Who Lives Where in the City? Amenities, Commuting and Income Sorting*

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

May 15, 2020

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. The residential equilibrium involves the spatial separation of households sharing similar incomes. Under Stone-Geary preferences, amenities and commuting are subsumed into a location-quality index. Hence, the assignment of households to locations becomes one-dimensional. Since residential choices are driven by the location-quality index, the income mapping may be fully characterized. Using a rich micro-dataset on the Randstad, the Netherlands, we show that household income sorting is indeed driven by amenities and commuting times.

Keywords: cities, social stratification, income, amenities, commuting

JEL classification: R14, R23, R53, Z13.

*We are grateful to N. Baum-Snow, J. Brinkman, P.-P. Combes, M.S. Delgado, R.J.G.M. Florax (†), M. Fujita, J.V. Henderson, M. Miyake, Y. Murata, H.G. Overman, P. Picard, G. Ponzetto, S. Riou, F. Robert-Nicoud, 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, the Higher School of Economics in Moscow, Purdue, AMSE, the Workshop on Public Policies, Cities and Regions, the Utrecht School of Economics, Laval University, the 2017 Public Economic Theory conference, the summer school "Sustainable Cities", the 64th AFSE Congress, 32d EEA Congress, and the 7th North American Meeting of the Urban Economics Association, for their comments. The authors acknowledge, respectively, financial supports from the Agence Nationale de la Recherche (ANR-12-INEG-0002), the Institut Universitaire de France, the Netherlands Organisation for Scientific Research (VENI research grant), and the Russian Academic Excellence Project ‘5-100’.

†INRA, UMR1302, SMART-LERECO Rennes (France) and CREATE, Laval University, Quebec (Canada).

‡Department of Spatial Economics, Vrije Universiteit Amsterdam (The Netherlands), HSE University (Russia) and CEPR.

§CREM (Condorcet Center), Université de Rennes 1.

¶CORE-UCLouvain (Belgium), HSE University (Russia) and CEPR.

1 Introduction

Residential segregation by income is a prominent and growing feature of many metropolitan areas. While residential segregation generates negative and persistent effects on economic development and social cohesion, it both reflects and reinforces interpersonal and social inequalities. For these reasons, we find it important to study the spatial sorting of heterogeneous consumers. We also believe that such an effort is warranted if policy actions against spatial segregation are taken to favor more social cohesion.

The canonical monocentric city model leads to a fairly extreme prediction: households are sorted by increasing income order as the distance to the central business district rises. This result contradicts empirical evidence (Glaeser *et al.*, 2008; Rosenthal and Ross, 2015). One missing key explanation is the existence of *amenity endowments*, such as historic buildings and architecture, scenic landscapes, river and sea proximity. That such amenities matter in residential choices has been well documented (Brueckner *et al.*, 1999; Glaeser *et al.*, 2001; Koster and Rouwendal, 2017; Lee and Lin, 2018; Cuberes *et al.*, 2019). Other facilities that appeal to people, such as restaurants and shops, are often located together with households. Such endogenous amenities are ignored in this paper because they are likely to be a consequence of residential choices, and thus the result of spatial income sorting.

This paper proposes a new approach that combines housing consumption and heterogeneity in incomes and locations. To be precise, we propose a new approach in which cities are "featureful" in that locations are distinguished by two vertically differentiated attributes, that is, commuting costs 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. Our aim is then to study how local amenity endowments and commuting costs interact to determine the spatial sorting of income-heterogeneous households. To check the relevance of our approach, we estimate the so-obtained income mapping by using rich datasets from the Netherlands.

The study of income sorting when locations are differentiated by more than one attribute brings about new, difficult issues. Indeed, the determination of a residential equilibrium has the nature of a matching problem between landlords and households in which land at specific locations is differentiated by two characteristics and households by one characteristic. Multidimensional matching problems typically assume quasi-linear preferences and indivisible choices (Chiappori, 2017). We square the circle of endogenous and income-dependent housing consumption by using the bid-rent approach.

What are our main contributions? First, we model space as a *topological network*. Attempts made to move away from the standard one-dimensional framework of urban economics typically use the Euclidean plane. Such an approach neglects one the most fundamental features of a space-economy, that is, the shape and density of its transportation networks. Residences and jobs are thus distributed over the topological network.

Second, recent surveys, such as Duranton and Puga (2015) and Behrens and Robert-Nicoud (2015), express some skepticism about the ability of the bid-rent approach to deal with a continuum of heterogeneous households to be mapped on a continuum of locations. We demonstrate that the bid-rent approach is still applicable to such settings and show that the interaction between amenities and commuting gives rise to turning points in the spatial income distribution. In other words, provided that amenities are distributed unevenly across space, a greater geographical distance between households no longer implies a wider income gap.

Third, when commuting costs vary with income, characterizing the equilibrium income mapping requires *non-homothetic* preferences. Indeed, empirical evidence suggests that housing expenditure shares decline with income (Albouy *et al.*, 2016). Furthermore, homothetic preferences such as the Cobb-Douglas or the CES must be ruled out because they generate multiple equilibria in a model that comprises a continuum of heterogeneous households and a continuum of locations. Using Stone-Geary preferences, we establish the following two results. In the first place, there exists a unique spatial equilibrium. Households with different incomes always choose locations with different characteristics, but households sharing the same income may live in separated neighborhoods. In the second one, we find a *location-quality index* whose behavior reflects the properties of the equilibrium income mapping. This index is built from the primitives of the model, as its value at any particular location is pinned down by the amenity endowments and households' commuting behavior. It is worth stressing that these results are not specific to Stone-Geary preferences. They hold true for other non-homothetic preferences; what changes is the functional form of the location-quality index.

The upshot is that the bliss point is the global maximizer of the location-quality index over the topological network, thus implying that this location is occupied by the affluent because they propose the highest bid. As one moves away from this location along all admissible directions, households are sorted by decreasing incomes until a local minimizer of the location-quality index is reached where poorest households are located. Around this minimizer, household income starts rising. As a result, households get more exposure to other social groups when the number of turning points of the location-quality index rises.

Fourth, our model is flexible enough to determine analytically the equilibrium outcome when incomes and the location-quality index are Fréchet-distributed. The so-obtained mapping between the location-quality index and income can be straightforwardly estimated using real-life data.

We aim to test the income mapping as to investigating whether the two main forces in our theoretical model – amenities and commuting costs – also drive income sorting in practice. We estimate a reduced-form version of our model for the Netherlands. The choice of the Netherlands is motivated by (i) the availability of large disaggregated data sets and (ii) the fact that the public services that underpin social cohesion (e.g., education and health) are centrally financed and/or administered (Ritzen *et al.*, 1997). First, we have access to rich microdata for more than 10 million households covering the years 2010 to 2015 on incomes, residential and job locations at the household level, employment accessibility, as well as land values and amenities at each location. Second, the central provision of education implies that the quality of schools, which is a major determinant of residential choices in many U.S. metropolitan areas, is much less of an issue in the Netherlands.

Dutch cities, which were established long ago, are known to offer a high quality of life, which is at least partly due to the presence of amenity endowments. Despite being a small country, the Netherlands hosts no less than 8 UNESCO world heritage sites, which is almost as much as London and Paris together, while it hosts $61,908$ listed buildings, which is more than three times the number of listed buildings in Greater London. With a population density of 407.4 pop/sq km, the Netherlands is almost as dense as the San Francisco Bay area whose area is similar to that of the Netherlands. This is an important feature in spatial settings where density economies matter. By focusing on the whole country rather than a subset of metropolitan areas, we are able to capture relationships that are deployed within the entire Dutch urban system, from the large cities to the small villages.

It is well known that measuring the quality of amenities is a hard task. In this paper, we use a proxy suggested by Ahlfeldt (2013) and Saiz *et al.* (2018): the number of outside geocoded pictures taken by *residents* at a certain location. One key advantage of this index is that it lets consumers choose the aesthetic quality of buildings and locations they like best by "voting with their clicks" (Carlino and Saiz, 2019). This allows us to move beyond the approach of defining amenities implicitly, as in Ahlfeldt *et al.* (2015) and Albouy (2016). We show the robustness of our results by using a completely different proxy for amenities based on a hedonic price approach developed by Lee and Lin (2018), using variations in housing prices.

Since there is no proxy that perfectly captures the full amenity potential at a certain location, amenities are measured with error. Employment accessibility is also likely to be endogenous due to correlation with unobservable household characteristics and agglomeration economies – the latter being more prevalent in dense areas where commutes are shorter. We address the endogenous nature of amenities and accessibility in our econometric analysis in several ways, e.g., by obtaining Oster's (2019) bias-adjusted estimates and by constructing historic instruments. Since the strategy of using instruments based on historic data raises several issues, we devote considerable attention to the validity of such an identification strategy.

In a first step, we show that picture density is strongly correlated to the presence of amenity endowments, such as the density of historic listed buildings and the presence of water bodies. We then show that both amenities and commuting costs are important in determining the spatial income mapping. More specifically, we find that *doubling the amenity level attracts households whose incomes are* 2.6% *higher*, while *doubling commuting time leads to households whose incomes are* 15% *lower*. Hence, the impact of commuting time seems to be somewhat stronger than the impact of amenities. We also find a strong impact on land prices: *doubling amenities leads to an increase in land prices of* 13%, which is sizable.

Thus, our results unambiguously suggest that both amenities and commuting costs are important in determining the spatial income distribution. Nevertheless, commuting costs seem to be a more important driver of income sorting than amenities. Hence, our results support the emphasis put on commuting costs in standard models of urban economics. However, our results also show that the featureless model of urban economics is far too restrictive to explain the city structure. Yet, focusing on amenities only is unwarranted because commuting costs are too important to be disregarded. A relevant theory of the space-economy must account explicitly for both amenities *and* commuting costs.

Related literature. Suggesting the complexity of the issue, 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 in the monocentric city. 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 in a featureless monocentric city when commuting costs do not depend on income.

In an important paper, Ahlfeldt *et al.* (2015) highlight the role of amenities, agglomeration economies and commuting in residential location choices in their study of the internal structure of Berlin. Our paper differs from theirs in two fundamental aspects. First, these authors do not provide any properties of the spatial equilibrium. This should not come as a surprise as characterizing the equilibrium outcome is problematic under a finite location set. Second, Ahlfeldt *et al.* (2015) find that the elasticity of amenities with respect to residential density is 0.15, which is quite high. This is so mainly because amenities are measured as 'structural residuals', meaning that it is unclear what these amenities actually capture (e.g., they may capture housing characteristics or sorting on unobserved household characteristics).

Diamond (2016) considers two skill-groups of workers endowed with Cobb-Douglas preferences to study skill sorting across cities in a Roback-like model. However, she disregards workers' residential choices within cities. Tsivanidis (2019) also considers two skill-groups in a closed-city setting. Like us, he uses Stone-Geary preferences because the observed Engel curves are nonlinear. However, unlike us, these authors do not aim to characterize the spatial equilibrium of their models. Using a dynamic setting, Lee and Lin (2018) showed that richer households are anchored in neighborhoods with better natural amenities. We differ from them in at least one fundamental aspect: in their setting people are assumed to work where they live. In our setting, households are free to choose where to live and where to work, while accounting explicitly for commuting costs between the residence and the workplace.

Couture *et al.* (2019) undertake a rich analysis of the impact of income inequality on the internal structure of American cities with different types of amenities. We depart from them by assuming a variable lot-size, which allows us to determine the spatial income mapping in a finer way. Note also the link with the literature on local public goods and schools which focuses on household sorting across heterogeneous jurisdictions (Epple and Sieg, 1999). However, these contributions disregard commuting within jurisdictions, and thus do not study the trade-off between amenities, commuting to jobs and housing prices, which occupies center stage in our approach.

The remainder of the paper is organized as follows. We provide a detailed description of our model in Section 2. In Section 3, we 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. Data are discussed in Section 5. Section 6 presents identification strategy and reduced-form empirical results on the impact of amenities and commuting time on income sorting. In Section 7, we summarize our main results and discuss the role of endogenous amenities.

2 The model

2.1 The spatial economy

The economy involves a unit mass of income-heterogeneous households. 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.[1]

The economy involves two normal consumption goods: *(i)* land h, which is a proxy for housing, and *(ii)* a homogeneous final consumption good q. Shipping the final good within the economy is costless. Therefore, its price is the same across locations. This good is used as the

[1] In Gaigné *et al.* (2020), we discuss how incomes can be determined from the interaction between individual skills and local productivity factors.

numéraire. The land density at each location of the network is 1 while the opportunity cost of land is given by the constant $R_0 \geq 0$.

The map formed by streets, roads, highways, and railway junctions (in a city, region or country) is modeled by means of a topological network. A topological arc, denoted a_τ, is the image in $\mathbb{R}^2$ of a compact interval of $\mathbb{R}$ by a continuous one-to-one mapping. Clearly, any arc linking two distinct locations contains a continuum of locations. A *topological network* $N = \cup_{\tau=1}^{T} a_\tau$ is defined as the union of a finite number of topological arcs. Each arc has a finite length. Furthermore, N is such that for any two points x_1 and x_2 belonging to N there is at least one concatenation of arcs and subarcs of N that links these two points. The distance $d(x_1, x_2)$ between x_1 and x_2 is given by the length of the shortest path that connects these locations. Clearly, $d(\cdot)$ is a metric defined on N. The endpoints of the arcs are called vertices. We assume that these vertices are not colinear, so that (N, d) is *not* a one-dimensional metric space. An example of transportation networks similar to ours can be found in Allen and Arkolakis (2014). When all the vertices are colinear, the network boils down to the standard linear space. In what follows, we assume that all functions are differentiable along each arc of the network N, except maybe at the vertices of N.

2.2 Consumption

Households share the same utility function. Since households prefer more amenities than less, we consider a preference structure similar to the one used in models of vertical product differentiation:

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

where b denotes the amenity level, q the costlessly traded numéraire, and h the land consumption while u is a well-behaved utility function. Let $b(x) > 0$ be a given function whose value expresses the amenity level (or, equivalently, an aggregator of distinct amenities) available at $x \in N$, which are exogenous and intrinsic to a location. In the featureless city of urban economics, $b(x)$ is constant across locations. In this paper, $b(x)$ varies with x. Preferences (1) imply that the amenity b 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 b, 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. However, as will be seen, the single-crossing condition between incomes and locations does not hold: richer households need not always choose locations with more amenities.

There is a given and finite number of employment centers $i \in N$ with $i = 1, ..., n$. The individual loss due to commuting is modeled as an *iceberg* cost. More specifically, if a household resides at x and works at i, we denote by $0 < t_i(x) \leq 1$ its effective number of working units, which decreases with the distance $d(x, i)$. As a result, 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 $d(x, i)$. 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). An iceberg commuting cost is also consistent with the empirical literature that shows that these costs increase with income because the opportunity cost of time increases with income (Small, 2012).

A household earns the same gross income ω regardless of its employment center i. When a household chooses her residential location x , she knows the corresponding effective number of working time units given by a function $0 < t(x) < 1$. For example, if the household is free to choose where to work, she chooses the employment center that maximizes her working time, that is, $t(x) \equiv \max_i t_i(x)$.[2] Alternatively, the household may know her employment location j before selecting her residential location x. In this case, we have $t(x) = t_j(x)$.[3]

Regardless of the specification chosen for $t(x)$, the household's budget constraint is 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).

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:[4]

$$u_h \left[\omega t(x) - R(x)h^*, h^*\right] - R(x)u_q \left[\omega t(x) - R(x)h^*, h^*\right] = 0. \tag{3}$$

[2] In this case, 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.

[3] Our model can be extended to the case where ω-households' net incomes are given by $\omega t_i(x)\nu_{kxi}$, where the random variables ν_{kxi} are i.i.d. shocks on commuting which are specific to the household k and locations x and i. In this case, the function $t(x)$ is given by $\mathbb{E}\left[\max_i t_i(x)\nu_{kxi}\right]$ and households located at x distribute their commuting among employment centers according to a gravity equation (Gaigné *et al.*, 2020).

[4] 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).

2.3 The residential equilibrium

Given an amenity function $b(x)$, a given mass of heterogeneous households choose where to live, how much land and how much of the composite good to consume. The ω-households may be distributed over several locations. Let $\omega(x)$ be the mapping from N to $[\underline{\omega}, \bar{\omega}]$ that specifies which ω-households are located at $x \in N$ while $s(x, \omega) \in [0, 1]$ is the share of the ω-households who reside at x. The land market clearing condition holds if $\omega(x)$ satisfies the following condition:

$$|s(x, \omega(x))f(\omega(x))h^*(x, \omega(x))\mathrm{d}\omega| = \mathrm{d}x. \tag{4}$$

This condition states that the amount of land available between any x and $x + \mathrm{d}x > x$ along a topological arc while the subarc occupied by the households whose income varies from ω to $\omega + \mathrm{d}\omega$ are the same. Since $\omega(x)$ need not be monotonic, the land market clearing condition is expressed in absolute value.

Last, the population constraint implies

$$\int_N s(x, \omega(x))f(\omega(x))h^*(x, \omega(x))\mathrm{d}x = 1. \tag{5}$$

A 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:

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

with $x \in N$, such that

$$b(x) \cdot u[q^*(x, \omega^*(x)), h^*(x, \omega^*(x))] \geq b(y) \cdot u[q^*(y, \omega^*(x)), h^*(y, \omega^*(x))] \qquad \text{for all } x, y \in N$$

holds under the budget constraint (2), the land market clearing condition (4), and population constraint (5).

If the above inequality is strict for all $y \neq x$, then all $\omega^*(x)$-households are located at x ($s^*(x, \omega^*(x)) = 1$). Otherwise, there exist at least two locations x_1 and x_2 such that the $\omega^*(x)$-households are indifferent between the locations x_1 and x_2. In this case, we have $0 < s^*(\cdot, s^*(x)) < 1$ at x_1 and x_2, while the sum of the shares is equal to 1. Hence, there is *spatial splitting* of identical households.

