBD (graph 9). Leeds has the highest rate at nearly every quintile but also the steepest decline across distance. Variation in social housing is greatest at the closest and furthest distances from the CBD (graph 10). Manchester has the largest amount of social housing at these two extremes, and Bristol the lowest; this reflects the amount of social housing overall in these cities (24% for Manchester and only 15% for Bristol). Council Tax bands display the most heterogeneity across cities. These closely reflect house values and in Birmingham, Bristol, Liverpool and Sheffield they largely increase with distance from the CBD. In Leeds, Manchester and Nottingham the spatial Council Tax band profile is fairly flat. Newcastle displays a u-shaped pattern, with the lowest average Council Tax bands at most distances, except for the furthest quintile where there is a steep increase.

In order to more fully understand household location, we also consider a number of household characteristics (defined in Appendix Table A1). In particular, we control for age, sex, and education of the head of household; Albouy and Lue (2015) suggest that these wage predicting characteristics are a key driver of within city residential sorting in the US. We also control for whether or not the head of household is born in the UK, home ownership, car ownership and single adult households. In preliminary analysis we considered ethnicity rather than migrant status, but this was never statistically significant. In sensitivity analysis (Section 5.1) we also consider how our results vary according to whether or not the household head is employed, and whether the household has children. Table 2 shows the distribution of household characteristics by distance from the CBD. This reveals a fairly uniform distribution of the number of UKHLS households by distance, and shows that mean household income increases with distance. Household heads closer to the city centre are (on average) younger, less likely to be male, more likely to be a migrant and more likely to be not employed. The pattern for education is not as clear; the proportion with higher education is almost exactly the same in the nearest and furthest quintile, with the highest proportion seen in the mid quintiles. Households nearer the CBD are less likely to own a car and more likely to be single adults. They are less likely to own their own home and more likely to rent, both from the Local Authority and private sector.

Appendix Figure A1 shows the distribution of household characteristics by distance for each city. There are similar patterns across cities for age and sex of household head, proportion of homeowners, private renting and car ownership (graphs 1, 2, 6, 8 and 9). However, for the other characteristics some interesting differences are revealed. Most cities see a gradual decline in the proportion of migrants with distance from the CBD (graph 3). In Birmingham (which as the highest proportion overall at 23%) the decline is particularly steep with 52% migrant households nearest to the CBD and only 4% in the furthest quintile. In Bristol (with only 12% migrant population) the pattern is an inverted u-shape with the lowest proportion of migrants in quintile 4. Similarly most cities display a decline in the number of not-employed household heads as we move further from the CBD (graph 4). Liverpool and Manchester are different; with a u-shape in Manchester dipping to only 18% of household heads not in work in quintile 3, with the opposite inverted u-shape in Liverpool where 47% are not employed in quintile 3. The pattern of household heads with higher education (graph 5) also varies a lot across cities. There is an increase with distance for Birmingham, Liverpool and Nottingham, and a decrease in Bristol and Sheffield, which is particularly sharp for the latter. Leeds and Manchester display an inverted u-shape, and in Newcastle the pattern is quite flat, which may reflect the deliberate urban planning policies discussed above. The pattern of households living in social housing also varies a lot (graph 7). It is fairly flat in four of our cities, although the absolute levels vary a great deal. In three of our cities (Birmingham, Leeds and Nottingham) it declines steadily with distance, and in Bristol there is a u-shaped pattern. The pattern of single person households also varies by city (graph 10). It largely declines with distance from the CBD but again Bristol displays a strong u-shape with the lowest proportion of single person households in quintile 3.

5. Empirical Models and Results

To analyse household location we estimate the following model:

$$logD_{ijkt} = \alpha + \beta logY_{ijkt} + \gamma_1 \, Z_{jkt} + \gamma_2 \, X_{ijkt} + u_j + \tau_t + \varepsilon_{ijt} \qquad (4)$$

where i is household, j is city, k is neighbourhood, t is wave, D is our location variable, distance from the CBD (and other ATCAs),[17] Y is household income, u_j is a city fixed-effect, τ_t is a wave fixed-effect, and X is a set of household characteristics. Z is a vector of eleven neighbourhood amenities. We estimate this model using OLS with Conley standard errors adjusted for spatial and serial correlation using the Stata program *ols_spatial_HAC* (Hsiang, 2010).[18]

A positive value of β indicates that households locate further from the CBD the higher their income. We explore the extent to which income is significantly associated with distance when neighbourhood amenities and other household characteristics are taken into account. γ_1 shows how amenities vary with households' distance from the CBD. For most of our amenities (except the Council Tax band) we expect this coefficient to be negative; the more amenities in the neighbourhood the closer that household is likely to be to the centre. γ_2 gives the association between household characteristics and location. It is important to emphasise that we do not interpret our estimates as causal. Establishing convincing causal relationships would require an exogenous change in our explanatory variables through some type of natural experiment, or require us to find instrumental variables for each of our regressors; arguably an impossible task.

Table 3 shows results for four different specifications of equation (4), which include in columns: (1) household income (Y) but no further controls; (2) income and other household characteristics (X); (3) income and amenities (Z); (4) income, household characteristics and amenities. There is a positive association between household income and distance from the CBD but this is reduced substantially, in both size and significance, as controls are added. The inclusion of household characteristics alone reduces the income coefficient from 0.169 to 0.055. In model (4) the income coefficient is small and not statistically significant. This is consistent with the finding of Rosenthal and Ross (2015) for the US. Looking at the effect of other household characteristics we see that all except owner-occupier status are significantly associated with distance from the CBD. Households with a male head, older people and car owners tend to live further from the CBD, while those with higher education, single person households and migrants live closer. The effect of migrant status is particularly large suggesting that on average migrants live 25% closer to the CBD than non-migrants; this compares,

[17] Distance to the CBD is the dependent variable, so the cost of commuting is implicit in the model. This cost increases with distance, and the opportunity cost of travel time varies with income, which is controlled for in our model. Consistent with this we find a high correlation between commuting distance and time (Appendix Table A3).

[18] Estimating the model by simple OLS with standard errors clustered at the city and household levels gives smaller standard errors but makes very little difference to the results reported here.

for example, to 10% closer for household heads with higher education.[19] This result is consistent with the work of Schuetz et al. (2018), who find that in US cities, tracts near the CBD have larger minority population shares. The association between amenities and households' distance to the CBD is largely negative, reflecting the fact that amenities tend to be more available closer to the centre. For example, the negative coefficient on public transport, is a similar finding to Glaeser et al. (2008), who show that in the US reliance on public transit generally declines sharply with distance from the city centre.[20] Our results suggest that an increase in 10 public transport access points (such as bus or tram stops) is associated with a 2% reduction in distance from the CBD. Similarly, old houses are more prevalent near the CBD and we find a negative effect of old property on household distance. We also see in Table 3 that other amenities like public services, eating-out establishments and outdoor recreational facilities are negatively associated with distance to the CBD. The exception to this are the amenities provided by the home itself (proxied by the Council Tax band) as larger homes are more prevalent in the suburbs.[21]

Recognising the complexity of these relationships we also consider a number of interactions between the variables in our model. These are shown in models (5a) to (5c), where subscripts are suppressed for ease of exposition and all variable definitions are as for equation (4) above.

$$logD = \alpha + \beta logY + \gamma_1\ Z + \gamma_2\ X + \partial_1 Y.Z + u_j + \tau_t + \varepsilon \qquad (5a)$$

$$logD = \alpha + \beta logY + \gamma_1\ Z + \gamma_2\ X + \partial_2 X.Z + u_j + \tau_t + \varepsilon \qquad (5b)$$

$$logD = \alpha + \beta logY + \gamma_1\ Z + \gamma_2\ X + \partial_3 X.Y + u_j + \tau_t + \varepsilon \qquad (5c)$$

Model (5a) includes interactions between income (*Y*) and amenities (*Z*). The total effect of income on distance in the presence of amenities is given by the sum of β and ∂_1. Model (5b) includes interactions between household characteristics (*X*) and amenities (*Z*). The total effect of household characteristics in the presence of amenities is given by the sum of γ_2 and ∂_2. Model (5c) explores

[19] In a model of Y = a+bD+e, where Y is in log form and D is a 0/1 dummy variable, if D switches from 0 to 1 the % impact on Y is 100[exp(b)-1].
[20] They also show that in the US car travel is faster than public transit, on average, which is also true for our data (Appendix Table A2).
[21] In Table 3 model (4) historic and retail amenities are not significantly associated with distance. They are each significant if included individually, but they are highly collinear with correlations of > 0.8. Also the city-by-city results in Table 4 reveal a lot of heterogeneity in the relationship between retail and historic amenities and distance.

interactions between household characteristics (*X*) and income (*Y*). The total effect of income in the presence of household characteristics is given by the sum of β and ∂_3.

All models are estimated with a full set of household characteristics and amenities as in column (4) of Table 3 and for conciseness we simply summarise these results here.[22] For model (5a) in all but one case the main income effect ($\hat{\beta}$) is statistically insignificant (as in Table 3 column 4). There are only two significant interactions between household income and amenities; with sports facilities and social housing. In the former case the main income coefficient is positive and significant (0.085) and the interaction is small and negative (-0.009), suggesting that while on average higher income households live further from the CBD, this distance is attenuated in the presence of sports facilities, which are more prevalent closer to the CBD. For social housing the main income coefficient is not statistically significant, and the interaction is small and positive (0.004); suggesting that in the presence of social housing (which is more prevalent closer to the CBD) higher income households will live further away.

