Amenities and the Social Structure of Cities*

Carl Gaigné[†] Hans R.A. Koster[‡] Fabien Moizeau[§]
Jacques-François Thisse[¶]

March 15, 2017

Abstract

We develop a new model of a "featureful" city in which locations are differentiated by *two attributes*, that is, the distance to employment centers and the accessibility to given amenities, and we show how heterogeneous households in income are sorted out across the urban space according to these attributes. Under Stone-Geary preferences, the spatial income distribution is governed by a location-quality index which reflects the interaction between the amenity and commuting cost functions. The residential equilibrium typically involves the spatial separation of households sharing similar incomes. Using data on Dutch cities, we show that there is a causal relationship between the amenity level and consumer income, suggesting that richer households sort themselves into high amenity locations. We do not find strong evidence that employment accessibility leads to income segregation, suggesting that the standard monocentric city model without amenities is a poor predictor of the social structure of cities.

Keywords: social stratification, income, amenities, commuting
JEL classification: R14, R23, R53, Z13.

*We are grateful to M.S. Delgado, R.J.G.M. Florax, M. Fujita, J.V. Henderson, H.G. Overman, S.L. Ross, K. Schmidheiny, H. Takatsuka, P. Ushchev and J.N. van Ommeren for insightful suggestions. We also thank participants to seminars at the LSE, Purdue, the 64th AFSE congress and the 7th North American Meeting of Urban Economics Association for their comments. The authors acknowledge, respectively, financial supports from the Agence Nationale de la Recherche (ANR-12-INEG-0002), the Netherlands Organisation for Scientific Research (VENI research grant), and the the Russian Academic Excellence Project '5-100'.

[†]INRA, UMR1302, SMART Rennes (France) and CREATE, Laval University, Quebec (Canada).
[‡]Department of Spatial Economics, Vrije Universiteit Amsterdam (The Netherlands).
[§]CREM (Condorcet Center), Université de Rennes 1 and Institut Universitaire de France.
[¶]CORE-UCL (Belgium), NRU Higher School of Economics (Russia) and CEPR.

1 Introduction

Rising economic inequalities make the determinants of residential segregation between the affluent and the poor increasingly important. For example, Bischoff and Reardon (2013) observed that, over the last 40 years, the rise in income inequality was accompanied by an increase in the residential sorting by income in most of the 117 largest US metropolitan areas. In France, the decile ratio D9/D1 in the 230 largest metropolitan areas increased from 6.6 to 7.1 between 2007 and 2011 (Floch, 2014). What may explain this increasing income disparities in cities? One key explanation may be the rising importance of amenities (Glaeser *et al.*, 2001). This paper ascertains how *the trade-off between amenities and the proximity to jobs may explain the social structure of cities when households are differentiated by income.*

The canonical model of urban economics in which locations are differentiated only by their distance to the central business district (CBD) predicts that, as the distance to the CBD rises, households are sorted by increasing income order because the desire to consume more space leads the affluent to reside further away from the city center (Hartwick *et al.,* 1976; Fujita, 1989; Kamecke, 1993). If the city center has a large amenity advantage while amenities fall rapidly with distance, then the pattern may be reversed: the affluent live in the city center and the poor live in the suburbs (Brueckner *et al.*, 1999). In both cases, the residential equilibrium implies the same relationship between locations and incomes: *the wider the income gap between two households, the greater the distance between their residential locations*, and vice versa. Using a "featureless" city as a baseline setup has a consequence which is overlooked in the literature: there is no need to appeal to externalities for complete spatial segregation to occur. This is not to say that such externalities do not exist, but a competitive land market may be sufficient to generate perfect income sorting across space. Since this result is obtained in the absence of externalities, sorting need not be inefficient. Furthermore, perfect spatial sorting contradicts the empirical evidence. Most metropolitan areas display pronounced U-shaped or W-shaped spatial income distributions (Glaeser *et al.,* 2008; Rosenthal and Ross, 2015; Section 2 below).

We propose a new approach in which cities are "featureful" in that locations are distinguished by *two vertically differentiated attributes*, that is, the distance to employment centers and the accessibility to given and dispersed amenities. While the demand for amenities has a tendency to rise with income, high-income commuters bear higher costs than low-income commuters. As these two attributes are generally not perfectly correlated, residential choices are the outcome of the interplay between amenities and commuting costs through households' housing demands. In this context, we show that, regardless of the nature of the functional form of amenity, commuting cost and income distributions, households' spatial sorting is *imperfect* provided that amenities are distributed *unevenly* across the city. More specifically, a greater distance between two households no longer implies a wider income gap. Furthermore, there is never bunching: households with different incomes always choose locations with different characteristics. However, households sharing the same income need not live in two locations near each other. Instead, they may live in separated neighborhoods. In sum, there is imperfect spatial sorting *between* and *within* income classes.

The determination of a residential equilibrium with heterogeneous households has the nature of a matching problem between landlords and households in which land at specific locations is differentiated by two characteristics - the distance to employment centers and the accessibility to amenities - and households by one characteristics - their incomes. By using the bid rent function of urban economics, we show within a very general framework that the interaction between amenities and commuting gives rise to turning points in the spatial income distribution. This is to be contrasted to existing studies, which often depend on specific functional forms or fine details about some key parameter values (Duranton and Puga, 2015). Nevertheless, to investigate empirically how the trade-off between amenities and commuting shapes the social structure of cities, it is convenient to reduce our two-to-one matching to a one-to-one matching. We show that specific preferences add sufficient structure to the model for the matching to become one-to-one and for the equilibrium conditions to yield a *location-quality index* that pins down the spatial income distribution: households ranked by increasing income are mapped onto locations with an increasing location-quality index. As a consequence, the interaction between amenities and commuting generates a wealth of possible residential patterns which are not explored in urban economics. Fortunately, the properties of the location-quality index may be used to test the predictions about how amenities affect the residential pattern.

The upshot is that the city's bliss point is not the CBD or the city limit anymore. The global maximizer of the location-quality index is now the most-preferred location for all households, implying that this location is occupied by the affluent. As one moves away from this location, households are sorted by decreasing incomes until a minimizer of the location-quality index is reached, where low-income households, but not necessarily the poorest, are located. Beyond this minimizer, household income starts rising. In other words, we have perfect sorting by increasing or decreasing incomes between two adjacent extrema of the location-quality index. Since the sign of the income gradient changes at any extremum of the location-quality index, households get more exposure to other income-groups when the number of turning points of the quality index rises. In other words, the city's two vertically differentiated attributes are encapsulated into a single attribute which is generally not vertically differentiated because two households having the same income may choose spatially separated residences.

To assess the empirical relevance of our approach, we test the key equation of the theoretical sections, that is, *the income assignment mapping* by using microdata on incomes, amenities, residential and job locations for the two largest cities in the Netherlands (Amsterdam and Rotterdam). More specifically, we test whether and how amenities and commuting costs determine the sorting pattern of heterogeneous households within the city by exploiting the structure of the theoretical model developed in the previous sections. We have chosen to study Dutch cities because these cities are well known for providing different types of high-quality amenities that were developed long ago. However, there might be reverse causality, meaning that the location of amenities and jobs is determined by the spatial income distribution. To mitigate endogeneity issues, we use ancillary detailed data on demographic, housing and site characteristics.

We digitize extensive urban planning maps from the 1930s and use the historic and planned land use patterns to instrument for amenities and commuting costs. Since the strategy of using instruments based on historic city plans raises several issues, we devote considerable attention to the validity of such an identification strategy. The results unequivocally suggest that amenities are important in determining the spatial income distribution. However, we find that differences in commuting costs hardly account for spatial income differences, suggesting that the classical monocentric model without amenities is a fairly poor predictor of the social structure of cities. Furthermore, by using information on land prices and lot sizes we are able to estimate preference parameters and to undertake counterfactual analyses. Under a multimodal amenity distribution, our calculations confirm that the resulting spatial income distribution also displays several peaks and valleys.

We focus on exogenous and persistent amenities, such as natural geographic features and historic buildings. Recent empirical evidence suggests that such amenities play a major role in the way households are sorted out by income (Brueckner *et al.*, 1999; Lee and Lin, 2015; Koster *et al.*, 2016; Koster and Rouwendal, 2017). Historic buildings and open spaces are numerous in many European cities, whereas the central-city infrastructure of American cities seldom offer appreciable and lasting aesthetic benefits. We recognize that the provision of local public goods is decentralized and endogenous (Epple and Nechyba, 2004). To a large extent, this is because the institutional context that prevails in the US implies that the quality of schools and other neighborhood characteristics are often determined by the average income in the neighborhood (Bayer *et al.*, 2007). This is to be contrasted with what we observe in many other countries where local public goods such as schools are provided by centralized bodies. As a consequence, school performances are often less uneven than in the US (Ritzen *et al.*, 1997). In addition, racial disparities are often less of a concern than in the US. For example, Dutch people have a preference for living among people of one's own ethnic group, but in a sufficiently diverse neighborhood (Bakens *et al.*, 2016). Thus, we believe that the impact of local public goods and race, though relevant, is less important for studying the social structure of non-US cities. Finally, households also care about the proximity to private facilities such as shops, restaurants and theaters. We address the endogenous nature of such amenities in our econometric analysis.

Related literature The emphasis in on-going research is on skill sorting across cities (Behrens *et al.*, 2014; Eeckhout et *al.*, 2014; Davis and Dingel, 2015; Diamond, 2016). By contrast, the issue of household sorting within cities is largely overlooked. Suggesting the complexity of the exercise, only a handful of papers in urban economics have studied the social stratification of cities with heterogeneous households. Beckmann (1969) was the first attempt to take into account a continuum of heterogeneous households. Unfortunately, the assignment approach used by Beckmann was flawed (Montesano, 1972). Hartwick *et al.* (1976) and Fujita (1989) proposed a rigorous analysis of the residential pattern for a finite number of income classes. When commuting costs are distance-dependent and income-independent, these authors show that income-classes are ranked by increasing income as the distance to the CBD rises. Kamecke (1993) extends this result to a continuum of heterogeneous households by showing that there is perfect sorting moving out from the CBD by increasing incomes. When the lot size is fixed, Miyake (2003) showed that the bid rent approach may be viewed as the limit of a two-sided matching market between landlords and households.

Behrens *et al.* (2013) revisit Beckmann (1969) in a setup where households differ along two dimensions, that is, income and per unit distance commuting cost. To use assignment techniques, Behrens *et al.* assume a fixed lot size. However, they do not study the residential equilibria with imperfect sorting. We differ from them in that the differentiation of location attributes allows us to determine from the equilibrium conditions the rule that assigns incomes to locations without guessing a priori what this rule is, as in Sattinger (1993). Other papers related to ours include Määttänen and Terviö (2014) and Landvoigt *et al.* (2015), who study the relationship between the income and house price distributions as a matching problem in which houses are indivisible and heterogeneous by quality. Our paper is closer to Lee and Lin (2015) who show that richer households are anchored in neighborhoods with better natural amenities, whereas we stress the importance of aesthetic amenities. Very much like us, their results support the importance of natural amenities for the persistence of the social structure of cities. However, the last three papers differ from us in that people are assumed to work where they live. The absence of commuting vastly simplifies the theoretical analysis. Finally, Couture and Handbury (2016) document the revival of CBD areas in several large US cities and confirm the role played by amenities in this reversal of fortune.

Note, finally, the link with the vast literature on Tiebout and the sorting of households among jurisdictions. Local public finance disregards the trade-off between amenities, proximity to jobs and housing prices (see Epple *et al.*, 2010, for an attempt at synthesizing the Tiebout-like literature with urban economics). The main focus of the literature on local public finance is on stratification by income (Epple and Nechyba, 2004). For example, Henderson and Thisse (2001) show how households endowed with different incomes belong to a finite number of distinct jurisdictions which compete directly with the adjacent communities on the income line. In this case, there is perfect sorting of households through community formation. In a way, our model may be viewed as a setup involving a continuum of jurisdictions where households cross borders between jurisdictions. De Bartolome and Ross (2007) is a noticeable exception. These authors show that income mixing may arise in a city formed by a central and a peripheral jurisdiction. When households are heterogeneous in incomes and in their preferences for local public goods, Epple and Sieg (1999) show that several jurisdictions may host residents having the same income. Unlike Epple and Sieg and others, we show in this paper that households sharing the same preferences and the same income may live far apart and have different housing consumptions.

The remainder of the paper is organized as follows. In Section 2, we compare the income, amenity and housing price gradients in Amsterdam and Rotterdam. Since our setup is used in the empirical part of the paper, we provide a detailed description of our model in Section 3 and to show how the bid rent function may be used to determine the social stratification of the city. In Section 4, we study the properties of the residential pattern for preferences that generate a location-quality index. Since the equilibrium is undetermined under homothetic preferences, we illustrate our results for Stone-Geary preferences. In Section 5, we determine analytically the market outcome when income and the location-quality index are Fréchet-distributed. Data are discussed in Section 6. In Section 7, we study the causal relationship between incomes, job accessibility and amenities, while we present the results of our counterfactuals in Section 8. Section 9 concludes.

2 Are urban gradients monotonic?

In this section, we provide some empirical evidence showing that income, land value and amenity gradients are not necessarily monotonic. The main empirical analysis focuses on the two largest cities in the Netherlands: Amsterdam and Rotterdam on incomes, commuting times, house price and amenities. Both Amsterdam and Rotterdam have a population of about one million. Amsterdam was the economic capital of Europe in the 17th century while Rotterdam is the largest seaport in Europe for a very long time. In Appendix A.1, we gather descriptive data on London which show similar patterns.

European cities differ markedly from many US cities. The first distinctive feature is that Europe's urban history is much longer than in the US (Bairoch, 1988). The current urban structure is thus heavily influenced by the past, for example through the presence of historic buildings in city centers and the pattern of canals in Amsterdam. In contrast to many US cities, the presence of historic buildings implies that the amenity levels in inner cities are much higher. An exception is Rotterdam, which was completely bombed out during World War II. Thus, its city center does not host historic amenities and displays an urban structure somewhat comparable to a monocentric US city with a CBD dominated by high-rise buildings.

Dutch cities are characterized by an extremely flat geography. Most cities in the Netherlands are just above (and sometimes below) sea level. In the selected cities, the highest point is 15 m, while the lowest is -8.9 m in an area that was reclaimed from the sea; the differences within cities are smaller. The expansion of a city is, therefore, not as constrained by geographical features. However, Dutch land-use planning strongly restricts urban growth. A well-known example is the Green Heart, which is in the vast space between the largest four cities in the Netherlands. The Green Heart policy dates back to 1954: new constructions have essentially been prohibited there since then (Koomen *et al.*, 2008). A second distinctive characteristic of Dutch cities is that school quality is unlikely to play a major role in household residential choices because school performances are higher and much less uneven in the Netherlands than in the US, suggesting that the search for better school districts is less of a concern in the former. Moreover, households are free to choose their most preferred school.

We use information on incomes, land values, and employment locations. We also use geocoded outside pictures by residents, which were kindly provided to us by Eric Fisher, as a proxy for amenities. More information on the data, the procedures used to estimate land values, the proxy for amenities and commuting costs, and the way in which we address the endogenous nature of amenities, is provided in Section 6 and Appendix C. Because cities are highly non-symmetric around the city center, plotting the amenity index against the distance to the city center masks a substantial portion of the spatial variations in the above-mentioned variables. Nevertheless, for illustration Figure 1 provides a one-dimensional gradient for a strip of 2 km wide which stretches along the rivers Amstel and Maas for Amsterdam and Rotterdam, respectively (two-dimensional maps are provided in Appendix A.2). For the two cities, *income gradients are highly non-monotonic.* The land price gradient is non-monotonic for Rotterdam because the attractive neighborhoods are located outside the bombed area, while this gradient is decreasing in Amsterdam.

[Figure 1 about here]

This paper argues that such contrasted patterns may be (at least, partially) explained by the presence of amenities. At this point, we cannot infer any causal effect of amenities on incomes because the provision of some amenities may be the result of the presence of high-income households. Nevertheless, it is informative to investigate whether the amenity and employment density gradients are also non-monotonic. Figure 2 shows that *amenity levels are* *always higher near the city center*, probably because the city centers provide historic amenities (Brueckner *et al.*, 1999). For Amsterdam the amenity gradient is essentially decreasing and increasing somewhat after 7 km from the city enter. This also holds for Rotterdam, but we observe a sharper increase beyond 7 km from the city center, implying that *amenity gradients are non-monotonic*. Despite the bombing of its historic city center, Rotterdam has a high amenity level in the center, probably due to the high architectural quality of the new-built buildings.[1] Figure 2 shows that *commuting costs are also non-monotonic*.

[Figure 2 about here]

3 The model and preliminary results

3.1 The model

Although the model can be developed for a two-dimensional asymmetric city, it is notationally convenient to consider a linear city spread over a subset X of the real line. Location and distance from the origin are identically denoted by $x \in X$. The city is polycentric; jobs are available in the CBD located at the origin $e_0 = 0$ or in secondary business districts located at $e_i > 0$, $i = 1, ...n$. The city is endowed with a unit mass of inhabitants. There are two normal consumption goods: (i) land (h), which is a proxy for housing, and (ii) a composite non-spatial good (q), which is used as the numéraire. The opportunity cost of land is given by the constant $R_A \geq 0$, while the amount of land available at each location x is normalized to 1. Let $a > 0$ be the level or, equivalently, the utility of amenities, which is common to all households. Consumers' preferences are separable in amenities and consumption goods (q, h):

$$U(q, h; a) = a \cdot u(q, h), \tag{1}$$

where u is strictly increasing in the numéraire and land consumption, strictly quasi-concave in (q, h), and indifference curves do not cut the axes. The utility (1) implies that the amenity a and the private goods (q, h) are substitutes. In addition, the utility level associated with the consumption of a given bundle (q, h) increases with the amenity level a, while the utility derived from consuming amenities rises with income. Hence, a high-income household needs more numéraire than a low-income household to be compensated for the same decrease in amenity consumption. As a result, the single-crossing condition between incomes and amenities holds (Schmidheiny, 2006). However, as will be seen, the single-crossing condition between incomes and locations does not hold: richer households need not choose locations with more amenities because they care about their commuting costs.

[1] As an illustration, Rotterdam was ranked in the list 'top cities in the world' by Lonely Planet's Best in Travel 2016.

A household's gross income is given by ω units of the numéraire, with $\omega \in [\underline{\omega}, \bar{\omega}]$ and $0 < \underline{\omega} \leq \bar{\omega}$. The income c.d.f. $F(\omega)$ and density $f(\omega)$ are continuous over $[\underline{\omega}, \bar{\omega}]$. We are agnostic about the reasons that explain inequality in earnings. In this paper, we model the individual loss due to commuting as an *iceberg* cost. More specifically, if a household residing at x works at e_i, we denote by $0 < t_i(|x - e_i|) \leq 1$ her effective number of working units, which decreases with the distance $|x - e_i|$ while $t_i(0) = 1$. As a consequence, the individual's net income is equal to $\omega t_i(x)$ and her commuting cost by $c_i(\omega, x) = \omega - \omega t_i(x)$, which increases with both her earning ω and the distance $|x - e_i|$. In other words, commuting is considered here as an income loss. This modeling strategy captures the fact that individuals who have a long commute are more prone to being absent from work, to arrive late at the workplace and/or to make less work effort (van Ommeren and Gutiérrez-i-Puigarnau, 2011).

If a household located at x earns the same income in each employment center e_i, she chooses to work at the closest employment center. In this case, the function $t(x)$ is given by the maximum working time, which is not monotonic in x when the city is polycentric.[2] We will see in Section 7 that our results do not hinge on this particular definition of $t(x)$. The household's budget constraint is thus as follows:

$$\omega t(x) = q + R(x)h, \tag{2}$$

where $R(x)$ is the land rent at x. In line with the literature, we assume that the land rent is paid to absentee landlords (Fujita, 1989).