Our problem has the nature of a matching problem between landlords and households. However, 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*. For example, for the same given housing consumption, a household can be indifferent between living close to its employment center while having a low level of amenities, or living far from the center 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.

Since $x(\omega)$ 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, we will show that those who reside at the same location x must share the same income. Therefore, we may define and characterize the *income mapping* $\omega^*(x)$ from the location set N 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 The residential equilibrium with amenities

Since u is strictly increasing in q, the equation $u(q,h) = U/b(x)$ has a single solution $Q(h, U/b(x))$, which describes the consumption of the numéraire when the utility level is $U/b(x)$ 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. Formally, the bid rent function is defined as follows:

$$\begin{aligned}\Psi(x,\omega,U) &\equiv \max_{q,h}\left\{\frac{\omega t(x)-q}{h}\,\middle|\,\text{s.t. } b(x)\cdot u(q,h)=U\right\}\\ &= \max_{h}\frac{\omega t(x)-Q(h,U/b(x))}{h}. \end{aligned} \tag{6}$$

When space is differentiated, since the bid rent $\Psi(x, \omega, U)$ is such that the ω-households are indifferent across locations, (3) implies that the Alonso-Muth condition for the ω-households is as follows:

$$h^*(x,\omega)R_x(x,\omega,U) - \omega t_x(x) = \frac{b_x(x)}{b(x)}\frac{u(q^*(x,\omega),h^*(x,\omega))}{u_q(q^*(x,\omega),h^*(x,\omega))}.$$

Since each household treats the utility level parametrically, applying the first-order condition to (6) yields the equation:

$$Q(h,U/b(x)) - hQ_h(h,U/b(x)) - \omega t(x) = 0 \tag{7}$$

whose solution, denoted $H(\omega t(x), U/b(x))$, 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).[5]

The budget constraint implies that the bid rent function may be rewritten as follows:

$$\Psi(x,\omega,U) \equiv \frac{\omega t(x) - Q(\omega t(x), U/b(x))}{H(\omega t(x), U/b(x))}. \tag{8}$$

[5] As shown by Fujita (1989), (7) may have several solutions. Our results hold true for any solution.

This expression shows that a household's bid rent at x depends separately on both $b(x)$ and $t(x)$ while its housing consumption H also varies with these two attributes of location x. 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 housing consumption at x of a ω-household when 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:

$$\frac{\partial\Psi(x,\omega,U^*(\omega)))}{\partial\omega} = 0, \tag{9}$$

while the second-order condition implies $\partial^2\Psi/\partial\omega^2 < 0$. Characterizing the class of utilities $u(q,h)$ and commuting costs $t(x)$ for which this condition holds would require technical developments which are beyond the scope of this paper. Importantly, we will show that this assumption holds for Stone-Geary preferences.

Totally differentiating (9) with respect to x yields:

$$\frac{\mathrm{d}\omega^*}{\mathrm{d}x} = -\left[\frac{\partial^2\Psi(x,\omega,U^*(\omega))}{\partial\omega^2}\right]^{-1}\cdot\frac{\partial^2\Psi(x,\omega,U^*(\omega))}{\partial\omega\partial x}, \tag{10}$$

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

Set

$$B(x) \equiv \frac{b_x(x)}{b(x)} \qquad T(x) \equiv -\frac{t_x(x)}{t(x)},$$

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.[6]

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

$$\phi(x,\omega) \equiv \left(1 - \frac{\varepsilon_{H,\omega} + \varepsilon_{u_q,\omega}}{\varepsilon_{U,\omega}}\right)B(x) - (1-\varepsilon_{H,\omega})T(x) \tag{11}$$

is positive (negative) *at this location.*

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

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

The expression (11) 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 the benchmark case of a monocentric, linear and featureless city $[0, L]$ where $b(x) = 0$ and $t_x(x) > 0$ for all x since the CBD is at $x = 0$. We know from Fujita (1989) that household locations are determined by ranking the bid rent slopes with respect to income. It follows from (11) 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 featureful monocentric and linear city ($B(x) \neq 0$). 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 (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 $[0, L]$ because, as shown in Section 2, the slope of the income gradient changes, so that the bid rent functions may intersect several times. In this case, there is imperfect sorting: household income rises over some range of sites and falls over others. We develop this argument in more detail in the next section.

The expression (11) 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 (11) may be used to determine how households are distributed according to the behavior of $B(x)$ and $T(x)$ by calculating those elasticities. In the limit, when the elasticities are constant, the sign of (11) 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 $\left(1-\left(\varepsilon_{H,\omega}+\varepsilon_{u_q,\omega}\right)/\varepsilon_{U,\omega}\right)B(x)$.

4 The social structure under Stone-Geary preferences

4.1 Non-homothetic 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.

It is common place to work with homothetic preferences, as they include the CES, Cobb-Douglas and translog. Here, there are at least two reasons to rule out such preferences. First, using homothetic preferences implies that the elasticity of the housing demand is constant in price and income, which contradicts its variability across heterogeneous consumers (Albouy *et al.*, 2016). Second, when the utility u is homogeneous linear, we show in Appendix A.2 that $\varepsilon_{U,\omega}=\varepsilon_{H,\omega}=1$ and $\varepsilon_{u_q,\omega}=0$, so that (11) yields $\phi(x,\omega)=0$ for all x. In other words, *there exists a continuum of residential equilibria under homothetic preferences.*

For these reasons, we consider Stone-Geary preferences, which obviate the multiplicity of equilibria and account for the plausible assumption of a minimum lot size $\overline{h}>0$:

$$u(q,h)=q^{1-\mu}(h-\overline{h})^{\mu}, \tag{12}$$

where $0<\mu<1$. Maximizing (12) with respect to q and h subject to (2) leads to the following demands for the numéraire:

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

and for housing:

$$h^*(x,\omega)=(1-\mu)\overline{h}+\mu\frac{\omega t(x)}{R(x)}. \tag{13}$$

Thus, the demand for housing at any location x increases less than proportionally with income (Albouy *et al.*, 2016). Note that the housing demand is perfectly inelastic when $\mu = 0$, which corresponds to the assumption of a fixed lot size.

Under (12), we show in Appendix A.3 that the sign of $\Psi_{x\omega}$ is the same as that of

$$B(x) - (1 - \mu)T(x), \tag{14}$$

which accounts for the presence of amenities and commuting costs, as well as for the relative intensity of preferences for housing through the parameter μ.

Set

$$\Delta(x) \equiv b(x)[t(x)]^{1-\mu}, \tag{15}$$

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. In other words, as the intensity of preference for housing increases, commuting matters less than the accessibility to amenities. More importantly, differentiating (15) shows that $\Delta_x(x)$ and $\phi(x, \omega)$ have the same sign. Hence, $\phi(x, \omega)$ changes sign at any extrema of the location-quality index. We assume without much loss of generality that $b(x)$ and $t(x)$ are such that $\Delta(x)$ is never flat on a positive measure interval.

Although we assume Stone-Geary preferences, our results hold true whenever the location-quality index $\Delta(x)$ is a function of $b(x)$ and $t(x)$ which is independent of ω. To illustrate, consider $u(q, h) = q^{\rho_1} + h^{\rho_2}$ with $0 < \rho_i < 1$ and $\rho_1 \neq \rho_2$. The elasticity of substitution between land and the numéraire is variable and equal to $1/(1 - \delta_1\rho_1 - \delta_2\rho_2)$ where δ_i is the expenditure share on good $i = 1, 2$. When $\rho_1 > \rho_2$, i.e., the composite good matters more than land, it can be shown that the above preferences generate the index $\Delta(x) \equiv [b(x)]^{1/\rho_1}\, t(x)$, which is similar to (15).

4.2 The spatial income distribution

Our objective is now to determine the mapping $\omega^*(x)$ from N to $[\underline{\omega}, \bar{\omega}]$ that specifies which ω-households are located at x under (12). 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 incomes are distributed within the location set N according to the values of the location-quality index. To show this, we first rank the values of $\Delta(x)$ by increasing order and denote by $G(\Delta)$ the corresponding c.d.f. defined over the domain $[\underline{\Delta}, \overline{\Delta}]$ where $\underline{\Delta}$ ($\overline{\Delta}$) is the minimum (maximum) value of $\Delta(x)$ over N.

The following proposition is proved in Appendix A.4.

Proposition 2. *Assume Stone-Geary preferences. Then,* (i) *each location hosts at most one household type;* (ii) *there exists a unique residential equilibrium;* (iii) *the equilibrium income mapping* $\omega^*(x)$ *and the location-quality index* $\Delta(x)$ *vary together with* x; *and* (iv) *the equilibrium income mapping is given by*

$$\omega^*(x) = F^{-1}[G(\Delta(x))]. \tag{16}$$

In Appendix A.4, we also show that the equilibrium utility level satisfies the Spence-Mirrlees condition, thus implying the existence of a positive assortative matching between incomes and the values of the location-quality index. Consequently, it is sufficient to study how $\Delta(x)$ varies, rather than $b(x)$ and $t(x)$ separately, to determine the properties of the residential equilibrium: the initial two-to-one matching is reduced to a one-to-one matching. Stated differently, there is a unique one-to-one and increasing relationship between ω and Δ (Chiappori, 2017). Hence, the highest income households locate where the location-quality index Δ reaches its maximum. As $\Delta(x)$ starts decreasing with x, the income level of the corresponding residents also decreases. The lowest income households reside at a global minimizer of the location-quality index. Around this location, the income level rises together with Δ. As a result, income sorting does not translate into spatial sorting because the function $\Delta(x)$ is in general not monotonic in x. In other words, we have:

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

For example, in a monocentric city, a wider income gap is no longer matched with a greater distance between two households.

In the absence of amenity effects ($b(x)$ is constant), households' residential choices are driven only by the distance to employment locations. Since the more productive workers have a higher bid rent, they will disproportionately choose residential locations in the neighborhood of the employment centers. By contrast, unevenly distributed historic and natural amenities are likely to attract the high-income people away from the employment centers. In sum, *the spatial sorting of income-heterogeneous households is governed by the location-quality index.*

To illustrate, consider Figure 1. The centrality of the city is described by the unique global maximizer $x = 0$ of $\Delta(x)$ over $[0, L]$ 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 $(x_2, L]$, the function $\Delta(x)$ falls again, which means that households' income decreases with x.

Since $\Delta(0) > \Delta(x_2) > \Delta(L) > \Delta(x_1)$, the intermediate value theorem implies that z_1 in $[0, x_1)$, z_2 in (x_1, x_2) and z_3 in $(x_2, L]$ 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, there is spatial splitting because *the 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 1 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 1 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 $d(x_2, z) \neq d(x_1, z)$ 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. Therefore, unlike Tiebout's prediction, *identical or similar households may live in spatially distinct areas.*

The equilibrium values of the shares $s(x, \omega(x))$ are determined as follows. 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 $j = 2, ..., n$. Since $H(\omega t(x), U/b(x)) \equiv H(\omega, \Delta(x), U)$ depends only on Δ under Stone-Geary preferences and $f(\omega^*(z_1)) = f(\omega^*(z_j))$ for $j = 2, ..., n$, we get (see Appendix A.4):

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

Furthermore, it must be that

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

The unique solution to these n linear equations is $s(z_i, \omega^*(z_i)) = 1/n$ for $i = 1, ...n$ (see Appendix A.4 for a more detailed proof). That is, the households who share income $\omega^*(z_1)$ are equally split across the locations that generate the same location-quality index $\Delta(z_1)$.

Before proceeding, it is worth explaining in more details how the arguments developed above are used to determine the residential equilibrium along the network N. Amenities and commuting costs are distributed along each topological arc a_τ. The location-quality index at $x \in a_\tau$ is specific to each topological arc and given by $\Delta(x; a_\tau) = b(x; a_\tau)[t(x; a_\tau)]^{1-\mu}$. A household chooses the arc a_τ and the location $x \in a_\tau$ that maximize her utility. Households ordered by decreasing incomes are assigned to arcs and locations by decreasing values of the location-quality index. Note that ω-households can reach their highest utility for several pairs of arcs and locations. In other words, households sharing the same income may occupy separated locations along the same topological arc or locations belonging to different arcs.

4.3 Land rent

It remains to characterize the equilibrium land rent. We show in Appendix A.5 that the equilibrium land rent is given by the following expression:

$$R^*(x) = \frac{\omega^*(x)t(x)}{H[\Delta(x), \omega^*(x), U^*(\omega^*(x))]}\left[1 - \frac{1-\mu}{\varepsilon_{U,\omega}(x)}\right], \tag{17}$$

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 (17) and rearranging terms, we obtain:

$$R^*(x) = \frac{\mu\omega^*(x)t(x)}{H[\Delta(x), \omega^*(x), U^*(\omega^*(x))] - (1-\mu)\overline{h}} > 0, \tag{18}$$

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

By totally differentiating (17) with respect to x, we obtain (see Appendix A.5):

$$R^*_x(x) = \frac{\omega^*(x)t(x)}{H[\Delta(x), \omega^*(x), U^*(\omega^*(x))]}\left[\frac{1}{\varepsilon_{U,\omega}(x)}B(x) - T(x)\right]. \tag{19}$$

Since $\varepsilon_{U,\omega}(x) > 1$, the above expression implies that the land rent gradient is always negative if $B(x) - T(x) < 0$ for all x. As x rises, the decreasing land rent compensates the households $\omega^*(x)$ located at x 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 $B(x) = 0$ and $T(x) > 0$ the land rent gradient is always negative.

When $B(x) - T(x) > 0$ 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 $B(x) - \varepsilon_{U,\omega}(x)T(x) > 0$. In this case, moving toward locations with more amenities ($B(x) > 0$) is sufficient for the land rent to increase. In short, the interaction between amenities, commuting and income sorting may give rise to a wealth of land rent profiles, which generally differ from that of the location-quality index alone.

4.4 From theory to data

To derive testable predictions about the effects of amenities and commuting costs on the income distribution within the city, we have to determine the explicit form of the income mapping $\omega^*(x) = F^{-1}[G(\Delta(x))]$. 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 $s_\omega > 0$: $F(\omega) = \exp\left[-(\omega/s_\omega)^{-\gamma_\omega}\right]$ over $[0, \infty)$. 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. $G(\Delta) = \exp\left[-(\Delta/s_\Delta)^{-\gamma_\Delta}\right]$ over $[0, \infty)$; the density is denoted $g(\Delta)$. The location-quality index covers a wider range of values when γ_Δ decreases.[7]

Using (16), the mapping $\omega^*(x)$ can then be retrieved from the condition:

$$\int_{\omega^*}^{\infty} f(y)\mathrm{d}y = 1 - \exp(-(\omega^*/s_\omega)^{-\gamma_\omega}) = \int_{\Delta}^{\infty} g(\delta)\mathrm{d}\delta = 1 - \exp(-(\Delta/s_\Delta)^{-\gamma_\Delta}),$$

which is the counterpart in the Δ-space of the land market clearing condition (4).

Setting $\gamma \equiv \gamma_\Delta/\gamma_s$ and solving the above equation yields the equilibrium income mapping:

$$\omega^*(x) = s_\omega \left[\frac{\Delta(x)}{s_\Delta}\right]^{\gamma}. \tag{20}$$

Last, we show in Appendix A.6 that the equilibrium land rent at x is given by

$$R^*(x) = \mu(1-\mu)^{\frac{1-\mu}{\mu}} k^{-\frac{1}{\mu}} t(x) \left[\Delta(x)\right]^{\frac{1}{\mu}} \left[\frac{\mu t(x)}{R^*(x)} + \frac{(1-\mu)\overline{h}}{\omega^*(x)}\right]^{\frac{1}{(1-\mu)\mu\gamma}}, \tag{21}$$

where k is a positive constant.

Toward an econometric specification. In the data, amenities and commuting costs are functions defined over a two-dimensional space. However, after having calculated the values of these functions at each location, we can collapse the two dimensions into one and order locations along the real line. In doing so, we run the risk of attributing different values of amenities and commuting costs to the same location. Since the number of locations in the dataset is discrete, the probability of such an event is zero. In this case, we can use the equilibrium mapping (20) to quantify the sorting consequences of the spatial distribution of amenities.

Assume that labor time is given by $t(x) \equiv \psi(x)[\tau(x)]^{-\theta}$, where $\tau(x)$ is the commuting time, $\theta > 0$ is the elasticity of labor time with respect to commuting time, and $\psi(x)$ is the

[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.

given number of working hours per year of the household residing at x.[8] In this case, the location-quality index becomes:

$$\Delta(x) = [b(x)]^{\beta} [\psi(x)]^{1-\mu} [\tau(x)]^{-\theta(1-\mu)}. \tag{22}$$

It follows from (20) that the income mapping is given by the following expression:

$$\omega^*(x) = s_\omega \left\{ \frac{[b(x)]^{\beta} [\psi(x)]^{1-\mu} [\tau(x)]^{-\theta(1-\mu)}}{s_\Delta} \right\}^{\gamma}.$$

Let $\tilde{\omega}(x)$ be the gross hourly income of a household observed in the data:

$$\tilde{\omega}(x) = (\omega^*(x)/\psi(x))\xi(x) \tag{23}$$

where the $\xi(x)$ are hourly labor income shocks that are independently and identically distributed according to some given distribution defined on $[0, \infty)$. Taking the log of (23), we obtain:

$$\log \tilde{\omega}(x) = \alpha_0 + \alpha_1 \log b(x) + \alpha_2 \log \tau(x) + \alpha_3 \log \psi(x) + \tilde{\xi}(x), \tag{24}$$

where $\alpha_0 \equiv \log(s_\omega / s_\Delta^\gamma)$, $\alpha_1 \equiv \beta\gamma$, $\alpha_2 \equiv -\theta(1-\mu)\gamma$ and $\tilde{\xi}(x) \equiv \log \xi(x)$. Note that we refrain from interpreting α_3 as being equal to $(1-\mu)\gamma$ because the number of working hours may have a direct effect on income, e.g., by working hard in order to get a wage raise (Bell and Freeman, 2001). At the same time, changes in wages may affect labor supply.