For model (5b) there are a number of interactions between household characteristics and amenities. Firstly, age interacts with the amount of old property; in this case there is no main age effect but there is a positive interaction with old property, such that on average older people live further from the CBD the higher the level of old property in the neighbourhood; this may suggest that older people have a stronger preference for older property, which seems intuitively reasonable but we can find no other evidence to verify this. Secondly, car ownership interacts with the presence of historic and cultural, retail and eating out amenities, old property and the property crime rate. In all cases there is no main effect of car ownership (in contrast to the result where no interactions are considered and the effect of car ownership is strongly positive), and the interactions are all small and positive. This means that while on average there is no systematic relationship between car ownership and distance, car owners tend to live further from the CBD the higher the presence of these amenities in their local neighbourhood. Thirdly, education interacts with the crime rate and social housing; the main effect of education is relatively large and negative while the interactions are small and positive. Where household heads have higher education they tend to live closer to the CBD, but this distance is increased in the presence of higher levels of crime and social housing, which are more prevalent closer to the CBD. One of the most important household characteristics when interacted with amenities is whether the head of the household is an immigrant. Migrant status interacts positively

[22] Note that for models (5a) and (5b) all interactions are estimated individually due to collinearity, alongside the full set of main effects. In (5c) all interactions are included in the same model as there are no major collinearity problems.

with a number of amenities including public transport, historic, eating out and sports. Migrants on average live closer to the CBD but this distance is greater the higher the level of these amenities in the local area. The importance of immigration status contrasts with studies based on US cities which tend to focus on race instead (see for example, Bayer et al., 2004; Waldfogel, 2008). In our analysis we did explore the effect of race but it was never a significant variable while immigration status is significant in all cases.

Finally, for model (5c) there are two significant interactions between household characteristics and income, a negative one with age of head of household and a positive one with migrant status. In both cases the interactions are small relative to the main effects. Older households on average live further from the CBD but this distance is attenuated the higher their income. The opposite is true for migrants who tend to live closer to the CBD, but this distance is increased the higher their income.

The regression results reported so far are based on pooled data for our eight cities so implicitly assume homogeneous effects across these cities. However, the descriptive statistics presented in Section 4 suggest that this may not be the case. The equivalent of Table 3 model (4) estimated for each of our eight cities individually is shown in Table 4. We have also tested a number of cross-city restrictions for each coefficient estimate.[23] In all cases the hypothesis that all 8 cities have equal coefficients is rejected for all variables. In order to see which cities differ in each case we have tested each estimate in a set of pairwise comparisons i.e. city 1 with cities 2, 3, …, 8; city 2 with cities 3, 4, …, 8 etc.; as well as testing different sub-groups of cities. Given the large number of tests we report only the key results here.

Consistent with Figure 3 a positive income-distance relationship is statistically significant only for Birmingham and Leeds; our two largest cities. In contrast, there is a (significant) negative relationship in Newcastle, and no significant relationship for the remaining cities. Cross-city restriction tests for the income coefficient reveal that statistically all cities except Birmingham can be pooled. Generally, where household characteristics are important they reflect the patterns displayed for the pooled city analysis presented in Table 3 with a few key exceptions. While there is a large degree of consistency across cities in the strong negative relationship between migrant status and distance, Bristol is the one

[23] Restrictions are tested using the 'seemingly unrelated' SUEST estimator in Stata v15.1 with no adjustment for spatial and serial autocorrelation. The coefficient estimates are the same as with a spatial estimator but the standard errors are smaller so these test results may slightly exaggerate the amount of heterogeneity across cities.

exception to this where there is evidence (at the 10% level of significance) that migrants live further from the CBD than non-migrants. Restriction tests confirm that Bristol cannot be pooled with the other cities for this coefficient. For higher education and home ownership status, Sheffield is different to the other cities, with a much larger negative gradient for the former and positive for the latter. For car ownership, there appear to be two subgroups; Bristol, Liverpool, Manchester, Newcastle and Sheffield where there is no effect of car ownership on distance to the CBD, and Birmingham, Leeds and Nottingham where it is positive.

The results also reveal variation in the influence of amenities, consistent with Figure 5. For example, looking at public transport amenities, in Liverpool, Nottingham and Sheffield these are negatively associated with household distance from the CBD, but the opposite is true for Bristol and there seems to be no systematic relationship in the remaining cities. Similarly public service provision is negatively associated with distance in most cities but is positive in Bristol; statistically, Sheffield is an outlier with a larger negative gradient. Retail amenities are positively associated with distance from the CBD in five of our cities, perhaps reflecting the presence of large 'out-of-town' shopping centres; such as Meadowhall in Sheffield and the Metrocentre in Newcastle. Only in Manchester do retail amenities have a negative association with distance, which reflects Manchester's identity as a shopping destination city in the North of England. Outdoor recreation facilities are distributed heterogeneously, with a positive relationship with distance for Birmingham, Liverpool and Sheffield and a negative relationship for all other cities. The size of the effect varies; for example one more recreation facility is associated with a 10% reduction in distance to the CBD in Leeds and only 4% in Newcastle. Property crime and social housing are both negatively associated with distance for all cities. The crime gradient is particularly steep in Bristol; here a 1 percent increase in the crime rate is associated with being 13% closer to the CBD, compared to Birmingham where the same 1 percent decrease is found 2% closer to the CBD. The relationship between Council Tax and distance is present (and positive) only for Leeds, Newcastle and Sheffield.

In further analysis not reported here we have also explored the interaction models of equations (5a) to (5c) by city. While there are a small number of differences in comparison to those reported above, the main interactions reported above largely remain important in the city-by-city models.

5.1 Robustness checks: sub-group analysis

To explore the robustness of our results, and to uncover some of the possible underlying mechanisms, in Table 5 we estimate equation (4) for a number of different sub-groups. We discuss the most relevant results here. Firstly, in the analysis presented so far we have excluded the 429 households who move home during the period. Compared to non-movers, these movers are younger, more likely to have higher education, less likely to be home owners and more likely to be single person households. The proportion of migrants is very similar in both groups. If we include these movers in the analysis the results are virtually identical to those presented so far. However, column (1) of Table 5 presents the equivalent to Table 3 column (4) for movers only and we see some important differences, which are reflected in differences in coefficient size, not simply statistical significance due to the smaller sample size. Unlike non-movers there is a positive and significant income gradient for movers. There is also a positive effect of home ownership, but no effect from car ownership; also the migrant-distance gradient is much smaller for movers. Access to public transport and public services appears not to be important for movers, and the Council Tax band of the home has no effect.

Since commuting is a prominent feature of the dominant theories in urban economics, in columns (2a) and (2b) we split the sample according to whether or not the household head is in employment. Neither group displays an income-distance gradient, but two key differences are that transport options (via car ownership and public transport access) are important only for the employed group, and local retail opportunities are only important for the not-employed group. Clearly an important difference between employed and not-employed households will be the level of household income; average real equivalised monthly income is £1748 for the employed households and only £1135 for the not-employed. In a separate analysis not reported here we split the sample between those above and below the poverty line (defined as 60% of median income); there are 8510 households in the former group and 1639 in the latter. The results are very similar to those reported in columns (2a) and (2b), public transport access is only important for those households above the poverty line, and retail outlets are only significant for those below it. These results seem to contrast with those from the US literature, where it has been argued that the poor prefer public transport, because car ownership is too expensive for them and also because they have a lower opportunity cost for commuting time (see for example LeRoy and Sonstelie (1983), Glaeser et al. (2008) and Pathak et al. (2017)). One possible explanation for the insignificance of transport in the location model for the not employed in England is that these households are less likely to be involved in active job search than those in the US; either as a preference or due to constraints. Transport costs may prohibit search and previous work has documented the very low geographic mobility of these groups in the UK citing cost as a factor (Cass et al., 2005; Kelly, 2013) The fact that local retail amenities only correlate with the location of the poor or not-employed add support to this, suggesting that they shop locally in a way that the better off and working population do not. It is also the case that some of the not-employed are prevented from working by health problems; 41% of our not employed group report poor health compared to only 13% of our employed group. Further, 17% of our not-employed group are on disability benefits, and this is much higher than generally found in the US.[24]

In columns (3a) to (4b) we distinguish groups who, as well as potentially having different location preferences, are also expected to be differentially constrained in their location choices. In (3a) and (3b) home owners are likely to be more constrained than renters, and similarly in (4a) and (4b) households with children are likely to be more constrained than those who do not have children. Car ownership and access to public services are significant for owners but not renters; although both groups share a similar effect of public transport access. In contrast, eating out opportunities are only significant for renters and historical and cultural amenities only for owners. Similarly, households with children are influenced by access to public services (schools and hospitals), whereas those without children are more influenced by eating out and sports amenities. The final two columns reveal differences between younger and older household heads. Only the younger age group (aged below 40) are influenced by access to public transport, public services and the Council Tax band of the property. In contrast the older group (aged 40 and above) are affected by eating out amenities and have a much larger effect from the amenities provided by old property. Another perhaps surprising result that emerges from Table 5 is that there is a large degree of consistency in the effects of the amount of outdoor recreational amenities, the property crime rate and the amount of social housing regardless of which sub-group we look at. Finally, and not reported here for conciseness, we excluded full-time students from our analysis in order to explore whether our results are affected by the growth in student numbers in these cities resulting from the expansion of higher education (Tallon and Bromley, 2004). This made very little difference to the results in Table 3, column (4), suggesting that the expansion of student numbers in city centres is not driving our results.

