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

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

November 2, 2018

Abstract

We study the sorting of income-heterogeneous consumers within cities. We allow for non-homothetic preferences and locations that are differentiated by their distance to employment centers and accessibility to exogenous amenities. The residential equilibrium is driven by the properties of an amenity-commuting aggregator obtained from the primitives of the model. Using micro-data on the Randstad (the Netherlands), we find that doubling the amenity level, resp. commuting time, attracts households whose incomes are $1-2.5\%$ higher, resp. $6-17.5\%$ lower. Using the model's estimated parameters, we predict the impact of changes in accessibility to jobs and amenities on the social structure of the Randstad.

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

JEL classification: R14, R23, R53, Z13.

*We are grateful to N. Baum-Snow, J. Brinkman, P.-A. Chiappori, 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, 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 Science Foundation under the grant N°18-18-00253.

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

1 Introduction

The sorting of workers *between* cities along the skill dimension is a major trend of our economies (Moretti, 2012; Davis and Dingel, 2018). Equally important is the sorting of workers by income *within* cities as residential segregation seems to generate negative and persistent effects on individual development and social mobility (Topa and Zenou, 2015; Chetty *et al.*, 2016; Chetty and Hendren, 2018). This spatial fragmentation is said to undermine the working of the city and to threaten social cohesion as a whole, which is likely to trigger strong and disruptive political effects (Becker *et al.*, 2017; Inglehart and Norris, 2016). Furthermore, over the last 40 years, the rise in income inequality was accompanied by an increase in the residential sorting by income in most of the 117 largest US metropolitan areas (Bischoff and Reardon, 2013; Glaeser *et al.*, 2008a). In particular, large cities, which are often the most productive, disproportionally attract both high and low-income workers (Eeckhout *et al.*, 2014). In the same vein, the spatial separation between the affluent and the poor seems to be on the rise in European cities (Musterd *et al.*, 2017). Therefore, we find it important to study the various forces that underpin the sorting of income-heterogeneous households within cities and to identify tools that may curb this trend. This is where we hope to contribute by developing a simple, but rich enough, model that can reproduce different sorting patterns, while being able to be tested by using microdata.

Studying the social stratification of cities is challenging for at least two reasons. First, one needs a setup developed within the realm of urban economics to account for what this literature considers as the main drivers of residential choices, that is, commuting and housing consumption. Second, the setup must be able to reflect the wide variety of situations observed in real-world cities. In this respect, the canonical model of urban economics in which locations are differentiated only by their distance to the central business district (CBD) leads to a fairly extreme prediction: households are sorted by increasing income order as the distance to the CBD rises. The reason is that households' desire to consume more space leads the affluent to reside further away from the city center (Hartwick *et al.,* 1976; Fujita, 1989; Kamecke, 1993). Hence, in a featureless city the residential equilibrium implies perfect sorting: the wider the income gap between two households, the greater the distance between their residential locations, and vice versa. Note that this extreme form of spatial sorting is not the outcome of social externalities; it is produced by competition on the land market alone. Furthermore, this pattern is not what we observe in the real world where several metropolitan areas display pronounced U-shaped or W-shaped spatial income distributions (Glaeser *et al.,* 2008b; Rosenthal and Ross, 2015).

##One missing key explanation is the existence of historic and natural amenities, such as scenic landscapes, rivers, historic buildings and architecture. These are specific features of many cities.## That such amenities matter in residential choices should not come as a surprise (Brueckner *et al.*, 1999; Glaeser *et al.*, 2001; Koster *et al.*, 2016; Ahlfeldt, 2016; Koster and Rouwendal, 2017; Lee and Lin, 2018). Since amenities and commuting time are generally not perfectly correlated, residential choices are the outcome of the interplay between three fundamental forces: amenities, commuting and housing. What we do not know is how the interaction between these forces determines the social stratification of the city. We thus propose a new approach in which cities are "featureful" in that locations are distinguished by two vertically differentiated attributes, that is, the distance to employment centers and the accessibility to *given* and *dispersed* amenities. While the demand for amenities has a tendency to rise with income, high-income commuters bear higher costs than low-income commuters. More specifically, we show that a featureful urban space is sufficient to replicate different spatial patterns of income sorting.