Let $\eta(y) > 0$ be a given function whose value expresses the level of amenities available at $y \in X$, such as scenic landscapes, rivers, historic buildings and architecture. The amenities are intrinsic to a location and exogenous. Therefore, the corresponding utility level is negatively affected by the distance between the household and the place where these amenities are available. As a result, a household at x ascribes the value $\varphi(|x - z|)\eta(z)$ to the amenity provided at $z \neq x$, which decreases with the distance $|x - z|$ between x and z, with $\varphi(0) = 1$. In other words, $\varphi(\cdot) \geq 0$ has the nature of a distance-decay function. As shown by (1), households' well-being at x depends on the total amenity level defined by the following expression:

$$a(x) \equiv \int_X \varphi(|x - z|)\eta(z)\mathrm{d}z. \tag{3}$$

We do not impose any functional restriction on $\eta(\cdot)$ and $\varphi(\cdot)$. Therefore, (3) includes the specifications used by Fujita and Ogawa (1982), Lucas and Rossi-Hansberg (2002) and others to

[2] Note that the function $t(x)$ is not differentiable at the intersection points between any two functions t_i and t_j. If the equilibrium arises at a point where t is not differentiable, the first-order conditions must be rewritten by using the tools of subdifferential calculus. This does not affect the meaning of our results but renders the exposition heavy. For this reason, we will assume throughout that all functions are as many times continuously differentiable as necessary.

describe various spatial externalities. In the featureless city of urban economics, $a(x)$ is constant across locations. In this paper, we focus on the case where $a(x)$ varies with x. Hence, the city is *differentiated* in the sense that preferences depend on household locations.

Maximizing the utility U of a ω-household residing at x and working at e_i with respect to q and h subject to (2) yields the *numéraire* demand

$$q^*(x,\omega) \equiv q(R(x),\omega t(x)) = \omega t(x) - R(x)h(R(x),\omega t(x))$$

and the *housing* demand $h^*(x,\omega) \equiv h(R(x),\omega t(x))$, which is the unique solution to the equation:[3]

$$u_h\left[\omega t(x) - R(x)h^*, h^*\right] - R(x)u_q\left[\omega t(x) - R(x)h^*, h^*\right] = 0. \tag{4}$$

Let $s(x,\omega) \in [0,1]$ be the share of the ω-households who reside at x. The land market clearing condition holds if $\omega(x)$ satisfies the following condition:

$$|s(x,\omega(x))f(\omega(x))h^*(x,\omega)\mathrm{d}\omega| = \mathrm{d}x, \tag{5}$$

which says that the amount of land available between any x and $x+\mathrm{d}x > x$ and the area occupied by the households whose income varies from ω to $\omega+\mathrm{d}\omega$ are the same. Since $\omega(x)$ needs not be monotonic, the land market clearing condition is expressed in absolute value. Hence, the endogenous city limit B is given by

$$\int_0^B s(x,\omega(x))f(\omega(x))h^*(x,\omega)\mathrm{d}x = B. \tag{6}$$

The residential equilibrium is such that no household has an incentive to move, all households sharing the same income have the same maximum utility level, and the land market clears. Formally, a *residential equilibrium* is defined by the following vector:

$$(\omega^*(x), s^*(x,\omega^*(x)), R^*(x), h^*(x,\omega^*(x)), B^*)$$

such that

$$a(x)\cdot u[q^*(x,\omega^*(x)), h^*(x,\omega^*(x))] \geq a(y)\cdot u[q^*(y,\omega^*(x)), h^*(y,\omega^*(x))] \qquad 0 \leq y \leq B \tag{7}$$

holds under the constraints (2), (5) and (6). For simplicity, we assume that the amenity function $a(x)$ is such that there is no vacant land at the residential equilibrium. By implication, $R^*(x) > R_A$ for $x < B^*$ and $R^*(x) = R_A$ for $x \geq B^*$.

In equilibrium, households sharing the same income are *not* indifferent across locations because space is differentiated. In particular, if the inequality is strict in (7) for all $y \neq x$, then

[3] For any function $f(y,z)$, let f_y (resp., f_{yz}) be the partial (cross-) derivative of f with respect to y (resp., y and z).

all $\omega^*(x)$-households are located at x $(s^*(x, \omega^*(x)) = 1)$. Otherwise, there exist at least two locations x and y such that the $\omega^*(x)$-households are indifferent between x and y. In this case, we have $0 < s^*(\cdot, \omega^*(x)) < 1$ at x and y, while the sum of the shares is equal to 1.

The problem consists in assigning households having a particular income to specific locations within the city. It seems that the residential equilibrium can be obtained by appealing to methods developed in matching theory. This raises two types of difficulties, which are typically ignored in this theory. First, a household's housing consumption, whence its equilibrium utility level, varies with both its income and location. Second, for the matching between households and locations to be imperfect, the rule $x(\omega)$ which assigns a particular income to locations must be a *correspondence*. Indeed, there is imperfect matching if households sharing the same income live far apart. For example, for the same given housing consumption, a household can be indifferent between living close to the CBD while having a low level of amenities, or living far from the CBD while enjoying a high level of amenities. Therefore, apart from special cases, there is no one-to-one correspondence between the income and location sets. Because $x(\omega)$ is a correspondence, it seems hopeless to guess what the equilibrium assignment could be, as typically done in Sattinger-like assignment models. By contrast, it is yet unnoticed that the reverse problem can be solved. Indeed, because households bid for locations, those who reside at the same place must share the same income. Therefore, we may define and characterize the *income mapping* $\omega^*(x)$ from the location set $[0, B]$ to the income set $[\underline{\omega}, \bar{\omega}]$ that specifies which ω-households are located at x. Evidently, this income is that of the households who make the highest bid (Fujita, 1989).

3.2 The residential equilibrium with amenities

Since u is strictly increasing in q, the equation $u(q, h) = U/a(x)$ has a single solution $Q(h, U/a(x))$, which describes the consumption of the numéraire when the utility level is $U/a(x)$ and the land consumption h. The *bid rent* $\Psi(x, \omega; U)$ of a ω-household is the highest amount it is willing to pay for one unit of land at x when its utility level is given and equal to U. The bid rent function is defined as follows:

$$\begin{aligned}\Psi(x, \omega, U) &\equiv \max_{q,h} \left\{ \frac{\omega t(x) - q}{h} \middle| \text{s.t. } a(x) \cdot u(q, h) = U \right\} \\ &= \max_{h} \frac{\omega t(x) - Q(h, U/a(x))}{h}. \end{aligned} \tag{8}$$

When space is differentiated, since the bid rent $\Psi(x, \omega, U)$ is such that the ω-households are indifferent across locations, (4) implies that the Alonso-Muth condition for the ω-households is as follows:

$$h^*(x, \omega)\Psi_x(x, \omega, U) - \omega t_x(x) = \frac{a_x(x)}{a(x)} \frac{u(q^*(x, \omega), h^*(x, \omega))}{u_q(q^*(x, \omega), h^*(x, \omega))}.$$

Since each household treats the utility level parametrically, applying the first-order condition to (8) yields the equation:

$$Q(h, U/a(x)) - hQ_h(h, U/a(x)) - \omega t(x) = 0 \tag{9}$$

whose solution, denoted $H(\omega t(x), U/a(x))$, is the quantity of land consumed at x when the household pays its bid rent $\Psi(\cdot)$, which is called the *bid-max lot size* (Fujita, 1989). The budget constraint implies that the bid rent function may be rewritten as follows:

$$\Psi(x, \omega, U) \equiv \frac{\omega t(x) - Q(\omega t(x), U/a(x))}{H(\omega t(x), U/a(x))}. \tag{10}$$

This expression shows that a household's bid rent at x depends separately on both $a(x)$ and $t(x)$ while its housing consumption H also varies with these two attributes of location x (we will return to this issue in Section 5). Since land is allocated to the highest bidder, the equilibrium land rent is given by the upper envelope of the bid rent functions:

$$R^*(x) = \max\left\{\max_{\omega \in [\underline{\omega}, \bar{\omega}]} \Psi(x, \omega, U^*(\omega)), R_A\right\},$$

where $U^*(\omega)$ denotes the maximum utility reached by the ω-households at the residential equilibrium. The bid rent function implies that all households residing at a particular location x share the same income $\omega^*(x)$. Hence, two households endowed with different incomes choose to reside in two different locations. Note that $H(\cdot)$ is the equilibrium consumption at x of a ω-household if its bid rent is equal to the land rent.

Since land is allocated to the highest bidder, the income $\omega^*(x)$ of the households who locate at x must solve the utility-maximizing condition:

$$\frac{\partial \Psi(x, \omega, U^*(\omega)))}{\partial \omega} = 0, \tag{11}$$

while the second-order condition implies $\partial^2\Psi/\partial\omega^2 < 0$. Totally differentiating (11) with respect to x yields:

$$\frac{\mathrm{d}\omega^*}{\mathrm{d}x} = -\left[\frac{\partial^2 \Psi(x, \omega, U^*(\omega))}{\partial \omega^2}\right]^{-1} \cdot \frac{\partial^2 \Psi(x, \omega, U^*(\omega))}{\partial \omega \partial x}, \tag{12}$$

which implies that $\Psi_{x\omega}(x, \omega, U^*(\omega))$ and $\mathrm{d}\omega^*(x)/\mathrm{d}x$ have the same sign.

Set

$$A(x) \equiv \frac{a_x(x)}{a(x)} \qquad T(x) \equiv -\frac{t_x(x)}{t(x)},$$

and

$$\varepsilon_{U,\omega} \equiv \frac{\omega}{U^*} U^*_\omega \qquad \varepsilon_{H,\omega} \equiv \frac{\omega}{H}\left(H_\omega + H_U U^*_\omega\right) \qquad \varepsilon_{u_q,\omega} \equiv \frac{\omega}{u_q}\frac{\partial u_q}{\partial \omega}.$$

We are now equipped to characterize the equilibrium income mapping.[4]

[4] When no ambiguity may arise, we do not specify the independent variables in the following equations.

Proposition 1. *The equilibrium mapping* $\omega^*(x)$ *is increasing* (decreasing) *at* x *if*

$$\phi(x,\omega) \equiv \left(1 - \frac{\varepsilon_{H,\omega} + \varepsilon_{u_q,\omega}}{\varepsilon_{U,\omega}}\right) A - (1 - \varepsilon_{H,\omega})T \tag{13}$$

is positive (negative) *at this location.*

Proof. The proof is given in Appendix B.1. ■

The expression (13) shows that for any given function u the interaction between the amenity and commuting cost functions determines the social stratification of the city through the behavior of the function ϕ. It also shows that the sign of ϕ, whence the slope of the spatial income distribution, changes at any solution $\omega^*(x)$ to the equation $\phi(x,\omega) = 0$ if $\phi_x(x,\omega) \neq 0$.

To illustrate, consider first the benchmark case of a monocentric and undifferentiated city, that is, $A(x) = 0$ and $T(x) > 0$ for all x. We know from Fujita (1989) that household locations are determined by ranking the bid rent slopes with respect to income. It follows from (13) that the sign of $\phi(x,\omega)$ depends on whether the income elasticity of the bid-max lot size is smaller or larger than 1 (Wheaton, 1977). Since the empirical evidence shows that the expenditure share allocated to housing declines as income rises, the income elasticity of housing is smaller than 1 (Albouy *et al.*, 2016). Therefore, when income increases, the slope of the bid rent function gets steeper. A longer commute shifts the utility of a high-income household downward more than that of a low-income household because the former has a higher opportunity cost of time than the latter. However, this effect is not offset by the higher housing consumption because the income elasticity of housing is smaller than 1. By implication, at the residential equilibrium, households are sorted by decreasing order of income as the distance to the CBD increases.

Consider now the case of a differentiated monocentric city ($A(x) \neq 0$). Owing to the existence of amenities, even when the bid rent functions are downward sloping, the equation $\phi(x,\omega) = 0$ may have several solutions. In this case, there is *imperfect sorting*, that is, *greater income differences are not mapped into more spatial separation.* The following three cases may arise.

(i) Assume that $\phi(x,\omega) > 0$ for all x. As ω rises, the bid rent curve becomes flatter. Since the bid rent of a high-income household is always flatter than that of a low-income household, individuals are sorted out by increasing income. In other words, *the richer the household, the closer to the city limit.* Consumers are willing to pay more to reside at a distant location because the corresponding hike in amenity consumption is sufficient to compensate them for their longer commute. In this case, $x = B^*$ is the most-preferred city location, whereas $x = 0$ is the worst. This configuration is the one obtained in standard urban economics when commuting costs are income-independent, space is undifferentiated, and housing consumption is variable (Hartwick *et al.*, 1976; Fujita, 1989).

(ii) If $\phi(x,\omega) < 0$ for all x, the bid rent curve becomes steeper as the income ω rises.

Therefore, the bid rent curves associated with any two different incomes intersect once and, for each ω, there exists a unique $x(\omega)$ such that $s(x(\omega), \omega) = 1$. In this case, $x = 0$ is the most-preferred city location. To put it differently, the utility loss incurred by an increase in distance to the workplace is exacerbated by a drop in the consumption of central amenities (Brueckner *et al.*, 1999).

(iii) The most interesting case arises when $\phi(x, \omega)$ changes its sign over $[0, B]$ because, as shown in Section 2, the slope of the income gradient changes. In other words, there is imperfect sorting: household income rises over some range of sites and falls over others. We develop this argument in more detail in Section 4.

The expression (13) highlights the fact that the impact of the amenity and commuting cost functions on the sign of $\phi(x, \omega)$ depends on the elasticities $\varepsilon_{H,\omega}$, $\varepsilon_{u_q,\omega}$ and $\varepsilon_{U,\omega}$. When the utility $u(q, h)$ is specified, the condition (13) may be used to determine how households are distributed according to the behavior of $A(x)$ and $T(x)$ by calculating those elasticities. In the limit, when the elasticities are constant, the sign of (13) is independent of income, and thus there is perfect sorting. Note that $T(x) = 0$ when commuting is not accounted for, like in most models of local public finance. In this event, the sign of $\phi(x, \omega)$ is determined by the sign of $(\varepsilon_{U,\omega} - \varepsilon_{H,\omega} - \varepsilon_{u_q,\omega})A$ only. In particular, there is a continuum of equilibria when $a(x)$ is constant because $A = 0$.

Finally, when the city is polycentric, the function $T(x)$ displays several extrema because $t(x)$ is no longer monotonic. Hence, even when $a(x)$ is constant across locations, *the decentralization of jobs favors income mixing.* As a consequence, to what extent the behavior of the income mapping is determined by amenities or commuting time is an empirical question.

Modeling commuting costs In the standard monocentric city model, the individual working time is supposed to be constant ($t_x = 0$), which implies commuting costs are given by an increasing function $c(x)$. In this case, we show in Appendix B.2 that $\phi(x, \omega)$ is given by the following expression:

$$\phi(x, \omega) = \left(1 - \frac{\varepsilon_{H,\omega} + \varepsilon_{u_q,\omega}}{\varepsilon_{U,\omega}}\right) A + \frac{c_x}{\omega t} \varepsilon_{H,\omega}. \tag{14}$$

Consequently, since $c_x > 0$ there is sorting by increasing income from the CBD for any utility u when $A(x) = 0$ (Duranton and Puga, 2015). However, this need not be true when $A(x) \neq 0$.

Note the following difference between (13) and (14). In the former, households are sorted by decreasing income order, whereas in the latter households are sorted by increasing income order. Absent amenities, *there is perfect sorting under the two modelling approaches*, but the ranking of households is reverse. Thus, how to model commuting costs matters for the order in which households are ranked. Since there is ample evidence that suggests that the opportunity cost of time varies significantly when households are heterogeneous in incomes (Koster and Koster, 2015), using income-independent commuting costs seems restrictive for studying the residential choices of such households. This is why we have chosen to use an iceberg commuting cost.

4 The city social structure under Stone-Geary preferences

In this section, we characterize the equilibrium mapping $\omega^*(x)$ and the equilibrium land rent $R^*(x)$ for preferences $u(q,h)$ that reduce the dimensionality of the matching problem.

Consider the following options. First, it seems natural to start with homothetic preferences, as they include the CES, Cobb-Douglas and translog. If the utility u is homogeneous linear, we show in Appendix B.3 that $\varepsilon_{U,\omega} = \varepsilon_{H,\omega} = 1$ and $\varepsilon_{u_q,\omega} = 0$. As a result, (13) can be reduced to $\phi(x,\omega) = 0$ for all x. In other words, *there is a continuum of residential equilibrium under homothetic preferences*. This should not come as a surprise as the income distribution does not play any role in the location of consumption and production under the combination of homothetic preferences and iceberg trade costs (Fajgelbaum *et al.*, 2011). Second, quasi-linear preferences are non-homothetic and simple to handle: $u(q,h) = v(h) + q$ where v is strictly increasing and concave. In this case, we have $\varepsilon_{H,\omega} = \varepsilon_{u_q,\omega} = 0$, so that (13) reduces to $\phi(x,\omega) = A - T$. This expression suggests that quasi-linear preferences are a good candidate to study the residential equilibrium. Unfortunately, assuming quasi-linear preferences is counterfactual as housing is a normal good ($\varepsilon_{H,\omega} > 0$).

Third, a well-known example of non-homothetic utility is Stone-Geary's:

$$u(q,h) = q^{1-\mu}(h - \overline{h})^{\mu}, \tag{15}$$

where $0 < \mu < 1$ and $\overline{h} > 0$ is the minimum lot size, which is supposed to be sufficiently low for the equilibrium consumption of the numéraire to be positive. Maximizing (15) with respect to q and h subject to (2) leads to the linear expenditure system:

$$q^*(x,\omega) = (1-\mu)[\omega t(x) - R(x)\overline{h}], \tag{16}$$

$$h^*(x,\omega) = (1-\mu)\overline{h} + \mu\frac{\omega t(x)}{R(x)}. \tag{17}$$

The housing demand at any location x increases less than proportionally with income, which is in line with Albouy *et al.* (2016).[5] Assuming $\mu = 0$ corresponds to the fixed lot size assumption while $\overline{h} = 0$ implies that (17) reduces to the housing demand under Cobb-Douglas preferences.

[5] Davis and Ortalo-Magné (2011) provide evidence that the expenditure share on housing is constant over time and across U.S. metropolitan areas. This does not mean that households having different incomes spend the same share of their incomes on housing *within* cities.

We show in Appendix B.4 that $\phi(x,\omega) = A - (1-\mu)T$. Set

$$\Delta(x) \equiv t(x)[a(x)]^{\frac{1}{1-\mu}}, \tag{18}$$

which subsumes the amount of time devoted to work and the amenity level at x into a single scalar, which has the nature of a *location-quality index*. Note that this index depends on location x but not on income $\omega^*(x)$. The higher μ, the stronger the preference for housing. Therefore, as the intensity of preference for housing increases, amenities matter more than the accessibility to jobs. Moreover, differentiating (18) shows that $\Delta(x)$ and $\phi(x,\omega)$ have the same sign. Hence, $\phi(x,\omega)$ changes sign at any extrema of the location-quality index.

Finally, consider the following non-homothetic CES utility: $u(q,h) = q^{\rho_1} + h^{\rho_2}$ with $0 < \rho_i < 1$ (u is quasi-linear for $\rho_1 = 1$). The elasticity of substitution $\sigma(q,h)$ between housing and the numéraire is now variable and equal to $1/(1 - s_1\rho_1 - s_2\rho_2)$ where s_i is the expenditure share on good $i = q,h$. When $\rho_1 > \rho_2$, it can be shown that the above preferences generate the index $\Delta(x) \equiv t(x)\,[a(x)]^{1/\rho_1}$, which is similar to (18).

In what follows, we work with Stone-Geary preferences because they are amenable to a simple analytical solution which can be brought to the data. However, our results hold true whenever the location-quality index $\Delta(x)$ is a function of $a(x)$ and $t(x)$ that is independent of income.

4.1 The spatial income distribution

Our objective is now to determine the mapping $\omega^*(x)$ from $[0, B^*]$ to $[\underline{\omega}, \bar{\omega}]$ that specifies which ω-households are located at x under (15). Since housing consumption is chosen optimally at each x, what makes a site attractive to households is both its amenity level and the corresponding working time. The next proposition shows that households distribute themselves across the city according to the values of the location-quality index.