5.2 Robustness checks: alternate distance measures and ATCAs

In this section we explore the robustness of our results to different distance measures and also compare the assumption of monocentricity to that of polycentricity. Appendix Table A4 reports the results of estimating equation (4) with five different definitions of the dependent variable chosen from the 16 measures we described in Section 4.1.[25] These results can be compared with Table 3 column (4) which uses linear distance to the main rail station as the dependent variable. The first three columns of Table A4 also use the rail station as the definition of the CBD, but instead of linear

[24] Less than 5% of the working age population of the US were on disability benefit is 2016. www.ssa.gov/policy/docs/statcomps/di_asr/2016/sect01.html

[25] The remaining measures all provide results that are very similar to those reported here.

distance for each household we use Google Maps data on the actual driving distance (1), as well as the car (2) and bus (3) journey times. Columns (4) to (6) relax the monocentric city assumption and consider Google Maps driving distance (4), as well as the car (5) and bus (6) journey times to the household's nearest ATCA. The overriding message from Table A4 is that regardless of which distance (or journey time) measure we use our main story does not change. There is no significant income association with distance (or journey time) in all cases, but household characteristics and neighbourhood amenities are important correlates of household location. Looking across the columns of Table A4 there are some important quantitative differences in coefficient estimates, depending on whether we assume monocentricity (columns 1 to 3) or polycentricity (columns 4 to 6). Most household characteristics and amenities seem to have larger (in magnitude) effects under polycentricity. Further, comparing the distance measures in columns (1) and (4) with the journey time measures in the remaining columns, all variables seem to have smaller effects when journey times are used. One, perhaps surprising, result is that the differences between using distance and time measures are greater than those between monocentric and polycentric assumptions. We have also explored using these alternative dependent variables in the city-by-city versions of equation (4).[26] The key results reported in Table 4 are largely unchanged (in substance) by the use of different distance measures, and the relaxing of the monocentricity assumption. Two exceptions are that in Leeds and Manchester, where the public transport amenity was not significant under the linear distance to the main CBD assumption, it is significant (and negative) if actual travel distance is used instead.

In further analysis not reported here we also estimate equation (4) using area level averages, rather than data on individual households. We averaged the data across the 2400 LSOAs present in our data and also the larger medium layer super output areas (of which 887 are represented). Like the other robustness checks reported here, this area level analysis does not change our results substantively. There is no systematic relationship between area level income and distance, but amenities and average household characteristics are important.[27]

6. Discussion and Conclusions

This paper is the first to test the amenity-based sorting theory of BTZ for cities in England. We have tested this under both monocentric and polycentric assumptions about city structure. On average for the 8 English 'core cities' we find that there is no systematic relationship between income and

[26] These results are not reported here for conciseness.

[27] Given our data is from a sample survey the cell sizes behind the area level averages for the personal characteristics, including household income, are very small, especially for the LSOA analysis.

household distance to the CBD once neighbourhood amenities and other household characteristics are taken into account. Amenities like public transport, retail outlets and the local crime rate are important correlates of household location. Moreover household heterogeneity is important, and as well as influencing location directly, there are some important interactions between other household characteristics and income, and between these characteristics and local amenities. These main results hold under both monocentric and polycentric assumptions. We also find that there are important differences between cities in England (for example higher income households seem to live further from the CBD in Birmingham, but closer to it in Newcastle, even after full account is taken of other influences). This suggests one should be careful in drawing conclusions when single city studies are used to validate a theoretical model, as is often the case in the literature (for example Agarwal et al., 2012; Pathak et al., 2017; Schmidheiny, 2006).

Our results reveal some important differences to the US evidence that has dominated this literature. Migrant status is important in England, and, in all cities except Bristol, migrants live much closer to the CBD than non-migrants (although this tendency is attenuated the higher their income), but race per se is not important to household location in England. Also it appears that in England only the employed (and those above the poverty line) are influenced by the availability of public transport; the location of the not-employed (and the poor) does not seem to depend on public transport access. This is in direct contradiction to the US evidence of Glaeser et al. (2008) and Pathak et al. (2017). This is an important finding and may suggest that the unemployed in England are less likely to be involved in active job search either due to better welfare benefits in UK than the US, or because they are constrained geographically.

As well as the theoretical importance of this topic for urban economics, our findings also inform the policy debate on urban poverty in Britain. Overall we conclude that the standard urban land use model provides a partial explanation of how households sort by income in cities, but that the role of amenities and household heterogeneity is large and warrants more attention. Importantly for urban planning debates, there is a need to recognise the complexity of household location choices and it is important to consider not only the location of amenities, but also the characteristics of the local population.

References

Agarwal, A., Giuliano, G., Redfearn, C.L., 2012. Strangers in our Midst: the Usefulness of Exploring Polycentricity. The Annals of Regional Science 48(2), 433-450.

Ahlfeldt, G.M., 2011. Blessing or Curse? Appreciation, Amenities and Resistance to Urban Renewal. Regional Science and Urban Economics 41(1), 32-45.

Ahlfeldt, G.M., Holman, N., 2016. Distinctively Different: A New Approach to Valuing Architectural Amenities. The Economic Journal 128(608), 1-33.
Ahlfeldt, G.M., Maennig, W., 2009. Arenas, Arena Architecture and the Impact on Location Desirability: The Case of `Olympic Arenas' in Prenzlauer Berg, Berlin. Urban Studies 46(7), 1343-1362.
Ahlfeldt, G.M., Maennig, W., 2010a. Impact of sports arenas on land values: evidence from Berlin. The Annals of Regional Science 44(2), 205-227.
Ahlfeldt, G.M., Maennig, W., 2010b. Stadium Architecture and Urban Development from the Perspective of Urban Economics. International Journal of Urban and Regional Research 34(3), 629-646.
Ahlfeldt, G.M., Moller, K., Waights, S., Wendland, N., 2017. Game of Zones: The Political Economy of Conservation Areas. The Economic Journal 127(605), F421-F445.
Albouy, D., 2008. Are Big Cities Bad Places to Live? Estimating Quality of Life across Metropolitan Areas. NBER Working Paper No. 14472
Albouy, D., 2016. What Are Cities Worth? Land Rents, Local Productivity, and the Total Value of Amenities. The Review of Economics and Statistics 98(3), 477-487.
Albouy, D., Lue, B., 2015. Driving to opportunity: Local rents, wages, commuting, and sub-metropolitan quality of life. Journal of Urban Economics 89, 74-92.
Alonso, W., 1964. Location and Land Use: Toward a General Theory of Land Rent. Cambridge, Massachusetts: Harvard University Press.
Anas, A., 1990. Taste Heterogeneity and Urban Spatial Structure: the Logit Model and Monocentric Theory Reconciled. Journal of Urban Economics 28(3), 318-335.
Asabere, P.K., Huffman, F.E., 1991. Historic Districts and Land Values. The Journal of Real Estate Research 6(1), 1-7.
Atack, J., Margo, R.A., 1998. Location, Location, Location! The Price Gradient for Vacant Urban Land: New York, 1835 to 1900. Journal of Real Estate Finance and Economics 16(2), 151-172.
Baum-Snow, N., Marion, J., 2009. The Effects of Low Income Housing Tax Credit Developments on Neighborhoods. Journal of Public Economics 93(5-6), 654–666.
Bayer, P., McMillan, R., Rueben, K.S., 2004. What Drives Racial Segregation? New Evidence using Census Microdata. Journal of Urban Economics 56(3), 514-535.
Bowden, M.J., 1971. Downtown Through Time: Delimitation, Expansion, and Internal Growth. Economic Geography 47(2), 121-135.
Brown, S., 1987. The Complex Model Centre of City Retailing: an Historical Application. Transactions of the Institute of British Geographers 12(1), 4-18.
Brueckner, J.K., Rosenthal, S.S., 2009. Gentrification and Neighborhood Housing Cycles: Will America's Future Downtowns Be Rich? Review of Economics and Statistics 91(4), 725-743.
Brueckner, J.K., Thisse, J.-F., Zenou, Y., 1999. Why Is Central Paris Rich and Downtown Detroit Poor? An Amenity-Based Theory. European Economic Review 43(1), 91-107.
Cameron, S., 2003. Gentrification, Housing Redifferentiation and Urban Regeneration: 'Going for Growth' in Newcastle upon Tyne. Urban Studies 40(12), 2367-2382.
Carlino, G.A., Saiz, A., 2008. City Beautiful. IZA discussion papers 3778.
Cass, N., Shove, E., Urry, J., 2005. Social Exclusion, Mobility and Access. The Sociological Review 53(3), 539-555.
Cheshire, P., Hilber, C., Montebruno, P., Sanchis-Guarner, R., 2018. Take Me to the Centre of Your Town! Using Micro-geographical Data to Identify Town Centres. CESifo Economic Studies 64(2), 255-291.
Cheshire, P., Magrini, S., 2006. Population Growth in European Cities: Weather Matters - but only Nationally. Regional Studies 40(1), 23-37.
Couture, V., 2015. Knowledge spillovers in cities: An auction approach. Journal of Economic Theory 157, 668-698.

DCLG, 2011. Updating the Evidence Base on English Cities: Final Report. London: Department for Communities and Local Government.

Duranton, G., Puga, D., 2014. The Growth of Cities. In: Aghion, P., Durlauf, S. (Eds.). Handbook of Economic Growth, vol. 2. Elsevier North-Holland, Amsterdam, pp. 781-853.

Epple, D., Platt, G.J., 1998. Equilibrium and Local Redistribution in an Urban Economy when Households Differ in both Preferences and Incomes. Journal of Urban Economics 43(1), 23-51.