5 Data and descriptives

5.1 Datasets

We have gained access to various nationwide non-public microdata from *Statistics Netherlands* between 2010 and 2015. Unlike the United States or the United Kingdom, the Netherlands does not undertake censuses to register their population, but the register is constantly updated when people move or when there are changes in the household composition. The first dataset we use is the *Sociaal Statistisch Bestand* (*SSB*), which provides basic information on demographic characteristics, such as age, country of birth, marital status, and gender. We only keep people that could be part of the working population, that is, those between 18 and 65 years and aggregate these data to the household level. Importantly, the *SSB* data enable us to determine where households reside, up to the postcode level. Hence, space is discrete in the plane.

[8] We are agnostic about the reasons that explain why ψ varies with x. According to Rosenthal and Strange (2008), professionals work longer in denser areas, while Black *et al.* (2014) observe that married women work more when commutes are shorter.

The data on yearly income of households is are obtained from the *Integraal Huishoudens Inkomen* panel dataset. These data are based on the tax register, which provides information on taxable income, tax paid, as well as payments to or benefits from property rents or dividends. The income data also provide information on whether households are homeowners or renters. More than 65% of the rental sector applies to 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 70% of the data.[9]

To estimate the commuting time for each household, we use the tax register information, which provides information on individual jobs and the number of hours worked in each firm for each year. Using data on location information on each establishment from *ABR Regio* and network travel time from *SpinLab*, we calculate for each household the average commuting time. More information on how we calculate the commuting time between locations is provided in Appendix B.1.

A location is given by a *neighborhood* (which is a postcode 4-digit location) as defined by Statistics Netherlands. There are 4,033 neighborhoods; the median size of a neighborhood is only 528ha while the average population is about 4,000. Information on land values and lot sizes is not directly available. As is common practice, we infer them from data on housing transactions, provided by the *Dutch Association of Real Estate Agents* (*NVM*). The methodology used to calculate land values and lot sizes is described in Appendix B.2. The *NVM* data contain 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 m^2), 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.[10] We also know whether the house is a listed building.

We are interested in the impact of amenities on income sorting and land prices. We proxy the amenity level by the *picture density* in a neighborhood. More specifically, 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. We show in Appendix B.3 that there is a strong positive correlation

[9] We furthermore obtain information on the educational level of adults in the household. This is available for only 75% of the population, but our main specifications will not use these data, so this does not appear to be an issue.

[10] 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$m^2$of inside floor space, are smaller than 25m^2, or have lot sizes above 5000m^2. These selections consist of less than one percent of the data and do not influence our results.

between picture density and historic amenities or geographical variables, such as access to open water or open space. There are, however, several issues with using geocoded pictures as a proxy for amenities.[11]

First, to avoid the possibility of inaccurate geocoding, we keep only one geocoded picture per location defined by its geographical coordinates.[12] This reduces the number of pictures by about 50%. Second, one may argue that the patterns of pictures taken by tourists and residents may be very different. Since we have information on users' identifiers, we can distinguish between residents' and tourists' pictures by keeping users who take pictures for at least 6 consecutive months between 2004 and 2015 in the Randstad. It seems unlikely that tourists stay for 6 consecutive months in the area. Note that the correlation between residents' and tourists' pictures is 0.653, which is rather low.

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, which comprises information on the universe of buildings. 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 0.223. Last, we recognize that people who take pictures may belong to a specific socio-demographic group (e.g., young people with a smartphone) by including demographic controls and using instrumental variables.

Though imperfect, we believe that the picture density is probably the best proxy available for the relative importance of urban amenities at a certain location because it captures both the heterogeneity in aesthetic quality of buildings and residents' perceived quality of a certain location. Nevertheless, we test the robustness of our results using a completely different, hedonic, amenity index in the spirit of Lee and Lin (2018) (see Appendix B.3 for more details). The hedonic index aggregates the average impact of several proxies of amenities, such as the locations of historic buildings, proximity to open space and water bodies, by testing their joint impact on house prices. We also construct historic instruments. Knol *et al.* (2004) have scanned and digitized maps of the land use in 1900 into 50 by 50 meter grids and classified each grid into 10 categories, including built-up areas, water, sand, and forest. We aggregate these 10 categories into 3 categories: built-up areas, open space, and water bodies and calculate the share of the area used for each type in each neighborhood. We further gather data from the 1909 census on occupations and employment in each municipality. Those ones were much smaller than current

[11] Ahlfeldt (2013) shows that in Berlin and London the picture density is strongly correlated to the number of restaurants, music nodes, historic amenities and architectural sites, as well as parks and water bodies.

[12] In a continuous space, the probability that several pictures are taken at *exactly* the same location is zero. Hence, observing multiple pictures at the same location is likely caused by inaccurate geocoding.

ones and about 4 times the size of the current neighborhoods.

For each occupation we obtain the required skill level. This enables us to calculate the share of households who are medium and high-skilled. We gather additional data on the railway network in 1900 and the stations which by then existed (see Appendix B.4 for more information), enabling us to calculate employment accessibility in 1909.

5.2 Descriptive statistics

Figure 2A provides a map of the Netherlands, the studied area, where we indicate the most important cities. The conurbation formed by the four largest cities, i.e., Amsterdam, Rotterdam, The Hague, and Utrecht is known as the Randstad, which has a population of about 7.1 million. Figure 2B displays the commuting pattern across neighborhoods and shows that the Dutch urban structure is really polycentric as many commuting flows occur between different cities. This underlines the need for a model that allows for location choices in the whole country. Figure 2C is a map of the most important roads and railways that form the transportation network in the Netherlands.

[Figure 2 about here]

We report descriptive statistics of the 10, 213, 524 households of our sample in Table 1. The average (median) yearly income is €91, 535 (€86, 732). Incomes are approximately Fréchet distributed (see Appendix B.5).[13] The average land price in the sample is €1, 312 , but there are stark spatial differences. For example, in the capital Amsterdam, it is €3, 046, while in the rural province of Friesland it is only €716. As expected, the correlation between the estimated land price and lot size is negative ($\rho = -0.245$). The average lot size is 364m^2. However, in Amsterdam it is only 253m^2, which corresponds to the higher land values in this city. About 15% of households occupy apartments and the correlation between occupying an apartment and the land price is positive ($\rho = 0.153$).

[Table 1 about here]

The picture density, i.e., the proxy for amenity endowments, range from 0 to 231 pictures per hectare. Only 0.2% of the households live in neighborhoods that do not have any pictures. We will disregard those households. The average picture density in Amsterdam (22.7) is much higher than in Rotterdam (9.63), The Hague (6.17), and Utrecht (7.66). Recall that we only use pictures *outside* a building taken by *residents* in determining the amenity index. It appears that 80% of the pictures are taken outside a building while about 60% of the pictures are

[13]We report maps and histograms of income and land prices in Appendix B.5.

taken by local residents. Going back to Table 1, we see that the average commuting time is 26 minutes, which is very close to statistics provided by other sources (Department of Transport, Communications and Public Works, 2010). The unconditional correlation of picture density with income is close to zero ($\rho = 0.0533$), but this is not very informative as we do not control for household characteristics. The correlation of picture density with land prices is substantially higher ($\rho = 0.431$). Finally, *households that have a short commute do not seem to live in high amenity locations* as the correlation between picture density and commuting time is low ($\rho = -0.0454$).

The descriptive of the historic instruments that we use are described in Table B.6 of Appendix B.4.

6 Reduced-form analysis

In this section, we provide reduced-form evidence in support of the model's qualitative predictions that amenities and commuting time affect the income distribution in and between cities; and hence that exogenous amenities are an important determinant of sorting of households. The analysis is complemented by a wide range of controls that provide evidence against alternative possible explanations.

6.1 Econometric framework and identification

We consider the *income mapping*, which plays a key role in our model, and provide *reduced-form* evidence that the picture density is related to observed proxies for amenities and, in turn, we show that sorting by incomes is indeed related to our proxy for amenities and commuting time – the variables that constitute the location-quality index (see (22)). In line with (24), set

$$\log \tilde{\omega}_{ik}(x) = \alpha_1 \log \tilde{b}(x) + \alpha_2 \log \tilde{\tau}_i(x) + \alpha_3 C_k + \nu_{\mathcal{A}}(x) + \upsilon_i + \xi_{ik}(x), \tag{25}$$

where $\tilde{\omega}_{ik}(x)$ is the observed income of household k living at x and working in i; $\tilde{b}(x)$ is the density of geocoded pictures –our proxy for amenities–, $\tilde{\tau}(x)$ is the observed commuting time to workplace i, C_k are household characteristics, $\nu_{\mathcal{A}}(x)$ are travel-to-work-area $\mathcal{A}$ fixed effects, υ_i are workplace fixed effects, and $\xi_{ik}(x)$ is an error term. The parameters α_1, α_2, α_3 , $\nu_{\mathcal{A}}(x)$ and υ_i are estimated.

There are several issues when using (25) to identify the causal impact of $\tilde{b}(x)$ and $\tilde{\tau}(x)$ on sorting on the basis of income. First, regarding commuting time $\tilde{\tau}(x)$, unconditional correlations between incomes and commuting times are generally positive rather than negative (see Susilo and Maat, 2007 for the Netherlands). There are several reasons for that. Higher income and educated people are more specialized and, therefore, operate in 'thinner' labor markets.

Given that there is a strong idiosyncratic component to residential location choices (people are strongly attached to a location and usually dislike moving), this will imply that people with higher incomes live further away from their workplace (see, e.g., Manning, 2003). Another reason for a bias is that labor markets may not be fully competitive as households may bargain over to get an income compensation for living further away. Hence, observed incomes $\tilde{\omega}_{ik}(x)$ may be higher when people live further away. Note that about 15% of the costs of a longer commute is paid by the employer (Mulalic *et al.,* 2013).

Second, a more general concern about α_1 and α_2 as measures of the impacts of amenities and commuting time on the spatial income distribution is that there is an omitted variable bias due to sorting, heterogeneity in preferences for housing quality, agglomeration economies at the workplace, and unobserved spatial features. More specifically, households may not only sort on the basis of income, but also on the basis of other household characteristics. Households with children, for example, may aim to locate in neighborhoods with a large amount of green space. The variables $\tilde{b}(x)$ and $\tilde{\tau}(x)$ could also be correlated with unobserved housing attributes because households with different incomes may have different preferences for housing quality, such as the age of the housing stock (Brueckner and Rosenthal, 2009). For example, a large share of the housing stock in the city center of Amsterdam takes the form of apartments. This may imply that the affluent are not willing to locate there because they eschew apartment living (Glaeser *et al.*, 2008).

Third, there may be reverse causality between $\tilde{\omega}_{ik}(x)$ and $\tilde{b}(x)$ and between $\tilde{\omega}_{ik}(x)$ and $\tilde{\tau}(x)$. For example, the provision of amenities may be a direct result of the presence of high-income households. Indeed, anecdotal evidence suggests that cultural and leisure services are often abundantly available in upscale neighborhoods (Glaeser *et al.*, 2001). Similarly, high income neighborhoods may attract employers that are in need of specialized and highly educated labor. Last, since we do not observe the 'exact' amenity level, there may be a measurement error in $\tilde{b}(x)$, which may lead to a downward bias of α_1 when the error is random.[14]

The first step to mitigate the biases associated with these concerns is first to 'purge' *household*, *job* and *housing* characteristics, C_k, from neighborhood characteristics. For example, C_k captures the members of the households who work full-time or part-time, the size of the household and the age of the adults, while housing attributes are, for example, housing type and construction year. This approach reduces the likelihood that we measure sorting on the

[14] As suggested by the literature on local public goods, there might be reverse causality, meaning that the location of local public goods and jobs is determined by the spatial income distribution. To a large extent, this is because the institutional context that prevails in the U.S. 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.

basis of household characteristics other than incomes. We also control for travel-to-work-area fixed effects to address the concern of aggregate sorting effects between urban areas, as well as differences in spatial policies that are often spatially autocorrelated. Furthermore, by including workplace fixed effects υ_i we control for productivity differences (e.g., due to agglomeration economies) at the workplace.

The second step is to instrument for commuting times by constructing a measure of accessibility to employment that is unrelated to characteristics of individual households (e.g., the level of human capital). We define employment accessibility as:

$$a(x) = \sum_{i=1}^{I} F(\tau_i(x)) n_i. \tag{26}$$

In other words, at location x we weight the number of jobs n_i at i by the share of people whose commute is at most equal to $\tau_i(x)$.

Working with an endless string of controls will not fully address the endogeneity concerns raised above. Unfortunately, our data do not allow us to exploit quasi-experimental or temporal variation in $\tilde{b}(x)$ and $\tilde{\tau}(x)$. Therefore, to investigate the importance of the omitted variable bias we analyze coefficient movements after including controls. Oster (2019) shows that coefficient movements together with changes in the R^2 can be used to estimate biased-corrected coefficients. We outline this procedure and discuss the results in detail in Appendix C.2.

The omitted variable bias is not the only endogeneity issue. Our proxies may also suffer from measurement error and reverse causality. We will, therefore, rely on instrumental variables. Our first set of specifications uses contemporary instruments, while our second set of specifications appeals to historic instruments. Regarding contemporary instruments for amenities, we use a set of observed, arguably exogenous, proxies for amenities, such as the listed building density, the share of the neighborhood x that is in a historic district, as well as the share of built-up areas and water bodies. By using other proxies for amenities, the measurement error of $\tilde{b}(x)$ is likely to be mitigated. One may argue that the contemporary instruments do not convincingly address the issue of unobserved locational and household characteristics that may be correlated with $\tilde{b}(x)$. Moreover, they do not address the potential endogeneity of accessibility $\tilde{\tau}(x)$.

Alternatively, we exploit the fact that $\tilde{b}(x)$ and $\tilde{\tau}(x)$ are autocorrelated. First, land use in 1900 is used as an instrument. We expect aesthetic amenities to be positively correlated to the share of built-up area in 1900. For example, the historic city center of Amsterdam has many buildings that have been built before 1900, which 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. As an instrument for commuting time, we count the total number of households $E_{x,1909}$ in 1909 within a commuting distance by using the railway network in 1900:

$$a_{1909}(x) = \sum_{i=1}^{n} F(\tau_i(x))n_{i,1909}, \tag{27}$$

where $\tau_i(x)$ is the commuting time between x and employment location $i = 1, ..., n$, while $F(\tau_i(x))$ is the share of people who commute at most τ minutes in the sample (see Appendix B.1). Hence, $F(\tau_i(x))$ represents the aggregate cumulative distribution of commuting times, while $n_{i,1909}$ is the total employment at i in 1909. Because of temporal autocorrelation, we expect that a better employment accessibility in 1909 also implies a better employment accessibility today and, therefore, a lower commuting time.

Historic instruments can be criticized because of the (strong) identifying assumption that past unobserved locational features are correlated to current unobserved locational endowments. However, these instruments are more likely to be valid in the context of income sorting 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 as amenities. For example, the canals in Amsterdam were essentially open sewers (Geels, 2006). Therefore, locations near a canal often repelled high-income households who located in lush areas just outside the city. It was also before cars became the dominant mode of transport. People around 1900 often walked to their working place, so that commuting distances were short. However, the rich could afford to live outside the city and take the train to their workplace. The cities in 1900 were not yet influenced by (endogenous) planning regulations, as the first comprehensive city plans date from the 1930s.

Still, one may be concerned that the measure of amenities is itself determined by the wealth of individuals who locate there. The reason is that unobservables that determine the concentration of wealthy individuals in the past also determine the locations of landmarks today, and thus determine where pictures are taken. Moreover, one may argue that historic employment accessibility, which is correlated to current commuting time, makes it easier to find jobs for all household members, and thus increases household income due to better matching, rather than shorter commutes. We address these concerns in several ways.

1. We estimate specifications where we control for the current share of built-up areas and population density. Locations that were attractive in the past attracted people and consequently have a high share of built-up area in 1900. The share of built-up areas in 1900 is likely to be correlated to the current population density and to shares of built-up areas nowadays. By controlling for the current share of built-up areas and population density we mitigate the issue that our proxy for amenity just captures contemporary population density, rather than a higher amenity level because of the historic buildings.

2. We gather data from the 1909 census on occupations and skills in each municipality. We then control in various ways for the average skill level of households in 1909 as a proxy for the income in the past. Controlling for the skill level should also address the issue that employment density in 1909 may be correlated to better matching opportunities. Since this proxy may be imperfect, we also use the share of Protestants in 1899 at the municipality level as another proxy for income/skill. Indeed, at that time Protestants had a higher education level and were wealthier.

3. We also consider another instrument for employment accessibility. From the 1899 census, we gather data on the share of locally born people (i.e., within the same municipality). If the (lack of) mobility of households is correlated over time, the share of locally born people should be correlated positively to current commuting times because immobile households have to commute on average longer to their jobs.

4. Finally, we estimate specifications where we exclusively focus on areas of reclaimed land since 1900. These are areas that are reclaimed from the sea (about 5% of the land) just before and after World War II. As these reclaimed locations are otherwise identical, and as no one was living in those locations at that time, we address reverse causality, and strongly mitigate any remaining omitted variable bias.

6.2 Reduced-form results

Baseline results. We first check that picture density is a good proxy for amenity endowments. The results in Appendix B.3 show the expected signs: there is a higher picture density areas where there are many historic buildings, in built-up areas, and in areas with more water bodies (e.g., the Amsterdam canal district). The same holds if we control for sorting on other household characteristics (e.g., sorting based on household composition) and commuting time.

Table 2 reports the baseline reduced-form results for the income mapping. Column (1) shows a simple regression of log income on log amenities and log commuting times, while we only control for demographic characteristics and year fixed effects. This shows that more amenities and a better accessibility are associated with higher incomes. Doubling amenities implies an increase in neighborhood income of $(\log 2 - \log 1) \times 0.0387 = 2.7\%$. Doubling of commuting time seems to attract households whose incomes are 4.6% *higher* – contrary to the expectations. In column (2), we add a wide array of controls related to housing quality and add travel-to-work-area fixed effects to address spatial heterogeneity. Although the R^2 increases by about 30%, the coefficients related to amenities and accessibility are hardly affected. This suggests that amenities and commuting time are not so much correlated to building quality and aggregate characteristics of the urban area.