Proposition 2. *Assume Stone-Geary preferences. At the residential equilibrium, the income and location-quality index vary in the same way:* $\omega^*(x) = F^{-1}[G(\Delta(x))]$.

Proof. (i) We show in Appendix B.5 that the bid-max lot size

$$H(\omega t(x), U/a(x)) \equiv H(\Delta(x), \omega, U) \tag{19}$$

depends on $a(x)$ and $t(x)$ only through the location-quality index (18).

(ii) We show in the same appendix that the equilibrium condition $\Psi_\omega = 0$ is equivalent to the differential equation:

$$\frac{\mathrm{d}U^*}{\mathrm{d}\omega} = (1-\mu)\,\Delta(H-\overline{h})^{\frac{\mu}{1-\mu}}(U^*(\omega))^{-\frac{\mu}{1-\mu}} \tag{20}$$

whose solution varies with Δ. Hence, the equilibrium utility level depends on the two attributes of location x, that is, $a(x)$ and $t(x)$, only through the location-quality index Δ. Proposition 1 implies that the equilibrium mapping is given by the expression:

$$\omega^*(x) = F^{-1}[G(\Delta(x))],$$

with $\underline{\omega} = F^{-1}[G(\underline{\Delta})]$ and $\overline{\omega} = F^{-1}[G(\overline{\Delta})]$. ■

Proposition 2 states that it is sufficient to study how $\Delta(x)$ varies to determine the properties of the residential equilibrium, rather than $a(x)$ and $t(x)$ separately. This shows *how the initial two-to-one matching is reduced to a one-to-one matching.* In addition, the functions $\omega^*(x)$ and $\Delta(x)$ have the same extrema.

Furthermore, we show in Appendix B.6 that $\mathrm{d}U^*(\omega)/\mathrm{d}\omega$ is an increasing function of Δ:

$$\frac{\partial}{\partial \Delta}\frac{\mathrm{d}U^*}{\mathrm{d}\omega} > 0, \tag{21}$$

meaning that the equilibrium utility is supermodular in income and the location-quality index. As a consequence, the residential equilibrium involves a positive assortative matching between incomes and the values of the location-quality index (Legros and Newman, 2002). In other words, households ordered by increasing incomes are assigned to locations endowed with rising values of the location-quality index.

Let us rank the values of $\Delta(x)$ by increasing order and denote by $G(\Delta)$ be the corresponding c.d.f. defined over the domain $[\underline{\Delta}, \overline{\Delta}]$ where $\underline{\Delta}$ ($\overline{\Delta}$) is the minimum (maximum) of $\Delta(x)$ over $[0, B]$. Note that $G(\Delta)$ describes how the values of the location-quality index are distributed irrespective of location.

Since the function (18) is in general not monotonic, we have:

$$\frac{\partial}{\partial x}\frac{\mathrm{d}U^*}{\mathrm{d}\omega} \gtrless 0,$$

Therefore, income sorting is not mapped into spatial sorting, which means that *the income gradient need not be a monotonic function of the distance from the CBD.* As a consequence, the affluent (or the poor) do not necessarily locate at the CBD or the city limit. Rather, the richest locate where the location-quality index is maximized, whereas the poorest reside in locations with the lowest location-quality index.[6]

[6] As for the housing demand, we have $\partial H/\partial \Delta < 0$, for otherwise the utility level U of the ω-households would increase. Furthermore, we also know that $\partial H/\partial \omega > 0$ and $\partial H/\partial U > 0$ hold because housing is a normal good (Fujita, 1989). Given Proposition 2, $\partial H/\partial \Delta < 0$, $\partial H/\partial \omega > 0$ and $\partial H/\partial U > 0$ imply that the sign of $\mathrm{d}H/\mathrm{d}x$ is ambiguous. Indeed, when the location-quality index rises with x, the income of the corresponding residents also rises. Because housing is a normal good, this income hike incites households to consume more housing. However, those households also enjoy a higher location-quality index, which tends to reduce their housing consumption. How the housing consumption varies with the distance to the CBD is thus undetermined.

To illustrate, consider Figure 3. The centrality of the city is described by the unique global maximizer $x = 0$ of $\Delta(x)$ over $[0, B^*]$ because this site is endowed with the best combination of amenities and commuting costs. Proposition 2 implies that this location is occupied by the richest households, while households are sorted by decreasing income over $[0, x_1)$ where x_1 is a minimizer of Δ. Since x_1 is the unique global minimizer of $\Delta(x)$, this location is occupied by the poorest households. As the distance to the CBD rises, $\Delta(x)$ increases. This implies that households are now sorted by increasing income up to x_2 where $\Delta(x)$ reaches a local maximum. Over the interval $(x_2, B^*]$, the function $\Delta(x)$ falls again, which means that households' income decreases with x.

Since $\Delta(0) > \Delta(x_2) > \Delta(B^*) > \Delta(x_1)$, the intermediate value theorem implies that z_1 in $[0, x_1)$, z_2 in (x_1, x_2) and z_3 in $(x_2, B^*]$ exist such that $\Delta(z_1) = \Delta(z_2) = \Delta(z_3)$. Proposition 2 implies that the households residing at these three locations have the same income. In other words, *households sharing the income* $\omega^*(z_i)$ *do not live in the same neighborhood.* On the contrary, they are spatially separated by households having lower incomes in (z_1, z_2) and higher incomes in (z_2, z_3). Roughly speaking, Figure 3 depicts a spatial configuration where the middle class is split into two spatially separated neighborhoods with the poor in between, while the affluent live near the city center. Such a pattern describes more accurately the spatial distribution of incomes in "old" US cities and in many European cities, than the homogeneous monocentric city model (Glaeser *et al.*, 2008).

[Figure 3 about here]

More generally, assume that the location-quality index has n extrema. If $n = 2$ there is perfect sorting because Δ has a unique maximizer and a unique minimizer. When $n > 2$, the spatial separation between households is no longer the mirror image of their income differences. The residential pattern is partitioned into neighborhoods whose borders are defined by the adjacent extrema of the location-quality index and size depends on the behavior of the index. When z is a maximizer of Δ, then the locations $x_1 < z < x_2$ with $\Delta(x_1) = \Delta(x_2) < \Delta(z)$ are in general such that $x_2 - z \neq z - x_1$ because Δ is not symmetric. In other words, the households whose income is $\omega^*(x_1) = \omega^*(x_2)$ are not located equidistantly about z. The same holds when z is a minimizer of Δ. Moreover, when the location-quality index is multimodal, there is perfect income sorting *within* each neighborhood, but income mixing *between* neighborhoods. Unlike Tiebout's prediction, *identical or similar households may live in spatially separated areas.*

To determine the residential distribution of households, it remains to find the equilibrium values of the shares $s(x, \omega(x))$. If $z_1 \neq z_2 ... \neq z_n$ exist such that $\Delta(z_1) = \Delta(z_j)$ for $j = 2, ..., n$, it follows from Proposition 2 that $\omega^*(z_1) = \omega^*(z_j)$ for $j = 2, ..., n$. Using (5) and (19), we also have:

$$|s(z_i,\omega^*(z_i))f(\omega^*(z_i))H\{\Delta(z_i),\omega^*(z_i),U^*(\omega^*(z_i))\}\,\mathrm{d}\omega| = \mathrm{d}x \qquad i=1,...,n.$$

Since $H\{\omega^*(z_1),\Delta(z_1),U^*(\omega^*(z_1))\} = H\{\Delta(z_j),\omega^*(z_j),U^*(\omega^*(z_j))\}$ and $f(\omega^*(z_1)) = f(\omega^*(z_j))$ for $j=2,...,n$, we get:

$$s(z_1,\omega^*(z_1)) = s(z_j,\omega^*(z_j)) \qquad j=2,...,n.$$

Furthermore, it must be that

$$\sum_{i=1}^{n} s(z_i,\omega^*(z_i)) = 1.$$

It follows from these n equations that $s(z_i,\omega^*(z_i)) = 1/n$ for $i = 1,...n$. That is, the households who share income $\omega^*(z_1)$ are equally split across the locations that generate the same location-quality index $\Delta(z_1)$.

Back to commuting costs. When commuting costs are income-independent, the net income at x is given by $\omega - c(x)$. Using (B.5.1) in Appendix B.5 in which the net income $t(x)\omega$ is replaced by $\omega - c(x)$ and repeating the above argument, we obtain the following aggregator:

$$\Delta(x,\omega) = (\omega - c(x))[a(x)]^{1/(1-\mu)},$$

which depends on both x and ω. Since $\Delta(x,\omega)$ increases with ω, the ω-households reside at location x which maximizes $\Delta(x,\omega)$. This implies that the location-quality index is now given by the upper-envelope function:

$$\Phi(x) = \max_{\omega} \Delta(x,\omega).$$

The spatial distribution of heterogeneous households may then be determined by applying the approach developed above to $\Phi(x)$. The same approach can be used to cope with a city that expands both to the left and right of the CBD and where $\Delta(x)$ need not be equal to $\Delta(-x)$. In this case, the location-quality index is given $\Phi(x) = \max\{\Delta(x),\Delta(-x)\}$.

4.2 Land rent

It remains to characterize the equilibrium land rent. We show in Appendix B.7 that the equilibrium land rent is given by the following expression:

$$R^*(x) = \frac{\omega^*(x)t(x)}{H[\Delta(x),\omega^*(x),U^*(\omega^*(x))]}\left[1 - \frac{1-\mu}{\varepsilon_{U,\omega}(x)}\right], \tag{22}$$

where

$$\varepsilon_{U,\omega}(x) = (1-\mu)\frac{\omega^*(x)t(x)}{q^*(x)} = \frac{\omega^*(x)t(x)}{\omega^*(x)t(x) - \overline{h}R^*(x)} > 1.$$

Substituting $\varepsilon_{U,\omega}(x)$ in (22) and rearranging terms, we obtain:

$$R^*(x) = \frac{\mu\omega^*(x)t(x)}{H[\Delta(x), \omega^*(x), U^*(\omega^*(x))] - (1-\mu)\overline{h}} > 0, \tag{23}$$

where we assume that $\mu > 0$ for the numerator and denominator to be strictly positive.

By totally differentiating (22) with respect to x, we obtain (see Appendix B.7):

$$R_x^*(x) = \frac{\omega^*(x)t(x)}{H[\Delta(x), \omega^*(x), U^*(\omega^*(x))]} \left[\frac{1}{\varepsilon_{U,\omega}(x)} A(x) - T(x)\right]. \tag{24}$$

Since $\varepsilon_{U,\omega}(x) > 1$, the above expression implies that the land rent gradient is always negative if $A(x) - T(x) < 0$ for all x. As x rises, the decreasing land rent compensates the $\omega^*(x)$-households for bearing higher commuting costs and being farther away from places endowed with more amenities. For example, in the standard monocentric city model in which $A(x) = 0$ and $T(x) > 0$ the land rent gradient is always negative. When $A(x) - T(x) > 0$ over some interval $[x_1, x_2]$, the land rent gradient can be positive or negative according to the value of $\varepsilon_{U,\omega}(x)$. Since household income increases over $[x_1, x_2]$, commuting costs also increase over this interval. Therefore, the land rent is a priori neither monotonic nor the mirror image of the spatial income distribution. However, $R^*(x)$ is upward sloping when $A(x) - \varepsilon_{U,\omega}(x)T(x) > 0$. In this case, moving toward locations with more amenities ($A(x) > 0$) is sufficient for the land rent to increase. In sum, the interaction between amenities and commuting may give rise to a variety of land rent profiles, which generally differ from that of the location-quality index. What drives the land rent is essentially an empirical issue.

5 The residential equilibrium under Fréchet distributions

To derive testable predictions about the effects of income inequality on the residential equilibrium, we have to determine the explicit form of the income mapping $\omega^*(x) = F^{-1}[G(\Delta)]$. For this, we must specify the distributions F and G. Earning distributions are skewed to the right and the Fréchet distribution is a good candidate to capture this. Equally important, the Fréchet distribution leads to an analytical solution of our model. In the following, we assume that incomes are drawn from a Fréchet distribution with the shape parameter $\gamma_\omega > 0$ and the scale parameter $s_\omega > 0$: $F(\omega) = \exp\left[-((\omega - \underline{\omega})/s_\omega)^{-\gamma_\omega}\right]$ over $[\underline{\omega}, \infty)$. An increase in γ_ω leads to less income inequality. It is analytically convenient to assume the values of Δ are also drawn from a Fréchet distribution with the c.d.f. $G(\Delta) = \exp\left[-((\Delta - \underline{\Delta})/s_\Delta)^{-\gamma_\Delta}\right]$ over $[\underline{\Delta}, \infty)$; the density is denoted $g(\Delta)$. The location-quality index covers a wider range of values when γ_Δ

decreases.[7]

It follows from Proposition 2 that households ranked by decreasing incomes are assigned to locations having a decreasing location-quality index. Hence, the mapping $\omega^*(\Delta)$ can be retrieved from the condition:

$$\int_{\omega}^{\infty} f(\varphi)\mathrm{d}\varphi = 1 - \exp\left[-((\omega - \underline{\omega})/s_\omega)^{-\gamma_\omega}\right] = \int_{\Delta}^{\infty} g(\varphi)\mathrm{d}\varphi = 1 - \exp\left[-((\Delta - \underline{\Delta})/s_\Delta)^{-\gamma_\Delta}\right],$$

which is the counterpart in the Δ-space of (5). Solving the above equation yields the equilibrium mapping $\omega^*(x)$ defined over the x-space:

$$\omega^*(x) = s_\omega \left[\frac{\Delta(x) - \underline{\Delta}}{s_\Delta}\right]^{\frac{\gamma_\Delta}{\gamma_\omega}} + \underline{\omega}. \tag{25}$$

This expression shows that the spatial income distribution, shifted downward by the minimum income $\underline{\omega}$, is a power of the location-quality index, also shifted downward by the minimum value of this index. Observe that what matters in the equilibrium mapping (25) is the ratio $\gamma \equiv \gamma_\Delta/\gamma_\omega$ and not the values of the two shape parameters. From now on, therefore, we work with γ.

It remains to determine the equilibrium land rent. To obtain an analytical solution, we make one more assumption: $\mu = 0$ so that the lot size is fixed and given ($h = 1$). In this case, the utility function (1) becomes $U = a \cdot q$, so that the location-quality index is now given by $\Delta(x) = a(x)t(x)$ while $B^* = 1$. The fixed lot size assumption does not allow one to replicate the fact that the population density is higher in the central city than at the city outskirts. However, as it accounts for the basic trade-off between long/short commutes and low/high land rents, this assumption is widely used in applied models of urban economics.

Setting $\mu = 0$ in (20) yields

$$\frac{\mathrm{d}U^*}{\mathrm{d}\omega} = \Delta(x), \tag{26}$$

which says that, in equilibrium, the marginal utility of income of households at x is equal to the corresponding location-quality index. Using (26), we show in Appendix B.8 that the equilibrium land rent is given by the following expression:

$$R^*(x) = \frac{\Delta(x) - \underline{\Delta}}{\Delta(x)} \cdot t(x) \left\{\frac{s_\omega}{1+\gamma}\left[\frac{\Delta(x) - \underline{\Delta}}{s_\Delta}\right]^{\gamma} + \underline{\omega}\right\} + \frac{a(\underline{x})}{a(x)} R_A, \tag{27}$$

where $\underline{x}$ is the distance at which $\Delta(\underline{x}) = \underline{\Delta}$. Hence, the land rent at x capitalizes a hike or a fall in the corresponding location-quality index. However, as the relationship (25) is non-linear in Δ, the land rent and the location-quality index do not vary proportionally. The expression (27) shows that even in the simple case of a fixed lot size $R^*(x)$ is a *not* a simple function of

[7] Note that we obtain similar expressions with a Pareto distribution. The main difference is that the Fréchet gives us one more degree of freedom than the Pareto in the estimations.

$a(x)$ and $t(x)$. This confirms our point that amenities and commuting interact in fairly complex ways to determine the residential equilibrium.[8]

6 Data

6.1 Datasets

We use several datasets. The first is obtained from *EDM*, a marketing service provider. Data are available at the household level. Using an extensive representative survey of about 10% of the Dutch population, *EDM* gathers information on a number of key household characteristics, such as gross yearly household income (in 7 classes), the average age of the adults in the household, the household size and type (e.g. single, family). Since households may include full-time and/or part-time workers, we calculate the income per hour by dividing the household's gross income by the total number of working hours per year.[9] The survey consists of three waves: $629,511$ observations for 2004-2007, $640,998$ observations for 2008-2012 and $502,514$ observations for 2013-2014. It is not possible to trace households over time, so the data are cross-sectional. The *EDM* dataset also provides detailed geocoded information on locations, which are described by their geographic coordinates. We focus on households whose locations are within 10 km from the city center. For the first wave, we have locations at the postal code level.[10] For the second and third waves, we have information on the exact location of the property in which a particular household lives.

The *EDM* dataset also provides information on (self-reported) housing attributes, such as housing type (e.g. apartment, terraced), the construction decade and whether the house is owned by the occupants. In the Netherlands, about 90% of the properties in the rental sector is public housing. Public housing is rent controlled and there are often long waiting lists for public housing, so households are not entirely free to choose their utility-maximizing location. Therefore, we will focus on owner-occupied housing, which means that we keep about 60% of the data. Although *EDM* provides information on some important housing attributes, it does

[8] We show in Appendix B.9 that (27) reduces to well-known expressions in the case of homogeneous households.

[9] In the Netherlands there are about 260 working days. Dutch people have on average 26.5 vacation days. Hence, the number of effective working weeks is 33.63. The total number of working hours per week in the household is obtained by assuming that a full-time job implies a working week of 38 hours and a part-time job of 24 hours. If people are unemployed or non-employed, we assume that the working hours per week is 38 but will include a dummy for whether people are unemployed or non-employed. Further, in the sensitivity analysis we show that our results are robust to excluding unemployed, non-employed and self-employed people.

[10] Postcode areas are small and usually one side of a street consisting of about 20 households, so that housing is typically very homogeneous within the same postal code.

not provide information on house size. This might be important as richer households are likely to prefer larger properties. We, therefore, use information on the size of all properties from the *GKN* (*Gebouwkenmerken Nederland*) dataset. Using the address information provided in the *EDM* dataset, we obtain the size of each residential property. Since the location is only available at the postal level before 2008, we use the median size of residential properties in the postcode. To link the observations in the *EDM* data to the locations of amenities and jobs we calculate the travel time between any two locations. We obtain information on the street network from *SpinLab*, which provides information on average free-flow speeds per short road segment (the median length of a segment is 96 m), which are usually lower than the speed limit. More information on how we calculate the travel time between observations is provided in Appendix C.1.

Information on land values and lot sizes is not directly available. Therefore, we infer them from data on housing transactions provided by *NVM* (*Dutch Association of Real Estate Agents*). Specifically, we adopt the procedure developed by Rossi-Hansberg *et al.* (2010), which means that we can only use information on residential properties with land. The methods used to calculate land values and lot sizes are described in Appendix C.2. *NVM* contains information on the large majority (about 75%) of owner-occupied house transactions between 2000 and 2015. We know the transaction price, the lot size, inside floor space size (both in square meters), the exact address, and a wide range of housing attributes such as house type, number of rooms, construction year, garden, state of maintenance, and whether a house is equipped with central heating.[11] The construction year controls for a range of house attributes are difficult to observe (e.g. building quality, architectural style). We also know whether the house is a listed building.