Fujita, M., Ogawa, H., 1982. Multiple Equilibria and Structural Transition of Non-Monocentric Urban Configurations. Regional Science and Urban Economics 12(2), 161-196.

Gabriel, S.A., Rosenthal, S.S., 2004. Quality of the Business Environment versus Quality of Life: Do Firms and Households Like the Same Cities? The Review of Economics and Statistics 86(1), 438-444.

Giuliano, G., Small, K.A., 1991. Subcenters in the Los Angeles Region. Regional Science and Urban Economics 21(2), 163-182.

Glaeser, E.L., Kahn, M.E., Rappaport, J., 2008. Why Do the Poor Live in Cities? The Role of Public Transportation. Journal of Urban Economics 63(1), 1-24.

Glaeser, E.L., Kolko, J., Saiz, A., 2001. Consumer City. Journal of Economic Geography 1(1), 27-50.

Gyourko, J., Kahn, M.E., Tracy, J., 1999. Quality of Life and Environmental Comparisons. In: Cheshire, P., Mills, E.S. (Eds.). Handbook of Regional and Urban Economics, vol. 3. Elsevier North-Holland, Amsterdam, pp. 1413-1454.

Handbury, J., 2013. Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across U.S. Cities. Wharton School Research Paper No. 71

Hsiang, S.M., 2010. Temperatures and Cyclones Strongly Associated with Economic Production in the Caribbean and Central America. Proceedings of the National Academy of Science of the USA 107(35), 15367-15372.

Kantor, Y., Rietveld, P., Van Ommeren, J.N., 2014. Towards a General Theory of Mixed Zones: The Role of Congestion. Journal of Urban Economics 83, 50-58.

Kelly, B., 2013. The Process of Socio-Economic Constraint on Geographical Mobility: England 1991 to 2008. University of Manchester CCSR Working Paper No. 2013-08.

Koster, H.R.A., Rouwendal, J., 2017. Historic Amenities and Housing Externalities: Evidence from the Netherlands. The Economic Journal 127(605), F396-F420.

Koster, H.R.A., Van Ommeren, J.N., Rietveld, P., 2016. Historic amenities, income and sorting of households. Journal of Economic Geography 16(1), 203-236.

Lee, S., Lin, J., 2018. Natural Amenities, Neighbourhood Dynamics, and Persistence in the Spatial Distribution of Income. Review of Economic Studies 85(1), 663-694.

LeRoy, S., Sonstelie, J., 1983. Paradise Lost and Regained: Transportation Innovation, Income, and Residential Location. Journal of Urban Economics 13(1), 67-89.

Madariaga, R., Martori, J.C., Oller, R., 2014. Income, distance and amenities. An empirical analysis. Empirical Economics 47(3), 1129-1146.

Mills, E.S., 1967. An aggregative model of resource allocation in a metropolitan area. American Economic Review 57(2), 197-210.

Muth, R.F., 1969. Cities and Housing. The Spatial Pattern of Urban Residential Land Use. Chicago and London: The University of Chicago Press.

Nathan, M., Urwin, C., Champion, A., Morris, J., 2005. City People: City Centre Living in the UK. London: Institute for Public Policy Research.

Ng, C.F., 2008. Commuting Distances in a Housing Location Model with Amenities. Journal of Urban Economics 63(1), 116-129.

ODPM, 2004. Producing Boundaries and Statistics for Town Centres, England and Wales 2000 (Interim Report). London: Office of the Deputy Prime Minister.

ODPM, 2005. Technical Report: Using Town Centre Statistics to Indicate the Broad Location of Retail Development - Initial Analysis. London: Office of the Deputy Prime Minister.

ODPM, 2006. State of the English Cities Volume 1. London: Office of the Deputy Prime Minister.

Pathak, R., Wyczalkowski, C.K., Huang, X., 2017. Public Transit Access and the Changing Spatial Distribution of Poverty. Regional Science and Urban Economics 66, 198-212.

Picard, R., 2010. GEODIST: Stata Module to Compute Geodetic Distances. Boston College Department of Economics Statistical Software Components S457147, revised 22 Feb 2012.

Rae, A., 2013. English Urban Policy and the Return to the City: a Decade of Growth, 2001-2011. Cities 32, 94-101.

Rappaport, J., 2007. Moving to Nice Weather. Regional Science and Urban Economics 37(3), 375-398.

Rappaport, J., 2008. Consumption Amenities and City Population Density. Regional Science and Urban Economics 38(6), 533-552.

Rappaport, J., Sachs, J., 2003. The United States as a Coastal Nation. Journal of Economic Growth 8(1), 5-46.

Redding, S.J., Rossi-Hansberg, E., 2017. Quantitative Spatial Economics. Annual Review of Economics 9, 21-58.

Roback, J., 1982. Wages, Rents, and the Quality of Life. The Journal of Political Economy 90(6), 1257-1278.

Rosen, S., 1974. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. Journal of Political Economy 82(1), 34-55.

Rosenthal, S.S., Ross, S.L., 2015. Change and Persistence in the Economic Status of Neighborhoods and Cities. In: Duranton, G., Henderson, V., Strange, W. (Eds.). Handbook of Regional and Urban Economics, vol. 5B. Elsevier North-Holland, pp. 1047-1120.

Schiller, R., 2001. The Dynamics of Property Location. London: Verso.

Schmidheiny, K., 2006. Income segregation and local progressive taxation: Empirical evidence from Switzerland. Journal of Public Economics 90(3), 429-458.

Schuetz, J., Larrimore, J., Merry, E.A., Robles, B.J., Tranfaglia, A., Gonzalez, A., 2018. Are central cities poor and non-white? Journal of Housing Economics 40, 83-94.

Tallon, A., Bromley, R., 2004. Exploring the Attractions of City Centre Living: Evidence and Policy Implications in British Cities. Geoforum 35(6), 771-787.

University of Essex, 2016. Understanding Society: Waves 1-7, 2009-2016 8th Edition. Institute for Social and Economic Research, NatCen Social Research, Kantar Public. UK Data Service. SN: 6614.

Van Duijn, M., Rouwendal, J., 2013. Cultural Heritage and the Location Choice of Dutch Households in a Residential Sorting Model. Journal of Economic Geography 13(3), 473-500.

Waldfogel, J., 2008. The Median Voter and the Median Consumer: Local Private Goods and Population Composition. Journal of Urban Economics 63(2), 567-582.

Wheaton, W.C., 1977. Income and Urban Residence: An Analysis of Consumer Demand for Location. American Economic Review 67(4), 620-631.

Table 1: Number of households and mean income by city and wave.

City	2011 Population[1]	No. of LSOAs	Wave 3 2011/12 N	Income[2]	Wave 4 2012/13 N	Income	Wave 5 2013/14 N	Income	Wave 6 2014/15 N	Income	Wave 7 2015/16 N	Income
Birmingham	1073045	524	478	1432	455	1510	425	1506	381	1560	473	1481
Bristol	428234	227	184	1677	181	1707	161	1767	160	1829	175	1784
Leeds	751485	212	196	1638	172	1630	153	1614	139	1568	159	1512
Liverpool	466415	180	187	1588	172	1529	149	1468	144	1545	139	1563
Manchester	503127	316	303	1578	269	1658	247	1528	223	1647	288	1526
Newcastle	280177	286	291	1605	273	1571	251	1533	246	1592	231	1614
Nottingham	305680	319	326	1463	301	1549	280	1538	245	1651	234	1566
Sheffield	552698	336	356	1543	296	1543	289	1516	264	1686	253	1634
Total (mean)	**4360861**	**2400**	**2321**	**(1543)**	**2119**	**(1575)**	**1955**	**(1546)**	**1802**	**(1630)**	**1952**	**(1568)**

Note: [1] Population figures from the 2011 Census. [2] Income is measured as monthly equivalised net income in 2012/13 prices.

Table 2: Household (HH) and head of household (HoH) characteristics at different distances from the CBD.

distance percentile from CBD	Number of households	Mean Distance from CBD (km)	Mean monthly HH income (£)[1]	Average age of HoH (years)	HoH is male (%)	HoH is a migrant (%)	HoH not employed (%)	HoH has Higher Education (%)	HH owns home (%)	rents social housing (%)	rents private sector housing (%)	HH has a car (%)	Single adult HH (%)
<= 20	2,047	2.5	1,378	42	53	30	35	36	47	18	35	63	34
20-40	2,014	4.6	1,558	46	59	21	32	41	68	11	20	79	26
40-60	2,028	6.7	1,602	45	58	13	28	38	65	12	18	78	25
60-80	2,035	9.7	1,640	47	61	5	29	31	70	11	19	81	25
>80	2,025	18.6	1,675	47	69	4	24	35	74	10	16	88	21
All	10,149	8.5	1,571	45	60	15	30	36	65	13	22	79	26

Note: [1] Income is measured as monthly equivalised net income in 2012/13 prices.

<u>Table 3</u>: Household distance from the CBD, all cities pooled.

Dependent variable is log distance from CBD.