[Table 2 about here]

In column (3), we instrument for commuting time to address the issue that more skilled and specialized people often work in thinner labor markets and therefore commute longer. Our instrument is then the accessibility to jobs given the aggregate commuting time distribution (see (26)). The first-stage results reported in Appendix C.1 show that employment accessibility is a strong instrument: a 10% increase in employment accessibility is associated with a reduction in commuting time of 1.8-2.9%. Going back to Table 2, we observe in column (3) that this addresses the upward bias of the commuting time effect. The coefficient implies that when commuting time doubles, this attracts households whose incomes are 26% lower. It seems that the effect of amenities is now somewhat lower: a 100% increase in the amenity level attracts households whose incomes are 0.6% higher. Column (4) includes job characteristics and workplace fixed effects to address any productivity effects and agglomeration economies arising at the workplace, which may make workers more productive. This indeed explains part of the negative commuting effect, as the coefficient is now more than 60% lower.

Despite the inclusion of controls, travel-to-work-area and workplace fixed effects, one may argue that we do not convincingly address the omitted variable bias. We deal with this issue by estimating bias-corrected regressions following Oster (2019) in Appendix C.2. We show that when we choose the appropriate maximum attainable R^2 (as only part of the variation in incomes can be explained by variables varying at the neighborhood level), the estimates are very close to the IV estimates we present below (so a positive amenity elasticity of around 0.025 and a commuting time elasticity of -0.05--0.45). This strongly suggests that the omitted variable bias is not a major issue in the preferred estimates.

In column (5) we further aim to address potential measurement error in the picture density and address reverse causality issues by instrumenting amenities and commuting time with historic variables. The instruments are the shares of water bodies and of built-up area in 1900 within a neighborhood x, within 500m and between 500 and 1000m, as well as the number of households within commuting distance in 1909 using the railway network in 1900.[15] In Appendix C.2, we report the corresponding first-stage results. The share of built-up area, the share of water bodies in 1900 are strongly and positively correlated to the current amenity level, while employment accessibility in 1909 is negatively related to current commuting time. Overall, the Kleibergen-Paap F-statistic is above the rule-of-thumb value of 10 in all specifications, suggesting that the instruments are sufficiently strong.

Going back to Table 2, the coefficient of amenities is now somewhat higher: doubling amenities attracts households whose incomes are 2.6% higher; doubling commuting time implies an effect of -15%. Column (6) is the most comprehensive specification where we include job

[15] Since we have more instruments than endogenous variables, one might object that two-stage least squares estimates are biased (Angrist and Pischke, 2009). Hence, we also have experimented with other estimators that are (approximately) median unbiased, such as LIML or GMM estimators. The results are virtually identical. For this reason, we do not report them in the paper.

characteristics and workplace fixed effects. We consider this as the preferred specification. Doubling amenities attracts households whose incomes are 1.6% higher. Doubling commuting time leads to households whose incomes are 14% lower. Hence, the impact of commuting time seems to be somewhat stronger than the impact of amenities.

To sum up, the results unequivocally indicate that *the impact of amenities on income sorting is positive and highly significant.* As for the commuting time, *its effect on income sorting is negative and strong.* To compare the effects of commuting and amenities, it is informative to look at a standard deviation of a log change in commuting time or amenities. In the preferred specification, a standard deviation increase in log amenities attracts households whose incomes are about $1.656 \times 0.023 = 3.7\%$ higher. On the other hand, a standard deviation increase in log commuting time attracts households whose incomes are $0.645 \times 0.207 = 13.3\%$ lower. Hence, commuting time seems to be a more important driver of income sorting than amenities. However, the impact of amenities is far from negligible. For example, if we compare an area with the lowest amenity level (some rural area in the north) with the area with the highest amenity level (the center of Amsterdam), the predicted income difference is 28.6%, which is clearly non-negligible.

Alternative proxies for amenities. One may worry that our results hinge on the particular choice of the amenity index. We therefore consider an alternative proxy for amenity endowments. Following Lee and Lin (2018), we construct an aggregate hedonic amenity index that describes the amenity provision at every location using house prices. The procedure is described in Appendix B.3. We report the results in Table 3, which replicates the specifications in Table 2 but replaces picture density by the log of the hedonic amenity index.

[Table 3 about here]

The effect of commuting time is very much comparable to the results reported in Table 2. Regarding the hedonic amenity index, once we control for housing and location characteristics, we find an elasticity of about 0.8, implying that a 100% increase in the hedonic amenity index leads to an increase in income of 55%. We find a somewhat lower amenity elasticity in column (6) when we include control variables, workplace and travel-to-work-area fixed effects, and instrument for amenities and commuting time with historic variables. Doubling the amenity level is then associated with an increase in incomes of 36%.

To make the results comparable, we rescale the hedonic amenity index in such a way that the standard deviation of the log of the hedonic amenity index is the same as that of the log of the picture index. The estimated elasticity of the rescaled hedonic amenity index then varies between 1.6% and 2.4%, which is very much comparable to the impact of picture density.

Effects on land prices. We also investigate the reduced-form impacts of amenities and commuting times on land prices, so we estimate a reduced-form version of (18). In our setup the signs of the effects of amenities and accessibility on land prices and incomes are the same (although magnitudes may differ). Therefore, we now estimate the effects of amenities and commuting time on land prices. The results are reported in Table 4.

[Table 4 about here]

We start in column (1), Table 4, with a simple OLS specification including amenities and commuting time, while controlling for demographic characteristics. This leads to a strong positive effect of amenities on land prices: doubling amenities implies a land price increase of 13%, while doubling accessibility leads to land prices that are 2.5% higher. When we instrument for commuting time with accessibility to employment, control for workplace and travel-to-work fixed effects, the coefficient of commuting time becomes considerably stronger, while the impact of amenities is somewhat less strong. Column (4) shows that doubling commuting decreases land prices by 50%, while a 100% increase in the amenity level is associated with an increase in land prices of 4.0%. Like in the baseline results, the impact of amenities become considerably stronger once we use historic instruments in columns (5) and (6). The preferred specification in column (6) indicates that doubling amenities leads to an increase in land prices of 13% , which is sizable. The impact of commuting is somewhat imprecisely estimated, but the coefficient is negative. The point estimate indicates that a 100% increase in commuting time is associated with a decrease in land prices of 27%. Hence, the reduced-form effects on land prices do indeed have the same signs as the effects on income, but are stronger in magnitude. For example, if we compare the land price differential between the location with the lowest and the highest amenity level, it is 238%, which is considerable.

Other sensitivity checks. Appendix C shows that our results still hold for a wide range of alternative robustness checks and sample selections. To the extent one is still worried that endogeneity plagues our estimates, we strongly advise the reader to consult Appendices C.2 and C.3. More specifically, in Appendix C.2 we report Oster's bias-adjusted estimates, leading to similar results. In Appendix C.3 we show that our results hold if we (i) only focus on the urban area of the Randstad, (ii) include municipality fixed effects, (iii) control for current land use and population density, (iv) control for sorting based on skills in 1909, (v) use alternative (historic) instruments, and (vi) only use observations on land that was reclaimed from the sea.

Further robustness analyses minimize any measurement error regarding accessibility and workplace productivity, by running specifications where we only keep households (i) with a single job, (ii) with a single job in a single-plant firm, and (iii) with a company car who are more likely to use the car for commuting. We further test whether our results change when using the share of highly educated adults in the household, which is a measure that is very much correlated to income, but arguably is measured with little error. We find very similar effects, both in terms of sign and magnitude, which confirms that looking at income or skill levels is more or less equivalent. We also use commuting time by rail instead of commuting time over the road and focus on areas close to city centers of large cities. Overall, *the impact of amenities and commuting time on income sorting choice is robust.*

7 Concluding remarks

In this paper, we used a new setup in which any location is differentiated by two attributes, i.e., amenity endowments and commuting costs. The bid rent function of urban economics may be used to show that the uneven distribution of amenity endowments is sufficient to break down the perfect sorting of households across the space-economy. More specifically, our analysis suggests that *promoting equal access to amenities is likely to favor residential segregation*, whereas *a multi-modal provision of amenities across the city fosters income mixing.* Under Stone-Geary preferences, there exists a location-quality index that blends amenities and commuting costs into a single aggregate whose behavior drives households's residential choices. Studying this index allows us to gain insights about how governments and urban planners can design policies whose aim is to redraw the social map of cities. For example, the higher the index of a particular location, the higher the income of households who choose to locate there. The relevance of local amenity endowments and commuting costs to explain the residential choices of heterogeneous consumers is confirmed by the empirical analysis of where both effects are found to be significant.

Our analysis also suggests that the provision of local public goods (LPGs) may affect the urban structure in a way that appears socially desirable by choosing adequately the location of LPGs. That said, we now aim to develop a simple setting in the hope of shedding light on the role of *endogenous* amenities. Households at x choose their consumption level $c(x)$ of LPGs, which, like in U.S. cities, are financed by a property tax $\gamma(x)$. Hence, under the assumption of a fixed lot size ($\mu = 0$ and $h = 1$), we have $c(x) = \gamma(x)R(x)$. Assuming that amenities and LPGs are bundled into a Cobb-Douglas aggregate, preferences become $U = b^{\alpha}c^{1-\alpha}q$, with $0 < \alpha < 1$. Solving the utility-maximizing condition for the equilibrium tax rate for a ω-household at x yields $c^*(x) = (1-\alpha)\left[\omega t(x) - (1+\gamma)R(x)\right]$. Using $c^*(x)$ and applying the same approach as in Section 3, it is readily verified that

$$\Psi_{\omega x}(x, \omega, U^*(\omega)) = t(x)\left[\frac{\alpha}{2-\alpha}B(x) - T(x)\right].$$

In this case, the location-quality index becomes $\Delta(x) = [b(x)]^{\alpha/(2-\alpha)}\, t(x)$. Comparing this condition to (14) where $\mu = 0$ shows that, *the decentralized provision of LPGs weakens the impact of amenity endowments in individual residential choices.* In particular, when households do not value much historic or natural amenities (α is small) or when the level of such amenities is almost constant across space, residential choices will be mainly driven by commuting costs.

References

[1] Ahlfeldt, G.M. (2013) Urbanity. *London School of Economics*, mimeo.

[2] Ahlfeldt, G., S.J. Redding, D. Sturm, and K. Wolf (2015) The economics of density: Evidence from the Berlin Wall. *Econometrica* 83: 2127-89.

[3] 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.

[4] Albouy, D., G. Ehrlich and Y. Liu (2016) Housing demand, costs-of-living inequality, and the affordability crisis. *NBER Working Paper* 22816.

[5] Allen, T. and C. Arkolakis (2014) Trade and the topography of the spatial economy. *Quarterly Journal of Economics* 129: 1085-140.

[6] Altonji, J.G., T.E. Elder and C.R. Taber (2005) An evaluation of instrumental variables strategies for estimating effects of Catholic schooling. *Journal of Human Resources* 40: 791-821.

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

[9] Brinkman, J. and J. Lin (2019) Freeway Revolts! *Federal Reserve Bank of Philadelphia*, mimeo.

[10] 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.

[11] 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.

[12] Carlino, G.A. and A. Saiz (2019) Beautiful city: Leisure amenities and urban growth. *Journal of Regional Science* 59: 369-408.

[13] Chiappori, P.-A. (2017) Matching with Transfers. The economics of love and marriage. *Princeton: Princeton University Press.*

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

[15] Couture, V., C. Gaubert, J. Handbury and E. Hurst (2019) Income growth and the distributional effects of urban spatial sorting. *UC Berkley*, mimeo.

[16] Cuberes, D., J. Roberts and C. Sechel (2019) Household location in English cities. *Regional Science and Urban Economics* 75: 120-135.

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

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

[19] 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.

[20] Dustmann C., J. Ludsteck, and U. Schönberg (2009) Revisiting the German wage structure. *Quarterly Journal of Economics* 124: 843-81.

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

[22] Gaigné, C., H. Koster, F. Moizeau, and J.-F. Thisse (2020) Amenities and the social structure of cities. *CEPR Discussion Paper* 11958.

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

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

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

[26] 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.

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

[28] 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.

[29] Koopmans, C., P. Rietveld and A. Huijg (2012) An accessibility approach to railways and municipal population growth, 1840–1930. *Journal of Transport Geography* 25: 98-104.

[30] Koster, H.R.A. and T. Pinchbeck (2019). How do households value the future? Evidence from property taxes. *CEP Discussion Paper* 1571.

[31] Koster, H.R.A. and J. Rouwendal (2017) Historic amenities and housing externalities. Evidence from the Netherlands. *Economic Journal* 127: F396-F420.

[32] Lee, S. and J. Lin (2018) Natural amenities, neighborhood dynamics, and persistence in the spatial distribution of income. *Review of Economic Studies* 85: 663-94.

[33] Manning, A. (2003) The real thin theory: Monopsony in modern labour markets. *Labour Economics* 10: 105-31.

[34] Montesano, A. (1972) A restatement of Beckmann's model of distribution of urban rent and residential density. *Journal of Economic Theory* 4: 329-54.

[35] Mulalic, I., J.N. van Ommeren and N. Pilegaard (2013) Wages and commuting: Quasi-natural experiments' evidence from firms that relocate. *Economic Journal* 124: 1086-105.

[36] Oster, E. (2019) Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business and Economic Statistics* 37: 187-204.

[37] 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.

[38] Rosenthal, S.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, 1047-1120.

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

[40] Saiz, A., A. Salazar and J. Bernard (2018) Crowdsourcing architectural beauty: Online photo frequency predicts building aesthetic ratings. *PLOS One* 13: e0194369.

[41] Tsivanidis, N. (2019) Evaluating the impact of urban transit infrastructure: Evidence from Bogotá's TransMilenio. *UC Berkeley*, mimeo.

Tables

Table 1 – Descriptive statistics

	(1) mean	(2) sd	(3) min	(4) max
Gross income *(in €)*	91,535	53,683	3,589	999,897
Land price *(€per m^2)*	1,312	752.2	0.00753	22,418
Lot size *(m^2)*	364.3	923.8	25	24,824
Pictures per ha	2.189	8.840	0	231.9
Commuting time in minutes	26.39	17.18	0	120.0
Hedonic amenity index	2.821	0.0915	2.723	3.885
Share historic district	0.0347	0.139	0	1
Listed building	0.0941	0.699	0	17.06
Share built-up land	0.449	0.298	0.000856	1
Share water	0.0496	0.0738	0	0.813
Employment accessibility	624,940	275,990	14,427	1,347,124
Total hours worked in household	2,159	913.1	416.1	6,239
Household has company car	0.149	0.356	0	1
Works at single-establishment firm	0.443	0.497	0	1
Number of jobs in household	1.511	0.968	1	18
Person is male	0.521	0.215	0	1
Person is foreigner	0.0718	0.217	0	1
Age of person	41.99	9.008	18	64
Apartment	0.153	0.360	0	1
House built <1945	0.192	0.394	0	1

Notes: The number of observations is 10,213,540. For land price and lot size the number of observations is 2,196,280. Because of confidentiality restrictions the minimum and maximum values refer to the 0.01% and 99.99% percentile. This implies that we exclude the bottom and top 1,024 observations.

Table 2 – Baseline regression results

(Dependent variable: the log of pictures per ha)

		+ Controls, fixed effects	Contemporary instrument for commuting time		Historic instruments for amenities and commuting time	
	(1) OLS	(2) OLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
Pictures per ha *(log)*	0.0387***	0.0294***	0.0086***	0.0082***	**0.0368*****	**0.0226*****
	(0.0017)	(0.0013)	(0.0030)	(0.0016)	(**0.0075**)	(**0.0065**)
Commuting time *(log)*	0.0670***	0.0694***	**-0.3767*****	**-0.1536*****	**-0.2179****	**-0.2065*****
	(0.0013)	(0.0011)	(**0.0553**)	(**0.0220**)	(**0.0895**)	(**0.0657**)
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
Housing controls	No	Yes	Yes	Yes	Yes	Yes
Job controls	No	No	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Travel-to-work-area fixed effects	No	Yes	Yes	Yes	Yes	Yes
Workplace fixed effects	No	No	No	Yes	No	Yes
Number of observations	10,213,540	10,213,540	10,213,540	10,213,540	10,213,540	10,213,540
R^2	0.1996	0.2594				
Kleibergen-Paap F-statistic			134.3	364.9	14.29	17.81

Notes: **Bold** indicates instrumented. Household controls include household size, mean age of adults, mean gender, household type (couple, single, kids), the share of the household that is foreign-born. Job controls are the total hours worked, whether the household has a company car, the share of full-time contracts, the share of permanent contracts. Housing controls include house type, height of the building, construction year dummies and whether a building is listed. Standard errors are clustered at the neighborhood level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Table 3 – A hedonic amenity index

(Dependent variable: the log of household gross income)

		+ Controls, fixed effects	Contemporary instrument for commuting time		Historic instruments for amenities and commuting time	
	(1) OLS	(2) OLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
Hedonic amenity index *(log)*	1.2668***	0.8009***	0.8422***	0.7148***	**0.8061*****	**0.5216*****
	(0.1217)	(0.0978)	(0.0931)	(0.0673)	(**0.2020**)	(**0.1482**)
Commuting time *(log)*	0.0539***	0.0635***	**-0.4453*****	**-0.1965*****	**-0.4691*****	**-0.3317*****
	(0.0015)	(0.0011)	(**0.0470**)	(**0.0191**)	(**0.0543**)	(**0.0386**)
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
Housing controls	No	Yes	Yes	Yes	Yes	Yes
Job controls	No	No	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Travel-to-work-area fixed effects	No	Yes	Yes	Yes	Yes	Yes
Workplace fixed effects	No	No	No	Yes	No	Yes
Number of observations	10,233,133	10,233,133	10,233,133	10,233,133	10,233,133	10,233,133
R^2	0.1904	0.2553				
Kleibergen-Paap F-statistic			202.5	497.5	29.75	32.15