The study of income sorting when locations are differentiated by more than one attribute brings about new and difficult issues. At first sight, the determination of a residential equilibrium seems to have the nature of a matching problem between landlords and households in which land at any specific location is differentiated by two characteristics and households by one characteristic. This raises two types of difficulties. First, a household's housing consumption varies with both its income and location while it is exogenous in matching theory (Chiappori, 2017). Landvoigt *et al.* (2015) and Määttänen and Terviö (2014) provide examples of this approach in their studies of the housing market. Thus, we cannot appeal to the techniques of matching theory to solve our problem and have to develop an alternative approach. Second, if the sorting between incomes and locations is imperfect, the rule $x(\omega)$ which assigns a particular income ω to locations x must be a *correspondence*. For example, for the same given housing consumption, a household can be indifferent between living close to the CBD while having a low level of amenities, or living far from the CBD while enjoying a high level of natural 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 figure out what the equilibrium assignment could be. By contrast, it went unnoticed that the reverse problem can be solved. Indeed, because households bid for locations, those who reside at the same place must share the same gross income. Therefore, we may define and characterize the *income mapping* from the location set to the income set, which specifies which households are located at a particular location. As known from urban economics, this income is that of the households who make the highest bid (Fujita, 1989).

What are our main findings? Even though there is some skepticism about the ability of the standard methods of urban economics to deal with heterogeneous households (Duranton and Puga, 2015), we show that the bid-rent approach is applicable to the case of income-heterogeneous consumers. By using this approach we can pin down the properties of the equilibrium income mapping. We then show within a general framework that the interaction between amenities and commuting gives rise to turning points in the spatial income distribution. This is to be contrasted to existing studies, which often depend on specific functional forms or fine details about some key parameter values (Duranton and Puga, 2015). Regardless of the nature of the functional form of the amenity, commuting cost and income distributions, households' spatial sorting is *imperfect* provided that amenities are distributed *unevenly* across the city and/or employment is *decentralized*. In other words, a greater distance between two households no longer implies a wider income gap.

To investigate further how amenities, commuting and housing interact to shape the city and to test empirically our conclusions, we need a full characterization of the equilibrium income mapping. Homothetic preferences must be ruled out because they are associated with multiplicity of equilibria. Under Stone-Geary preferences, the residential equilibrium can be predicted from the spatial distribution of a *location-quality index*, which blends the two sources of location heterogeneity considered in this paper, i.e., amenities and commuting costs. This index is built endogenously from the primitives of the model because the Spence-Mirrlees condition appears to hold in the location-quality index and the household income space. However, our approach is *not* equivalent to working with a single source of location heterogeneity as it disregards interactions between amenities and commuting costs, which appear to be empirically relevant. We also show that there exists a unique residential equilibrium. Note that these results are not specific to the Stone-Geary preferences. They hold true for other non-homothetic preferences: what changes is the functional form of the location-quality index.

Furthermore, households sharing similar incomes need not live in the same neighborhood. Instead, they may live in separated neighborhoods where they enjoy similar values of the location-quality index. Hence, there is imperfect spatial sorting between income classes because of the spatial splitting of individuals belonging to the same income class. Equally important, the properties of the location-quality index may be used to test the predictions about how amenities affect the income residential pattern. The upshot is that the city's bliss point is not the CBD or the city limit anymore. The global maximizer of the location-quality index is now the most-preferred location for all households, implying that this location is occupied by the affluent because they have the highest bid. As one moves away from this location, households are sorted by decreasing incomes until a *local* minimizer of the location-quality index is reached where low-income households, but not necessarily the poorest, are located. Beyond this minimizer, household income starts rising. In other words, we have perfect sorting by increasing or decreasing incomes between two adjacent extrema of the location-quality index. Since the sign of the income gradient changes at any extremum of the location-quality index, *households get more exposure to other income-groups when the number of turning points of the quality index rises*. Yet, the tyranny of the bid rent remains implacable: the affluent still occupy the best locations.