We are interested in the impact of amenities and employment accessibility on income sorting and land prices. To this end, we gather data from Eric Fisher's *Geotagger's World Atlas*, which contain all geocoded pictures on the website Flickr. The idea is that locations with an abundant supply of aesthetic amenities will have a high picture density.[12] Using data on the exact location of historic, natural and consumption amenities, we show in Appendix C.3 that there is a strong positive correlation between picture density and historic amenities, open water, the presence of shops, cafés and restaurants, cultural and leisure establishments. There are, however, some

[11] We exclude transactions with prices that are above € 1 million or below € 25,000 and have a price per square meter which is above € 5,000 or below € 500. We furthermore leave out transactions that refer to properties that are larger than 250 square meters, are smaller than 25 square meters, or have lot sizes above 5000 square meters. These selections consist of less than one percent of the data and do not influence our results.

[12] Ahlfeldt (2014) shows that for the cities of Berlin and London the photo density is strongly correlated to the number of restaurants, music nodes, historic amenities and architectural sites, as well as parks and water bodies.

issues with using geocoded pictures as a proxy for amenities.

First, one third of these pictures are automatically geocoded. To avoid the possibility of inaccurate manual geocoding, we delete most of these pictures and keep only one geocoded picture for each location defined by its geographical coordinates. Unlike Ahlfeldt (2014), we do not aggregate locations and use a measure of picture density that is continuous over space.

Second, one may argue that the patterns of pictures taken by tourists and residents may be very different. We make a distinction between residents' and tourists' pictures by keeping users who take pictures for at least 6 consecutive months between 2004 and 2015. Note that the correlation between residents' and tourists' pictures is above 0.8.

Third, many recorded pictures may not be related to amenities but to ordinary events in daily life occurring inside the house. Hence, we only keep pictures that are taken outside buildings, using information on all the buildings in the Netherlands from the *GKN* dataset. Furthermore, if pictures are not related to amenities, one would expect almost a one-to-one relationship with population density. However, if we calculate the population density in the same way as we calculate the amenity level, the correlation is only about 0.5. In the sensitivity analysis we show that using pictures as a proxy for amenities is superior to using population density.

Fourth, one may argue that pictures are disproportionately clustered at public transport nodes because tourists usually arrive there. Therefore, we estimate specifications where we control for the distance to public transport nodes.

Although imperfect, we believe that the picture density is probably the best proxy available for the relative importance of urban amenities at a certain location. Nevertheless, we test the robustness of our results using an alternative hedonic amenity index in the spirit of Lee and Lin (2015) (see Appendix C.3 for more details).

Let $\tilde{a}(x)$ be the observed amenity level at location x in city c as defined in (3):

$$\tilde{a}(x) = \varphi \sum_{z=1}^{Z(c)} e^{-\varphi \tau(x,z)} n(z), \tag{28}$$

where $Z(c)$ is the number of locations z in city c, $n(z)$ the number of pictures taken at z and $\varphi > 0$ a decay parameter. If pictures were uniformly distributed across space, the expected travel time to each picture location would be $E[\tau(x, z)] = 4\pi/\varphi^2$. The burgeoning literature on the economic effects of amenities suggests that the effect of amenities is much more localized than the effects of employment (Legget and Bockstael, 2000; Turner *et al.*, 2014). We, therefore, assume that people take into account amenities within 20 minutes drive. Thus, $E[\tau(x, z)] = 20$ yields $\varphi = 0.793$. In the sensitivity analysis we show that our results are robust for this assumption.

We gather information on job locations from the *GKN* for all commercial and public buildings in the Netherlands in 2015. Using the total building space and based on the average space usage per employee, we come up with a proxy for the number of employees at each location. Since we know the residence but not the workplace of people, we cannot measure $t(x)$ as assumed in Section 3. Rather, we use the *expected* commuting time $\tilde{\tau}(x)$ of a household at x defined as the weighted sum of the commuting times for a given set $N_x(c)$ of employment centers e_i that can be reached within a given travel time from x in city c:

$$\tilde{\tau}(x) \equiv \frac{\sum_{e_i \in N_x(c)} m(e_i)\tau(x, e_i)}{\sum_{e_i \in N_x(c)} m(e_i)}, \tag{29}$$

where $\tau(x, e_i)$ is the travel time between x and the employment location e_i while $m(e_i)$ is the number of jobs at e_i. If all jobs are located in the CBD, then $\tilde{\tau}(x) = \tau(x, e_0)$. Given that the average commute in the Netherlands is about 30 minutes (Department of Transport, Communications and Public Works, 2010), we only include jobs that are within a 45 minutes commute from location x.[13]

We also construct historic instruments. Knol *et al.* (2004) have scanned and digitized maps of land use in 1900 into 50 by 50 meter grids and classified these maps into 10 categories, including built-up areas, water, sand and forest. We aggregate these 10 categories into built-up, open space and water bodies. We also use information on municipal population density in 1900 from *NLGIS*. We proxy the local population in 1900 by using the location of buildings and assuming that the population per building is the same within each municipality. We then calculate the population at each location using a formula in the spirit of (29). Similar instruments based on land use in 1832 obtained from *HISGIS* are also constructed.

For each city, we also digitize hardcopy historic planning maps from around 1930. We calculate the share of residential land, of planned residential development, of commercial land, of planned commercial land, of parks, water bodies and open space within 500 m of each location x. Since the planning maps only partly capture the urban areas we are interested in, in the IV-regressions we keep approximately 60% of the observations. More information, maps of land use patterns in 1900 and historic planning maps are provided in Appendix C.4.

6.2 Descriptive statistics

We report descriptive statistics of the 54,279 observations of our sample in Table 1. We calculate the per hour income by dividing the gross income by the total hours worked per year. The total number ℓ of hours worked for each household are approximated by taking into account

[13] Our results are robust to this assumption, as we will show in the sensitivity analysis.

the fact that members of the household work full-time or part-time.[14] The average per hour wage is € 34.66. The median is about € 6 lower (€ 27.78). About 5% of the households belong to the lowest income class. It appears that the logarithm of incomes is approximately normally distributed.

The average land price in Amsterdam is substantially higher (about € 2000) than in Rotterdam (about € 1000). As expected, the correlation between the estimated land price and lot size is negative ($\rho = -0.598$).

[Table 1 about here]

The amenity index based on pictures range from almost zero to $2,358$. The underlying data on pictures is reported in Table C.3 in Appendix C.3. It is observed that 65% of the pictures are taken in Amsterdam. The average amenity value in Amsterdam (273) is much higher than in Rotterdam (113). Table C.3 also shows that 74% of the pictures are taken outside a building. About 44% of the pictures are taken by local residents. This share is a bit higher in Rotterdam (55%), suggesting that Amsterdam receives more tourists. Recall that we only use pictures *outside* a building taken by *residents* in determining the amenity index. Going back to Table 1, we see that the expected travel time to employment is on average 19 minutes. Given the national average of 30 minutes, and taking into account that commutes in cities are generally shorter, this seems a reasonable value. The unconditional correlation of the overall amenity index with the income level is close to zero ($\rho = 0.0521$), but this is not very informative yet. The correlation of the amenity index with land prices is substantially higher ($\rho = 0.686$). Finally, the locations that are well accessible do not necessarily have a high amenity level, as the correlation between the amenity level and the travel time to employment is only -0.241.

The average house size is 103 m^2. However, in Amsterdam houses are only about 95 m^2, which corresponds to the higher land values and house prices in this city. About 40% of households occupy apartments, which is not too surprising given that we focus on the two largest cities in the Netherlands where land is very expensive. However, the correlation between occupying an apartment and the land price is not very high ($\rho = 0.220$).

The descriptives of the instruments that we will use are described in Table C.5 in Appendix C.4.

[14] In the Netherlands there are about 260 working days. Dutch people have on average 26.5 vacation days. Hence, the number of effective working weeks is 33.63. The total hours worked per week in the household is obtained by assuming that a full-time job implies a working week of 38 hours and a part-time job of 24 hours. If people are unemployed or non-employed, we assume that the working hours per week is 38 but will include a dummy for whether people are unemployed or non-employed. Further, in the sensitivity analysis we show that our results are robust to excluding unemployed, non-employed and self-employed people. We also test robustness to the assumption that wages are continuously distributed.

7 Reduced-form estimation

7.1 Econometric framework

We use the equilibrium mapping (25) to quantify the sorting consequences of the spatial distribution of amenities. We have seen in Sections 4 and 5 that the Stone-Geary utility function and Fréchet distributions are good candidates to estimate the equilibrium spatial income distribution. Therefore, we assume that

$$U = [\tilde{a}(x)]^{\beta} q^{1-\mu} (h - \overline{h})^{\mu}$$

where $\tilde{a}^{\beta}$ is the utility associated with amenities with $\beta > 0$ (the $\tilde{}$ indicates observed values). Furthermore, we assume that labor time is given by $t(x) = \ell[\tilde{\tau}(x)]^{-\theta}$, where $\tilde{\tau}(x)$ is the expected commuting time, $\theta > 0$ is the elasticity of labor time with respect to commuting time and ℓ is the number of working hours per year.[15] Under this configuration, the location-quality index becomes

$$\Delta(x) = [\tilde{a}(x)]^{\beta} [\tilde{\tau}(x)]^{-\theta(1-\mu)}. \tag{30}$$

Assuming for simplicity that $\underline{\omega} = 0$ and $\underline{\Delta} = 0$, it follows from (25) that the income mapping is given by the following expression:

$$\omega^{*}(x) = s_{\omega} \left\{ \frac{[\tilde{a}(x)]^{\beta} [\tilde{\tau}(x)]^{-\theta(1-\mu)}}{s_{\Delta}} \right\}^{\gamma}, \tag{31}$$

where s_{Δ} is the scale parameter of the location quality index distribution and s_{ω} the scale parameter of the income distribution.

Let $\tilde{\omega}(x)$ be the gross hourly income of a household residing at x. We assume that $\omega^{*}(x) \equiv \ell\tilde{\omega}(x)/\xi(x)$ where the $\xi(x)$ are labor income shocks that are independently and identically distributed according to some given distribution defined on $(0, \infty]$. Taking the log of (31), we obtain:

$$\log \tilde{\omega}(x) = \alpha_0 + \alpha_1 \log \tilde{a}(x) + \alpha_2 \log \tilde{\tau}(x) + \tilde{\xi}(x), \tag{32}$$

where $\alpha_0 \equiv \log\left(s_{\omega} / \left[\ell^{(1-\mu)\gamma-1} s_{\Delta}^{\gamma}\right]\right)$, $\alpha_1 \equiv \beta\gamma$, $\alpha_2 \equiv -\theta(1-\mu)\gamma$ and $\tilde{\xi}(x) = \log \xi(x)$.

We are interested in causal estimates of α_1 and α_2. There are several problems associated with identifying these parameters. The first is that the expected commuting time $\tilde{\tau}(x)$ may capture productivity rather than income sorting effects. We do not believe this to be a major problem here because we have information on the income at the residence. Nevertheless, in some specifications, we control for employment density, which is known to be highly correlated with agglomeration economies (Combes *et al.*, 2008). The second concern is that $\tilde{a}(x)$ and $\tilde{\tau}(x)$

[15] Since $\tilde{a}(x)$ and $\tilde{\tau}(x)$ are indices, we can only identify them up to a scaling factor. However, because we take logs, this is not a problem here.

are correlated with $\tilde{\xi}(x)$. Households may not only sort on the basis of income, but also on the basis of other household characteristics. For example, households with children may aim to locate in neighborhoods with a large amount of green space. The variables $\tilde{a}(x)$ and $\tilde{\tau}(x)$ could also be correlated with unobserved housing attributes because households with different incomes may have different preferences for housing quality. This is important as Brueckner and Rosenthal (2009) showed that housing quality is unevenly distributed over space. For example, a large share of the housing stock in the city center of Amsterdam takes the form of apartments. This may imply that affluent households are not willing to locate there because they eschew apartment living (Glaeser *et al.*, 2008). The third issue is reverse causality between $\tilde{\omega}(x)$ and $\tilde{a}(x)$ because the provision of amenities may be a direct result of the presence of high-income households. For example, cultural and leisure services are often abundantly available in upscale neighborhoods (Glaeser *et al.*, 2001).

The first step to mitigate the biases associated with these concerns is to add control variables. Most importantly, we control for *household* characteristics, $D(x)$. For example, we can control for whether people are employed, are full-time or part-time workers, the size of the household and the age of the adults. This reduces the probability that we measure sorting on basis of household characteristics other than incomes. We also control for *housing* attributes, $C(x)$, such as house size, type and construction year.

Furthermore, we estimate specifications where we include *location* attributes, $L(x)$. These attributes capture the characteristics of the local housing stock, such as the share of owner-occupied housing and the mean construction year in the vicinity. More importantly, we also control for accessibility to transit. As shown by Glaeser *et al.* (2008) and Rosenthal and Ross (2015), access to transit matters more to poor households. Transit stations are mainly located close to the city center, which may imply a correlation with $\tilde{a}(x)$ and $\tilde{\tau}(x)$. We, therefore, count the number of train stations, metro stations and bus/tram stops within $0-250$ and $250-500$ m distance bands and include those as separate control variables.

Finally, we include a city dummy ($\eta_1(x)$) and year fixed effects ($\eta_2(y)$). We then estimate:

$$\log\tilde{\omega}(x) = \alpha_0 + \alpha_1\log\tilde{a}(x) + \alpha_2\log\tilde{\tau}(x) + \alpha_3 D(x) + \alpha_4 L(x) + \alpha_5 C(x) + \eta_1(x) + \eta_2(z) + \tilde{\xi}(x),$$

where α_3, α_4, α_5, are additional vectors of the parameters to be estimated.

It is unlikely that working with an endless string of controls will fully address the endogeneity concerns raised above. Our data do not allow us to exploit quasi-experimental or temporal variation in $\tilde{a}(x)$ and $\tilde{\tau}(x)$. In this paper, we exploit the fact that $\tilde{a}(x)$ and $\tilde{\tau}(x)$ are autocorrelated. First, we use land use patterns in 1900 as instruments. We expect aesthetic amenities to be positively correlated to the share of built-up area in 1900. The historic city center of Amsterdam has many buildings that are from (or before) 1900 and are now listed buildings.

Furthermore, we also expect water bodies available in 1900 to be correlated to current water bodies, which are often considered as an amenity (Legget and Bockstael, 2000).

Historic instruments are often criticized because of the strong identifying assumption that past unobserved locational traits are uncorrelated to current unobserved locational endowments. This assumptions is indeed hard to defend when analyzing sorting patterns *between* cities. However, *within* a city this assumption is much less restrictive because the patterns of income sorting within each city have considerably changed throughout the last century. Around 1900, open water and densely built-up areas were not necessarily considered amenities. For example, the canals in Amsterdam were open sewers (Geels, 2006). Local governments made ordinances against this practice, but with little effect. Therefore, locations near a canal often repelled high-income households who located in lush areas just outside the city. It was also before the time when cars became the dominant mode of transport. People around 1900 usually walked to their working place, and thus commuting distances were very short. However, the rich could afford to live outside the city and take the coach to their working place. The cities in 1900 were not yet influenced by (endogenous) planning regulations, as the first comprehensive city plans date from the 1930s. Hence, unobserved reasons that may cause the clustering of high-income people in the past are unlikely to be correlated to current amenities.

The main threat to the validity of the instrument is that built-up areas in 1900 are correlated to current unobservable attributes of the housing stock. To address this issue we estimate specifications where we only use the share of water in 1900 as an instrument and directly control for the share of built-up area in 1900. Furthermore, in the spirit of Arzaghi and Henderson (2008), as an informal test of the validity of the instrument, we include the instruments directly in an OLS specification, together with the endogenous variables. If the impact of the latter variables declines, this may call into question the validity of the historic instruments.

As an alternative, we appeal to land use as stipulated in large scale historic extension plans from the 1930s. Because of the accelerating population growth at that time, many cities called for comprehensive extension plans. The idea behind these plans was to accommodate demand over a long period. For the case of Amsterdam, these plans were expected to lay-out the city until the year 2000. However, these plans were developed at a time where the automobile played only a marginal role in commuting and other daily trips. Therefore, the land use extension plans are unlikely to accommodate unexpected demand shocks that occurred in the decades that followed their design. Large parts of the land use extension plans have been realized (e.g. the whole Western part of Amsterdam). However, because they did not foresee the enormous rise in car use and the large population increase after World War II, another large part of these plans was revised. We use the land use patterns as outlined in the plans to predict the current spatial distribution of jobs and amenities. The exclusion restriction permits that historic land use patterns are uncorrelated to unobserved reasons why households sort, but are correlated to the current level of amenities and jobs. We think this may be a reasonable assumption because the plans were not based on the past or expected income distribution, so that the issue of reverse causality is addressed. However, one may again argue that the unobserved characteristics of the current building stock may be correlated to the land use extension plans. For example, the parts of the city that were already built have other characteristics that attract or repel affluent households. To mitigate this problem we control for the mean construction year per neighborhood. We also estimate specifications where we control for the share of land that was occupied by residential buildings that already exist at the time the plans were laid out and the share of planned residential land use.

We perform a two-stage least squares estimation. We first determine the predicted values $\hat{a}(x)$ and $\hat{\tau}(x)$ by estimating:

$$\{\log \tilde{a}(x), \log \tilde{\tau}(x)\} = \alpha_0^1 + \alpha_1^1 I(x) + \alpha_2^1 D(x) + \alpha_3^1 L(x) + \alpha_4^1 C(x) + \eta_1^1(x) + \eta_2^1(y) + \epsilon(x),$$

and then estimate the income mapping in which we have inserted the predicted values $\hat{a}(x)$ and $\hat{\tau}(x)$:

$$\log \tilde{\omega}(x) = \alpha_0^2 + \alpha_1^2 \log \hat{a}(x) + \alpha_2^2 \log \hat{\tau}(x) + \alpha_3^2 D(x) + \alpha_4^2 C(x) + \alpha_5^2 L(x) + \eta_1^2(x) + \eta_2^2(y) + \lambda_1^2(x) + \lambda_2^2(x) + \tilde{\xi}(x).$$

In line with (22), we also repeat the above specifications when income is replaced by land prices. It follows immediately from (27) that a simple double log equation does not identify any structural parameter of the model. By contrast, we show in Section 8 that combining the income and land price equations allows us to separately identify β, γ and θ and, therefore, to predict $\Delta(x)$.

7.2 Empirical results

Ordinary least squares estimates. We analyze the effects of amenities and commuting time on income sorting. Table 2 reports the results. Column (1) reports a naive regression of log income on log amenities, log commuting time to the CBD and city and year fixed effects. The results show that amenities and commuting time to the city center do not seem to attract high-income households. However, these results may be severely biased because we do not control for household characteristics and housing attributes.

[Table 2 about here]

In column (2), we control for household characteristics to address the issue that sorting may be caused by household composition rather than income. Although the coefficient of amenities is somewhat larger, it is still statistically insignificant. Column (3) investigates the importance of an uneven spatial distribution of housing quality. This issue is important: the results show that incomes are now 2.8% higher when the amenity level doubles. The effect of commuting time to the city center is again statistically insignificant.

Because many cities are characterized by dispersed employment patterns, the accessibility to the city center is likely a very poor proxy for employment accessibility. We, therefore, use the expected commuting time to all jobs within 45 minutes travel time in column (4) instead of commuting time to the CBD. The coefficient related to commuting time is now negative and statistically significant. Doubling the commuting time would decrease incomes by 8.8%. Moreover, doubling the amenity level implies an income increase of 2.9%, which is very similar to the previous specification.[16]

In column (5), we include the share of owner-occupied housing in $0-250$ m and $250-500$ m distance bands to control for potential negative externalities of public housing and for local differences in the supply of owner-occupied housing. We further control for the average construction year of buildings in $0-250$ m and $250-500$ m distance bands, which may pick up differences in housing supply constraints: when the average construction year is high, this indicates that new construction is permitted. Finally, we control for access to transit, by separately controlling for the number of railway stations, metro stations in bus stops in $0-250$ m and $250-500$ m distance bands. Those variables do not seem to be correlated to amenities or the expected commuting time, as the coefficients are essentially identical to the previous specification.