Notes: **Bold** indicates instrumented. Household controls include household size, mean age of adults, mean gender, household type (couple, single, kids), the share of the household that is foreign-born. Job controls are the total hours worked, whether the household has a company car, the share of full-time contracts, the share of permanent contracts. Housing controls include house type, height of the building, construction year dummies and whether a building is listed. Standard errors are clustered at the neighborhood level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Table 4 – Effects on land prices

(Dependent variable: the log of land price)

		+ Controls, fixed effects	Contemporary instrument for commuting time		Historic instruments for amenities and commuting time	
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	2SLS	2SLS	2SLS	2SLS
Pictures per ha *(log)*	0.1874***	0.0960***	0.0481***	0.0575***	**0.2052*****	**0.1910*****
	(0.0054)	(0.0032)	(0.0059)	(0.0042)	(**0.0212**)	(**0.0208**)
Commuting time *(log)*	-0.0357***	-0.0223***	**-1.0866*****	**-0.7154*****	**-0.3114**	**-0.3854***
	(0.0033)	(0.0016)	(**0.1095**)	(**0.0622**)	(**0.2518**)	(**0.2063**)
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
Housing controls	No	Yes	Yes	Yes	Yes	Yes
Job controls	No	No	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Travel-to-work-area fixed effects	No	Yes	Yes	Yes	Yes	Yes
Workplace fixed effects	No	No	No	Yes	No	Yes
Number of observations	2,196,324	2,196,324	2,196,324	2,196,324	2,196,324	2,196,324
R-squared	0.2257	0.6007				
Kleibergen-Paap F-statistic			154	307.1	11.56	12.43

Notes: **Bold** indicates instrumented. Household controls include household size, mean age of adults, mean gender, household type (couple, single, kids), the share of the household that is foreign-born. Job controls are the total hours worked, whether the household has a company car, the share of full-time contracts, the share of permanent contracts. Housing controls include house type, height of the building, construction year dummies and whether a building is listed. Standard errors are clustered at the neighborhood level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Figures

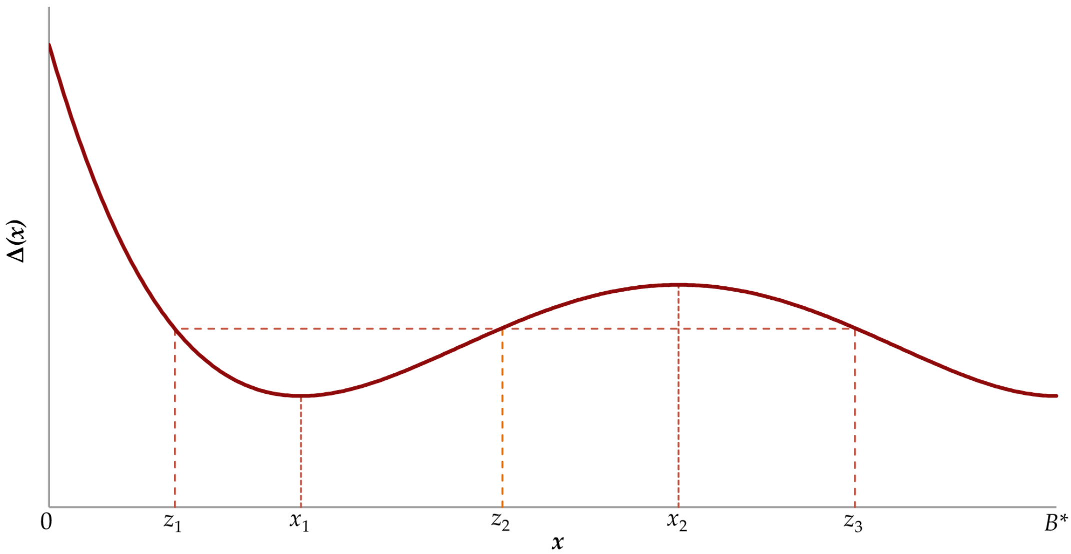
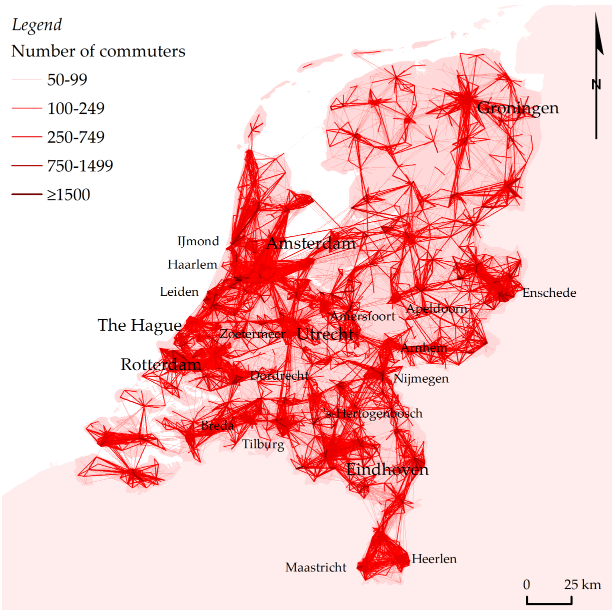
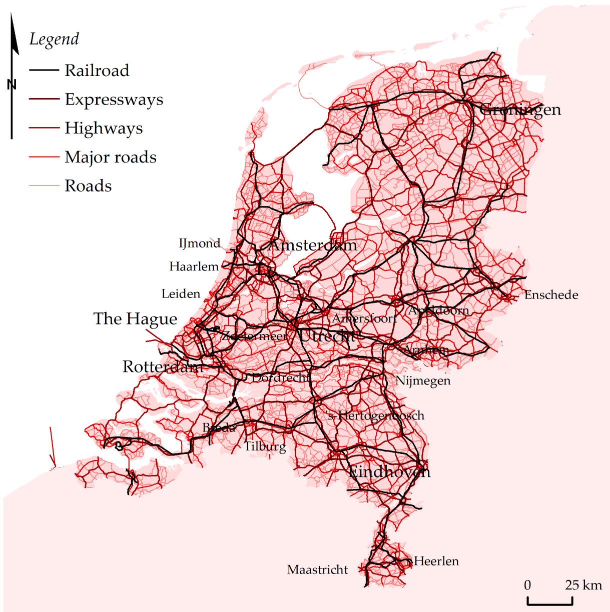


FIGURE 1 – SORTING AND LOCATION QUALITY

(A) Overview map



(B) Commuting networks



(C) Transport Network

Figure 2 – The Netherlands

Appendix of "Income Sorting Across Space: The Role of Amenities and Commuting Costs"

Abstract

In this Appendix we provide proofs of Proposition 1 and 2, the derivation of the land rent under Stone-Geary preferences. We also provide more detailed information on the data and discuss a wide range of robustness checks for the reduced-form income mapping.

Keywords: cities, social stratification, income, amenities, commuting

JEL classification: R14, R23, R53, Z13.

Appendix A. Theory

A.1 Proof of Proposition 1

Differentiating (8) with respect to x and using (7), we obtain:

$$\Psi_x(x,\omega,U^*(\omega)) = \frac{\omega t}{H}\left(\frac{t_x}{t} - \frac{Q_b}{\omega t}b_x\right) \tag{A.1}$$

Differentiating (A.1) with respect to ω and rearranging terms yields the following expression:

$$\begin{aligned}\Psi_{x\omega}(x,\omega,U^*(\omega)) \quad = \quad & \frac{t}{H}\left\{\frac{t_x}{t}\left[1-\frac{\omega}{H}(H_\omega + H_U U^*_\omega)\right]\right. \\ & \left. +\frac{b_x}{t}\left[\frac{H_\omega + H_U U^*_\omega}{H}Q_b - (Q_{bH}(H_\omega + H_U U^*_\omega) + Q_{bU}U^*_\omega)\right]\right\}. \end{aligned} \tag{A.2}$$

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

$$\begin{aligned} Q_b &= -\frac{U}{b^2 u_q} \\ Q_{bU} &= -\frac{1}{b^2 u_q} + \frac{U}{b^2 u_q^2}u_{qq}Q_U \\ Q_{bH} &= \frac{U}{b^2 u_q^2}\left(u_{qq}Q_H + u_{qh}\right). \end{aligned}$$

Assume that the ω-households are located at x. Differentiating $u = U^*(\omega)/b$ with respect to ω and using the budget constraint $Q = \omega t(x) - H\Psi$ and (9), we obtain:

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

Since

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

at the residential equilibrium, we have:

$$t = \frac{U^*_\omega}{b u_q}. \tag{A.3}$$

Plugging this expression, Q_b, Q_{bU} and Q_{bH} in (A.2), we get

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

which is equivalent to

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

Using

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

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

$$\Psi_{x\omega}(x,\omega,U^*(\omega))=\frac{t}{H}\left[\left(1-\frac{\varepsilon_{H,\omega}+\varepsilon_{u_q,\omega}}{\varepsilon_{U,\omega}}\right)B-(1-\varepsilon_{H,\omega})\,T\right], \tag{A.4}$$

which proves Proposition 1.

A.2 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:

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

Taking the total derivative of this expression with respect to ω yields:

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

so that

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

In short, we have $\varepsilon_{U,\omega} = 1$, $\varepsilon_{H,\omega} = \varepsilon_{h,\omega} = 1$ and $\varepsilon_{u_q,\omega} = 0$.

A.3 Stone-Geary preferences

It is readily verified from (12) that

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

It follows from this expression that

$$\begin{aligned}Q_U &= \frac{1}{1-\mu}U^{\frac{1}{1-\mu}-1}\left[\frac{1}{b(h-\overline{h})^{\mu}}\right]^{\frac{1}{1-\mu}} = \frac{1}{(1-\mu)}\frac{Q}{U}, \\ Q_{Ub} &= -\frac{Q_U}{(1-\mu)b}, \\ Q_b &= -\frac{U}{b}Q_U, \\ Q_h &= -\frac{\mu}{1-\mu}\left[\frac{1}{(h-\overline{h})}\frac{U}{b}\right]^{\frac{1}{1-\mu}} \\ Q_{bH} &= \frac{U}{b}\frac{\mu}{1-\mu}(h-\overline{h})^{-1}Q_U.\end{aligned}$$

Plugging Q_b, Q_{bH} and Q_{bU} into (A.2) and rearranging terms leads to

$$\begin{aligned}\Psi_{x\omega}(x,\omega,U^*(\omega)) &= \frac{t}{H}\left\{\frac{t_x}{t}\left[1 - \frac{\omega}{H}(H_\omega + H_U U^*_\omega)\right]\right. \\ &\quad + \frac{b_x}{b}\left[\frac{H_\omega + H_U U^*_\omega}{H}\left(-\frac{U}{t}Q_U\right)\left(\frac{h-(1-\mu)\overline{h}}{(1-\mu)(h-\overline{h})}\right)\right. \\ &\quad \left.\left. + \frac{Q_U}{(1-\mu)t}U^*_\omega\right]\right\}. \qquad \text{(A.5)}\end{aligned}$$

Plugging Q_h and Q in (7) and solving the corresponding equation yields

$$\frac{h-(1-\mu)\overline{h}}{(1-\mu)(h-\overline{h})} = \omega t\left[\frac{b}{U}(h-\overline{h})^{\mu}\right]^{\frac{1}{1-\mu}}. \qquad \text{(A.6)}$$

Given the expression of Q_U, it turns out that

$$\left(-\frac{U}{t}Q_U\right)\left[\frac{h-(1-\mu)\overline{h}}{(1-\mu)(h-\overline{h})}\right] = -\frac{\omega}{1-\mu}. \tag{A.7}$$

Differentiating (8) with respect to ω and using (7), we obtain:

$$\Psi_\omega(x,\omega,U^*(\omega)) = \frac{t}{H}\left(1-\frac{Q_U}{t}U^*_\omega\right) \tag{A.8}$$

which is equal to 0 if

$$U^*_\omega = \frac{t}{Q_U}. \tag{A.9}$$

Using (A.7) and (A.9), (A.5) can be rewritten as follows

$$\Psi_{x\omega}(x,\omega,U^*(\omega)) = \frac{t}{H}\cdot\left[1-\frac{\omega}{H}(H_\omega + H_U U^*_\omega)\right]\cdot\frac{1}{1-\mu}\cdot\left[(1-\mu)\frac{t_x}{t}+\frac{b_x}{b}\right]. \tag{A.10}$$

Applying the implicit function theorem to (A.6) yields

$$H_U = \frac{(h-(1-\mu)\overline{h})(h-\overline{h})}{U\mu h}$$

and

$$H_\omega = -\frac{t(1-\mu)^2}{\mu h}U^{-\frac{1}{1-\mu}}b^{\frac{1}{1-\mu}}(h-\overline{h})^{1+\frac{1}{1-\mu}}.$$

Given Q_U, (A.9) can be expressed as the following differential equation:

$$U^*_\omega = t\cdot(1-\mu)\left[b\cdot(h-\overline{h})^\mu\right]^{\frac{1}{1-\mu}}(U^*(\omega))^{-\frac{\mu}{1-\mu}}. \tag{A.11}$$

We thus obtain

$$H_\omega + H_U U^*_\omega = t\cdot(1-\mu)(h-\overline{h})\left[\frac{b}{U}(h-\overline{h})^\mu\right]^{\frac{1}{1-\mu}}$$

Therefore, by implication of (A.6), we have:

$$1-\frac{\omega}{H}(H_\omega + H_U U^*_\omega) = \frac{(1-\mu)\overline{h}}{h}.$$

Substituting this expression into (A.10) yields:

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

A.4 Proof of Proposition 2

The proof involves four steps.

(i) Existence and uniqueness of the residential equilibrium and the bid-max lot size. From the definition of the location-quality index given by (15), (A.6) can be rewritten as follows:

$$\frac{H-(1-\mu)\overline{h}}{(1-\mu)(H-\overline{h})} = \omega\Delta^{\frac{1}{1-\mu}}\left[\frac{(H-\overline{h})^{\mu}}{U}\right]^{\frac{1}{1-\mu}}, \tag{A.12}$$

which implies $H(\omega t(x), U/b(x)) \equiv H(\Delta(x), \omega, U)$ so that the bid-max lot size depends on $b(x)$ and $t(x)$ only through the location-quality index $\Delta(x)$.

The LHS of (A.12) is decreasing and tends to $+\infty$ when $H \to \overline{h}$ and to $1/(1-\mu) > 0$ when $H \to +\infty$. The RHS of (A.12) is increasing in H. It tends to 0 when $H \to \overline{h}$ and to $+\infty$ when $H \to +\infty$. Therefore, (A.12), equivalently (7), has a single solution $H(\omega t(x), U/b(x))$, which implies that the housing demand is uniquely determined.

Applying the implicit function theorem to (A.12) yields

$$\frac{\partial H}{\partial \Delta} = -\left[U^{\frac{1}{1-\mu}}(H-\overline{h})^{-\frac{1}{1-\mu}-1}\frac{\mu H}{(1-\mu)}\right]^{-1}\omega\Delta^{\frac{\mu}{1-\mu}} < 0. \tag{A.13}$$

Given (7), expression (8) reduces to

$$\Psi(x,\omega,U) = -Q_H(H, U/b(x)).$$

Using Q_H leads to

$$\Psi(x,\omega,U) = \frac{\mu}{1-\mu}(H-\overline{h})^{\frac{-1}{1-\mu}}\left[\frac{U}{b}\right]^{\frac{1}{1-\mu}}. \tag{A.14}$$

(ii) Equilibrium utility level. Using the definition of the location-quality index, (A.11) implies that the equilibrium utility level is a solution to the differential equation in U^*:

$$U^*_{\omega} = \Delta^{\frac{1}{1-\mu}}(1-\mu)(H-\overline{h})^{\frac{\mu}{1-\mu}}(U^*(\omega))^{-\frac{\mu}{1-\mu}}, \tag{A.15}$$

so that $U^*(\omega)$ depends on Δ only.

(iii) Supermodularity of the equilibrium utility level. Differentiating (A.15) w.r.t. Δ, we obtain:

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

Using (A.13), this expression may be rewritten as follows:

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

From (A.12), the expression in the bracketed term is equivalent to

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

Therefore,

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

The Spence-Mirrlees condition thus holds, which implies the existence of a positive assortative matching between incomes and the values of the location-quality index. In other words, there is a unique one-to-one and increasing relationship between ω and Δ (Chiappori, 2017). Since a single value of Δ is associated with x, a unique value of ω must be associated with x. Therefore, the equilibrium conditions (9) has a unique solution, which means that $\omega^*(x)$ is a mapping.

Note that the supermodularity of $U^*(\omega)$ is equivalent to the inequality $\Psi_{\omega\Delta} > 0$. Indeed, differentiating (A.8) w.r.t. Δ and using (A.9) yield:

$$\begin{aligned}\Psi_{\omega\Delta}(x, \omega t(x), U^*(\omega))|_{\Psi_\omega=0} &= \frac{t}{H}\left[\frac{\partial(t/Q_U)/\partial\Delta}{U^*_\omega}\right] \\ &= \frac{t}{H}\frac{\partial U^*_\omega/\partial\Delta}{U^*_\omega} > 0.\end{aligned}$$

(iv) Uniqueness of the equilibrium shares. The proof follows Montesano (1972). Assume that there are $m \geq 2$ points $x_1 \neq x_2 ... \neq x_m$ exist such that $\Delta(x_1) = \Delta(x_j)$ for $j = 2, ..., m$. Using Step (i), we may rewrite (4) as follows:

$$|s(x, \omega(x))f(\omega^*(x))H\left\{\Delta(x), \omega^*(x), U^*(\omega^*(x))\right\}\mathrm{d}\omega| = \mathrm{d}x \qquad j = 1, ..., m. \tag{A.16}$$

If $\omega^*(x_1) = \omega^*(x_j)$, then $f(\omega^*(x_1)) = f(\omega^*(x_j))$ for $j = 2, ..., m$. Therefore,

$$H\left[\Delta(x_1), \omega^*(x_1), U^*(\omega^*(x_1))\right] = H\left[\Delta(x_j), \omega^*(x_j), U^*(\omega^*(x_j))\right].$$

Dividing the relationships (A.16) between themselves leads to

$$\frac{s(x_1, \omega^*(x_1))}{s(x_k, \omega^*(x_k))} = 1 \qquad k = 2, ..., m.$$

Since $\sum_{j=1}^{m} s(x_j, \omega) = 1$, the above system of equations has a unique solution given by

$$s^*(x_j, \omega^*(x_j)) = \frac{1}{m} \qquad j = 1, ..., m.$$

As $b(x)$ and $t(x)$ are never constant on a non-negligible subset of N, there is an integer M such that $m \leq M$ holds.