A major advantage of our model is that it is parsimonious and tractable as the main parameters of the model can be estimated using only the equilibrium income mapping and land rent ##equation##. We test the equilibrium income mapping under Stone-Geary preferences for the Randstad, i.e., the main polycentric metropolitan area in the Netherlands, which is inhabited by about 7.1 million people. We use rich microdata for 4.3 million households covering the years 2010 to 2015 on incomes, residential and job locations at the household level, as well as land values and amenities at each location. We also test whether amenities and commuting costs determine the sorting pattern of heterogeneous households within the Randstad by exploiting the structure of the theoretical model developed in the previous sections. Within the Randstad there is sufficient variation in accessibility to jobs, and so in commuting costs. Dutch cities, which were established long ago, are well known for providing different types of high-quality amenities. For example, the Randstad hosts more UNESCO world heritage sites than London and Paris (more on this in Section 2). In addition, the basic public services that underpin social cohesion are centrally financed and/or administered in the Netherlands. As a result, competition between jurisdictions that characterizes U.S. metropolitan areas is much less of an issue. We use a novel proxy for amenities: the number of outside geocoded pictures taken by residents at a certain location. Hence, this amenity index, because it is based on households' behavior, is likely to capture exogenous and persistent amenities, such as natural geographic features and historic buildings. This allows us to move beyond the approach of defining amenities implicitly, as in Ahlfeldt *et al.* (2015) and Albouy (2016).

Admittedly, households care about the proximity to private facilities such as shops, restaurants and theaters, which may be disproportionately located in upscale neighborhoods where many pictures are taken. In addition, since there is no proxy that perfectly captures the full amenity potential at a certain location, amenities are measured with error. Commuting time is also likely to be endogenous due to correlation with unobservable household characteristics (e.g., higher educated households who operate in thinner labor markets often have longer commutes) and agglomeration economies – the latter being more prevalent in dense areas where commutes are shorter. We address the endogenous nature of amenities and commuting time in our econometric analysis in several ways. First, we use ancillary detailed data on demographic, housing and site characteristic and include work location fixed effects that ensure that we control for any productivity effects. We further instrument amenities and commuting time with proxies for observed amenities, such as distance to historic districts, and a measure of employment accessibility. In addition, we digitize historic city maps from 1832 and 1900 and use the historic land use, amenity and population patterns to instrument for current amenities and commuting costs. 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. We further show the robustness of our results by using alternative proxies for amenities based on Lee and Lin (2018), who use variations in housing prices, and an amenity index based on the augmented reality game *Pokémon Go.*

The results unambiguously suggest that both amenities and commuting costs are important in determining the urban income distribution. We find that doubling the amenity level attracts households whose incomes are $1 - 2.5\%$ higher. Furthermore, when commuting costs are twice as high, households' incomes decrease by $6 - 17.5\%$. Hence, commuting costs are a more important driver of income sorting than amenities. In other words, 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, this is not the end of the story: focusing on amenities only is unwarranted because commuting costs are too important to be disregarded. A relevant theory of city structure must account explicitly for both amenities *and* commuting costs.

By using information on land prices and lot sizes we are also able to estimate the necessary parameters of the model and to undertake counterfactual analyses. For example, we assume that the Randstad has a spatial structure akin to that of a typical American polycentric metropolitan area without mass transit and historical amenities; jobs are concentrated in the centers of the four major cities. This spatial structure leads to much less social mixing, and to a drop of about 8.5% in ##aggregate## land rents. Moreover, people commute longer, but live on 25% larger plots. Furthermore, changes in travel modes may also have significant impacts on the social structure of cities. These exercises illustrate that our model can be readily used to predict changes in the social structure generated by specific changes in accessibility or amenities, e.g., due to policy interventions.

Our results show that, contrary to the general belief, promoting equal access to amenities/public services favors residential segregation by income, whereas *the uneven provision of amenities across the city fosters income mixing*. Therefore, place-based policies that aim to reduce amenity disparities may exacerbate spatial segregation. The intuition is easy to grasp. If amenities are equally distributed across locations and employment is all located in the CBD, we fall back on the standard monocentric city model with perfect sorting. On the contrary, if amenities are unequally distributed along the distance to the city center, higher-income households split into distant residential areas, so that households are, on average, closer to households with incomes very different from theirs. To be effective, such policies must create long-lasting change in neighborhoods with a low quality index. Our approach shows how one can direct the locations of the rich by improving the amenity level in specific neighborhoods. Of course, pro-income mixing policies are not sufficient to defeat the alleged negative consequences of residential segregation, but they are likely to alleviate them.