	(1)	(2)	(3)	(4)
Household characteristics				
income	0.169***	0.055**	0.031*	0.010
	(0.027)	(0.025)	(0.018)	(0.017)
sex		0.069**		0.049**
		(0.027)		(0.020)
age		0.006***		0.004***
		(0.001)		(0.001)
higher education		-0.089***		-0.104***
		(0.029)		(0.027)
owner occupier		0.122***		0.001
		(0.031)		(0.024)
single person		-0.092***		-0.050**
		(0.026)		(0.022)
car owner		0.194***		0.082***
		(0.033)		(0.022)
migrant		-0.505***		-0.282***
		(0.037)		(0.030)
Amenities				
public transport			-0.002*	-0.002**
			(0.001)	(0.001)
public services			-0.006***	-0.005***
			(0.002)	(0.002)
historic/culture			0.004	0.004
			(0.007)	(0.007)
retail			0.000	0.000
			(0.001)	(0.001)
eating out			-0.003**	-0.003**
			(0.001)	(0.001)
sports			-0.018***	-0.018***
			(0.004)	(0.004)
outdoor recreation			-0.028***	-0.028***
			(0.007)	(0.007)
old property			-0.329***	-0.272***
			(0.078)	(0.076)
property crime			-0.038***	-0.035***
			(0.008)	(0.007)
social housing			-0.020***	-0.018***
			(0.002)	(0.002)
Council Tax			0.016**	0.017**
			(0.008)	(0.007)
adj. R-sq	0.042	0.145	0.394	0.422

Models include city and wave dummies and a constant. n = 10,149. *,**,*** denote significance at 10, 5 and 1%. Conley standard errors adjusted for spatial and serial correlation. Estimation via Stata v15.1 *ols_spatial_HAC*

Table 4: Household distance from the CBD, by city.

Dependent variable is log distance from CBD.

	Birm'ham	Bristol	Leeds	Liv'pool	Man'ter	Newcastle	Nott'ham	Sheffield
Income	0.078***	-0.018	0.064*	0.009	0.033	-0.085**	-0.020	-0.054
	(0.019)	(0.044)	(0.035)	(0.032)	(0.024)	(0.040)	(0.038)	(0.036)
Sex	-0.028	0.050	-0.002	-0.046	0.037	0.075	0.233***	0.060
	(0.030)	(0.059)	(0.041)	(0.037)	(0.025)	(0.052)	(0.056)	(0.048)
Age	0.003***	0.005*	0.001	0.003*	0.004***	0.009***	0.003	0.001
	(0.001)	(0.003)	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)	(0.002)
higher educ	-0.019	-0.033	-0.117***	-0.089*	-0.027	-0.072	-0.063	-0.371***
	(0.032)	(0.056)	(0.044)	(0.047)	(0.034)	(0.047)	(0.078)	(0.058)
owner occr	-0.047**	0.010	0.011	-0.028	0.059	-0.040	-0.005	0.162***
	(0.022)	(0.062)	(0.038)	(0.051)	(0.036)	(0.044)	(0.076)	(0.054)
single person	-0.035	0.078	-0.033	-0.075**	-0.093***	-0.151***	-0.100	-0.018
	(0.031)	(0.052)	(0.047)	(0.037)	(0.025)	(0.056)	(0.064)	(0.066)
car owner	0.093**	-0.077	0.090**	-0.040	0.048	-0.027	0.148**	0.035
	(0.041)	(0.063)	(0.045)	(0.043)	(0.035)	(0.046)	(0.066)	(0.067)
migrant	-0.287***	0.106*	-0.154***	-0.189*	-0.209***	-0.347***	-0.313***	-0.317**
	(0.029)	(0.060)	(0.059)	(0.103)	(0.032)	(0.104)	(0.098)	(0.138)
public trans	-0.000	0.007*	0.001	-0.003**	0.000	-0.004	-0.011***	-0.006***
	(0.001)	(0.004)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
public serv	-0.013***	0.005*	-0.004	-0.012***	-0.003	-0.005*	0.004	-0.020***
	(0.002)	(0.003)	(0.004)	(0.004)	(0.002)	(0.003)	(0.005)	(0.004)
hist/cult	0.008	-0.003	0.007	-0.061***	-0.008	-0.013	-0.016	0.046***
	(0.012)	(0.007)	(0.011)	(0.013)	(0.006)	(0.008)	(0.032)	(0.010)
Retail	0.000	-0.000	0.004***	0.005***	-0.001***	0.006***	0.005***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.002)	(0.001)	(0.002)
eat out	-0.007***	-0.006**	-0.012***	0.002	-0.001	0.007***	-0.005	-0.011***
	(0.002)	(0.003)	(0.002)	(0.001)	(0.001)	(0.002)	(0.004)	(0.003)
sports	0.008**	-0.046***	0.004	-0.008*	-0.020***	-0.044***	-0.038***	0.011
	(0.004)	(0.008)	(0.005)	(0.005)	(0.005)	(0.006)	(0.009)	(0.007)
outdoor rec	0.021*	-0.066***	-0.101***	0.041***	-0.047***	-0.044***	-0.057***	0.024*
	(0.012)	(0.015)	(0.018)	(0.010)	(0.008)	(0.010)	(0.018)	(0.014)
old prop	-0.294***	0.645	-0.078	-0.959***	-0.026	-0.612***	-0.488**	0.170
	(0.076)	(0.565)	(0.129)	(0.138)	(0.086)	(0.169)	(0.237)	(0.181)
prop crime	-0.016***	-0.137***	-0.062***	-0.086***	-0.019**	-0.061***	-0.046***	-0.022
	(0.005)	(0.019)	(0.007)	(0.020)	(0.008)	(0.021)	(0.016)	(0.014)
soc housing	-0.021***	-0.017***	-0.039***	-0.015***	-0.013***	-0.015***	-0.010**	-0.012***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.001)	(0.003)	(0.004)	(0.003)
Council Tax	-0.008	-0.010	0.018*	0.001	-0.008	0.057***	-0.003	0.060***
	(0.007)	(0.016)	(0.011)	(0.010)	(0.007)	(0.012)	(0.023)	(0.019)
N	2212	861	819	791	1330	1292	1386	1458
adj. R-sq	0.637	0.694	0.682	0.628	0.567	0.433	0.438	0.383

Models include wave dummies and a constant. *,**,*** denote significance at 10, 5 and 1%. Conley standard errors adjusted for spatial and serial correlation. Estimation via Stata v15.1 *ols_spatial_HAC*

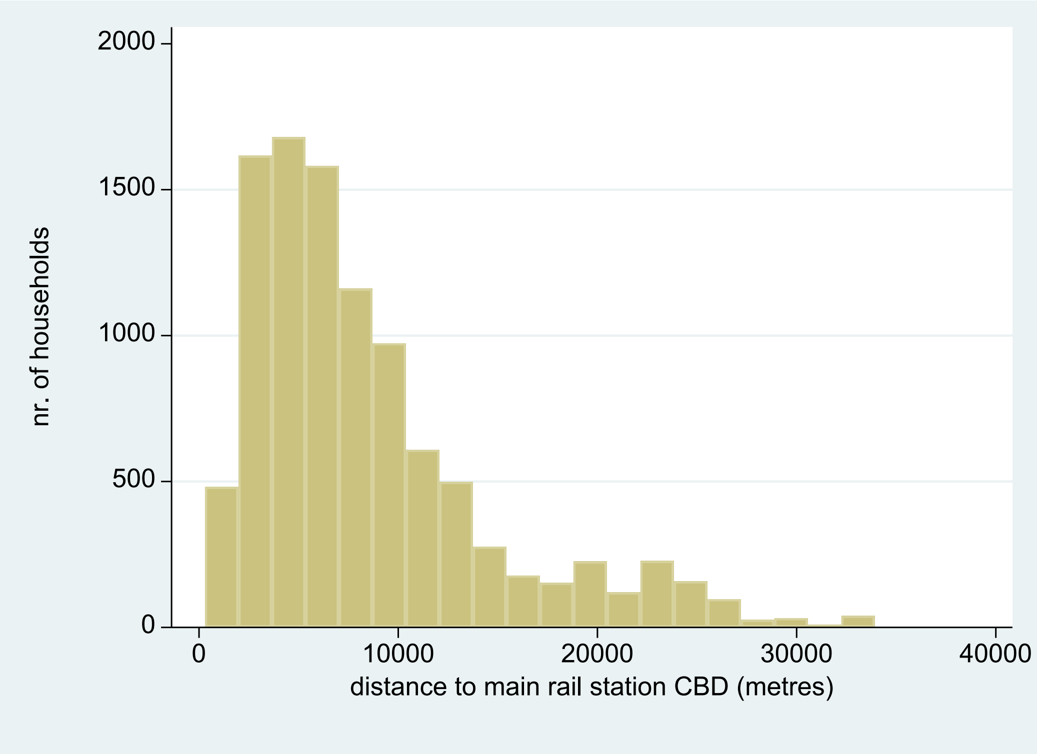
<u>Table 5</u>: Robustness checks: sub-group analysis

Dependent variable is log distance from CBD.