A.5 The land rent and land gradient

1. The expression (8) can be rewritten as follows:

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

Using (A.9) and plugging Q_U in the above expression leads to

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

2. By plugging Q_b into (A.1), we obtain:

$$\Psi_x(x,\omega,U^*(\omega)) = \frac{\omega t}{H}\left[\frac{UQ_U}{\omega t}B(x) - T(x)\right]$$

and substituting t by its expression given in (A.9) we obtain:

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

A.6 The equilibrium land rent under Fréchet distributions

Rearranging terms in (13) yields:

$$H - \overline{h} = \mu\left[\frac{\omega t}{\Psi(x,\omega,U)} - \overline{h}\right].$$

Plugging the above expression into (A.14) leads to

$$\Psi(x,\omega,U) = \mu^{-\frac{\mu}{1-\mu}}(1-\mu)^{-1}\left[\frac{\omega t}{\Psi(x,\omega,U)} - \overline{h}\right]^{\frac{-1}{1-\mu}}\left[\frac{U(\omega)}{b}\right]^{\frac{1}{1-\mu}}.$$

Dividing this expression by $t(x)$ and setting $\Phi \equiv \Psi/t$, we get

$$\Phi = \mu^{-\frac{\mu}{1-\mu}}(1-\mu)^{-1}\left(\frac{\omega}{\Phi} - \overline{h}\right)^{-\frac{1}{1-\mu}}\left[U(\omega)\right]^{\frac{1}{1-\mu}}\Delta^{\frac{-1}{1-\mu}}.$$

Rearranging terms, this expression becomes:

$$\Phi = \mu(1-\mu)^{\frac{1-\mu}{\mu}}\left(\omega - \Phi\overline{h}\right)^{\frac{1}{\mu}}\left[U(\omega)\right]^{-\frac{1}{\mu}}\Delta^{\frac{1}{\mu}}. \tag{A.17}$$

Applying the first-order condition to Φ yields the following differential equation in ω:

$$U^*_\omega(\omega) = \frac{1}{\omega - \Phi\overline{h}}U^*(\omega).$$

Let

$$U^*(\omega) = \left(\omega - \Phi\overline{h}\right)X(\omega) \tag{A.18}$$

be a solution to the above differential equation where $X(\omega)$ is determined below. Differentiating (A.18) with respect to ω, we obtain

$$U_\omega(\omega) = \left[\frac{1}{\omega - \Phi\overline{h}} - \frac{\overline{h}}{\omega - \Phi\overline{h}}\Phi_\omega + \frac{X_\omega(\omega)}{X(\omega)}\right]U(\omega).$$

Totally differentiating Φ leads to

$$\Phi_\omega \equiv \frac{\mathrm{d}\Phi}{\mathrm{d}\omega} = \frac{\partial\Phi}{\partial\omega} + \Phi_\Delta\Delta_\omega = \Phi_\Delta\Delta_\omega. \tag{A.19}$$

Differentiating (A.17) with respect to Δ yields:

$$\Phi_\Delta = \Phi\left[\frac{1}{\mu}\Delta^{-1} - \frac{1}{\mu}\Phi_\Delta\overline{h}\left(\omega - \Phi\overline{h}\right)^{-1}\right],$$

whose solution in Φ_Δ is

$$\Phi_\Delta = \frac{1}{\Delta}\frac{\Phi}{\mu}\left[\frac{\mu(\omega - \Phi\overline{h})}{\mu(\omega - \Phi\overline{h}) + \overline{h}\Phi}\right].$$

Therefore, we may rewrite (13) as follows:

$$H\Phi = \mu(\omega - \Phi\overline{h}) + \overline{h}\Phi. \tag{A.20}$$

Plugging (A.20) into Φ_Δ leads to

$$\Phi_\Delta = \frac{\omega - \Phi\overline{h}}{\Delta H}.$$

Using Φ_ω and Δ_ω, (A.19) becomes:

$$\Phi_\omega = \Phi_\Delta\Delta_\omega = \frac{1}{\gamma}\frac{\omega - \Phi\overline{h}}{\omega H} = \frac{1}{\gamma\mu}\frac{(H - \overline{h})\Phi}{\omega H} > 0.$$

Since $U_\omega(\omega)/U(\omega)$ is equal to $1/(\omega - \Phi\overline{h})$ in equilibrium, it must be that

$$\frac{X_\omega(\omega)}{X(\omega)} = \frac{\overline{h}}{\omega - \Phi\overline{h}}\Phi_\omega = \frac{\overline{h}}{\omega - \Phi\overline{h}}\frac{1}{\gamma\mu}\frac{(H - \overline{h})\Phi}{\omega H}.$$

Therefore, using (A.20) leads to the following differential equation in ω:

$$X_\omega(\omega) = \frac{1}{\gamma}\frac{\overline{h}}{\omega H}X(\omega),$$

whose solution is

$$X(\omega) = k\left(\frac{\omega}{H}\right)^{\frac{1}{(1-\mu)\gamma}}, \tag{A.21}$$

where $k > 0$ is the constant of integration. Indeed, differentiating the above equation with respect to ω leads to

$$X_\omega(\omega) = \frac{1}{(1-\mu)\gamma}\frac{H - \omega(H_\omega + H_U^* U_\omega)}{H^2}\frac{H}{\omega}X(\omega).$$

Using $1 - (\omega/H)\left(H_\omega + H_U U_\omega^*\right) = (1-\mu)\overline{h}/h$, we obtain:

$$X_\omega(\omega) = \frac{1}{(1-\mu)\gamma}\frac{(1-\mu)\overline{h}}{H}\frac{1}{\omega}X(\omega) = \frac{1}{\gamma}\frac{\overline{h}}{\omega H}X(\omega).$$

Substituting (A.21) into (A.18) yields:

$$U(\omega) = \left(\omega - \Phi\overline{h}\right) k \left(\frac{\omega}{H}\right)^{\frac{1}{(1-\mu)\gamma}}.$$

Plugging this expression into (A.17) and rearranging terms, we obtain the following implicit solution for the equilibrium land rent:

$$R^*(x) = \mu(1-\mu)^{\frac{1-\mu}{\mu}} k^{-\frac{1}{\mu}} t(x) \Delta^{\frac{1}{\mu}} \left[\frac{\mu t(x)}{R^*(x)} + \frac{(1-\mu)\overline{h}}{\omega^*(x)}\right]^{\frac{1}{(1-\mu)\mu\gamma}}. \tag{A.22}$$

Since the RHS of (A.22) is strictly decreasing and tends to 0 (∞) when $R(x) \to \infty$ (0), (A.22) has a unique solution in $R^*(x)$.

The lowest income in the sample, denoted by $\underline{\omega}$, is strictly positive. It follows from (20) that the lowest location-quality index associated with the poorest household is given by

$$\underline{\Delta} = s_\Delta \left(\frac{\underline{\omega}}{s_\omega}\right)^{1/\gamma} > 0.$$

The constant k may be obtained by evaluating $R^*(x)$ at the least enjoyable location $\underline{x}$ where $\Delta(x)$ reached its minimum $\underline{\Delta}$. We assume that $\underline{x}$ is unique. Furthermore, the land rent at $\underline{x}$ is equal to the opportunity cost of land, R_A. Therefore, it is readily verified that k is given by

$$k^{-\frac{1}{\mu}} = R_A \mu^{-1} (1-\mu)^{-\frac{1-\mu}{\mu}} \left[t(\underline{x})\right]^{-1} \underline{\Delta}^{-\frac{1}{\mu}} \left[\frac{\mu t(\underline{x})}{R_A} + \frac{(1-\mu)\overline{h}}{\underline{\omega}}\right]^{\frac{-1}{(1-\mu)\mu\gamma}}.$$

Plugging this expression into (A.22) yields the equilibrium land rent at x:

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

Note that this expression captures several effects: the commuting costs at x and $\underline{x}$, the location-quality index at x and $\underline{x}$, and the income mapping $\omega^*(x)$.

Appendix B. Data

In this appendix, we detail the construction of the various datasets. In Appendix B.1 we elaborate on how we calculate network distances and show the relationship with Euclidian distance. Appendix B.2 continues by explaining how we measure land prices and lot sizes for all locations. This is followed in B.3 by more information on our proxies for amenities: the picture index and the construction of the hedonic amenity index. In Appendix B.4 we introduce the historical data based on 1900 land use maps and the 1832 Census. Appendix B.5 reports distributions of the variables of interest.

B.1 Commuting and travel times

To estimate the commuting time for each household, we use the tax register information, which provides information on individual jobs and the number of hours worked in each firm for each year. From the *ABR Regio* dataset, we get information on all firms which provide information on each establishment in the Netherlands, such as its exact location, the industrial sector, and the *estimated* number of employees in each establishment. Since we do not know the exact establishment, only the firm people work for, we assume that they work at the nearest establishment of the firm. This assumption may be problematic for firms having a large number of establishments (e.g., supermarkets or large banks). Therefore, we keep only firms with a maximum of 15 establishments throughout the Netherlands. As many such firms have establishments in different cities, it is reasonable to assume that people work in the nearest establishment.[16] Overall, we are left with 95% of firms. To avoid miscoding and to exclude employment agencies (where people do not actually work), we consider establishments with no more than 10 thousand employees, which correspond to 0.1% of the total number of establishments and 1% of total employment. Hence, excluding those establishments is unlikely to affect our results.

We first calculate the commuting time from each home location x to each job location i for each year. Then, we determine the commuting time of each household by computing the average commuting time of each adult household member weighted by the number of hours (s)he worked. To calculate the travel time (as well as the time to travel to amenities) 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 96m), which are usually lower than the speed limit.

[16] Alternatively, we could consider a distance-decay average of distances to the firm's establishments. Instead, we test robustness by keeping households that have only one working-member who works during the whole year in a single-establishment firm. This leads to nearly identical results.

The dataset from *SpinLab* provides information on free-flow driving speeds for every major street in the Netherlands. The actual speeds are usually well below the free-flow driving speeds, due to traffic lights, roundabouts and intersections. For each neighborhood we calculate the straight-line distance to the nearest access points on 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 195m (the average is 326m). We assume that the average speed to get to the nearest access points is 10km/h. This is the speed based on the Euclidian distance. In reality, the distance to get to the network will be higher because streets are usually curved. Hence, the assumption of 10km/h seems reasonable as the minimum speed on roads in the network is 20km/h. 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, which is the sum of the driving time to access the network, the network driving time and the time it takes from the network to arrive at the destination. Alternatively, we calculate for each location pair the Euclidian distance and assume again an average speed of 10km/h.

We also calculate the travel time using the train, using a similar approach. The median distance of each centroid to the nearest station is 5.25km. We then choose the minimum of the travel time over the road, using the train or taking the Euclidian travel time.

[Figure B.1 about here]

The correlation between travel time and Euclidian distance is modest ($\rho = 0.643$). For short distances (< 10km) the correlation is, however, much higher ($\rho = 0.862$). We plot the relationship between distance and travel time in Figure B.1A. This relationship is monotonic. Figure B.1B shows the share of commuting people who travel at most τ minutes, which we use to calculate employment accessibility in 1900.

B.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 use information on home sales from *NVM* (The Dutch Association of Realtors), which comprises the large majority (about 75%) of owner-occupied house transactions between 2003 and 2017. We know the transaction price, the lot size, inside floor space size (both in m^2), 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. We make some selections to make sure that our results are not driven by outliers. First, 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 also leave out transactions that refer to properties that are larger than 250m^2 or smaller than

$25m^2$, or have lot sizes above $5000m^2$. These selections consist of less than 1% of the data and do not influence our results. We follow a similar procedure as Rossi-Hansberg *et al.* (2010), implying that we can only use information on residential properties with land. We are left with $1,337,445$ housing transactions.

Let $\mathcal{P}(x)$ denote the house price in year y, $H(\tilde{x})$ the observed lot size and $C(\tilde{x})$ the housing characteristics of property $\tilde{x}$. The log land rent $R(x)$ is equal to the fixed effects at the level of the postcode (about 15-20 addresses), which we denote by $\varsigma(x)$, while $\vartheta(y)$ denote year y fixed effects. We estimate:

$$\log \frac{\mathcal{P}(\tilde{x}, y)}{H(\tilde{x}, y)} = \eta_1 C(\tilde{x}, y) + \varsigma(x) + \vartheta(y) + \epsilon(\tilde{x}, y), \tag{C.2.1}$$

where $\epsilon(\tilde{x}, y)$ is an i.i.d. error term that is assumed to be uncorrelated to land rents and housing characteristics, while η_1 are parameters to be estimated. As $\log R_x$ are given by the very local fixed effects, we do not impose any structure on how land rents R_x vary across locations. For about 80% of the data we do not observe land prices directly, because either there were no multiple sales in our study period or because there is no owner-occupied housing in the respective postcode. We therefore also estimate the above equation with neighborhood fixed effects instead.

[Tables B.1 and B.2 about here]

Descriptive statistics for the housing sample are reported in Table B.1. Coefficients η_1 related to the housing attributes are reported in Table B.2. It appears that the house price per square meter of land is generally a bit lower when the property is larger. However, the house price per square meter of land of properties that are (semi-)detached is generally higher. Furthermore, when the maintenance state of a property is good, prices are about $502/1269 = 40\%$ higher. When a property has central heating, the price per square meter is about 5.1% higher. The dummies related to the construction decades show the expected signs. Properties constructed after World War II until 1970 generally have lower prices because this is a period associated with a lower building quality. Lot sizes are inversely related to pattern of land prices ($\rho = -0.245$). In other words, more expensive locations generally have smaller lots, which makes sense.

B.3 Amenities

Picture density. Here we investigate whether there is a meaningful correlation between picture density and observed proxies for amenities. We gather data on the share of a neighborhood in a historic district, the density of listed buildings, as well as the share of the neighborhood occupied by buildings or water. We use similar specifications as reported in Table 2, but take the picture density as dependent variable. Table B.3 reports the results.

[Table B.3 about here]

In column (1) we regress picture density on our proxies for amenities. We find that a standard deviation increase in listed building density is associated with an increase in picture density of 8%. When the share of the area in a historic district increases by 10% percentage points, this is associated with an increase in the picture density of 30%. Hence, the impact of historic amenities on the revealed amenity level is large. The share of built-up land and water bodies in the neighborhood are also positively related to the picture density: a 10% increase in the share of built-up land or water bodies is associated with an increase in the picture index of about 33%. In column (2) we include commuting time, demographic characteristics, housing attributes and travel-to-work-area fixed effects, leading to very similar estimates.

In column (3) we instrument for commuting time. We observe that commuting time does not seem to be statistically significantly correlated to the amenity level in a certain location. However, once we include job characteristics and workplace fixed effects, we find a weak relationship between commuting time and amenity level: a shorter commuting time is generally associated with a higher picture density. We think this makes sense as the most central locations generally provide a high amenity level and a somewhat better accessibility to jobs. However, we reiterate that this relationship is not statistically strong.

A hedonic amenity index. We test whether our results are robust to using an alternative hedonic amenity index, rather than relying on geocoded pictures. Following Lee and Lin (2018), we construct an aggregate amenity index that describes the amenity level in every neighborhood x.[17] We will make a distinction between *historic* amenities and *natural* amenities.

Let $\mathcal{A}(\tilde{x})$ be a set of variables that describe amenities of property $\tilde{x}$ (so the location is more detailed than the neighborhood x). For example, we calculate the share of historic districts, the number of listed buildings, water bodies and open space within 500m of each property. Let $\mathcal{P}(\tilde{x}, y)$ the house price, while $C(\tilde{x}, y)$ are housing characteristics of property $\tilde{x}$, and $\vartheta(y)$ are year y fixed effects. We also include neighborhood fixed effects $\varsigma(x)$, and thus we identify the effects of amenities on prices *within* neighborhoods. Specifically, we estimate the following equation:

[17] Albouy (2016) uses information on wages and housing costs to infer the level of amenities for U.S. cities. However, his approach is not applicable here because we are interested in *intra*-city variation in amenities and incomes.

$$\log \mathcal{P}(\tilde{x}, y) = \eta_0 \mathcal{A}(\tilde{x}) + \eta_1 C(\tilde{x}, y) + \vartheta(y) + \varsigma(x) + \epsilon(\tilde{x}, y), \tag{B.3.1}$$

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

$$\tilde{b}(x) = \frac{1}{N(x)} \sum_{\tilde{x}=1}^{N_x} \widehat{\eta}_0 \mathcal{A}(\tilde{x}), \tag{B.3.2}$$

where $\tilde{b}(x)$ is the (alternative) amenity value at x and $N(x)$ are the number of observations in neighborhood x. Hence, we take the mean amenity value of all locations $\tilde{x}$ within neighborhoods x.

We use data on the universe of housing transactions in the Netherlands between 2010 and 2015 from the *NVM*. Additional descriptive statistics of the *NVM* data are reported in Table B.4. We have 695,709 observations and the average house price is €229 thousand.