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. Unfortunately, the assignment approach used by Beckmann was flawed (Montesano, 1972). Fujita (1989) proposed a rigorous analysis of the residential pattern for a finite number of income classes. When commuting costs are distance-dependent and income-independent, Fujita shows that income-classes are ranked by increasing income as the distance to the CBD rises. Kamecke (1993) extended this result to a continuum of heterogeneous households by showing that there is perfect sorting moving out from the CBD by increasing incomes.

To the best of our knowledge, Roback (1982) was the first who studied how the labor and land market interact in the presence of amenities. She considered a setting in which a large number of households identical in terms of skills and income distribute themselves across a large number of communities differentiated by their given level of amenities. Our paper is closer to Lee and Lin (2018) who showed that richer households are anchored in neighborhoods with better natural amenities. Very much like us, their results support the importance of exogenous (natural) amenities for the persistence of the social structure of cities. Carlino and Saiz (2008) showed that the number of tourists visiting a city is a good predictor of the growth of US metropolitan areas in the 1990s. After controlling for a city's proximity to the ocean and its climate, they found that population and employment grew by about 2% more in a city with twice as many tourists as another comparable city, which confirms that amenities matter. These papers differ from us in one fundamental aspect: people are assumed to work where they live. By contrast, we account explicitly for commuting costs between the residence and the workplace.

In an important paper, Ahlfeldt *et al.* (2015) highlight the role of amenities in residential location choices in their study of the internal structure of Berlin. They find that the elasticity of amenities with respect to residential density is 0.15, which is quite high. However, this is 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). Hence, using residuals as a proxy for amenities mechanically leads to a high model fit. In our paper, we strive to show that amenities and commuting costs have a causal and significant impact on the urban structure. By concentrating explicitly on amenities and commuting that we know to be important, ##we are able to figure out how these factors contribute to explaining the residential equilibrium##. Differently from us, Ahlfeldt *et al.* consider an open city model in which the total city population is endogenous while households enjoy the same exogenous utility level. In contrast, we work with a closed city model in which the utility level is endogenous and varies across households heterogeneous in income. Last, but not least, apart from the existence and uniqueness of the spatial equilibrium, Ahlfeldt *et al.* do not provide any characterization of this equilibrium. In sum, our model does not belong to the family of spatial quantitative economics, but to the more traditional literature of urban economics. We do not see this as a flaw because the current state of the art in spatial quantitative economics does not allow us to answer the questions that serves as the main motivation of this paper. However, we are able to undertake counterfactuals that provide useful insights in the quantitative implications for the spatial income distribution and land rents when the location quality changes.

Note also the link with the vast literature on Tiebout and the sorting of households across different jurisdictions. The main focus of this literature is on stratification by income. Bénabou (1996) developed a model of community formation and human capital accumulation that highlights the roles played by preferences, capital markets, neighborhood effects, and local public goods on stratification. When households are heterogeneous in incomes and preferences for local public goods, Epple and Sieg (1999) showed that several jurisdictions may host residents having the same income. Unlike Epple and Sieg, we show that households sharing the same preferences and the same income may live far apart and have different levels of housing consumption. However, these contributions disregard commuting within jurisdictions, and thus do not study the trade-off between amenities, proximity to jobs and housing prices, which occupies center stage in our approach. O'Sullivan (2005) and De Bartolome and Ross (2007) are noticeable exceptions. Assuming fixed-lot size, O'Sullivan showed that income mixing occurs when consumers differ in their marginal disutility of travel. However, crime disrupt this pattern because high-income consumers have more to lose from crime. De Bartolome and Ross developed a Tiebout-like model where individuals have to commute to the CBD. Assuming fixed-lot size, they showed that commuting cost and the demand for public service may give rise to income mixing in a city formed by a central and a peripheral jurisdiction.

The remainder of the paper is organized as follows. In Section 2, we compare the income gradients in the Randstad (the Netherlands). We provide a detailed description of our model in Section 3 and to show how the bid rent function may be used to determine the social stratification of the city. In Section 4, we study the properties of the residential pattern for preferences that generate a location-quality index. Since the equilibrium is undetermined under homothetic preferences, we illustrate our results for Stone-Geary preferences. In Section 5, we determine analytically the market outcome when income and the location-quality index are Fréchet-distributed. Data are discussed in Section 6. In Section 7, we study the causal relationship between incomes, amenities, and commuting times, while we present the results of our counterfactuals in Section 8. In Section 9, we summarize our main results and discuss two possible extensions of our basic setup.