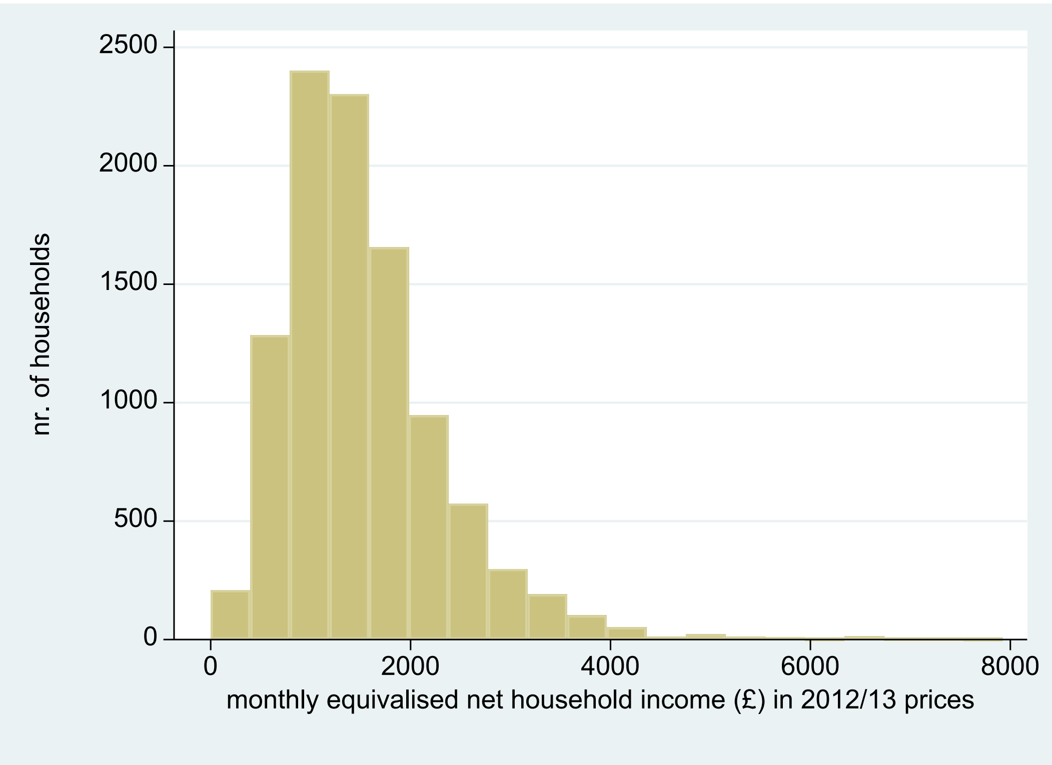
	(1) movers only	(2a) employed	(2b) not emp.	(3a) home owners	(3b) renters	(4a) children in household	(4b) no children	(5a) age under 40	(5b) age 40 plus
income	0.081**	0.019	-0.023	0.011	0.012	0.022	0.014	-0.004	0.038
	(0.039)	(0.027)	(0.019)	(0.022)	(0.021)	(0.028)	(0.016)	(0.020)	(0.026)
sex	-0.054	0.032	0.073**	0.027	0.110***	0.037	0.079***	0.038*	0.068**
	(0.053)	(0.023)	(0.034)	(0.022)	(0.033)	(0.027)	(0.026)	(0.023)	(0.032)
age	0.011***	0.004***	0.005***	0.003***	0.005***	0.001	0.006***		
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)		
high ed	-0.289***	-0.105***	-0.140***	-0.115***	-0.078**	-0.094***	-0.112***	-0.097***	-0.135***
	(0.062)	(0.026)	(0.045)	(0.030)	(0.035)	(0.034)	(0.029)	(0.029)	(0.035)
owner	0.146**	0.005	-0.000			-0.055*	0.059*	-0.008	0.022
	(0.058)	(0.026)	(0.043)			(0.029)	(0.030)	(0.028)	(0.037)
single	-0.116*	-0.025	-0.110***	0.013	-0.144***			-0.022	-0.103***
	(0.060)	(0.025)	(0.032)	(0.029)	(0.030)			(0.024)	(0.037)
car	0.038	0.132***	0.006	0.135***	0.019	0.102***	0.075***	0.064**	0.107***
	(0.054)	(0.029)	(0.032)	(0.034)	(0.028)	(0.039)	(0.027)	(0.026)	(0.036)
Migrant	-0.150**	-0.252***	-0.347***	-0.260***	-0.301***	-0.207***	-0.336***	-0.300***	-0.226***
	(0.075)	(0.032)	(0.043)	(0.032)	(0.041)	(0.031)	(0.041)	(0.036)	(0.045)
pub tran	0.002	-0.002**	-0.001	-0.002**	-0.003**	-0.003**	-0.002**	-0.003***	-0.001
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
pub serv	-0.003	-0.005**	-0.006***	-0.008***	-0.002	-0.008***	-0.002	-0.006***	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
hist/cult	0.006	0.010	-0.005	0.017**	-0.006	0.011	0.001	0.001	0.009
	(0.006)	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.009)
retail	0.002	-0.000	0.002**	0.000	0.001	0.000	0.001	0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
eat out	-0.007***	-0.003**	-0.004**	-0.002	-0.004***	-0.002	-0.004***	-0.000	-0.006***
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
sports	-0.011**	-0.020***	-0.016***	-0.021***	-0.011**	-0.009*	-0.025***	-0.021***	-0.011**
	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)
out rec	-0.025**	-0.028***	-0.031***	-0.021***	-0.034***	-0.040***	-0.020***	-0.028***	-0.033***
	(0.012)	(0.007)	(0.009)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.009)
old prop	-0.562***	-0.106	-0.542***	-0.102	-0.612***	-0.304***	-0.225**	-0.127	-0.541***
	(0.136)	(0.083)	(0.097)	(0.089)	(0.097)	(0.096)	(0.089)	(0.089)	(0.100)
crime	-0.017	-0.035***	-0.037***	-0.041***	-0.036***	-0.035***	-0.032***	-0.042***	-0.027***
	(0.010)	(0.007)	(0.013)	(0.010)	(0.007)	(0.012)	(0.008)	(0.009)	(0.009)
soc hous	-0.022***	-0.018***	-0.019***	-0.017***	-0.022***	-0.016***	-0.020***	-0.018***	-0.018***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
coun tax	-0.006	0.015**	0.021*	0.017**	0.020**	0.015*	0.019**	0.026***	-0.001
	(0.013)	(0.007)	(0.011)	(0.008)	(0.010)	(0.008)	(0.009)	(0.008)	(0.010)
N	1576	7149	3000	6572	3577	3931	6218	6955	3194
adj. R-sq	0.437	0.413	0.457	0.353	0.505	0451	0.424	0.417	0.435

Models include city and wave dummies and a constant. *,**,*** denote significance at 10, 5 and 1%. Conley standard errors adjusted for spatial and serial correlation. Estimation via Stata v15.1 *ols_spatial_HAC*

<u>Figure 1</u>: Households' linear distance to the CBD (main rail station).

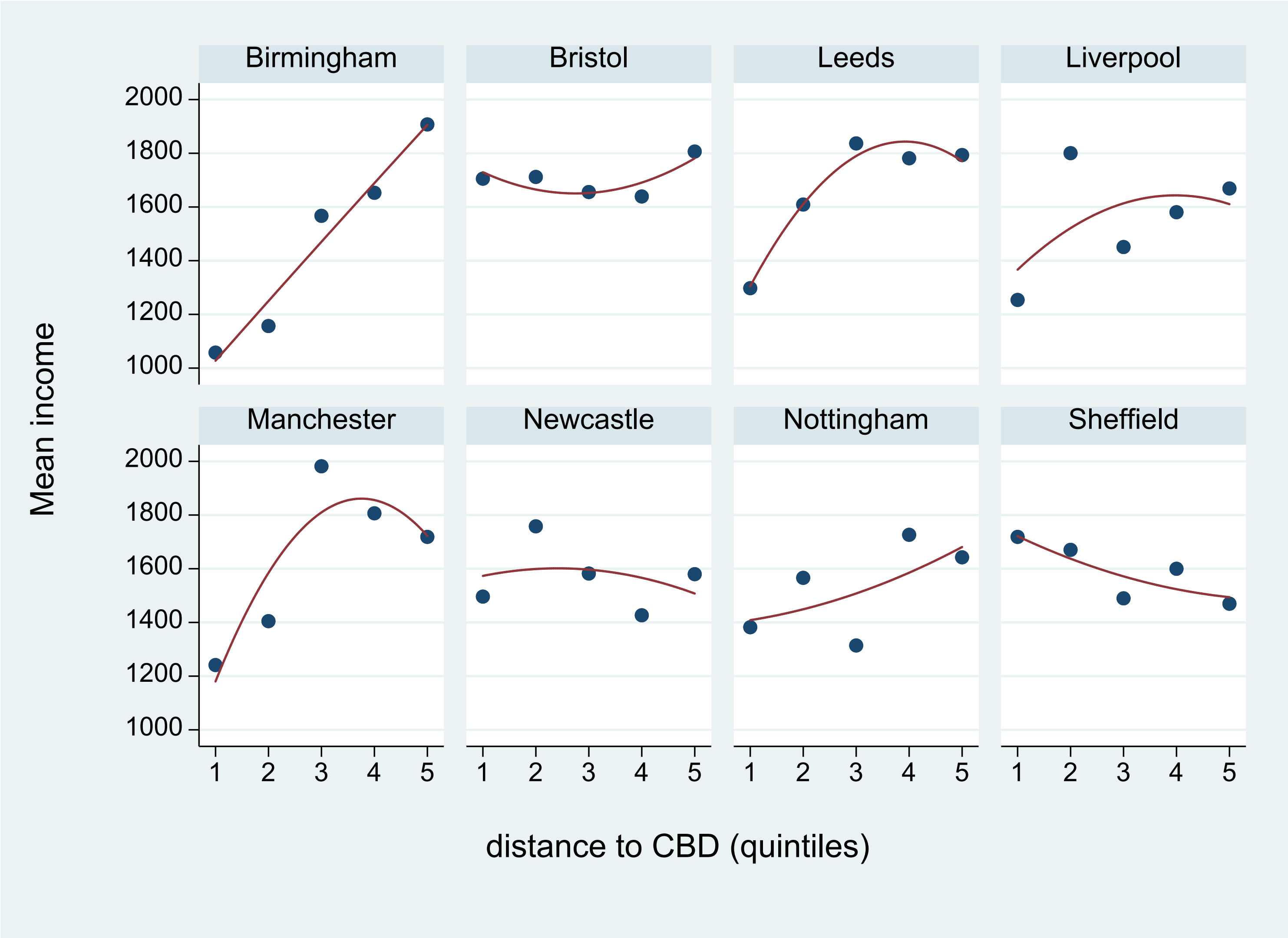


<u>Figure 2</u>: The distribution of household income



Note: The graph is censored at incomes > £8000.