[Tables B.4 and B.5 about here]

In Table B.5 we report the results of the regression of equation (B.3.1). We first investigate the impact of listed buildings. It can be seen that the share of historic districts leads to higher price. A 10 percentage point increase in the share of land designated as historic district increases prices by 1.8%. Listed buildings do have a small additional effect of 0.5% per listed building. In column (2) we investigate the impact of water bodies and open space. For a 10 percentage point increase in water bodies, prices rise by 3%. Moreover, a 10 percentage point increase in open space implies a price increases of 0.6%, so this effect is considerably smaller. When we put historic amenities and natural amenities together, the coefficients are essentially unaffected. We consider this as our preferred specification. In the last specification, we investigate whether the results change when we include endogenous amenities, such as shops, cafés, and leisure establishments. This appears not to be the case. Only hotels restaurants and cafés have a statistically significant impact on prices, which suggests that *exogenous amenities related to the built environment and land use are more important than endogenous amenities.*

B.4 Historic data

To instrument current amenity levels and commuting time we use information on land use, the railway network and amenities in 1900. For the 1900 land use maps, Knol *et al.* (2004) have scanned and digitized maps into 50 by 50 meter grids and classified these grids 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 in 1900 from *NLGIS*. Municipalities were much smaller at that time and about the size of a large neighborhood nowadays. 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 further use information on railway stations from Koopmans *et al.* (2012). We enrich these data by adding missing stations from various sources on the internet and create a network with travel times. To approximate the speed, we fit a regression of the length of (current) railway segments between stations on current travel time on the railway network. Based on historic sources, it appears that the average speed is about 50% of what it is currently (about 70km/h).

We show a map of land use and the railway network for the Netherlands in 1900 in Figure B.2. In Panel A it is shown that cities like Amsterdam, Rotterdam, The Hague, and Utrecht were already large in 1900. It can also be seen that some areas that have been reclaimed from the sea (e.g., to the northeast of Amsterdam) did not exist in 1900. The Panel B of Figure B.2 shows the railway network. In particular, places around Amsterdam and Utrecht have a high accessibility. The first railway line in the Netherlands was opened in 1839 between Amsterdam and Haarlem, soon followed by the openings of many other lines.

[Figure B.2 about here]

In Table B.6 we provide descriptives for all instruments. The average share of built-up area in 1900 was 4.3%, while it was 4.2% in 1832. However, this figure is a bit misleading because for 1832 we have more data near urban areas. On average about 38 thousand jobs and 89 thousand people could be reached within commuting distance in 1900. Not surprisingly, this was much lower (40 thousand) in 1832.

[Table B.6 about here]

B.5 Other descriptive statistics

In Figure B.3 we report the distributions of the log of income and the log of land price. The distributions of land prices is somewhat positively skewed.

[Figure B.3 about here]

In Figure B.4 we show maps of income and land price distributions across the Netherlands. As expected, land prices are generally higher in cities. The pattern for incomes is less clear, but generally speaking we find that wealthier households locate close to or in cities.

[Figure B.4 about here]

Appendix C. Other empirical results and sensitivity

This appendix reports various additional econometric results. First, we show first-stage results in Appendix C.1. Then we report bias-corrected estimates using Oster's (2019) methodology in Appendix C.2. We undertake additional robustness checks in Appendices C.3 and C.4.

C.1 Pictures and amenities: the first-stage results

We report first-stage estimates in Table C.1 related to the baseline results reported in Table 2.

In columns (1) and (2) we focus on the impact on picture density. We find strong positive impact of share of built-up land in 1900 and share water in 1900 on current picture density. For example, when we include workplace and travel-to-work-area fixed effects in column (2), increasing the share of built-up land in 1900 by 10 percentage points in the neighborhood is associated with an increase in the picture density of 42%. This makes sense: locations with many historic buildings are often locations which were already developed around 1900. The impact of water bodies is also positively correlated to the picture density, although the impact is somewhat smaller in magnitude: a 10 percentage point increase in the share of water in 1900 leads to an increase in the picture density of 9.9%.

[Table C.1 about here]

In columns (3)-(6), Table C.1 we take employment accessibility as a dependent variable. The instruments for accessibility are relevant. Contemporary employment accessibility is strongly correlated to commuting times. In the preferred specification with workplace fixed effects and job controls (column (4)), doubling employment accessibility reduces commuting times by 20.7%.

Columns (5) and (6) rely on historic instruments. We find a strong positive effect of the share of built-up land in 1900 between 500 and 1000m on accessibility. Likewise, employment accessibility in 1909 has a strong positive effect on current employment accessibility. More specifically, in column (6), doubling employment accessibility in 1909 is associated with an decrease in commuting times of 3.5%.

C.2 Bias-corrected estimates

Many non-experimental papers use coefficient movements after the inclusion of control variables to investigate whether omitted variable bias is important. Oster (2019) argues that coefficient movements alone are not a sufficient statistic to calculate bias. Instead, she argues that whether omitted variable bias is a concern depends on the variance of the added control variables, as well as coefficient movements. In other words, changes in the coefficient(s) of interest after adding controls should be scaled by the change in the R^2. Oster (2019) then derives an estimator to correct estimates for omitted variable bias under the assumption that the relationship between the variables of interest and unobservables can be recovered from the relationship between the variables of interest and observable control variables. In our context, this assumption makes sense as control variables that are added bear some potential relationship to unobservables. To be precise, we add many housing, demographic and job controls, as well as workplace and travel-to-work-area fixed effects, which are likely to be correlated to unobservables.

Oster (2019) then derives a GMM estimator to derive bias-corrected estimates of the impact of amenities and employment accessibility on incomes. There are two key input parameters that have to be determined. First, a parameter must be chosen that determines the relative degree of selection on observed and unobserved variables, which we denote by δ. Although this parameter is fundamentally unknown, Altonji *et al.* (2005) and Oster (2019) show that $\delta = 1$ is a reasonable (upper-bound) value. Second, there is the maximum R^2 from a hypothetical regression of income on amenities, accessibility and controls, which we denote as R^2_{max}. Because of measurement error and random variation in incomes, R^2_{max} is usually well below 1. Oster (2019) considers to set $R^2_{\text{max}} = \Pi\hat{R}^2$, where $\hat{R}^2$ is the R^2 obtained from the regression of log income on the variables of interest and all controls and fixed effects. She provides some evidence based on experimental studies that $\Pi \approx 1.3$. To be on the safe side, we consider $\Pi = 1.5$ and $\Pi = 2$. We report results in Table C.2

[Table C.2 about here]

In column (1) we consider $\Pi = 1.5$. We find that the bias-corrected $\alpha_1 = 0.026$, which is virtually identical to the preferred estimate with historic instruments. The bias-adjusted impact of commuting time is now negative ($\alpha_1 = -0.058$), in line with instrumented estimates reported in Table 2. However, the magnitude is somewhat smaller: doubling commuting time attracts households whose incomes are 4% lower. However, when the hypothetical R^2_{max} is higher, so that it is twice the R^2 obtained from the regression of log income on the variables of interest and all controls and fixed effects ($\Pi = 2$), we see that the impact of commuting times becomes considerably stronger. The coefficient in column (2) indicates that a 100% increase in commuting time attracts households whose incomes are 31% lower, which is very much comparable to the baseline results. The impact of amenities with $\Pi = 2$ is somewhat stronger: doubling the picture density attracts households whose incomes are 5.9% higher.

In column (3) of Table C.2 we further investigate whether these bias-adjusted estimates are robust to the inclusion of more detailed fixed effects. More specifically, we further include municipality fixed effects and assume $\Pi = 1.5$. The impact of amenities is very similar to the result reported in column (1) (as well as the baseline result). The coefficient regarding commuting time is about twice as strong, but somewhat in between the results reported in columns (1) and (2). Overall, the results are robust.

In other words, although we do not know the 'true' value for Π, these results strongly suggest that omitted variable bias is not a major issue, as the coefficients lead to very similar outcomes as the 2SLS estimates with historic instruments. We note that Oster's (2019) methodology does not account for measurement error in amenities or commuting time or reverse causality. It may, therefore, still be important to apply our instrumental variables strategy using historic instruments.

C.3 Sensitivity: identification revisited

We consider additional robustness analyses in Table C.3 that should increase confidence in the validity of our identification strategy. First, we show that our results are similar once we focus solely on urban areas. In column (1) we only include observations in the Randstad, i.e., the main polycentric metropolitan area in the Netherlands. This reduces the total number of observations by more than 50%. However, our results are very similar.

[Table C.3 about here]

In column (2), Table C.3, we estimate specifications where we again use instruments from 1900, but include municipality fixed effects. This implies that we identify commuting and amenity effects *within* municipalities. Municipality fixed effects absorb any effects related to municipal policies and tax differentials. We find very similar coefficients.

Column (3) controls for the current share of built-up areas and population density to make sure that our amenity proxy is not just capturing population density or built-up land. We find very similar effects for amenities, but the effect of commuting time is now stronger. This may be because we have weak instruments, with a Kleibergen-Paap F-statistic that is barely 10.

Another concern is that clusters of high-income households are autocorrelated, so that our instruments are correlated to the concentrations of high-income households in 1909. To investigate whether this is an issue, we calculate the share of medium and high-skilled households in 1909. Municipalities then were much smaller, so this is a rather fine-grained measure of skill sorting across space. We also gather data on the share of Protestants in each municipality in 1899 and control for population accessibility in 1900. Including those measures in column (4) does not materially impact our coefficients. Note that the share of high-skilled and medium-skilled households in 1909 is negatively correlated to current incomes, which suggests that the determinants of residential choices in the two periods are fairly different. In addition, conditional on commuting time, population accessibility in 1900 does not influence the current income distribution.

In column (5), we further study the sensitivity of our results by choosing another instrument for commuting time. We use the share of the population in 1909 born in the same municipality. If mobility of households is correlated over time, the share of locally born people should be negatively correlated to current accessibility, as the areas that host a high number of jobs (so have a better accessibility) are expected to attract workers from other places. Indeed, we find that the share of locally born people in 1909 is negatively associated with current employment accessibility. The Kleibergen-Paap F-statistic again indicates that these are strong instruments. We find a similar coefficient related to commuting time.

To the extent one is still concerned that household income sorting is autocorrelated, in column (6) we only include neighborhoods on reclaimed land. The Netherlands is well known for its large-scale projects that reclaim land from the sea. We consider the three main projects (Wieringermeer, Noordoostpolder, Oostelijk, and Zuidelijk Flevoland) that occurred between 1930 and 1968, but permission by the government to reclaim those areas was already given in 1930. Most of the land was intended for agriculture, but a few small settlements were planned on the newly reclaimed land. Moreover, Lelystad was planned to be the largest city in the area, but nowadays Almere is by far the largest one. Hence, the plans differ considerably from the current spatial distribution of activities. Since only a small share of the population lives in those areas, we only keep about 2.5% of the observations. The latter approach should address any remaining concerns related to reverse causality as no one was living in those locations at that time, and thus income was zero.

We then instrument for amenities and commuting time with the share of planned built-up and green areas and planned accessibility in column (6). The instruments are, unfortunately, very weak leading to imprecise coefficients. Still, the point estimates are again similar to the baseline estimates. Instrumenting either amenities or commuting time leads to similar outcomes, but with lower standard errors.[18]

C.4 Sensitivity: other checks

Table C.4 reports the results of additional robustness checks. Our dataset contains observations on households. When calculating the commuting elasticity and when including workplace fixed effects, we calculate average commuting times based on working hours of different jobs of adults in the household. In column (1) we only include households that are associated with one job (location). This does not lead to significant differences in the outcomes.

Recall that we calculate the commuting time to the nearest plant when the firm has multiple establishments. We test whether this introduces error by only including households that are associated with one job in a single plant firm in column (2). In this way we address any

[18] These results are available upon request.

measurement error in commuting time. Again, the estimates are very similar, although the number of observations is considerably lower.

[Table C.4 about here]

Our measure of commuting time relies on the minimum travel time by road and rail. However, in almost all cases travel time by car is shorter. To make sure that households actually consider this travel time, we only keep households having a company car in column (3). This does not change the results.

Column (4) replaces the dependent variable income by the share of adults in the household that have a college degree or more. We find very similar effects. For example, when the picture density doubles, this increases the share of highly educated households by 1.1 percentage points. Doubling commuting times decreases the share of highly educated households by 32.3 percentage points.

Column (5) tests whether the results are robust when using commuting time by rail instead of commuting time by road. The results are comparable.

Finally, in column (6), we only keep locations that are within 15km of a city center with at least $100,000$ inhabitants to make sure that our results are not driven by rural locations. We find that the coefficients are very similar, except that standard errors of the coefficients are considerably higher. Hence, *including locations throughout the Netherlands has the benefit that it leads to more precise estimates.*

Appendix tables

Table B.1 – Descriptives for *NVM* dataset

	(1) mean	(2) sd	(3) min	(4) max
House price *(in € per m^2)*	1,269	927.2	25	25,000
Lot size *(in m^2)*	445.7	1,189	25	25,000
Size of property *(in m^2)*	132.4	45.16	26	538
Number of rooms	4.944	1.363	0	25
Terraced property	0.417	0.493	0	1
Semi-detached property	0.370	0.483	0	1
Detached property	0.189	0.392	0	1
Private parking space	0.454	0.498	0	1
Garage	0.394	0.489	0	1
Garden	0.966	0.182	0	1
Number of bathrooms	0.929	0.483	0	8
Number of kitchens	0.677	0.484	0	5
Number of balconies	0.132	0.354	0	4
Number of roof terraces	0.0674	0.257	0	3
Number of floors	2.717	0.636	1	13
Internal office space	0.00444	0.0665	0	1
Maintenance score of the outside	0.758	0.131	0	1
Maintenance score of the inside	0.753	0.143	0	1
Number of types of insulation	2.381	1.831	0	5
Central heating	0.920	0.271	0	1
Listed building	0.00652	0.0805	0	1
Newly built property	0.0417	0.200	0	1
Construction year	1,967	34.95	1,362	2,017
Year of observation	2,011	4.389	2,004	2,017

Notes: The number of observations is 1,337,495. Because of confidentiality restrictions the minimum and maximum values refer to the 0.01% and 99.99% percentile. This implies that we exclude the bottom and top observations

Table B.2 – Estimating land prices and lot sizes

(Dependent variable: the log of land price per m^2)

	(1)
Rooms	-6.1664***
	(0.4506)
Terraced property	702.4875***
	(6.5087)
Semi-detached property	510.0447***
	(6.5516)
Detached property	360.7740***
	(6.7580)
Private parking space	-56.3558***
	(1.9988)
Garage	-42.8166***
	(2.0556)
Garden	47.5907***
	(2.8356)
Number of bathrooms	17.3274***
	(0.9885)
Number of kitchens	-7.2575***
	(1.0818)
Number of balconies	47.8147***
	(1.5204)
Number of roof terraces	109.0801***
	(1.8878)
Number of floors	94.9407***
	(1.0148)
(Internal) office space	-55.3454***
	(6.3595)
Maintenance score of the outside	29.5137***
	(6.3366)
Maintenance score of the inside	501.7345***
	(5.8143)
Number of types of insulation	8.3945***
	(0.3138)
Central heating	65.8404***
	(1.7719)
Listed building	27.9334***
	(6.2691)
Newly built property	-13.3758***
	(4.3108)
3^{th}-order polynomial of property size	Yes
Construction decade dummies	Yes
Year fixed effects	Yes
Postcode fixed effects	Yes
Observations	1,280,031
R^2	0.8295

Notes: Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.3 – Picture density and amenities

(Dependent variable: the log of household gross income)

		+ Controls, fixed effects	Contemporary instrument for commuting time	
	(1)	(2)	(3)	(4)
	OLS	OLS	2SLS	2SLS
Listed building	0.1155***	0.1252***	0.1165***	0.1171***
	(0.0338)	(0.0370)	(0.0387)	(0.0372)
Share historic district	3.0024***	2.2823***	2.1309***	2.1445***
	(0.2149)	(0.1988)	(0.2231)	(0.2016)
Share built-up land	3.2642***	2.5656***	2.1545***	2.2238***
	(0.0871)	(0.0860)	(0.3144)	(0.1742)
Share water bodies	3.3416***	2.4511***	2.4469***	2.3972***
	(0.3320)	(0.3364)	(0.3347)	(0.3199)
Commuting time *(log)*		-0.0270***	**-1.1971**	**-0.7256***
		(0.0070)	(**0.8590**)	(**0.4016**)
Household controls	No	Yes	Yes	Yes
Housing controls	No	Yes	Yes	Yes
Job controls	No	No	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Travel-to-work-area fixed effects	No	Yes	Yes	Yes
Workplace fixed effects	No	No	No	Yes
Number of observations	10,213,540	10,213,540	10,213,540	10,213,540
R^2	0.5134	0.5970		
Kleibergen-Paap F-statistic			55.92	232.2

Notes: **Bold** indicates instrumented. Household controls include household size, mean age of adults, mean gender, household type (couple, single, kids), the share of the household that is foreign-born. Job controls are the total hours worked, whether the household has a company car, the share of full-time contracts, the share of permanent contracts. Housing controls include house type, height of the building, construction year dummies and whether a building is listed. Standard errors are clustered at the neighborhood level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Table B.4 – Other descriptive statistics for *NVM* data

	(1)	(2)	(3)	(4)
	mean	sd	min	max
House price *(in €)*	229,238	116,074	25,000	1,000,000
Share land in historic district <500m	0.0695	0.192	0	1
Listed buildings <500m	0.179	0.894	0	19.53
Share water bodies <500m	0.0411	0.0713	0	0.920
Share open space <500m	0.244	0.217	0	1
Shops, <500m	0.254	0.394	0	4.711
Hotels, restaurants, cafés <500m	0.159	0.364	0	7.983
Leisure establishments <500m	0.0127	0.0215	0	0.318

Notes: The number of observations is 695,709. Because of confidentiality restrictions the minimum and maximum values refer to the 0.01% and 99.99% percentile. This implies that we exclude the bottom and top 70 observations.