2 Are income gradients monotonic?

In this section, we provide some graphical evidence showing that income gradients need not be monotonic. Our empirical analysis focuses on the Randstad. The Randstad comprises the four largest Dutch cities, i.e., Amsterdam (the capital, which was also the economic capital of Europe in the 17^{th} century), Rotterdam (which is the largest seaport in Europe for a very long time), The Hague (the seat of the national government and administration), Utrecht (historically the religious center of the Netherlands since the 8^{th} century), and their surrounding areas. With a population of 7.1 million and covering an area of 6.4 thousand km^2, it is one of the largest polycentric metropolitan areas in Europe, and comparable in size to the San Francisco Bay Area. The Randstad has a higher population density than comparable metropolitan areas elsewhere. With about 1100 inhabitants per km^2, it is about 50% denser than Greater Paris (722), New York (688) or Chicago (509). The Randstad is often considered as a single urban area due to the close proximity and functional relationships between its different cities (see Figure 1 for more details). For example, it takes only about 20 minutes to travel from Utrecht to Amsterdam, implying that many people work in one city and live in another.[1]

The current urban structure is heavily influenced by the past as Amsterdam, Rotterdam and The Hague have a history that dates back to the 13^{th} century, while Utrecht was founded around the turn of the 7^{th} century. Hence, the inner cities host many historic buildings and the pattern of canals in Amsterdam and Utrecht is still present. In contrast to many US cities, the presence of historic structures imply that the amenity levels in inner cities are much higher. As an illustration, the Randstad hosts 6 UNESCO world heritage sites (and 4 on a tentative list), which is more than London (5) and Paris (4). Furthermore, there are 153 national historic districts and $27,175$ national listed buildings located in the Randstad.[2] Rotterdam, which was completely bombed out during World War II, hosts few historical amenities. However, Rotterdam still has a high amenity level in the center, due to the high architectural quality of the new-built buildings.[3]

The Randstad is characterized by an extremely flat geography.[4] The expansion of a city is, therefore, not much constrained by geographical features. However, Dutch land-use planning strongly restricts urban growth. A well-known example is the Green Heart, which is the vast open space in the center of the Randstad, shown in Figure 1. The Green Heart policy dates back to 1954: ##new constructions have been mostly prohibited there since ## (Koomen *et al.*, 2008). Another distinctive characteristic of Dutch cities is that school quality is unlikely to play a major role in household residential choices because school performances are much less uneven in the Netherlands than in the U.S., suggesting that the search for better school districts is less of a concern in the former (Ritzen *et al.*, 1997). Moreover, households are free to choose their most preferred school, independent of their residential location. Furthermore, racial disparities are often less pronounced than in the U.S. Dutch people have a preference for living among people of one's own ethnic group, but in a sufficiently diverse neighborhood (Ong, 2017).

[Figure 1 about here]

In Figure 2, we report income gradients for each of the four largest cities of the Randstad. Since cities are not symmetric around their center, we depict the income gradient along one dimension by taking averages along successive rings. It appears that *income gradients are highly non-monotonic*, especially for Rotterdam and The Hague where the gradients are double peaked. In this paper, we argue that such patterns may be (at least, partially) explained by the combination of amenities and commuting costs. In Sections 7 and 8, we study the relationship between incomes, amenities,

[1] Figure A2 in Appendix A.2 shows the employment accessibility and cross-commuting patterns between the different areas in the Randstad. From this map it becomes clear that the Randstad should be considered as a single functional urban area. Nevertheless, in our empirical analysis, we also analyse the different cities in the Randstad separately. This leads to similar results.

[2] This is more than Greater London, which has $18,920$ listed structures.

[3] For example, Rotterdam was ranked in the list 'top cities in the world' by Lonely Planet's Best in Travel 2016.

[4] It is just above (and sometimes below) sea level: the highest point is 50 m, while the lowest is -8.9 m in an area that has been reclaimed from the sea.

and commuting costs in more detail. The spatial income differences seem to be rather small – up to 10%. However, plotting income as a function of distance to the city center hides substantial heterogeneity in predicted income patterns across space. See the maps given in Appendix A.1 for an illustration.