<u>Figure 3</u>: Spatial distribution of income, by city, with fitted quadratic line.



Note: Income is monthly equivalised net income (£) in 2012/13 prices.

Figure 4: Spatial distribution of area controls, with fitted quadratic line.

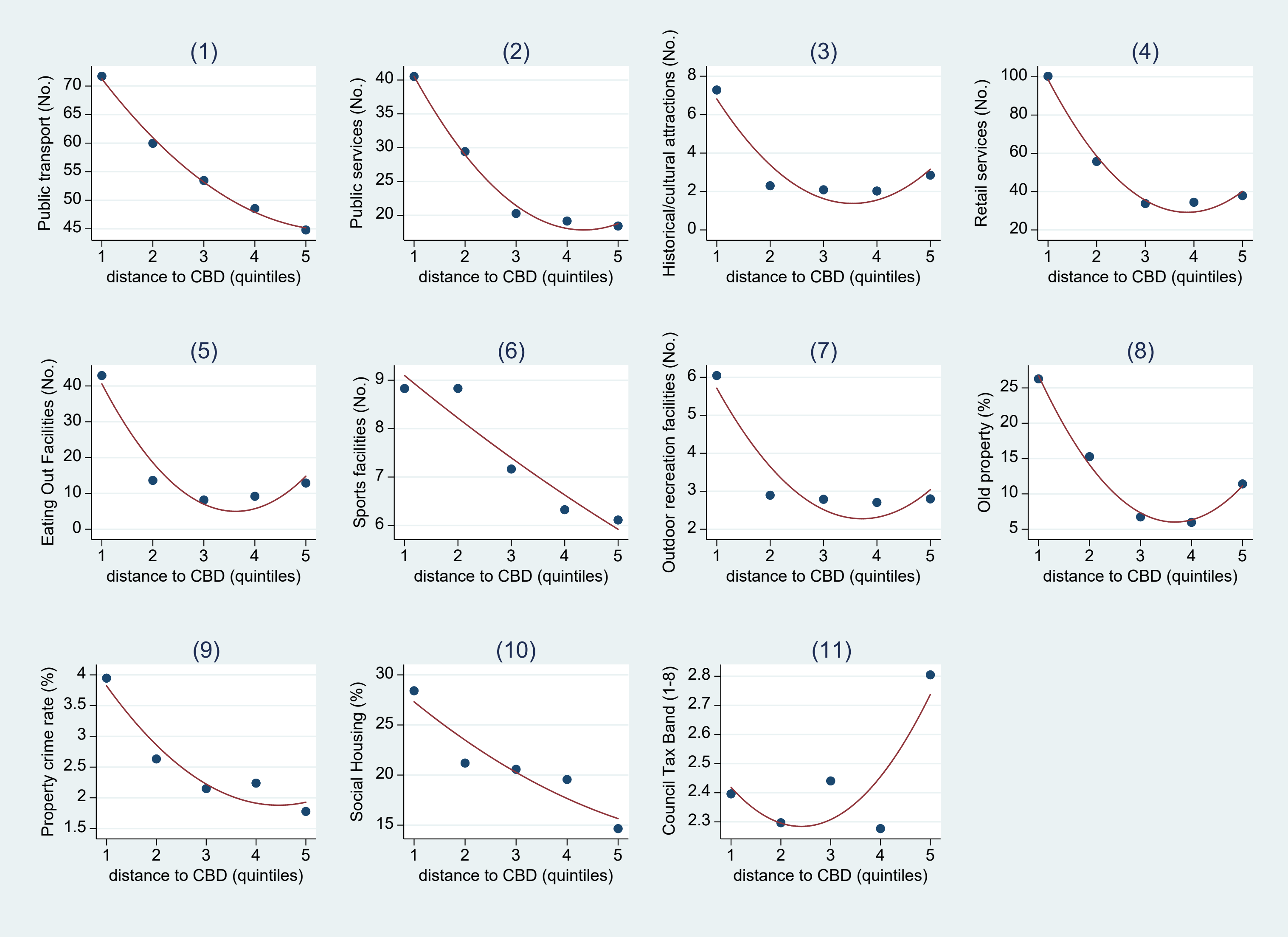
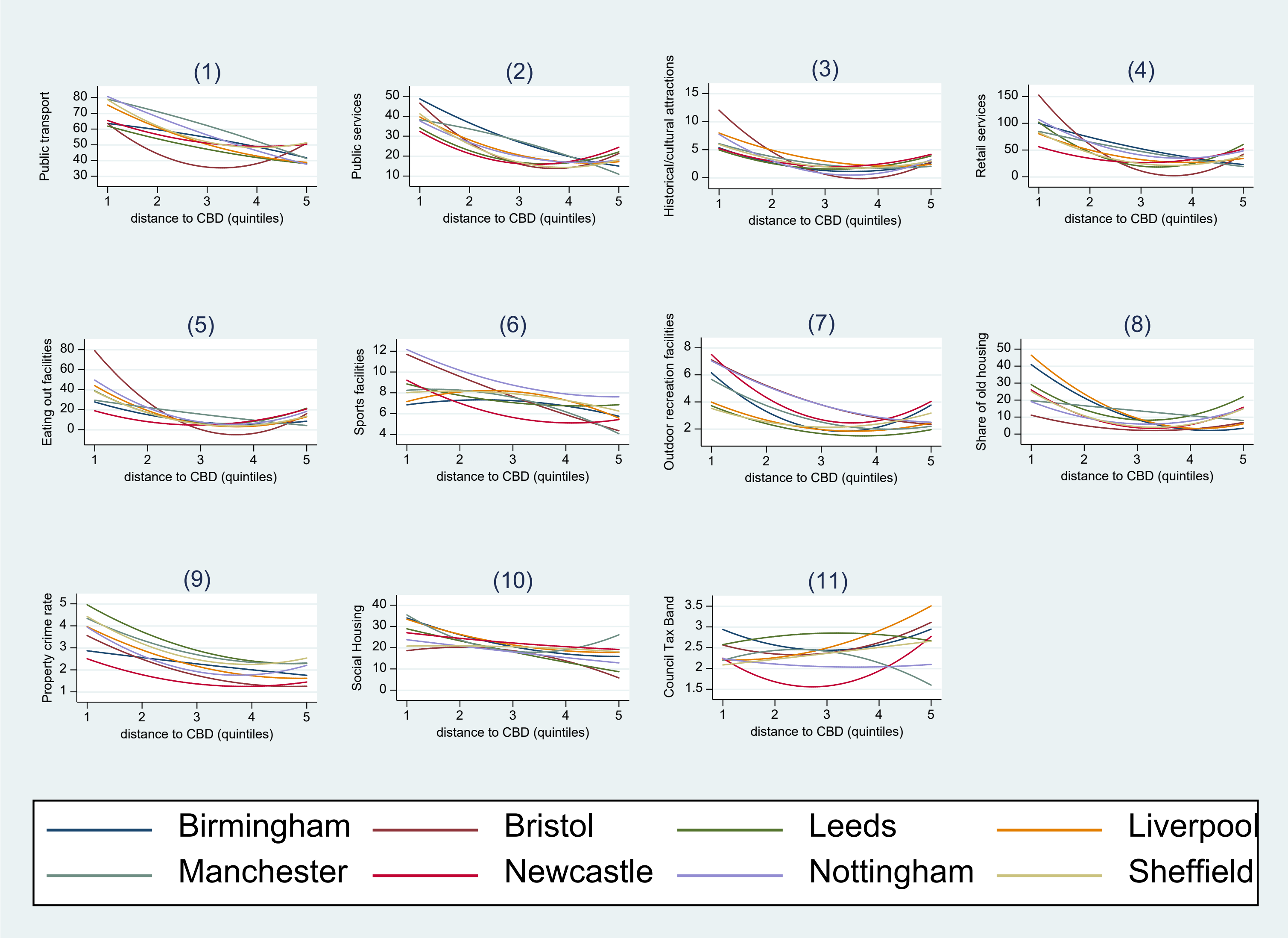


Figure 5: **Fitted quadratic lines for location of amenities by distance to the CBD, by city.**