We calculate the expected commuting time and the travel time to amenities using distance over the road network. One may argue that this may introduce measurement error when households are dependent on public transport. Furthermore, for people who are self- or unemployed, commuting costs may be less of an issue, so that the coefficient related to commuting would be downward biased. In column (6) we show that this is not a problem: when we exclude all those households, the results are remarkably similar, although the number of observations is much lower (about 35%). We also control for the impact of employment density to alleviate the concern that productivity effects arising in dense employment areas due to agglomeration economies might be correlated with amenities and expected commuting time. In the spirit of (28), we calculate the employment density assuming a decay parameter that corresponds to an expected travel time of 45 minutes. The effect of amenities and expected commuting time are once again very similar to the previous specifications. As for the employment density, its impact is small, negative and marginally significant.[17] This result is in line with recent empirical

[16] Note that doubling the amenity level is a much more likely event (given a standard deviation of the log amenity level of 1.310) than doubling the expected commuting costs (with a standard deviation of 0.129).

[17] Note that this result also holds for different values of the decay parameter.

evidence, which suggests that, conditional on commuting time, employment accessibility does not have an effect on incomes (Sanchis-Guarner, 2012).

In what follows, we estimate specifications where we use instruments based on the land use stipulated in historic city plans as instruments for the current distribution of amenities and commuting time. However, the area for which we have data is smaller than the 10 km radius from each city center. We, therefore, lose about 40% of the data. To make sure that the potentially different results between 2SLS and OLS are not driven by this sample selection, we repeat the baseline specification in column (5) for the smaller sample. This leads to very similar results (see column (8), Table 2), although the coefficient of amenities is a bit higher compared to the baseline specification. Also, the effect of commuting time is somewhat stronger.

Two-stage least squares estimates. Amenities and job locations may be endogenous due to omitted correlated variables or reverse causality. We address these issues in Table 3 by instrumenting for amenities and commuting time.

[Table 3 about here]

The first set of specifications in Table 3 relies on instruments that are constructed from land use in 1900. The instruments are the shares of water bodies and of built-up area within a distance of 500 m and the travel time to population in 1900.[18] In Table C.7 of Appendix C.4 we report first-stage results. Both the share of built-up area and the share of water bodies in 1900 are strongly and positively correlated to the current amenity level (the Kleibergen-Paap F-statistic is above the rule-of-thumb value of 10 in all specifications).

The second-stage results in column (1), Table 3, reveal that when we only instrument for amenities the impact of both amenities and expected commuting time is very similar to the baseline specification. However, the latter changes once we instrument for the expected commuting time with the travel time from each location to population in 1900 in column (2), Table 3. The coefficient implies that doubling the expected commuting time leads to a decrease in income of 39%, so the effect is much stronger. The results are similar once we instrument for both amenities and expected commuting time in column (3). Doubling amenities implies an income increase of 2.9%, while doubling expected commuting time implies an income decrease of 22%. The main concern regarding the use of historic instruments is that these instruments are correlated to current unobserved building characteristics. To mitigate this issue, in column (4) we directly control for the share of the built-up area within 500 m. The results suggest that doubling amenities increases income by 4.2%. The impact of expected commuting time is now

[18] We have experimented with other thresholds (e.g. 250 m or 1000 m), but this leaves the results essentially unchanged. These results are available upon request.

rather imprecise but the point estimate is similar to the baseline OLS estimate.

Alternatively, to test the robustness of our results for the use of a particular set of instruments, we exploit the information provided by the historic city plans. The instruments are the share of water bodies, parks, existing residential land use in the 1930s and planned residential land use within 500 m, and the travel time to existing and planned built-up land. The first-stage results are reported in Table C.7 of Appendix C.4. The instruments have again the expected signs: the share of open water, parks and existing residential land use are positively correlated with the picture index. The instruments seem to be less strong than the 1900 land use instruments. The average travel time to *existing* built-up land around 1930 is not statistically significant, whereas the travel time to *planned* built-up land is positive and strongly correlated to current expected commuting time.

The second-stage results are very much in line with the previous results: amenities attract higher incomes, while commuting time is negatively related to incomes. In column (5) we instrument for amenities only. Column (6), Table 3, shows that the effect of the expected commuting time is stronger and negative once we instrument for it, in line with the results obtained by using the instruments based on land use in 1900. In column (7) we instrument both for amenities and the expected commuting time, which leads to similar results. We again face the concern that the instruments are correlated with unobserved building attributes. For example, planned residential development may have resulted in currently unattractive buildings in the form of apartments. To mitigate this issue, in the final specification we directly control for the share of existing residential land use and planned residential land use in the city plans within 500 m. Again, the results are very similar and suggest that doubling amenities increases income by 4.8%. The impact of commuting time is somewhat lower than in the previous specification but again rather imprecise. However, we should be careful here because the Kleibergen-Paap F-statistic is low, suggesting that we have a weak instrument problem.

To sum up, we observe that the results unequivocally indicate that *the impact of amenities on income sorting is positive and highly significant.* As for the expected commuting time, *its effect on income sorting is negative but we observe more variability in the estimated coefficients.* There are concerns regarding the causality and whether the two variables we use are good proxies for amenities and commuting costs.

An hedonic approach to amenities. Following Lee and Lin (2015), we construct an aggregate hedonic amenity index that describes the amenity provision at every location using house prices. The procedure is described in Appendix C.3 and the regression results are reported in Table 4. In column (1), we show that this alternative index also has a strong impact on incomes. The effect of the expected commuting time is similar, albeit a bit stronger than in the baseline specification. To make the results comparable to the coefficients estimated for the picture index, we investigate the effect of a standard deviation in the log of the hedonic index, which implies an increase in incomes of 15.5%. A standard deviation in the log picture density implies a 5.2% increase in income in the baseline specification. The effect of picture density thus seems to be lower, implying that our estimates may be underestimates. In the next column, we include both the picture density and the hedonic amenity index. The effect of the picture density is very similar to the baseline specification, while the impact of the hedonic amenity index is now much lower but still statistically significant.

In column (3) we instrument for the hedonic amenity index and the expected commuting time. The impact of the hedonic amenity index becomes substantially stronger compared to the coefficient reported in column (1). When we include both the amenity index based on pictures and the hedonic amenity index and instrument them both, as well as the expected commuting time, a striking observation is that the picture index is highly significant and on a same order of magnitude as the baseline specification, while the hedonic amenity index is far from being statistically significant. This suggests that the picture index is superior to the hedonic amenity index, although we acknowledge that no proxy is perfectly correlated to actual amenities. Repeating this exercise in which we use historic city plan instruments leads to exactly the same conclusion (see columns (5) and (6), Table 4).

Sensitivity checks. We show in Appendix C.5 that our results still hold for a wide range of robustness checks. First, we add the instruments based on land use in 1900 as controls in an OLS specification. If the instruments are invalid and correlated to the error term, the impact of amenities and expected commuting times should decline because the instruments absorb part of the bias of the endogenous variables (Arzaghi and Henderson, 2008). We show that the coefficients of interest are not statistically different from the corresponding OLS specification. Second, using all our instruments and controlling for the current pattern of open spaces and water bodies lead to similar results. We also make sure that the results are not mainly driven by observations in historic districts, e.g. the canal district in Amsterdam.

Third, historic instruments cease to be valid when the unobserved characteristics of a location or building in the past are correlated with those in present time. By going back further in time, we may end up with weak instruments because the correlation between historic land use and current amenities and job locations will also be lower. Nevertheless, we exploit land use data from the census in 1832 because the data are 175 years before the sample period. The point estimates are again very similar but somewhat imprecise. Fourth, we use alternative values for the decay parameter and recognize the discreteness of self-reported incomes by estimating interval regressions. Finally, we estimate separate regressions for Amsterdam and

Rotterdam. We find that amenities have an impact on income sorting in both cities. The evidence for expected commuting costs is in line with the baseline outcomes for Amsterdam, while the evidence is somewhat mixed for Rotterdam. More tests and details are available in Appendix C.6.

Overall, *the impact of amenities is very robust.* The effect of commuting is negative when properly instrumenting for it, in line with the baseline specifications reported in Tables 2 and 3. However, there is more variability in the corresponding estimated coefficients.

Effects on land prices. Because land prices play a crucial role in the location decisions of households and because we postulated that the signs of the effects of amenities and expected commuting time on land prices and incomes should be the same, we estimate the effects of amenities and expected commuting time on land prices. We start in column (1), Table 5, with a simple specification including amenities and accessibility to the city center. This leads to a strong positive effect of amenities on the land price: doubling amenities implies a land price increase of 11.3%. Doubling the commuting time to the city center implies a price decrease of 12.1%. When we include housing attributes, the results are essentially identical (column (2)).

In column (3) we switch to expected commuting time, instead of the commuting time to the city center only. The results indicate that the effect of amenities is stronger, while the effect of commuting is now statistically insignificant. This is also the case if we include additional location attributes in column (4), Table 5. In columns (5) and (6), we instrument for amenities and expected commuting time with instruments based on, respectively, land use in 1900 and historic city plans. The results confirm that amenities are very important, while commuting costs matter less: doubling the amenity levels implies land price increases of 19.2 and 27.2%, respectively, while the coefficients of expected commuting costs are statistically insignificant.

8 Counterfactual analysis

8.1 Structural estimation

What happens to the residential equilibrium when the distribution of the location-quality index within the city changes? We cannot separately identify the parameters β, γ and θ by using the sole income mapping (32). Therefore, we use land prices to estimate those parameters, which we then use to determine the value of $\Delta(x)$ at each location.

Using $t(x) = [\tilde{\tau}(x)]^{-\theta}$, we can rewrite (23) as follows:

$$e(x) \equiv \frac{R^*(x)\left[h^*(x) - (1-\mu)\overline{h}\right]}{\mu\omega^*(x)} = \ell\tilde{\tau}(x)^{-\theta}, \tag{33}$$

which we refer to as the housing expenditure share. Since we observe income and have estimations of land prices and lot sizes, we may calculate $e(x)$ for given values of $\overline{h}$ and μ. We assume that $\overline{h} = 15$.[19] Since we are not able to identify μ, we must pick a particular value. Dutch households spend about one-third of their income on housing. According to Albouy *et al.* (2015), the income elasticity of housing demand is near two-thirds (from their preferred estimates). Therefore, we guesstimate the actual value of μ by using the income elasticity of (17):

$$\mu = \frac{R^*(x)h^*(x)}{\omega^*(x)t(x)} \times \varepsilon_{H,\omega t} \approx 1/3 \times 2/3 = 2/9.$$

Finally, the observed housing expenditure share is given by $\tilde{e}(x)/\upsilon(x)$, where $\upsilon(x)$ are shocks that are independently and identically distributed according to a distribution defined on $(0, \infty]$.

Taking the log of (33), we obtain:

$$\log \tilde{e}(x) = \zeta_0 + \zeta_1 \log \tilde{\tau}(x) + \tilde{\upsilon}(x), \tag{34}$$

where $\zeta_0 = \log \ell$, $\zeta_1 = -\theta$ and $\tilde{\upsilon}(x) = \log \upsilon(x)$. We will further include the controls $D(x)$, $C(x)$ and $L(x)$ as well as year fixed effects in this specifications.

We first estimate (34) to obtain $\hat{\theta}$ and, then, estimate (32) to obtain $\hat{\beta}$ and $\hat{\gamma}$:

$$\hat{\theta} = -\hat{\zeta}_1 \qquad \hat{\beta} = \frac{\hat{\alpha}_1(1-\mu)\hat{\zeta}_1}{\hat{\alpha}_2} \qquad \hat{\gamma} = \frac{\hat{\alpha}_2}{(1-\mu)\hat{\zeta}_1}.$$

Finally, using $\hat{\beta}$, $\hat{\gamma}$ and $\hat{\theta}$, we can estimate (30). By fitting a Fréchet distribution to $\hat{\Delta}(x)$, we obtain $\hat{s}_\Delta$ and $\hat{\gamma}_\Delta$. Using $\hat{\gamma}_\Delta$ and $\hat{\gamma}$, we estimate $\hat{\gamma}_\omega = \hat{\gamma}_\Delta/\hat{\gamma}$. Since $\tilde{a}(x)$ and $\tilde{\tau}(x)$ are indices we cannot find s_ω, implying that we identify the effects on incomes up to a proportionality constant. We have chosen to guesstimate s_ω by fitting a Fréchet distribution to the observed distribution of hourly incomes. Because $\hat{\beta}$ and $\hat{\gamma}$ are nonlinear combinations of estimated parameters, the standard errors are likely larger than for $\hat{\theta}$, which is only a function of $\hat{\zeta}_1$.

8.2 Experiments

In what follows, we undertake three counterfactual scenarios in which the values of the income parameters $\hat{\gamma}_\omega$ and $\hat{s}_\omega$ are the same. The first two scenarios dismiss spatial differences in either amenities or commuting costs. In the first , we assume that the amenity level is the same across the city. This implies that $\hat{\gamma}_\Delta$, whence $\hat{\gamma}$, becomes larger because there will be less variation in $\Delta(x)$. The second scenario is the mirror image of the first. More specifically, we predict the income assignment when commuting costs are the same regardless of households' locations.

[19] We assume $\overline{h} = 15$ because the legal minimum size of a self-contained apartment is 15 square meters in the Netherlands. This is likely on the high side, as the minimum lot required may be lower in the case of apartment buildings.

Our third scenario is very different in that we construct a new amenity index with $n > 2$ extrema, which has the same c.d.f. as $\tilde{a}(x)$.[20] This involves three steps. First, we normalize the values of the amenity distribution as follows:

$$\Omega(x) \equiv 2\frac{\tilde{a}(x) - \underline{a}}{\overline{a} - \underline{a}} - 1,$$

to map $[\underline{a}, \overline{a}]$ onto $[-1, 1]$. Second, we rank the values of $\tilde{a}(x)$ and $\Omega(x)$ by increasing order. Third, we set

$$\Lambda(x) \equiv \sin\left(\frac{n}{B^*}\pi x\right), \tag{35}$$

which maps $[0, B^*]$ onto $[-1, 1]$. The adjacent extrema of $\Lambda(x)$ are thus separated by a distance equal to B^*/n.

Let $X \equiv \{x_1, ..., x_m\} \subset [0, B^*]$ with $x_1 < x_2 < ... < x_m$ be the set of observations The highest value of $\tilde{a}(\cdot)$ over X is reached at the location $x_j \in X$, which determines the highest value of $\Omega(x_j)$ over X. We assign the value $\tilde{a}(x_j)$ to the solutions of the equation $\Lambda(x) = \Omega(x_j)$ in $[0, B^*]$. Since X contains a finite number of locations, we may select the second highest value of $\tilde{a}(x)$ over X and proceed in a similar way until x_m. By smoothing out the so-obtained pointwise curve defined on X, we obtain the *counterfactual amenity distribution* defined over $[0, B^*]$. We simulate results for $n = 6$ and $B^* = 10$.

We report the results of the estimations in Panel A of Table 6. In column (1)-(3), Table 6, we estimate the various parameters using the full sample, while columns (4)-(6) and (7)-(9) refer to Amsterdam and Rotterdam, respectively. In column (1) of Table 7 we find that $\hat{\beta}$ is 0.039. Because $\hat{\gamma}$ is close to one, the estimated preference parameter is similar to the reduced-form elasticities obtained in Tables 2 and 3, while $\hat{\theta}$ is equal to 0.154. Hence, the elasticity of the number of effective units of labor with respect to commuting time seems to be substantially lower than 1. Note, however, that the standard errors are quite large. In column (2) we instrument for amenities and the expected commuting time with land use in 1900. The main difference with previous estimates is that the elasticity of the number of effective units of labor with respect to commuting time is now substantially higher. By contrast, $\hat{\beta}$ and $\hat{\gamma}$ are comparable to the previous specifications. In column (3) we use instruments based on historic city plans. Since we have fewer observations, the estimates are less precise. In particular, $\hat{\beta}$ may seem higher but the estimation is not statistically significantly different from the OLS estimates. Observe that $\hat{\theta}$ is now not statistically significantly different from 1.

In columns (4)-(6) we repeat the same set of specification for the city of Amsterdam only. The estimates are similar, but they become imprecise once we instrument for amenities and expected commuting costs (see columns (5) and (6), Table 6). Columns (7)-(9) focus on Rotterdam. $\hat{\theta}$ is

[20] This ensures that the support of the predicted income distribution will be essentially the same as the baseline income mapping.

similar to the baseline estimate, but it is very imprecisely estimated. Note also that $\hat{\beta}$ and $\hat{\gamma}$ have large confidence intervals. This becomes even worse once we use 1900 land use instruments. This should not come as a surprise: the city that existed in 1900 was approximately the same part of the city that was bombed during World War II. Because the city was completely restructured after the bombing, land use in 1900 bears too low a correlation with current land use and amenities, implying that we have a weak instrument problem. For the historic city plan instruments, we do not have that problem because we have ample information on land use outside the bombed area. The results are remarkably similar to the specifications based on the full sample: $\hat{\theta}$ is not statistically significantly different from 1 at conventional significance levels while $\hat{\beta}$ is around 0.1.

We use the estimated parameters to construct $\hat{\omega}(x)$ for the three scenarios. In Panel B of Table 6 we check that for each scenario the support of the counterfactual income mapping is (almost) the same as in the baseline scenario. **This seems to be the case: for each scenario and each specification the 5th and 95th percentiles and the median are comparable.** The income difference between the 5th and 95th percentile generated by amenities and commuting is about € 4.50 (19% of the median). When using instruments based on land use in 1900 or historic city plans, the income differences between the 5th and 95th percentiles are € 5.14 (21%) and € 6.91 (28%), respectively.

We then plot predicted income per hour for both cities as a function of distance to the city center. The results corresponding to the OLS specifications are depicted in Panel A of Figure 4 where the bold line corresponds to the predicted income mapping for the baseline scenario. In this case, there is almost perfect sorting in distance to the city center.[21] In the first scenario, the predicted income mapping is almost flat (the correlation with predicted incomes in the baseline scenario is still 0.422). That is not to say that households are indifferent to their locations, but that is to say that the distance to the city center remains a poor predictor of the social city structure. In other words, at least for Dutch cities, using the featureless monocentric city model of urban economics would lead to very poor predictions.

[Figure 4 about here]

In the second scenario, we investigate what happens when we assume equal commuting costs across the city. The dotted line shows that the predicted income mapping is almost the same. The correlation of predicted incomes with the mapping in the baseline scenario is 0.997, much higher than for the first scenario. In other words, the lion's share of the spatial variation

[21] Since the city is asymmetric, the income gradient may be different for different radii (see Section 2). In Appendix C.6 we report the predicted location-quality index. Because they look very similar to the predicted income mappings, we only report the latter here.

in incomes is due to amenities, rather than commuting costs, which highlights the contribution of our model.

In the third scenario, we assume that amenities are unevenly distributed across space according to (35). The dashed-dotted line shows the predicted income mapping under the counterfactual amenity function. This confirms again that amenities, rather than commuting costs, are the main driver behind the sorting of households within the city. Comparing the bold and dashed-dotted lines shows that *providing amenities very unevenly across the city leads to the splitting of the affluent into spatially separated neighborhoods that need not be close to the city center or the city limit.*

In Panels B and C of Figure 4 we instrument for amenities and expected commuting costs with land use in 1900 and land use in historic city plans, respectively. It is straightforward to see that results are very similar. In line with the results reported in Table 3, the impact of expected commuting costs is a bit stronger: the correlation between the predicted assignment when amenities are uniformly distributed over space and the baseline mapping is around 0.5. However, the correlation of the predicted income mapping with the baseline income assignment when commuting costs are location-independent within the city is 0.99. Hence, the conclusions that amenities are the main reason for the sorting of the rich and the poor within the city is unchanged.

In Figure 5, we investigate whether the results differ for Amsterdam and Rotterdam. The results look very similar. In Amsterdam, the location-quality index causes income disparities up to € 7.29 (29%). For Rotterdam income differences due to amenities and commuting costs are smaller (€ 3.30, 14%), partly because amenity differences within the city are smaller. Commuting costs do not seem to play any role in Rotterdam, probably because Rotterdam's inner city is easily accessible by car.