[Figure 2 about here]

3 The model and preliminary results

3.1 The model

Consider a polycentric city spread over a space X. Jobs are available in exogenously set business districts located at $e_i \in X$, $i = 1, ...n$. The city is endowed with a unit mass of households/workers. There are two normal consumption goods: (i) land (h), which is a proxy for housing, and (ii) a composite non-spatial good (q), which is used as the numéraire. The opportunity cost of land is given by the constant $R_0 > 0$, while the amount of land available at each location x is normalized to 1. Let $a > 0$ be the amenity level (or, equivalently, an aggregator of distinct amenities), which is common to all households. Since all households prefer more amenities than less, we consider a preference structure similar to the one used in models of vertical product differentiation:

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

where u is strictly increasing in the numéraire q and land consumption h, strictly quasi-concave in (q, h), and indifference curves do not cut the axes. Preferences (1) imply that the utility level associated with the consumption of a given bundle (q, h) increases with the amenity level a, while the utility derived from consuming amenities rises with income since q and h are normal goods. 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, this condition does *not* hold between incomes and locations: richer households need not choose locations with more amenities because they also care about their commuting costs. In addition, we will see below that further restrictions must be imposed to u to rule out degenerate cases.

A household's gross income is given by ω units of the numéraire, with $\omega \in \mathbb{R}_+$. For example, the heterogeneity in incomes reflects differences in skills. The income c.d.f. $F(\omega)$ and density $f(\omega)$ are continuous over $\mathbb{R}_+$. For simplicity, we make the (heroic) assumption that a household's gross income is the same across employment centers. However, the model can be readily extended to the case where a household's gross income varies exogenously across employment centers (see the concluding section). We have chosen not to endogenize the income distribution because we aim to determine *how income disparities translate into urban inequalities*. Given this objective, it is hard to avoid working with an exogenous income distribution. By contrast, in Ahlfeldt *et al.* (2015), households are ex ante identical, but they earn ex post different incomes because employment centers' productivity is endogenous while households face idiosyncratic shocks. Thus, urban inequalities do not stem from the same source of heterogeneity as us.

To ease the exposition, we assume in the theoretical sections that X is the real line. By convention, location and distance from the origin are identically denoted by $x \in X$. However, our arguments can be extended to the case of a city network formed by a finite set of links that intersect at a finite number of nodes (see Appendix B.7). We model the individual loss due to commuting as an *iceberg* cost. If a household residing at x works at the employment center e_i, we denote by $0 < t_i(|x - e_i|) \leq 1$ her effective number of working units, which typically decreases with the distance $|x - e_i|$ while $t_i(0) = 1$.

The household's *net* income is equal to $\omega t_i(x)$. Therefore, her income depends on both her residential choice and her working place. Her commuting cost is given by $c_i(\omega, x) = \omega[1 - t_i(x)]$, which increases with both her earning ω and the length of her commute. In other words, commuting is considered 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 (Börjesson *et al.*, 2012).

Workers select the closest employment center because they earn their highest net income therein.[5] In other words, we have:

$$t(x) = \max_{i=1,\ldots,n} t_i(x).$$

However, for our purpose, it is sufficient to consider any given function $t(x)$ that specifies the working time of a worker residing at location x. Therefore, our results are independent of the mechanism that determines how workers are associated with a workplace through a particular labor contract or within a broader setting.

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).

Let $\eta(y) > 0$ be a given function whose value expresses the level of amenities available at $y \in X$. The amenities are intrinsic to a location and exogenous. Therefore, the corresponding

[5] 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.

utility level is negatively affected by the distance between the household and the place where these amenities are available. As a result, a household at x ascribes the value $\varphi(|x-z|)\eta(z)$ to the amenity provided at $z \neq x$, which decreases with the distance $|x-z|$ between x and z, with $\varphi(0)=1$. In other words, $\varphi(\cdot) \geq 0$ has the nature of a distance-decay function. As shown by (1), households' well-being at x depends on the amenity field defined by the following expression:

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

We do not impose any functional restriction on $\eta(\cdot)$ and $\varphi(\cdot)$. Therefore, (3) includes the functional forms used by Fujita and Ogawa (1982), Lucas and Rossi-Hansberg (2002) and others to describe spatial interactions. In the featureless city of urban economics, $a(x)$ is constant across locations because both φ and η are constant across space. In this paper, we focus on the case where $a(x)$ varies with x. Hence, the city is *differentiated* in the sense that preferences depend on household locations. However, our approach does not rely on social interactions because $a(x)$ is independent of households' locations. The use of the distance-decay function φ means that (3) can be viewed as a gravitational force.