Table B.5 – Determining the hedonic amenity index

(Dependent variable: the log of house price per m^2)

	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Share land in historic district <500m	0.1796***		0.1710***	0.1695***
	(0.0210)		(0.0204)	(0.0209)
Listed buildings <500m	0.0047**		0.0052**	-0.0043
	(0.0024)		(0.0024)	(0.0029)
Share water bodies <500m		0.3014***	0.2824***	0.2869***
		(0.0255)	(0.0253)	(0.0251)
Share open space <500m		0.0604***	0.0636***	0.0690***
		(0.0084)	(0.0084)	(0.0085)
Shops <500m				-0.0084
				(0.0074)
Hotels, restaurants, cafés <500m				0.0423***
				(0.0118)
Cultural establishments <500m				0.0480
				(0.0640)
Leisure establishments <500m				0.0232
				(0.0730)
Housing controls	Yes	Yes	Yes	Yes
Neighborhood fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Number of observations	695,709	695,709	695,709	695,709
R^2	0.8206	0.8207	0.8217	0.8219

Notes: Housing controls include house type, house size, whether the property has a garage, garden and/or central heating, the number of layers of insulation, the maintenance quality, the number of rooms, construction year dummies and whether a building is listed. Standard errors are clustered at the neighborhood level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Table B.6 – Descriptive statistics for historic data

	(1) mean	(2) sd	(3) min	(4) max
Employment accessibility in 1909	38,029	23,884	1,494	163,349
Share of high-skilled workers in 1909	0.0298	0.0285	0	0.197
Share of medium-skilled workers in 1909	0.216	0.128	0.00386	0.688
Population accessibility in 1900	89,184	62,641	3,008	422,544
Share built-up land in 1900	0.0432	0.103	0	0.930
Share water in 1900	0.0591	0.175	0	1
Share locals in 1899	0.643	0.102	0.217	0.950
Share protestants in 1899	0.518	0.337	0	0.998

Notes: The number of observations is 10,213,540. Because of confidentiality restrictions the minimum and maximum values refer to the 0.01% and 99.99% percentile. This implies that we exclude the bottom and top 1,024 observations

Table C.1 – Pictures and commuting time: first-stage regression results

	Dependent variable: the log of pictures per ha		*Dependent variable: the log of commuting time*			
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
Employment accessibility *(log)*			-0.1835***	-0.2945***		
			(0.0158)	(0.0154)		
Pictures per ha *(log)*			-0.0411***	-0.0465***		
			(0.0022)	(0.0019)		
Share built-up land in 1900	4.2960***	4.2413***			-0.3319***	-0.3119***
	(0.3192)	(0.3075)			(0.0364)	(0.0348)
Share built-up land in 1900, 0-500m	0.6415	0.7537			0.3573**	0.2261
	(1.2821)	(1.2279)			(0.1483)	(0.1482)
Share built-up land in 1900, 500-1000m	7.2817***	6.5271***			-1.0300***	-1.1923***
	(1.4426)	(1.3754)			(0.1837)	(0.1869)
Share water in 1900	0.8896***	0.9869***			0.1885***	0.1601***
	(0.3413)	(0.3222)			(0.0452)	(0.0402)
Employment accessibility in 1909 *(log)*	0.1472**	0.1624***			-0.0353***	-0.0499***
	(0.0593)	(0.0569)			(0.0084)	(0.0096)
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
Housing controls	Yes	Yes	Yes	Yes	Yes	Yes
Job controls	No	Yes	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Travel-to-work-area fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Workplace fixed effects	No	Yes	No	Yes	No	Yes
Number of observations	10,213,540	10,213,540	10,213,540	10,213,540	10,213,540	10,213,540
R^2	0.4907	0.5291	0.0423	0.1612	0.0357	0.1496

Notes: Household controls include household size, mean age of adults, mean gender, household type (couple, single, kids), the share of the household that is foreign-born. Job controls are the total hours worked, whether the household has a company car, the share of full-time contracts, the share of permanent contracts. Housing controls include house type, height of the building, construction year dummies and whether a building is listed. Standard errors are clustered at the neighborhood level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

TABLE C.2 – BIAS CORRECTED ESTIMATES

(Dependent variable: the log of household gross income)

	Travel-to-work-area fixed effects		Municipality fixed effects
	(1)	(2)	(3)
	Bias-adj	Bias-adj	Bias-adj
Pictures per ha *(log)*	0.0255***	0.0850*	0.0289**
	(0.0026)	(0.0497)	(0.0111)
Commuting time *(log)*	-0.0585***	-0.445***	-0.112***
	(0.0040)	(0.0306)	(0.0068)
Household controls	Yes	Yes	Yes
Housing controls	Yes	Yes	Yes
Job controls	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Travel-to-work-area fixed effects	Yes	Yes	Yes
Municipality fixed effects	No	No	Yes
Workplace fixed effects	Yes	Yes	Yes
δ	1.0	1.0	1.0
Π	1.5	2.0	1.5
Number of observations	10,213,540	10,213,540	10,213,540

Notes: Household controls include household size, mean age of adults, mean gender, household type (couple, single, kids), the share of the household that is foreign-born. Job controls are the total hours worked, whether the household has a company car, the share of full-time contracts, the share of permanent contracts. Housing controls include house type, height of the building, construction year dummies and whether a building is listed. Standard errors are clustered at the neighborhood level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Table C.3 – Identification revisited

(Dependent variable: the log of household gross income)

	Only Randstad	+ Municipality fixed effects	Control for current land use	1909 skills	Other instruments	Reclaimed land
	(1) 2SLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
Pictures per ha *(log)*	**0.0300*****	**0.0240*****	**0.0206****	**0.0229*****	**0.0305*****	**0.0191**
	(**0.0102**)	(**0.0072**)	(**0.0080**)	(**0.0069**)	(**0.0070**)	(**0.0143**)
Commuting time *(log)*	**-0.2427****	**-0.2230*****	**-0.5618*****	**-0.2962*****	**-0.1647****	**-0.0561**
	(**0.1146**)	(**0.0861**)	(**0.1121**)	(**0.0865**)	(**0.0790**)	(**0.1077**)
Share built-up land			-0.1632***			
			(0.0238)			
Population per ha *(log)*			-0.0234***			
			(0.0041)			
Share of medium-skilled workers in 1909				-0.2068***	-0.1747***	
				(0.0287)	(0.0265)	
Share of high-skilled workers in 1909				-0.1187	-0.0822	
				(0.0990)	(0.0956)	
Share protestants in 1899				-0.0008	-0.0072	
				(0.0113)	(0.0102)	
Population accessibility in 1900 *(log)*				-0.0053	0.0002	
				(0.0056)	(0.0051)	
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
Housing controls	Yes	Yes	Yes	Yes	Yes	Yes
Job controls	Yes	Yes	Yes	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Travel-to-work-area fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Municipality fixed effects	No	Yes	No	No	No	No
Workplace fixed effects	Yes	No	Yes	Yes	Yes	Yes
Number of observations	4,340,639	10,213,540	10,213,540	10,213,540	10,213,540	10,213,540
Kleibergen-Paap F-statistic	7.500	18.18	10.05	16.08	21.49	0.264

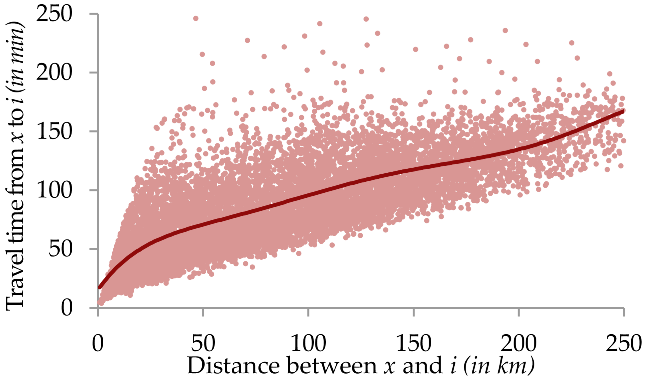
Notes: **Bold** indicates instrumented. Household controls include household size, mean age of adults, mean gender, household type (couple, single, kids), the share of the household that is foreign-born. Job controls are the total hours worked, whether the household has a company car, the share of full-time contracts, the share of permanent contracts. Housing controls include house type, height of the building, construction year dummies and whether a building is listed. Standard errors are clustered at the neighborhood level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Table C.4 – Other robustness checks

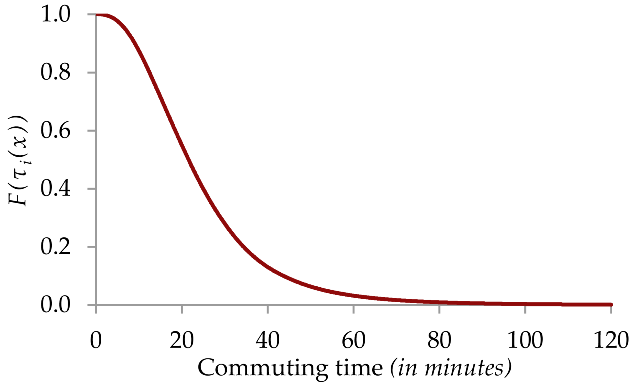
	One job households	+ Single plant firm	Company car	Education level	Commuting by rail	City center <15km
	(1) 2SLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
Pictures per ha *(log)*	**0.0239***** **(0.0065)**	**0.0230***** **(0.0068)**	**0.0207**** **(0.0085)**	**0.0165*** **(0.0090)**	**0.0294***** **(0.0061)**	**0.0229*** **(0.0119)**
Commuting time *(log)*	**-0.2127***** **(0.0643)**	**-0.1869***** **(0.0691)**	**-0.2285**** **(0.1009)**	**-0.4740***** **(0.0927)**		**-0.3173**** **(0.1440)**
Commuting time by rail *(log)*					**-0.1045**** **(0.0454)**	
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
Housing controls	Yes	Yes	Yes	Yes	Yes	Yes
Job controls	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Travel-to-work-area fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Workplace fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	6,706,524	3,532,906	1,523,567	7,626,355	10,213,540	6,023,886
Kleibergen-Paap F-statistic	17.63	14.56	7.026	17.02	10.15	5.657

Notes: **Bold** indicates instrumented. The dependent variable in columns (1)-(3) and (5)-(6) is the log of gross yearly income. In column (4) it is the share of the adults in the household that has a bachelor's degree or higher. Household controls include household size, mean age of adults, mean gender, household type (couple, single, kids), the share of the household that is foreign-born. Job controls are the total hours worked, whether the household has a company car, the share of full-time contracts, the share of permanent contracts. Housing controls include house type, height of the building, construction year dummies and whether a building is listed. Standard errors are clustered at the neighborhood level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Appendix figures

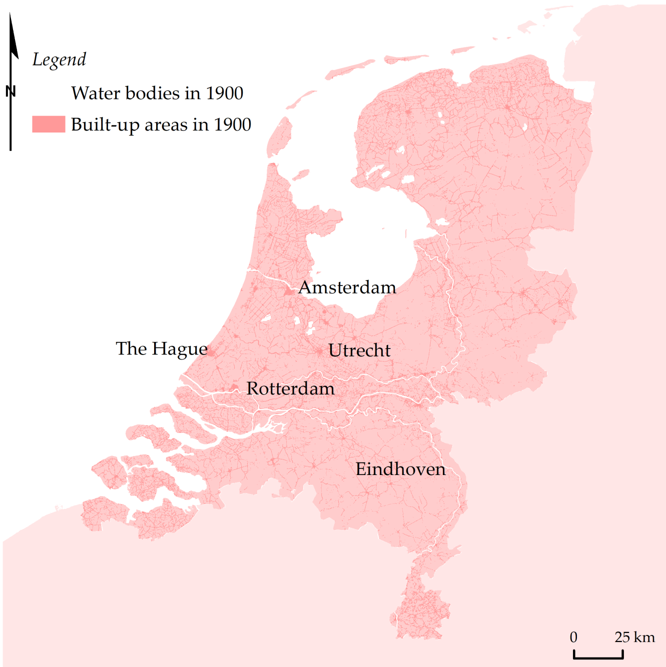


(A) Distance and travel time

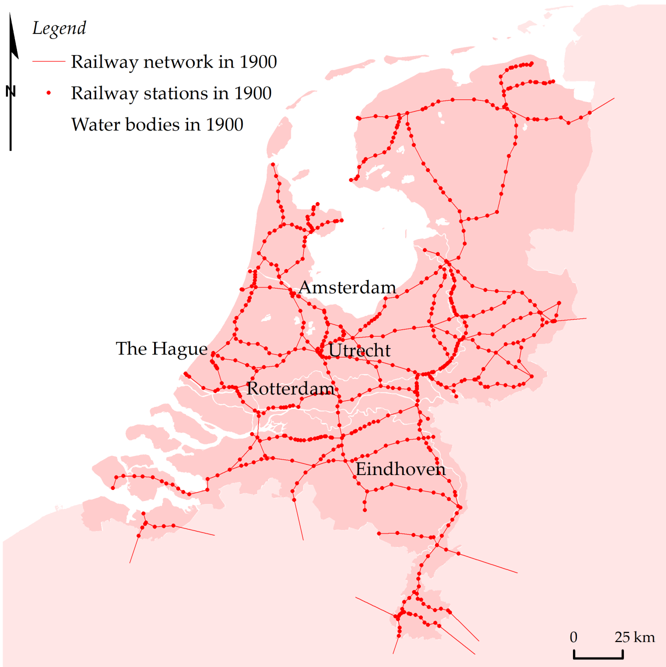


(B) Commuting time distribution

Figure B.1 – Calculation of travel time and speed

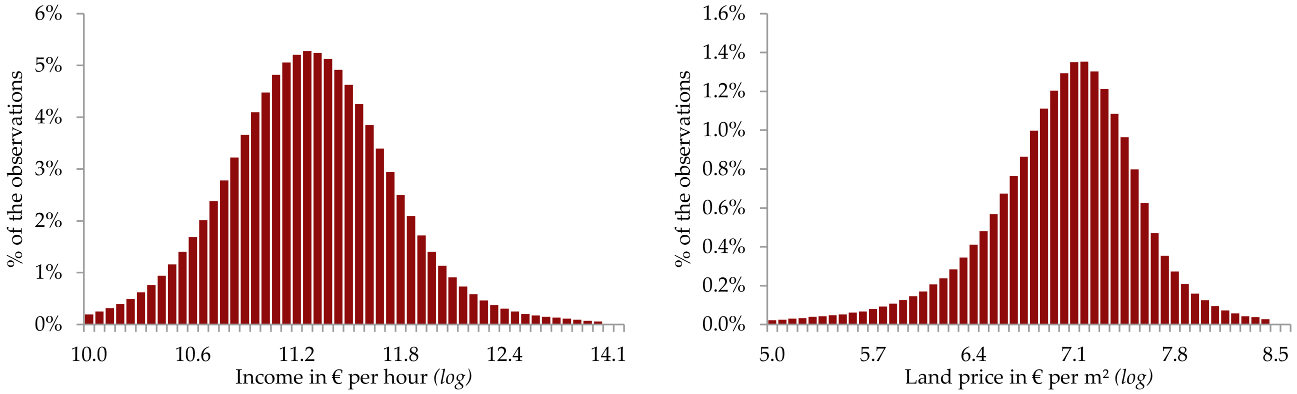


(A) Built-up land



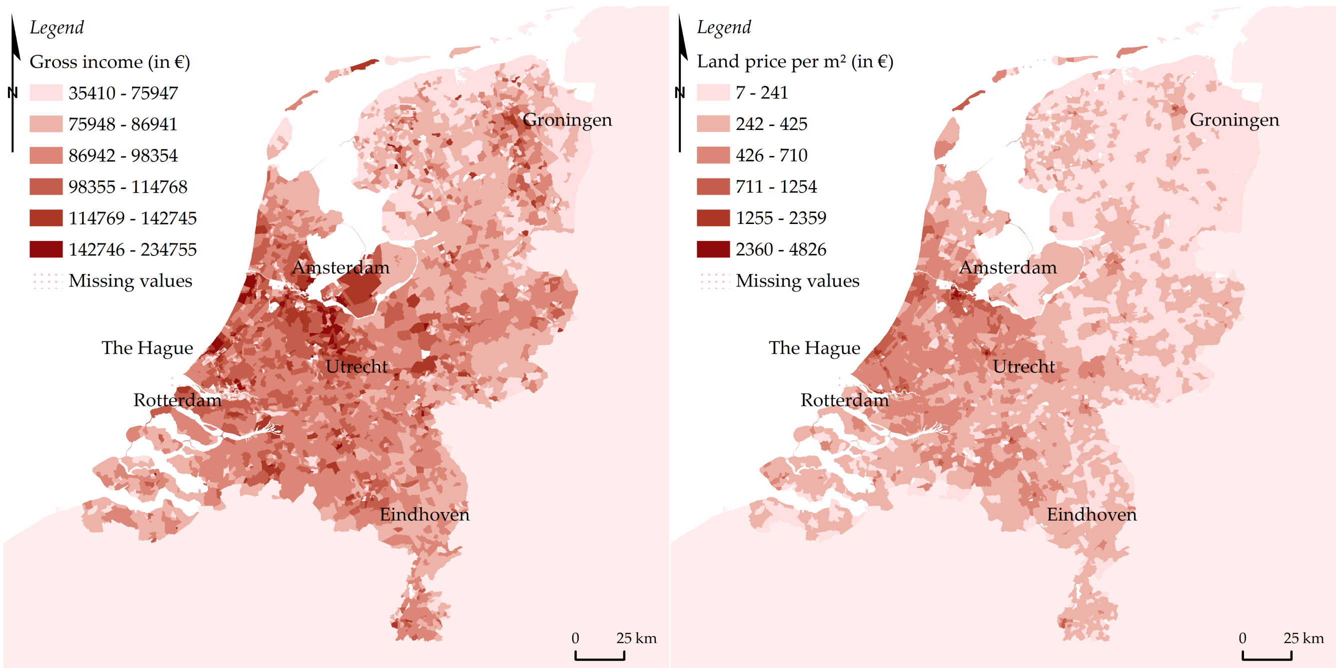
(B) The railway network and accessibility

Figure B.2 – Historic data from 1900



(A) Incomes (B) Land prices

Figure B.3 – Histograms for the variables of interest



(A) Average gross income *(in €)* (B) Average land prices per m^2 *(in €)*

Figure B.4 – Spatial distribution of variables of interest