[Figure 5 about here]

9 Concluding remarks

It is well documented that amenities have a significant impact on the distribution of households *between* cities. We, therefore, find it reasonable to expect that amenities also affect the distribution of households *within* cities. More specifically, we have shown that the attractiveness of a location depends on the interplay between the overall benefit generated by the amenity field at any particular location and the distance to the nearest employment center. In the case of Dutch cities at least, the former effect seems to overcome the latter, suggesting that using a monocentric city model to predict the social structure of cities is not warranted.

The existence of persistent neighborhood disparities in laissez-faire cities is often viewed as a source of strong social tensions within the city. Residential segregation seems to generate negative and persistent effects on individual development, thus undermining the social fabric of the city (Durlauf, 2004; Topa and Zenou, 2015; Chetty *et al.*, 2016). For this reason, we find it important to study the sorting of heterogeneous households *within* cities. Our analysis suggests that providing amenities at specific locations may foster a break-down of the spatial income gap into several sections. We also believe that such an effort is warranted if policy actions against spatial segregation are taken to strengthen social cohesion within the city. By shifting our attention from "featureless" to "featureful" cities, we are able to understand how amenities impair perfect sorting. This in turn suggests how governments and urban planners can design policies whose aim is to redraw the social map of cities. Contrary to the general belief, we show that *promoting equal access to amenities and public services favors residential segregation by income*, whereas *the uneven provision of amenities across the city fosters income mixing*. In other words, improving the provision of public services in deprived neighborhoods can attract higher-income households and reduce the social fragmentation of cities. However, to be effective, such policies must create long-lasting change in poor neighborhoods. To be sure, pro-income mixing policies are not sufficient to defeat the negative consequences of residential segregation, but they are likely to alleviate them.

References

[1] Ahlfeldt, G.M. (2016) Urbanity. London School of Economics. *Memo.*

[2] Albouy, D. (2016) What are cities worth? Land rents, local productivity, and the total value of amenities. *Review of Economics and Statistics* 98: 477 – 87.

[3] Albouy, D., G. Ehrlich and Y. Liu (2015) Housing demand and expenditures: How rising rent levels affect behavior and costs-of-living over space and time. *Memo.*

[4] Arzaghi, M. and J.V. Henderson (2008) Networking off Madison Avenue. *Review of Economic Studies* 75: 1011 – 38.

[5] Bairoch, P. (1988) *Cities and Economic Development: From the Dawn of History to the Present.* Chicago: University of Chicago Press.

[6] Bakens, J., R. Florax and P. Mulderc (2016) Ethnic drift and white flight: A gravity model of neighborhood formation. Tinbergen Institute Discussion Paper 2016-062/VIII.

[7] Bayer, P., F. Ferreira and R. McMillan (2007) A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy* 115: 588 – 638.

[8] Baum-Snow, N. (2007) Did highways cause suburbanization? *Quarterly Journal of Economics* 122: 775 – 805.

[9] Beckmann, M.J. (1969) On the distribution of urban rent and residential density. *Journal of Economic Theory* 1: 60 – 7.

[10] Behrens, K., G. Duranton and F. Robert-Nicoud (2013) Household sorting in the city. *Memo.*

[11] Behrens, K., G. Duranton and F. Robert-Nicoud (2014) Productive cities: Sorting, selection, and agglomeration. *Journal of Political Economy* 122: 507 – 53.

[12] Bischoff, K. and S. Reardon (2011) Growth in the residential segregation of families by income, 1970-2009. *American Journal of Sociology* 116: 1092 – 153.

[13] Brueckner, J., S.S. Rosenthal (2009) Gentrification and neighbourhood cycles: will America's future downtowns be rich? *Review of Economics and Statistics* 91: 725 – 43.

[14] Brueckner, J., J.-F. Thisse and Y. Zenou (1999) Why is Central Paris rich and Downtown Detroit poor? An amenity-based theory. *European Economic Review* 43: 91 – 107.

[15] Chetty, R., N. Hendren, and L.F. Katz (2016) The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review* 106: 855 – 902.

[16] Combes, P.P., G. Duranton and L. Gobillon (2008) Spatial wage disparities: Sorting matters! *Journal of Urban Economics* 63: 723 – 42.

[17] Couture, V. and J. Handbury (2016) Urban revival in America, 2000 to 2010. University of Pennsylvania. *Memo.*

[18] Davis, M. and F. Ortalo-Magné (2011) Household expenditures, wages, rents. *Review of Economic Dynamics* 14: 248 – 61.

[19] de Bartolome, C.A.M. and S.L. Ross (2007) Community income distributions in a metropolitan area. *Journal of Urban Economics* 61: 496 – 518.

[20] De Moel, H., J.C. Aerts and E. Koomen (2011) Development of flood exposure in the Netherlands during the 20th and 21st century. *Global Environmental Change* 21: 620 – 27.

[21] Department of Transport, Communications and Public Works (2010). *Mobiliteitsbalans 2010.* The Hague: Kennisinstituut voor Mobiliteit.

[22] Diamond, R. (2016) The determinants and welfare implications of US workers diverging location choices by skill: 1980-2000. *American Economic Review* 106: 479 – 524.

[23] Duranton, G. and D. Puga (2015) Urban land use. In G. Duranton, V. Henderson and W. Strange (eds), *Handbook of Regional and Urban Economics, Volume 5*. Amsterdam: Elsevier, 2063 – 117.

[24] Durlauf, S. (2004) Neighborhood effects. In: J.V. Vernon and J.-F. Thisse (eds.) *Handbook of Urban and Regional Economics, Volume 4*. Amsterdam: North-Holland, 2173 – 242.

[25] Eeckhout, J., R. Pinheiro and K. Schmidheiny (2014) Spatial sorting. *Journal of Political Economy* 122: 554 – 620.

[26] Epple, D. and Nechyba (2004) Fiscal decentralization. In V. Henderson and J.-F. Thisse (eds), *Handbook of Regional and Urban Economics. Volume 4.* Amsterdam: Elsevier, 561 – 624.Epple, D. and H. Sieg (1999) Estimating equilibrium models of local jurisdictions. *Journal of Political Economy* 107: 645 – 81.

[27] Epple, D., B. Gordon and H. Sieg (2010) Drs. Muth and Mills meet Dr. Tiebout: Integrating location-specific amenities into multi-community equilibrium models. *Journal of Regional Science* 50: 381 – 400.

[28] Fajgelbaum, P., G.M. Grossman and E. Helpman (2011) Income distribution, product quality, and international trade. *Journal of Political Economy* 119: 721 – 65.

[29] Floch, J.-M. (2014) Des revenus élevés et en plus forte hausse dans les couronnes des grandes aires urbaines. France, portrait social, INSEE.

[30] Fujita, M. (1989) *Urban Economic Theory.* Cambridge: Cambridge University Press.

[31] Fujita, M. and H. Ogawa (1982) Multiple equilibria and structural transition of non-monocentric urban configurations. *Regional Science and Urban Economics* 12: 161 – 96.

[32] Geels, F. W. (2006) The hygienic transition from cesspools to sewer systems (1840–1930): The dynamics of regime transformation. *Research Policy* 35: 1069 – 82.

[33] Glaeser, E.L., J. Kolko and A. Saiz (2001) Consumer city. *Journal of Economic Geography* 1: 27 – 50.

[34] Glaeser, E., M. Kahn and J. Rappaport (2008) Why do the poor live in cities? *Journal of Urban Economics* 63: 1 – 24.

[35] Hartwick, J.M., U. Schweizer and P.V. Varaiya (1976) Comparative statics of a residential economy with several classes. *Journal of Economic Theory* 13: 396 – 413.

[36] Henderson, J.V. and J.-F. Thisse (2001) On strategic community development. *Journal of Political Economy* 109: 546 – 69.

[37] Kamecke, U. (1993) Mean city - A consistent approximation of bid rent equilibria. *Journal of Urban Economics* 33: 48 – 67.

[38] Knol, W. C., H. Kramer and H. Gijsbertse (2004) *Historisch grondgebruik Nederland: een landelijke reconstructie van het grondgebruik rond 1900*. Wageningen: Alterra.

[39] Koster, P.R. and H.R.A. Koster (2015) Commuters' preferences for fast and reliable travel: A semiparametric estimation approach. *Transportation Research Part B: Methodological* 81: 289 – 301.

[40] Koster, H.R.A. and J. Rouwendal (2017) Historic amenities and housing externalities. Evidence from the Netherlands. *Economic Journal*, forthcoming.

[41] Koster, H.R.A., J.N. van Ommeren and P. Rietveld (2016) Historic amenities, income and sorting of households. *Journal of Economic Geography* 16: 203 – 36.

[42] Koomen, E., J. Dekkers and T. Van Dijk (2008) Open-space preservation in the Netherlands: Planning, practice and prospects. *Land Use Policy* 25: 361 – 77.

[43] Landvoigt, T., M. Piazzesi and M. Schneider (2015) The housing market(s) of San Diego. *American Economic Review* 105: 1371 – 407.

[44] Lee, S. and J. Lin (2015) Natural amenities, neighborhood dynamics, and persistence in the spatial distribution of income. Working Papers 15-46, Federal Reserve Bank of Philadelphia.

[45] Leggett, C.G. and N.E. Bockstael (2000) Evidence of the effects of water quality on residential land prices. *Journal of Environmental Economics and Management* 39: 121 – 44.

[46] Legros, P. and A.F. Newman (2002) Monotone matching in perfect and imperfect worlds. *Review of Economic Studies* 69: 925 – 42.

[47] Lucas, R.E. and E. Rossi–Hansberg (2002) On the internal structure of cities. *Econometrica* 70: 1445 – 76.

[48] Määttänen, N. and M. Terviö (2014) Income distribution and housing prices: An assignment model approach. *Journal of Economic Theory* 151: 381 – 410.

[49] Miyake, M. (2003) Precise computation of a competitive equilibrium of the discrete land market model. *Regional Science and Urban Economics* 33: 721 – 43.

[50] Racine, J.S., J. Hart and Q. Li (2006) Testing the significance of categorical predictor variables in nonparametric regression models. *Econometric Reviews* 25: 523 – 44.

[51] Ritzen, J.M.M., J. van Dommelen and F.J. de Vijlder (1997) School finance and school choice in the Netherlands. *Economics of Education Review* 16: 329 – 35.

[52] Rosenthal, S. and S.L. Ross (2015) Change and persistence in the economic status of neighborhoods and cities. In G. Duranton, V. Henderson and W. Strange (eds), *Handbook of Regional and Urban Economics, Volume 5*. Amsterdam: Elsevier Publisher.

[53] Rossi-Hansberg, E., P.D. Sarte and R. Owens III (2010) Housing externalities. *Journal of Political Economy* 118: 485 – 535.

[54] Sanchis-Guarner, R. (2012) Driving up wages: The effects of road construction in Great Britain. *LSE-SERC Discussion Paper* 120.

[55] Sattinger, M. (1993) Assignment models of the distribution of earnings. *Journal of Economic Literature* 31: 831 – 80.

[56] Schmidheiny, K. (2006) Income segregation from local income taxation when households differ in both preferences and incomes. *Regional Science and Urban Economics* 36: 270 – 99.

[57] Topa, G. and Y. Zenou (2015) Neighborhood versus network effects. In G. Duranton, V. Henderson and W. Strange (eds.) *Handbook of Regional and Urban Economics. Volume 5.* Amsterdam: Elsevier, 561 – 624.

[58] Turner, M.A., A. Haughwout and W. Van der Klaauw (2014) Land use regulation and welfare. *Econometrica* 82: 1341 – 403.

[59] van Ommeren, J.N. and E. Gutiérrez-i-Puigarnau (2011) Are workers with a long commute less productive? An empirical analysis of absenteeism. *Regional Science and Urban Economics* 41: 1 – 8.

[60] Waterdienst (2008) *R.W.S. Waterdienst, Risicokaart voor Overstromingen*. Lelystad, The Netherlands: Rijkswaterstaat Waterdienst.

[61] Wheaton, W.C. (1977) Income and urban residence: An analysis of consumer demand for location. *American Economic Review* 67: 620 – 31.

Appendix A

A.1 Descriptive evidence for London

In Section 2, we have presented evidence that urban gradients may be non-monotone. Using ancillary data on London we show that this is not a particular feature of Dutch cities, but a more general observation. We gather data on incomes from the London Datastore, created by the Greater London Authority, at the *Lower Super Output Area* (LSOA) level. Information on all housing transactions is obtained from the Land Registry. Furthermore, we calculate the picture index in the same way as we did for Amsterdam and Rotterdam. Finally, commuting data is compiled from census data in 2001.[22]

We calculate the gradients along the Thames and only include LSOAs within 1 km of the river. The left panel of Figure A.1 shows that incomes and house prices follow more or less the same pattern. Surprisingly, the income and house prices are somewhat lower close the center, which could be explained by the historically abundant provision of council (public) housing near the center. Beyond 10 km, the income tends to decrease relatively less with distance than house prices.

In the right panel of Figure A.1 we show that amenities are more or less monotonically decreasing in distance to the city center. However, over some segments (between 12.5 and 20 km), the amenity distribution is essentially flat. Commuting costs are generally also increasing in distance to the city center, but decrease beyond 20 km from the city center. Hence, these results show that in London gradients may be non-monotonic. However, because we lack detailed micro-data for London, in our empirical work we focus on Amsterdam and Rotterdam.

[Figure A.1 about here]

A.2 Maps for Amsterdam and Rotterdam

In Figure A.2 we plot the spatial income distribution for Amsterdam and Rotterdam. Clearly, they are non-symmetric in distance to the city center. Mainly neighborhoods situated to the south of the Amsterdam city center are occupied by the affluent. The opposite holds for Rotterdam, where mainly the north(east) attracts rich households. Figure A.3 displays land prices in both cities. Expensive land is much more concentrated in the city center than high-income

[22] Commuting data was unofrtunately not available at a fine spatial level for the 2011 census.

households. Hence, high-income households do not necessarily live in the most expensive locations. As mentioned earlier, the correlation between amenities and land values is quite high; likewise, amenities are strongly concentrated in the city centers (Figure A.4). In Amsterdam, they seem a bit more spread than in Rotterdam, which is in line with the observation that Amsterdam receives many more tourists than Rotterdam (approximately 9 million versus 1 million) whose aim to enjoy aesthetic amenities. In Figure A.5 we display the expected commuting costs. Not surprisingly, the commuting costs are lower in the city center, in particular near the south-western section of the highway ring. For Rotterdam, mainly locations close to the eastern section of the ring command lower commuting costs.

[Figures A.2-A.5 about here]

Appendix B

B.1 The cross-derivative of the bid rent function

Differentiating (10) with respect to ω and using (9), we obtain:

$$\Psi_\omega(x,\omega,U^*(\omega)) = \frac{t}{H}\left(1 - \frac{Q_U}{t}U^*_\omega\right) \tag{B.1.1}$$

which is equal to 0 if

$$U^*_\omega = \frac{t}{Q_U}. \tag{B.1.2}$$

Differentiating (B.1.1) with respect to x, using the envelop theorem and rearranging terms yields the following expression:

$$\begin{aligned}\Psi_{x\omega}(x,\omega,U^*(\omega)) \;=\;& \frac{t}{H}\left\{\frac{t_x}{t}\left[1 - \frac{\omega}{H}(H_\omega + H_U U^*_\omega)\right]\right. \\ & \left. + \frac{a_x}{t}\left[\frac{H_\omega + H_U U^*_\omega}{H}Q_a - (Q_{aH}(H_\omega + H_U U^*_\omega) + Q_{aU}U^*_\omega)\right]\right\}. \end{aligned} \tag{B.1.3}$$

Since Q is the solution to the equation $u(q,h) = U/a(x)$, the following expressions must hold:

$$\begin{aligned} Q_a &= -\frac{U}{a^2 u_q} \\ Q_{aU} &= -\frac{1}{a^2 u_q} + \frac{U}{a^2 u_q^2} u_{qq} Q_U \\ Q_{aH} &= \frac{U}{a^2 u_q^2}\left(u_{qq} Q_H + u_{qh}\right). \end{aligned}$$

Assume that the ω-households are located at x. Differentiating $u = U^*(\omega)/a$ with respect to ω and using the budget constraint $Q = \omega t(x) - H\Psi$ and (11), we obtain:

$$[t - (H_\omega + H_U U^*_\omega)\Psi]u_q + (H_\omega + H_U U^*_\omega)u_h = \frac{U^*_\omega}{a}.$$

Since

$$-u_q\Psi + u_h = 0$$

at the residential equilibrium, we have:

$$t = \frac{U^*_\omega}{au_q}.$$

Plugging this expression, Q_a, Q_{aU} and Q_{aH} in (B.1.3), we get

$$\begin{aligned}\Psi_{x\omega}(x,\omega,U^*(\omega)) \quad = \quad & \frac{t}{H}\left\{\frac{t_x}{t}\left[1 - \frac{\omega}{H}(H_\omega + H_U U^*_\omega)\right]\right. \\ & -\frac{a_x}{U^*_\omega}au_q\frac{H_\omega + H_U U^*_\omega}{H}\frac{U^*}{a^2u_q} \\ & \left. +\frac{a_x}{U^*_\omega}au_q - \left[\frac{U}{a^2u_q^2}\left(u_{qq}Q_H + u_{qh}\right)(H_\omega + H_U U^*_\omega) - \frac{U_\omega}{a^2u_q} + \frac{U}{a^2u_q^2}u_{qq}Q_U U^*_\omega\right]\right\},\end{aligned}$$

which is equivalent to

$$\begin{aligned}\Psi_{x\omega}(x,\omega,U^*(\omega)) \quad = \quad & \frac{t}{H}\left\{-T(x)\left[1 - \frac{\omega}{H}(H_\omega + H_U U^*_\omega)\right]\right. \\ & \left. +A(x)\left[1 - \frac{H_\omega + H_U U^*_\omega}{\omega H}\frac{\omega U}{U^*_\omega} - \frac{\omega U}{u_q\omega U^*_\omega}\left(u_{qq}Q_H + u_{qh}\right)(H_\omega + H_U U^*_\omega) - \frac{U}{u^*_q}u_{qq}Q_U\right]\right\}.\end{aligned}$$

Using

$$\frac{\partial u_q}{\partial \omega} = u_{qq}Q_H\left(H_\omega + H_U U^*_\omega\right) + u_{qq}Q_U U^*_\omega + u_{qh}\left(H_\omega + H_U U^*_\omega\right),$$

we can rewrite $\Psi_{x\omega}$ as follows:

$$\Psi_{x\omega}(x,\omega,U^*(\omega)) = \frac{t}{H}\left[\left(1 - \frac{\varepsilon_{H,\omega} + \varepsilon_{u_q,\omega}}{\varepsilon_{U,\omega}}\right)A - \left(1 - \varepsilon_{H,\omega}\right)T\right],$$

which proves Proposition 1.

B.2 Income-independent commuting costs

If commuting costs are given by $\omega\tau(x) + c(x)$, the bid rent function $\Psi(x,\omega;U)$ is given by

$$\Psi(x,\omega,U) \equiv \max_h \frac{\omega t(x) - c(x) - Q(h, U/a(x))}{h}.$$

The utility-maximizing condition implies that the bid rent may be rewritten as follows:

$$\Psi(x,\omega,U) \equiv \frac{\omega t - c - Q}{H}.$$

Consequently,

$$\Psi_x(x,\omega,U^*(\omega)) = \frac{1}{H}\left(\omega t_x - c_x + \frac{a_x}{a}\frac{u}{u_q}\right).$$

and

$$\Psi_{x\omega}(x,\omega,U^*(\omega)) = \frac{t_x}{H}(1-\varepsilon_{H,\omega}) + \frac{a_x}{aH}\frac{u}{\omega u_q}\left(\varepsilon_{U,\omega} - \varepsilon_{H,\omega} - \varepsilon_{u_q,\omega}\right) + \frac{\varepsilon_{H,\omega}}{H}\frac{c_x}{\omega}. \tag{B.2.1}$$

Using (13) and (B.1.1), (B.2.1) may be rewritten as follows:

$$\Psi_{x\omega}(x,\omega,U^*(\omega)) = \frac{t}{H}\left[\left(1 - \frac{\varepsilon_{H,\omega} + \varepsilon_{u_q,\omega}}{\varepsilon_{U,\omega}}\right)A - (1-\varepsilon_{H,\omega})T + \frac{c_x}{\omega t}\varepsilon_{H,\omega}\right],$$

which reduces to $\phi(x,\omega) = 0$ when $c_x = 0$. By contrast, when $t_x = 0$, we have $T(x) = 0$. Therefore, if $A(x) = 0$, we obtain:

$$\Psi_{x\omega}(x,\omega,U^*(\omega)) = \frac{c_x}{\omega t}\varepsilon_{H,\omega} > 0,$$

which implies perfect ranking by increasing income order.