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:[6]

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

which are both unique because u is strictly quasi-concave.

As will be discussed in Section 4, we focus on the case where there exists a mapping from $[0, B]$ to $\mathbb{R}_+$ that specifies which ω-consumers are located at x, where B is the endogenous city limit defined below. The ω-consumers may be distributed over several locations where $\sigma(x,\omega) \in [0,1]$ is the share of the ω-households who reside at x. The land market clearing condition holds if $\omega(x)$ satisfies the following condition:

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

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

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

[6]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).

The residential equilibrium is such that no household has an incentive to move, all households sharing the same income have the same maximum utility level, and the land market clears. Denote by $\omega^*(x)$ the *equilibrium income mapping* from the location set $[0, B]$ to $\mathbb{R}_+$ that specifies which ω-households are located at x. Formally, a *residential equilibrium* is defined by the following vector:

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

such that

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

holds under the constraints (2), (5) and (6). For simplicity, we assume that there is no agricultural land within the city. By implication, $R^*(x) > R_0$ for $x < B^*$ and $R^*(x) = R_0$ for $x \geq B^*$. In our setting, heterogeneous households reach different equilibrium utility levels. This is to be contrasted with Ahlfeldt *et al.* (2015) who assume that households share the same expected utility level, which is equal to the exogenous reservation utility that prevails in the rest of the economy. Note also that people cannot revise their choices once all uncertainty is resolved.

In equilibrium, households sharing the same income are *not* indifferent across locations because space is differentiated. In particular, if the inequality is strict in (7) for all $y \neq x$, then all $\omega^*(x)$-households are located at x ($\sigma^*(x, \omega^*(x)) = 1$). Otherwise, there exist at least two locations x and y such that the $\omega^*(x)$-households are indifferent between x and y. In this case, we have $0 < \sigma^*(\cdot, \omega^*(x)) < 1$ at x and y, while the sum of the shares is equal to 1. In this case, we say that there is *spatial splitting* of identical households.

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

3.2 The residential equilibrium with amenities

3.2.1 The equilibrium income mapping

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

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

where $Q(h, U/a(x))$ is the unique solution to $a(x) \cdot u(q,h) = U$ because u is strictly increasing in h and indifference curves do not cut the axes.

The bid rent $\Psi(x, \omega, U)$ is such that the ω-households are indifferent across locations because they reach the same utility level U. Therefore, (4) implies that the Alonso-Muth condition for the ω-households under a differentiated space becomes:

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

which boils down to the standard condition when $a(x)$ is constant.

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

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

whose solution, denoted $H(\omega t(x), U/a(x))$, is the quantity of land consumed at x by the ω-household at x if her bid rent is equal to the land rent; $H(\cdot)$ is called the *bid-max lot size* (Fujita, 1989). The equation (9) may have several solutions. In this case, there is multiplicity of equilibria. However, what we do in the remaining of this section applies to each solution which are utility-maximizers, hence to each equilibrium. Moreover, under Stone-Geary preferences, we show that the equilibrium is unique.

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

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

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

$$R^*(x) = \max\left\{ \max_{\omega\in\mathbb{R}_+} \Psi(x,\omega,U^*(\omega)), R_0 \right\},$$

where $U^*(\omega)$ denotes the maximum utility reached by the ω-households at the residential equilibrium.