Appendix

<u>Table A1</u>. Variable descriptions and data source

Variable	Variable description
Dependent Variables	
Linear distance	Linear distance from household to: (a) main rail station; (b) main M&S; (c) Areas of Town Centre Activity (ATCAs) calculated via Stata module 'geodist' (Picard, 2010) . Household location estimated using the population weighted LSOA centroid http://geoportal.statistics.gov.uk/datasets/lower-layer-super-output-areas-december-2011-population-weighted-centroids. See App Table A2 for full list of distance measures.
Google Maps distance and time	Google Maps API was used to estimate the distance and time by car and bus from each household to the points listed in (a) to (c) above. See Appendix Table A2 for full list.
Explanatory Variables	
UKHLS	www.understandingsociety.ac.uk
household income	Monthly equivalised net income (£) from all sources in the month preceding the interview in 2012/13 prices.
sex	=1 if head of household is male; zero otherwise.
age	Age of the household head.
higher education	Education status =1 if household head has higher education; zero otherwise.
home owner	Home ownership status =1 if house is owned outright or with a mortgage; zero otherwise.
migrant	=1 if household head was not born in the UK; zero otherwise.
access to car	=1 if household owns a car; zero otherwise.
single	Single person household = 1 if household is single adult; zero other wise
not employed	Employment status =0 if head of household is employed or self-employed; one otherwise.
Council Tax band	The Council Tax is a tax on domestic property. Each property is assigned to a band (1 to 8) bands based on the value of the property on 1 April 1991.
Ordnance Survey Place of Interest	Available via the Digimap service provided by Edina, University of Edinburgh (http://digimap.edina.ac.uk)
public transport	Bus, rail and tram stops and stations within 1000m of household.
public services	School facilities (primary, secondary and tertiary education) and health facilities (practitioners and establishments), within 1000m of household.
historical & cultural	Historical (and cultural) attractions include archaeological sites, art galleries, historic buildings, museums, cinemas, theatres and concert halls within 1000m of household, .
retail services	Clothing and accessories, food, drink, home, leisure, garden and multi-item retail, within 1000m of household.
eating out	Cafes, snack bars, tea-rooms, pubs, bars, inns, and restaurants within 1000m of household.
sport	Gymnasiums, sports halls, leisure centres, swimming pools, sports grounds, stadia, and pitches within 1000m of household.
Outdoor recreation	Commons, parks and gardens, picnic areas, and playgrounds within 1000m of household,
Other Sources	
share of old housing	Ratio of old dwellings (built pre 1900s) to total number of dwellings in LSOAs within 1000m of household. www.gov.uk/government/statistics/council-tax-property-attributes.
property crime rate	Ratio of number of property crimes to 100 resident population in LSOAs within 1000m of household. Crimes include burglary, vehicle crimes and other theft). Source: www.ukcrimestats.com
share of social housing	Proportion of the population who live in social housing in LSOAs within 1000m of household. Source: Census 2011 via NOMIS www.nomisweb.co.uk

Table A2: Summary statistics for alternative distance and travel time measures

Variable	Model assumption	Mean	Std. Dev.	Min	Max
Linear Distance (m)					
rail station	monocentric	8465	6146	296	33925
M&S	monocentric	8445	6123	222	34209
nearest ATCA	polycentric	4668	5389	90	32570
mean to any ATCA	polycentric	9324	5576	2907	35209
Google maps travel distance (m)					
by car to rail station	monocentric	11977	8570	694	48026
by car to nearest ATCA	polycentric	6330	6858	136	36305
mean by car to any ATCA	polycentric	13286	7981	4212	44476
by bus to rail station	monocentric	10998	7893	366	44239
by bus to nearest ATCA	polycentric	6677	7209	79	42821
mean by bus to any ATCA	polycentric	13157	7576	3650	46359
Google maps travel time (mins)					
by car to rail station	monocentric	1263	524	224	3218
by car to nearest ATCA	polycentric	658	446	41	2200
mean by car to any ATCA	polycentric	1243	424	538	2858
by bus to rail station	monocentric	2253	868	288	5987
by bus to nearest ATCA	polycentric	1536	1007	53	5818
mean by bus to any ATCA	polycentric	3031	963	1158	6664

Notes: ATCA = Area of Town Centre Activity (ODPM 2004, 2005). Monocentric urban models assume one main city centre. Polycentric models allow for multiple centres. The terms are used here to denote the measures we use under these different assumptions.

Table A3: Correlation matrix for alternative distance and travel time measures

	Linear Distance				**Google maps distance**						**Google maps travel times**			
	to rail station	to M&S	nearest ATCA	mean to any ATCA	by car to rail station	by car to nearest ATCA	mean by car to any ATCA	by bus to rail station	by bus to nearest ATCA	mean by bus to any ATCA	by car to rail station	by car to nearest ATCA	mean by car to any ATCA	by b to ra stati
Linear Distance														
to rail station (mono)														
to M&S (mono)	0.97													
to nearest ATCA (poly)	0.93	0.92												
mean to any ATCA (poly)	0.97	0.97	0.94											
Google maps distance														
by car to rail station (mono)	0.96	0.95	0.88	0.93										
by car to nearest ATCA (poly)	0.91	0.91	0.99	0.93	0.88									
mean by car to any ATCA (poly)	0.95	0.95	0.91	0.98	0.93	0.91								
by bus to rail station (mono)	0.98	0.97	0.90	0.94	0.94	0.89	0.93							
by bus to nearest ATCA (poly)	0.91	0.91	0.97	0.93	0.88	0.98	0.91	0.90						
mean by bus to any ATCA (poly)	0.96	0.96	0.92	0.98	0.92	0.92	0.97	0.95	0.94					
Google maps travel times														
by car to rail station (mono)	0.91	0.90	0.78	0.87	0.88	0.77	0.87	0.90	0.78	0.87				
by car to nearest ATCA (poly)	0.86	0.86	0.94	0.88	0.84	0.95	0.88	0.85	0.95	0.89	0.77			
mean by car to any ATCA (poly)	0.91	0.91	0.86	0.95	0.87	0.85	0.95	0.90	0.86	0.94	0.90	0.86		
by bus to rail station (mono)	0.78	0.77	0.65	0.72	0.75	0.64	0.73	0.80	0.67	0.74	0.77	0.63	0.72	
by bus to nearest ATCA (poly)	0.80	0.81	0.87	0.82	0.79	0.88	0.82	0.82	0.91	0.85	0.70	0.91	0.80	
mean by bus to any ATCA (poly)	0.85	0.85	0.79	0.88	0.82	0.79	0.90	0.86	0.82	0.91	0.81	0.79	0.90	

Notes: ATCA = Area of Town Centre Activity (ODPM 2004, 2005). Monocentric (mono) urban models assume one main city centre. Polycentric (poly) models allow for multiple centres. The terms are used here to denote the measures we use under these different assumptions

Table A4: Robustness checks: alternate distance measures and ATCAs

	(1)	(2)	(3)	(4)	(5)	(6)
	To main rail station			To nearest ATCA		
	driving distance	journey time car	journey time bus	driving distance	journey time car	journey time bus
income	0.007	0.002	0.006	-0.008	-0.009	-0.002
	(0.017)	(0.010)	(0.009)	(0.022)	(0.015)	(0.015)
sex	0.048***	0.028**	0.031***	0.066**	0.051**	0.046**
	(0.018)	(0.012)	(0.011)	(0.029)	(0.020)	(0.020)
age	0.004***	0.003***	0.003***	0.006***	0.004***	0.004***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
higher education	-0.114***	-0.062***	-0.039***	-0.123***	-0.088***	-0.095***
	(0.026)	(0.015)	(0.014)	(0.033)	(0.022)	(0.021)
owner occupier	-0.001	-0.008	0.000	0.000	-0.009	0.011
	(0.021)	(0.015)	(0.014)	(0.030)	(0.020)	(0.020)
single person	-0.055***	-0.037***	-0.030**	-0.083**	-0.060**	-0.053**
	(0.021)	(0.013)	(0.013)	(0.035)	(0.025)	(0.022)
car owner	0.065***	0.035***	0.041***	0.076**	0.050**	0.049*
	(0.019)	(0.013)	(0.013)	(0.034)	(0.024)	(0.026)
migrant	-0.225***	-0.139***	-0.104***	-0.236***	-0.133***	-0.144***
	(0.028)	(0.017)	(0.018)	(0.039)	(0.026)	(0.027)
public transport	-0.002**	-0.001	-0.002***	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
public services	-0.005***	-0.003**	-0.001	-0.017***	-0.013***	-0.011***
	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
historic/culture	0.004	0.001	-0.003	0.032***	0.020***	0.021***
	(0.005)	(0.004)	(0.004)	(0.009)	(0.006)	(0.005)
retail	-0.000	0.000	-0.000	-0.002***	-0.001	-0.002***
	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
eating out	-0.001	-0.001	-0.001	0.002	0.002*	0.002
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
sports	-0.019***	-0.011***	-0.011***	-0.027***	-0.017***	-0.015***
	(0.004)	(0.002)	(0.002)	(0.006)	(0.004)	(0.003)
outdoor recreation	-0.022***	-0.016***	-0.009***	0.018**	0.017***	0.011**
	(0.006)	(0.004)	(0.003)	(0.009)	(0.005)	(0.005)
old property	-0.230***	-0.109**	-0.192***	-0.290***	-0.142**	-0.244***
	(0.067)	(0.044)	(0.041)	(0.101)	(0.072)	(0.070)
property crime	-0.027***	-0.017***	-0.024***	-0.034***	-0.021***	-0.028***
	(0.006)	(0.004)	(0.004)	(0.008)	(0.006)	(0.006)
social housing	-0.016***	-0.011***	-0.006***	-0.011***	-0.007***	-0.005***
	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Council Tax	0.019***	0.008**	0.012***	0.026***	0.015**	0.016**
	(0.007)	(0.004)	(0.004)	(0.010)	(0.006)	(0.007)
N	10149	10149	10149	10149	10149	10149
Adjusted R2	0.391	0.362	0.369	0.318	0.278	0.303

Models include city and wave dummies and a constant. *,**,*** denote significance at 10, 5 and 1%. Conley standard errors adjusted for spatial and serial correlation. Estimation via Stata v15.1 *ols_spatial_HAC*

Figure A1: Fitted quadratic line for location of individual and household characteristics by distance to the CBD, by city.