Note that $\Psi_{x\omega}(x,\omega,U^*(\omega)) > 0$ holds for $\omega\tau(x) + c(x)$ when preferences are Cobb-Douglas because $\varepsilon_{H,\omega} = \varepsilon_{U,\omega} = 1$ and $\varepsilon_{u_q,\omega} = 0$.

B.3 Homothetic preferences

Assume that the utility $u(q,h)$ is homothetic, that is, homogeneous linear. Then, it must be that $\varepsilon_{h,\omega} = \varepsilon_{q,\omega} = 1$. The first-order condition for utility maximization implies

$$u_h = Ru_q.$$

It follows from Euler's theorem that

$$\begin{aligned} hu_h + qu_q &= u \\ &\Leftrightarrow\ h\frac{u_h}{u} + q\frac{u_q}{u} = 1, \end{aligned}$$

that is,

$$\varepsilon_{U,h} + \varepsilon_{U,q} = 1.$$

Since the income elasticity of utility is given by

$$\varepsilon_{U,\omega} = \varepsilon_{U,h}\cdot\varepsilon_{h,\omega} + \varepsilon_{U,q}\cdot\varepsilon_{q,\omega},$$

we get

$$\varepsilon_{U,\omega} = 1.$$

It remains to determine $\partial u_q/\partial\omega$. Using the first-order condition $u_h = Ru_q$, the budget constraint $Rh + q = \omega t$ and Euler's theorem, we obtain

$$u_q = \frac{u}{\omega t}.$$

Therefore,

$$\begin{aligned}\frac{\partial u_q}{\partial\omega} &= \frac{1}{t}\frac{u_\omega\omega - u}{\omega^2}\\ &= \frac{u}{\omega^2 t}(\varepsilon_{U,\omega} - 1)\\ &= \frac{u_q}{\omega}(\varepsilon_{U,\omega} - 1)\end{aligned}$$

so that

$$\varepsilon_{u_q,\omega} = 0.$$

To sum up, we have $\varepsilon_{U,\omega} = 1$, $\varepsilon_{H,\omega} = \varepsilon_{h,\omega} = 1$ and $\varepsilon_{u_q,\omega} = 0$.

B.4 Stone-Geary preferences

Differentiating (B.1.1) with respect to x and using (B.1.2) gives

$$\Psi_{\omega x}(x,\omega,U^*(\omega)) = \frac{1}{H}\cdot\left[\frac{t_x}{t}\frac{\mathrm{d}U^*(\omega)}{\mathrm{d}\omega}Q_U - \frac{\mathrm{d}U^*(\omega)}{\mathrm{d}\omega}\left(Q_{Ua}a_x + Q_{UH}H_\Delta\Delta_x\right)\right]. \tag{B.4.1}$$

It is readily verified from (15) that

$$Q(h, U/a(x)) = \left[\frac{1}{(h-\overline{h})^\mu}\frac{U}{a}\right]^{\frac{1}{1-\mu}}. \tag{B.4.2}$$

It follows from (B.4.2) that

$$\begin{aligned}Q_U &= \frac{1}{1-\mu}U^{\frac{1}{1-\mu}-1}\left[\frac{1}{a(h-\overline{h})^\mu}\right]^{\frac{1}{1-\mu}},\\ Q_{Ua} &= -\frac{Q_U}{(1-\mu)a},\\ Q_{Uh} &= -\frac{\mu Q_U}{(1-\mu)(h-\overline{h})},\\ Q_{UU} &= \frac{\mu}{1-\mu}\frac{Q_U}{U},\\ Q_a &= -\frac{U}{a}Q_U.\end{aligned}$$

Plugging Q_{Ua}, Q_{Uh} and (B.1.2) into (B.4.1) leads to

$$\Psi_{\omega x}(x,\omega,U^*(\omega)) = \frac{t}{H}\cdot\left[\frac{t_x}{t} + \frac{1}{1-\mu}\left(A + \mu\frac{\Delta_x H_\Delta}{H-\overline{h}}\right)\right]. \tag{B.4.3}$$

Using (B.5.3), we have

$$\frac{\Delta_x H_\Delta}{H-\overline{h}} = -\omega\Delta\frac{1-\mu}{\mu h}\left(\frac{H-\overline{h}}{U}\right)^{\frac{1}{1-\mu}}\left[A-(1-\mu)T\right].$$

Substituting this expression into (B.4.3) and using (B.5.2) yields

$$\Psi_{\omega x}(x,\omega,U^*(\omega)) = \frac{t}{H}\cdot\frac{\overline{h}}{H}\cdot\left[A-(1-\mu)T\right].$$

B.5 Proof of Proposition 2

The bid-max lot size. Differentiating (B.4.2) with respect to h, plugging this expression in (9) and solving the corresponding equation yields

$$\left[\frac{h-(1-\mu)\overline{h}}{(1-\mu)(h-\overline{h})}\right] Q(h,U/a(x)) - \omega t = 0. \tag{B.5.1}$$

Plugging (B.4.2) into (B.5.1), we obtain

$$U^{\frac{1}{1-\mu}}\cdot(h-\overline{h})^{-\frac{1}{1-\mu}}\left[\frac{h-(1-\mu)\overline{h}}{1-\mu}\right] = \omega\cdot\Delta, \tag{B.5.2}$$

which implies (19), so that the bid-max lot size depends on $a(x)$ and $t(x)$ through (18) only.

Equilibrium utility level. Applying the implicit function theorem to (B.5.2) yields

$$\frac{\partial H}{\partial\Delta} = -\left[U^{\frac{1}{1-\mu}}(H-\overline{h})^{-\frac{1}{1-\mu}-1}\frac{\mu H}{(1-\mu)^2}\right]^{-1}\omega < 0. \tag{B.5.3}$$

Differentiating (B.4.2) with respect to U and solving for Q_U, (B.1.2) may be rewritten as the following differential equation:

$$\frac{\mathrm{d}U^*}{\mathrm{d}\omega} = (1-\mu)\,\Delta(H-\overline{h})^{\frac{\mu}{1-\mu}}(U^*(\omega))^{-\frac{\mu}{1-\mu}}. \tag{B.5.4}$$

B.6 Supermodularity of the equilibrium utility level

Differentiating (B.5.4) with respect to Δ, we obtain:

$$\frac{\partial}{\partial\Delta}\frac{\mathrm{d}U^*}{\mathrm{d}\omega} = (H-\overline{h})^{\frac{\mu}{1-\mu}}(U^*(\omega))^{-\frac{\mu}{1-\mu}}\cdot\left[1-\mu+\mu\Delta(H-\overline{h})^{-1}\frac{\partial H}{\partial\Delta}\right].$$

Using (B.5.3), this expression may be rewritten as follows:

$$\frac{\partial}{\partial\Delta}\frac{\mathrm{d}U^*}{\mathrm{d}\omega} = (H-\overline{h})^{\frac{\mu}{1-\mu}}(U^*(\omega))^{-\frac{\mu}{1-\mu}}\cdot\left[1-\mu-(H-\overline{h})^{\frac{1}{1-\mu}}\frac{(1-\mu)^2\,\omega\Delta}{(U^*(\omega))^{\frac{1}{1-\mu}}H}\right].$$

Substituting (B.5.2) in the bracketed term, we obtain:

$$1 - (H - \overline{h})^{\frac{1}{1-\mu}} \frac{(1-\mu)\,\omega\Delta}{(U^*(\omega))^{\frac{1}{1-\mu}} H} = (1-\mu)\frac{\overline{h}}{H} > 0,$$

which implies

$$\frac{\partial}{\partial\Delta}\frac{\mathrm{d}U^*}{\mathrm{d}\omega} > 0.$$

B.7 The land rent and land gradient

1. The expression (10) can be rewritten as follows:

$$\Psi(x,\omega,U^*(\omega)) = \frac{\omega t}{H}\left(1 - \frac{Q}{\omega t}\right).$$

Using (B.1.2), (B.4.2) and Q_U, we obtain:

$$R^*(x) = \frac{\omega^* t}{H}\left[1 - (1-\mu)\frac{U^*}{\omega^* U^*_\omega}\right].$$

2. Differentiating

$$\Psi\left[x,\omega^*(x),U\left(\omega^*(x)\right)\right] = \frac{\omega^* t - Q}{H}$$

with respect to x leads to

$$\begin{aligned}\Psi_x\left[x,\omega^*(x),U\left(\omega^*(x)\right)\right] &= \left[\omega^*\frac{t_x H - t\left(\omega^*_x\left(H_\omega + H_U U^*_\omega\right) + H_\Delta \Delta_x\right)}{H^2} + \omega^*_x\frac{t}{H}\right]\cdot\left(1 - \frac{Q}{\omega t}\right)\\ &+\frac{\omega^* t}{H}\cdot\left[-\frac{Q_h\left(\omega^*_x\left(H_\omega + H_U U^*_\omega\right) + H_\Delta\Delta_x\right) + Q_a a_x + q_U U^*_\omega \omega^*_x}{\omega^* t} + \frac{\omega^*_x t + \omega^* t_x}{(\omega^* t)^2}Q\right].\end{aligned}$$

Using (9), we obtain

$$\begin{aligned}\Psi_x\left[x,\omega^*(x),U\left(\omega^*(x)\right)\right] &= \frac{\omega^* t_x + \omega^*_x t}{H}\cdot\left(1 - \frac{Q}{\omega^* t}\right)\\ &+\frac{\omega^* t}{H}\cdot\left[-\frac{Q_a a_x + Q_U U^*_\omega \omega^*_x}{\omega^* t} + \frac{\omega^*_x t + \omega^* t_x}{(\omega^* t)^2}Q\right].\end{aligned}$$

Rearranging terms,

$$\begin{aligned}\Psi_x\left[x,\omega^*(x),U\left(\omega^*(x)\right)\right] &= \frac{\omega^* t}{H}\cdot\left(\frac{t_x}{t} + \frac{\omega^*_x}{\omega^*} - \frac{Q_a a_x}{\omega^* t} - \frac{Q_U U_\omega \omega^*_x}{\omega^* t}\right)\\ &= \frac{\omega t}{H}\cdot\left(-\frac{Q_a a_x}{\omega t} - T\right) \qquad \text{(B.7.1)}\end{aligned}$$

where we have used (B.1.2). Substituting Q_a and (B.1.2) in (B.7.1), we obtain:

$$\Psi_x\left[x,\omega^*(x),U\left(\omega^*(x)\right)\right] = \frac{\omega^* t}{H}\left[\frac{1}{\varepsilon_{U,\omega}}A(x) - T(x)\right].$$

B.8 The equilibrium land rent under Fréchet distributions

Solving (25) with respect to Δ and substituting into (26), we obtain

$$\frac{\mathrm{d}U^*}{\mathrm{d}\omega} = s_\Delta \cdot \left(\frac{\omega - \underline{\omega}}{s_\omega}\right)^{1/\gamma} + \underline{\Delta}. \tag{C.8.1}$$

Integrating yields

$$U^*(\omega) = s_\Delta \cdot s_\omega \cdot \frac{\gamma}{\gamma+1}\left(\frac{\omega - \underline{\omega}}{s_\omega}\right)^{\frac{\gamma+1}{\gamma}} + \underline{\Delta}\omega + K, \tag{C.8.2}$$

where K is the constant of integration.

The bid-rent function of the ω-households at x may then be rewritten as follows:

$$\begin{aligned}\Psi(x,\omega,U^*(\omega)) &= \omega t(x) - \frac{U^*(\omega)}{a(x)} \\ &= \omega t(x) - \frac{1}{a(x)} \cdot s_\Delta \cdot s_\omega \cdot \frac{\gamma}{\gamma+1}\left(\frac{\omega - \underline{\omega}}{s_\omega}\right)^{\frac{\gamma+1}{\gamma}} - \frac{\underline{\Delta}}{a(x)}\omega - \frac{K}{a(x)}.\end{aligned}$$

To determine K, consider the global minimizer $\underline{x}$ (which we assume to be unique) of $\Delta(x)$ defined over the interval $[0, 1]$. Given the equilibrium mapping $\omega^*(x)$, the poorest households reside at $\underline{x}$. At this site, the equilibrium land rent reaches its global minimum value: $R^*(\underline{x}) = R_A > 0$. Indeed, $R^*(\underline{x}) > R_A$ would imply the existence of a site with a higher location-quality index and a lower rent where the poorest households would be better off, a contradiction. In other words, at the least-preferred site $\underline{x}$ the land rent is equal to R_A, very much like in the standard model where the land rent is equal to the opportunity cost of land at the city limit.

Hence, plugging (C.8.1) and (C.8.2) into (22) yields:

$$R_A = -\frac{K}{a(\underline{x})}.$$

By implication,

$$K = -a(\underline{x})R_A.$$

Using (25), the equilibrium land rent is then given by

$$R^*(x) = \frac{\Delta(x) - \underline{\Delta}}{\Delta(x)} \cdot t(x)\left\{\frac{s_\omega}{1+\gamma}\left[\frac{\Delta(x) - \underline{\Delta}}{s_\Delta}\right]^\gamma + \underline{\omega}\right\} + \frac{a(\underline{x})}{a(x)}R_A.$$

B.9 Homogeneous households

Assume for simplicity that ω and Δ are Pareto-distributed: $F(\omega) = 1 - (\underline{\omega}/\omega)^{\gamma_\omega}$ and $G(\Delta) = 1 - (\underline{\Delta}/\Delta)^{\gamma_\Delta}$ where $\underline{\omega} > R_A$ for the poorest households to be able to pay the opportunity cost of land. It is readily verified that the equilibrium mapping (25) is now given by

$$\omega^*(x) = \underline{\omega}\left[\frac{\Delta(x)}{\underline{\Delta}}\right]^\gamma,$$

where $\gamma \equiv \gamma_\Delta / \gamma_\omega$. Repeating the argument of Section 5.2, it can be shown that the equilibrium land rent is as follows:

$$R^*(x) = \frac{1}{\gamma+1} t(x) \underline{\omega} \left\{ \left[\frac{\Delta(x)}{\underline{\Delta}} \right]^\gamma - \frac{\underline{\Delta}}{\Delta(x)} \right\} + \frac{a(\underline{x})}{a(x)} R_A.$$

When γ_ω goes to infinity, the Pareto distribution boils down to the Dirac distribution in which all households share the same income $\underline{\omega}$. In this case, we get:

$$R^*(x) = t(x) \underline{\omega} \left[1 - \frac{\underline{\Delta}}{\Delta(x)} \right] + \frac{a(\underline{x})}{a(x)} R_A.$$

In the absence of amenities ($a(x) = a(\underline{x})$) for all x, this expression becomes:

$$R^*(x) = \underline{\omega} \left[t(x) - t(1) \right] + R_A.$$

We fall back on the following classical result of the standard monocentric city model under fixed lot size: up to the opportunity cost of land, the land rent is the mirror image of commuting costs. In other words, except in this simple case, in a featureful city the land rent and commuting costs are intertwined through a fairly involved relationship.

Appendix C

C.1 Network distances

We obtain information on network distances from the *SpinLab*. This dataset provides information on actual free-flow driving speeds for every major street in the Netherlands. The actual speeds are usually well below free-flow driving speeds the speed limits due to traffic lights, roundabouts and intersections. Because very local streets are missing from the dataset we first calculate for each observation the straight-line distance to nearest access point of the network and then calculate the network distance. The median distance from an observation in the dataset to the nearest access point of the network is 122 m (on average 153 m). We also calculate the Euclidian distance from every job and photo location in the city to the nearest access point of the network. Then, we assume that the average speed to get to the nearest access point is 15 km per hour. This seems reasonable as the minimum speed on roads in the network is 20 km. Furthermore, because of the dominance of the bicycle, this would be close to the average cycling speed. Using these information we calculate the total driving time using the network driving time plus the driving time to access the network and to arrive at the destination. Furthermore, we calculate for each location pair the Euclidian distance and assume again an average speed of 15 km. We then choose the lowest of the network travel time or Euclidian travel time for observations that are within 2.5 km of each other. This is because observations that are very close will not need to go via the network, only when it is faster.

The correlation between travel time and Euclidian distance is, not surprisingly, high. For example, the correlation between the distance to the city center and the travel time to the city center is 0.883. On average, it costs about 2.5 minutes to travel one kilometer, so the average driving speed is about 25 km per hour, which seems realistic in the cities that we study. Figure C.1 shows that for small distances (< 2.5 km), the Euclidian travel time is often lower than the actual travel time, which is not too surprising because travel time is higher when locations that are in the vicinity first have to drive to the network and then basically drive back. It is also seen that the marginal speed is increasing once distances between two locations become larger. This makes sense as it is more likely that people will be able to make use of highways.

C.2 Land prices and lot sizes

Information on land values and lot sizes is not directly available but may be inferred from data on home sales. We follow the procedure developed by Rossi-Hansberg *et al.* (2010), implying that we can only use information on residential properties with land.

Let $\mathcal{P}(x)$ denote the house price, $H(x)$ the lot size, $R(x)$ the land rent per square meter, $C(x)$ the housing characteristics at x, while $\eta_2(y)$ denote year y fixed effects. For each city, we estimate:

$$\frac{\mathcal{P}(x)}{H(x)} = R(x)e^{\eta_1 C(x)+\eta_2(y)+\epsilon(x)}, \tag{C.2.1}$$

where $\epsilon(x)$ is an identically and independently distributed error term that is assumed to be uncorrelated. to land rents and housing attributes, while $R(x)$, η_1 and η_2 are parameters to be estimated. We do not impose any structure on how land rents $R(x)$ vary across locations. Hence, (C.2.1) can be seen as a semi-parametric partially linear hedonic price function. We then employ (again) the Robinson (1988) procedure to estimate (C.2.1). The first step is to obtain consistent estimates of η_1 and η_2. We regress the log of $\mathcal{P}(x)$ and $C(x)$ nonparametrically on a geographically weighted constant. Then, we regress the residuals of log of house prices on the residuals of $C(x)$, which leads to $\sqrt{N}$consistent estimates of η_1 and η_2.

Set

$$\widehat{W}(x) \equiv \log \mathcal{P}(x) - \widehat{\eta}_1 C(x) - \widehat{\eta}_2(y).$$

To obtain an estimate of the land price at x, we calculate:

$$\log \widehat{R}(x, b) = \frac{\sum_{z=1}^{Z(c)} \mathcal{K}(z, b)\widehat{W}(y)}{\sum_{z=1}^{Z(c)} \mathcal{K}(z, b)}, \tag{C.2.2}$$

where $\mathcal{K}(\cdot)$ is the following Gaussian kernel function:

$$\mathcal{K}(x, b) = \frac{1}{b\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{d(x,z)}{b}\right)},$$

where $d(x, z)$ is the Euclidian distance between locations x and z. The city-specific bandwidth b determines the smoothness of the function to be estimated. We use the Silverman rule-of-thumb bandwidth, given by $1.06\widehat{\sigma}(d)Z^{-\frac{1}{5}}$, where $\widehat{\sigma}(d)$ is the sample standard deviation of the bilateral distances between x and z, for all x, z. Admittedly, we assume a particular functional form on how housing attributes affect house prices. Of course, more flexible functional forms may be used, but this does not have a substantial impact on the estimated land rents. Moreover, the η_1 are already city-specific, which allows for households with specific housing preferences to sort themselves in particular cities.[23]

We also estimate the average lot size at each location:

$$\log \widehat{H}(x, b) = \frac{\sum_{z=1}^{Z(c)} \mathcal{K}(z, b)\mathcal{H}(z)}{\sum_{z=1}^{Z(c)} \mathcal{K}(z, b)}, \tag{C.2.3}$$

where $\widehat{H}(x, b)$ is the estimated lot size at a given location x and $\mathcal{H}(z)$ the observed lot size at z.

Descriptive statistics for the housing sample are reported in Table C.1. Coefficients η_1 related to the housing attributes are reported in Table C.2. We see that the price per square meter is generally a bit higher when the property is larger or has more rooms. However, the land price of properties that are (semi-)detached is generally lower. Furthermore, when the maintenance state of a property is good, prices are between 6 and 20% higher. When a property has central heating, the price per square meter is about 8% higher. The impact of being listed is heterogeneous between cities, potentially because the quality of cultural heritage is also different between the two cities. The dummies related to the construction decades show the expected signs: in general, newer properties command higher prices. Properties constructed after World War II until 1970 generally are lower priced because this is a period associated with a lower building quality. In Figure A.3 we plot land rents for Amsterdam and Rotterdam. Land values are generally higher in the city center. In particular, Amsterdam command high house prices within the highway ring road. In Rotterdam, the highest land rent is not at the city center, but in neighborhoods such as Scheepvaartkwartier and Kralingen, which are reasonably close to the city center. The latter neighborhoods were not bombed out in World War II and still offer historic amenities.

[23] We use data on housing prices between 2000 and 2015. By including time fixed effects $\eta_2(z)$, the estimate of land values is time-invariant. If we would estimate (C.2.2) for each year, for some areas our data would be too 'thin' and the resulting land rents too imprecise.

C.3 Amenities

Picture index. We inspect the data underlying the picture index in Table C.3. It is observed that 65% of the pictures are taken in Amsterdam. It is also shown that 74% of the pictures are taken outside a building and about 44% of the pictures are taken by local residents. This share is a bit higher in Rotterdam (55%), suggesting that Amsterdam receives more tourists. Recall that we only use pictures *outside* buildings taken by *residents* in the determination of the amenity index.

[Tables C.3-C.5 about here]

We correlate observed proxies of amenities to the picture index. We make a distinction between historic amenities (historic districts, number of listed buildings within 500 m), natural amenities (share open space, water within 500 m, and maximum flood depth) and consumption amenities (shop, hotels, cafes, restaurants, cultural establishments, leisure establishments). Table C.4 presents descriptive statistics of those variables, whereas Table C.5 reports regression results with the picture density as dependent variable. Column (1) shows that all the proxies of historic amenities are positively correlated to the picture index. The picture density in historic districts is more than 200% higher than outside those areas. In column (2) we investigate the impact of natural amenities. The share of open space is negatively correlated with the picture index. This is likely due to the fact that people make fewer photos in and around inaccessible open space, such as farmland and forests. By contrast, the share of open water is positively correlated to the picture index, likely due to the many canals in Amsterdam and the many (small) rivers in Rotterdam that add to the appeal of both cities. Consumption amenities do generally have the expected positive signs (see column (3)). This holds in particular when we include all amenities in one specification in column (4). It is also worthwhile investigating the fit: with a simple linear specification we already explain about 70% of the picture index, suggesting that the picture index is highly correlated to observed proxies for amenities.

Hedonic amenity index. We also test whether our results are robust to using an alternative hedonic amenity index, rather than relying on geocoded pictures. Following Lee and Lin (2015), we therefore aim to construct an aggregate amenity index that describes the amenity provision at every location x.[24] We will make a distinction between *historic amenities* and

[24] Albouy (2016) uses information on wages and housing costs to infer, among other things, the level of amenities for US cities. However, his approach is not applicable here because we are interested in *intra*-city variation in amenities rather than *inter*-city variation. We hypothesized that households may sort on basis of amenities. Using Albouy's approach amenities would be a direct result of sorting of rich households in certain locations, while this is exactly the relationship we aim to test.

natural amenities. We ignore consumption amenities because they are potentially endogenous. However, it is worth noting that our results are similar when we also include consumption amenities.

Let $\mathcal{A}(x)$ be a set of variables that describe amenities (see Table C.4 for more details), $\mathcal{P}(x)$ the house price while $C(x)$ are housing characteristics at x, and $\eta_2(y)$ are year y fixed effects. Finally, we include neighborhood fixed effects $\eta_3(x)$. Neighborhoods in the Netherlands are rather small, so this should largely control for the accessibility to jobs and/or distance to the city center. We then estimate *for each city*:

$$\log \mathcal{P}(x) = \eta_0 \mathcal{A}(x) + \eta_1 C(x) + \eta_2(y) + \eta_3(x) + \epsilon(x), \tag{C.3.1}$$

where η_0, η_1 and η_2 are parameters to be estimated and $\epsilon(x)$ is an identically and independently distributed error term. We then use $\widehat{\eta}_0$ and $\mathcal{A}(x)$ to predict the amenity level in each location x:

$$\tilde{a}(x) = \widehat{\eta}_0 \mathcal{A}(x), \tag{C.3.2}$$

where $\tilde{a}(x)$ is the (alternative) amenity value at x.

C.4 Historic land use and first-stage results

We use information on land use in 1900 and historic city plans as instruments for current land use patterns. For the 1900 land use maps, Knol *et al.* (2004) have scanned and digitized maps into 50 by 50 meter grids and classified these maps into 10 categories, including built-up areas, water, sand and forest. We aggregate these 10 categories into built-up, open space and water bodies. Knol *et al.* document large changes in land use across the Netherlands from 1900 to 2000. For example, the total land used for buildings has increased more than fivefold. On the other hand, the amount of open space has decreased by about 10%. We also use information on municipal population density in 1900, which were much smaller at that time and about the size of a large neighborhood nowadays. We show a map of land use in 1900 in Figure C.2. We impute the local population distribution using the location of buildings and assuming that the population per building is the same within each municipality.

We furthermore will use data composed by *HISGIS*, which has compiled and digitized data from the first Dutch census in 1832. This dataset provides information on the land use of each parcel in the current inner cities of Amsterdam and Rotterdam. The *HISGIS* data also provide information on the land value of each parcel, which was used to determine the tax at that time.

We also use alternative instruments based on the land use stipulated in historic city plans. Before World War II all the larger cities developed large-scale extension plans to accommodate future growth in the population. It was also in the course of the 20th century that the government took the lead in new residential development because of the dire circumstances poor people had to live. In 1901 the first Housing Act was implemented that put limits to unhealthy housing. It was also the launch of a large scale plan to provide public housing. These plans usually had a long time-horizon of at least fifty years. However, the plans did of course not foresee the substantial rise in the use of the automobile and other technological improvements that had an impact on the urban structure. Therefore, these plans were only partly realized. Because of the bombing of Rotterdam's city center in World War II, many new developments in the city of Rotterdam occurred *in* the city center rather than around the city center, as stipulated in the plans. The plan of Amsterdam (1935) was designed by C. van Eesteren and the plan of Rotterdam (1928) by W.G. Witteveen. It is important to notice that the plans were not so detailed that they clearly outline what kind of housing had to be built. Moreover, because hardly any people were living in the new areas, the current spatial income distribution cannot have had any influence on the extension plans. We digitize the maps for each city. Figure C.3 displays the maps for Amsterdam and Rotterdam. It can be seen that the area for which we have information is somewhat smaller for Rotterdam. It can also be seen that there were also comprehensive plans for development of commercial areas, in particular of port areas. The total size of residential extension plans are the largest for Amsterdam. Large parts of the plans have been realized: the Western part of Amsterdam, for example, is almost exactly realized according to the plan. On the other hand, although the plans were expected to be able to accommodate population growth for the next 50-100 years, after World War II it became clear that the new housing construction as stipulated in the plans was not sufficient. Therefore, many other neighborhoods, in particular in Amsterdam have been built (e.g. Amsterdam Zuid-Oost). Hence, although there will likely be correlation of land use in the plans, the correlation will far from being equal to one.

[Figure C.2 and Figure C.3 about here]

In Table C.6 we provide descriptives for all instruments. The share of built-up area in 1900 was 10%, while it was maximally 8% in 1832. The travel time to existing population or residential and commercial land use is very comparable for all types of instruments (about 25 min).

[Table C.6 about here]

We report first-stage results in Table C.7. These are the specifications corresponding to the estimates in Table 3. The instruments generally have the expected signs. The share of built-up area in 1900 and the share of water bodies in 1900 are positively correlated to the current amenity level. The former effect can be explained by the fact that the presence of historic amenities is strongly correlated to the number of historic buildings in the vicinity. A similar positive correlation is observed for the share of existing and planned residential land use and the share of water bodies in historic city plans. Conditional on the share of built-up land and water bodies, travel time to population in 1900 is negatively associated with the presence of amenities, whereas positively correlated to current expected commuting time. The only somewhat unexpected result is the absence of a statistically significant effect of travel time to existing built-up land in 1930 on the current expected commuting time. However, the average travel time to planned built-up land is positively correlated to the current expected commuting time.

[Table C.7 about here]

C.5 Sensitivity analysis

Identification revisited. The validity of historical instruments may be questioned. We, therefore, estimate an additional set of specifications that should contribute to the belief that our identification strategy is valid. The results are reported in Table C.8. In column (1) we add the instruments based on land use in 1900 as controls in an OLS specification. If the instruments are invalid and correlated to the error term, the impact of amenities and expected commuting times should decline because the instruments absorb part of the bias of the endogenous variables (Arzaghi and Henderson, 2008). However, compared to the corresponding OLS specification in column (5), Table 2, the impact of amenities is not statistically different. The impact of expected commuting time is slightly lower, but not statistically different from the corresponding OLS specification. Moreover, the instruments are only marginally significant and negatively related to incomes. We replicate this specification by including the instruments based on historic city plans in column (2). The coefficients related to amenities and expected commuting costs are now somewhat smaller. Because they are not statistically different from the baseline OLS specification, we cannot conclude that the instruments are invalid. Since going back further in time is likely to generate more valid instruments, we will use census data in 1832 as an additional check.

[Table C.8 about here]

Before turning to those results, we estimate some specifications in which we include both historic city plan instruments and 1900 land use instruments, while controlling for the share of existing residential land use and planned residential land use in the city plans within a distance of 500 m and the share of built-up area in 1900 within 500 m. We also include a dummy whether instruments from historic city plans are missing in the second stage (as to keep the maximum number of observation). In column (3), Table C.8, we show that results are very similar to the corresponding results reported in columns (4) and (8), Table 3, but the effect of expected commuting costs is now stronger and somewhat more precisely estimated. In column (4) we replicate this specification, but exclude observations in historic districts, to test whether our results are not mainly driven by the historic amenities present in those neighborhoods. This does not appear to be the case as the effect of amenities becomes even somewhat stronger.

As said above, historic instruments cease to be valid when unobserved characteristics of a location or building in the past are correlated with those in present time, which is less likely with instruments based on land use further back in time. On the other hand, by going back further in time, we may end up with weak instruments because the correlation between historic land use and current amenities and job locations will also be lower. Nevertheless, we exploit land use data from the census in 1832 because the data are 175 years before the sample period. We have exact information on the land use of each parcel in 1832, as well as detailed information on land rents of each parcel at that time. If some past attractive features (e.g. housing attributes) are correlated to current sorting patterns, we expect this to be reflected in a positive coefficient of past land rents. Note that the 1832 data are available for the inner cities of Amsterdam and Rotterdam only, so that we have fewer observations. We again impute population per building using the municipal populations in 1832 and calculate the travel time of population within 45 minutes travel time, assuming an Euclidian travel speed of 5 km per hour. We further use the share of infrastructure land, the share of built-up area and the share of water bodies within 500 m as instruments.

The results displayed in column (5), Table C.8, show that the results are comparable to the baseline specification. The effect on incomes is now 2.3% for a 100% increase in amenities. The coefficient related to expected commuting costs is again negative, but marginally significant. The coefficient implies that incomes decrease with 28% when expected commuting times double. Interestingly, the coefficient related to the parcel price in 1832 is highly insignificant, suggesting that attractive locations in the past do not currently attract high-income households. We further observe that the Kleibergen-Paap F-statistic is rather low (3.313), confirming that going back further in time reduces the correlation between the endogenous variables and the instruments. Because we may have a weak instrument problem, we should be careful not to overinterpret the results. In column (6), Table C.8, we again directly control for the share of infrastructure land and share of built-up area within 500 m. The effect of amenities and expected commuting costs are very similar, although imprecise. Hence, given the larger confidence intervals, the results are not significantly different from the specifications in which we utilize instruments based on land use in 1900.

Other sensitivity checks. In Table C.9 we conduct some additional robustness checks. We consider the specifications in column (5), Table 2, and columns (3) and (7), Table 3, as the baseline specifications. One may wonder whether our results are primarily driven by the arbitrary choice of the decay parameter φ. We have set $\varphi = 0.793$, so that the expected travel time is 20 minutes when amenities were evenly distributed over space. We also run regressions when the expected travel time is, respectively, 10 and 40 minutes, which yield $\varphi = 1.121$ and $\varphi = 0.560$. Again, amenities have a positive and statistically significant effect, while the effect of commuting is negative (see columns (1) and (2), Table C.9). We also assumed that people only consider jobs within 45 minutes travel time. In column (3), Table 5, we reduce this to 30 minutes and in column (4) we increase this to 60 minutes. The results are very similar.

[Table C.9 about here]

In column (5), Table C.9, we include 25 municipality fixed effects to control for differences in municipal policies that may disproportionately attract the rich, such as low taxes on residential properties. The results are very similar to the baseline specification, although it is now harder to estimate the effect of expected commuting costs because the effect is to a large extent absorbed by the fixed effects. In other words, there is little variation in the expected commuting costs *within* municipalities.

In column (6) we consider an alternative dependent variable to correct for possible errors in self-reported incomes. We use the information on the educational level and create a dummy variable that is equal to one when an individual has at least a bachelor's degree. We then estimate linear probability models. The results are completely in line with what we find for incomes: people with a higher educational degree sort themselves near amenities and jobs. Column (7) further addresses the issue that we do not observe income as a continuously distributed variable: self-reported income, on which the variable income is based, is measured in 7 classes. Furthermore, to determine incomes we have to make assumptions on how many hours a full-time or part-time job takes (38 or 24 hours, respectively). We make different assumptions regarding what a full-time job is (36 or 40 hours), as well as a part-time job (16 or 32 hours, respectively). This implies that we observe incomes within an interval $[\tilde{\omega}_{\min}(x), \tilde{\omega}_{\max}(x)]$. Since there is a legal minimum wage in the Netherlands, we do not have left-censored observations, but we do have right-censored observations for the highest income category. Observed log incomes may be shown to be essentially normally distributed, so that we can use interval regression. Set

$$\mathcal{Z}(x) \equiv \alpha_0 + \alpha_1 \log \tilde{a}(x) + \alpha_2 \log \tilde{\tau}(x) + \alpha_3 D(x) + \alpha_4 L(x) + \alpha_5 C(x) + \eta_1(x) + \eta_2(y).$$

The log-likelihood is then given by:

$$\log \mathcal{L} = \sum_{x \in \mathcal{I}} \log \left[\Phi \left(\log \tilde{\omega}_{\max}(x) - \mathcal{Z}(x) \right) - \Phi \left(\log \tilde{\omega}_{\min}(x) - \mathcal{Z}(x) \right) \right] + \sum_{x \in \mathcal{R}} \log \left[1 - \Phi \left(\log \tilde{\omega}_{\min}(x) - \mathcal{Z}(x) \right) \right],$$

where $\mathcal{I}$ are the interval observations, $\mathcal{R}$ are the right-censored observations and $\Phi(\cdot)$ is the standard normal c.d.f. The results are shown to be essentially identical to the corresponding OLS and 2SLS specifications where we assume that incomes are continuously distributed.

Alternative proxies for amenities and commuting costs. We must check that our results do not hinge on the particular proxy we use for amenities. In Table C.10 we report some additional checks. We focus on OLS results; 2SLS results lead to similar conclusions and are available upon request. We first use population density as an alternative proxy for amenities. Column (1) shows that, when we calculate population density according to equation (28), increasing population density seems to have a positive effect on incomes. However, when we control for the picture density in column (2), the effect of population density becomes essentially zero and highly insignificant. This strongly suggests that using picture density for amenities is superior to using population density.

[Table C.10 about here]

In column (3), Table C.10, we also include pictures taken inside buildings in the calculation of the amenity index. The results suggest that the impact of amenities is slightly lower, albeit similar. In column (4) we use pictures taken by tourists instead of pictures taken by residents to address the issue that only a specific type of residents (e.g. young people) take pictures outside holidays. Again, the effect of amenities is very similar to the baseline specification.

One may also criticize our measure of commuting costs. We acknowledge that any measure of commuting costs will be imperfect because we do not have information on the actual job location of individual workers. To this end, we calculate the average commuting costs to all employment centers in the wider urban area. To determine the employment subcenters, we use McMillen's (2001) nonparametric approach. The resulting employment subcenters are reported in Table C.11. Column (5) in Table C.10 shows that this proxy performs poorly. The OLS-specification suggests that locations that are further away from employment subcenters will command *higher* incomes, which is counterintuitive. However, if we instrument for expected commuting costs, the coefficient becomes highly insignificant. Commuting costs proxied by expected commuting time seems to outperform the average commuting time. A likely explanation is that a few employment centers cannot capture the decentralized nature of employment patterns. This is more clearly illustrated in column (6), Table C.10, where we include both the commuting time to the nearest employment center, as well as the expected commuting time. The coefficient related to expected commuting costs is very similar to that obtained in the baseline specification.

[Table C.11 about here]

City-specific results. In Table C.12 we report reduced-form results for Amsterdam and Rotterdam treated separately. Columns (1)-(3) report results for Amsterdam, while columns (4)-(6) refer to Rotterdam. First, we estimate an OLS regression for Amsterdam, showing that the impacts of amenities and expected commuting costs are somewhat larger than the baseline estimates (column (1)). The results suggest that doubling amenities implies an income increase of 4.2%. Doubling expected commuting times implies an income decrease of 17%. The results are very much alike when using instruments based on land use in 1900 (column (2)). When using historic city plan instruments in column (3), the impact of amenities is somewhat stronger. The point estimate of expected commuting time is very similar to the previous specifications, but it is not statistically significant at conventional levels.

As for Rotterdam, column (4) shows that amenities are again highly statistically significant, while the coefficient related to commuting costs is statistically insignificant and close to zero. An explanation may be that due to the complete bombing of the city center in World War II, the inner city is well accessible by car. Hence, there is relatively little variation in expected commuting time close to the city center. This also holds if we instrument for amenities and expected commuting costs with land use in 1900 (column (5)). However, once we instrument with historic city plans, the coefficient of expected commuting costs is negative and highly statistically significant (column (6)). The impact of amenities is similar. In sum, for Rotterdam we also find that amenities are an important determinant of income sorting in the city, while the evidence for expected commuting costs is mixed.

[Table C.12 about here]

C.6 Structural estimation and counterfactual analyses

We report the predicted location quality indices for different scenarios in Figure C.4 as a function of the distance to the city center. Note again that Amsterdam and Rotterdam are non-symmetric cities, so that plotting $\Delta(x)$ as a function of distance to the city center masks a substantial portion of spatial variations. If we concentrate on the baseline case, the predicted location quality index is essentially monotonically decreasing in all specifications. This also holds if we keep expected commuting costs constant (counterfactual scenario 2). Because the location quality index looks so much like the second counterfactual scenario, this suggests that the location quality index is mostly determined by variation in amenities. This is confirmed by the outcomes of counterfactual scenario 1: if amenities are kept fixed over space, $\Delta(x)$ is essentially flat. Similarly, when we let amenities vary according to (35), the location quality index largely follows the counterfactual amenity distribution.