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

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

while the second-order condition implies $\partial^2 \Psi / \partial \omega^2 < 0$. So far, we have implicitly assumed that (10) has a single maximizer in ω. However, this equation may have several solutions in ω. Under a finite number of income classes and a featureless monocentric city, there exists a unique solution because one income group overbids all the others for almost all x (Fujita, 1989). Therefore, it seems natural to assume that the maximization problem

$$\max_{\omega \in \mathbb{R}_+} \Psi(x, \omega, U^*(\omega))$$

has a unique maximizer for almost all x, i.e., a single highest bidder. Characterizing the classes 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 (11) with respect to x yields:

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

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

Set

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

and

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

We are now equipped to characterize the equilibrium income mapping.[7]

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

$$\Psi_{x\omega}(x, \omega, U^*(\omega))) \equiv \frac{t(x)}{H}\left[\left(1 - \frac{\varepsilon_{H,\omega} + \varepsilon_{u_q,\omega}}{\varepsilon_{U,\omega}}\right) A(x) - (1 - \varepsilon_{H,\omega}) T(x)\right] \tag{13}$$

is positive (negative) *at this location.*

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

The expression (13) shows that for any given function $u(q, h)$ the interaction between the amenity and commuting cost functions determines the social stratification of the city through the behavior of the function $\Psi_{x\omega}$. Furthermore, inspecting (13) shows that the impact of the amenity

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

and commuting cost functions on the sign of $\Psi_{x\omega}$ is hard to predict because it depends on the values of A and T, the income ω, and how the elasticity of the housing demand varies with income. Last, the sign of $\Psi_{x\omega}$, whence the slope of the spatial income distribution, changes at any solution $\omega^*(x)$ to the equation $\Psi_{x\omega} = 0$.

To illustrate, consider first the benchmark case of a monocentric and featureless city, that is, $A(x) = 0$ and $T(x) > 0$ for all x. We know from Fujita (1989) that household locations are determined by ranking the bid rent slopes with respect to income. It follows from (13) that the sign of $\Psi_{x\omega}$ 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 city ($A(x) \neq 0$). Owing to the existence of amenities, even when the bid rent functions are downward sloping, the equation $\Psi_{x\omega} = 0$ may have several solutions in x. 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 $\Psi_{x\omega} > 0$ 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 $\Psi_{x\omega} < 0$ 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 $\sigma(x(\omega), \omega) = 1$. 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 $\Psi_{x\omega}$ changes its sign over $[0, B]$ because, as illustrated in Section 2, the slope of the income gradient changes. 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 Section 4.

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

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

The following remark is in order. In Appendix B.2 we compare income-dependent and independent commuting costs. In the former, households are sorted by decreasing income order, whereas in the latter households are sorted by increasing income order. Thus, how to model commuting costs matters for the order in which households are ranked. Since there is ample evidence that suggests that the opportunity cost of time increases with income (Small, 2012), using income-independent commuting costs seems restrictive for studying the residential choices of households. This is why we have chosen to use an iceberg commuting cost.

4 The city social structure under Stone-Geary preferences

Proposition 1 shows that the equilibrium income mapping is a linear combination of $A(x)$ and $T(x)$ weighted by coefficients that depend on the utility $u(q, h)$. Therefore, to characterize the equilibrium income mapping and to bring it to the data, we must determine what these coefficients are. This is what we accomplish in this section.

(i) It seems natural to start with homothetic preferences, as they include the CES, Cobb-Douglas and translog. If the utility u is homogeneous linear, we show in Appendix B.3 that $\varepsilon_{U,\omega} = \varepsilon_{H,\omega} = 1$ and $\varepsilon_{u_q,\omega} = 0$. As a result, (13) can be reduced to $\Psi_{x\omega} = 0$ for all x. In other words, *there is a continuum of residential equilibria under homothetic preferences*. Therefore, if the aim to to characterize the impact of income heterogeneity on residential choices, homothetic preferences must be ruled out.

(ii) Quasi-linear preferences are non-homothetic and simple to handle: $u(q, h) = v(h) + q$ where v is strictly increasing and concave. In this case, we have $\varepsilon_{H,\omega} = \varepsilon_{u_q,\omega} = 0$, so that the bracketed term of (13) reduces to $A(x) - T(x)$. This expression suggests that quasi-linear preferences are a good candidate to study the residential equilibrium. Unfortunately, assuming quasi-linear preferences is counterfactual as housing is a normal good ($\varepsilon_{H,\omega} > 0$).

(iii) A well-known example of non-homothetic utility is Stone-Geary's:

$$u(q, h) = q^{1-\mu}(h - \overline{h})^{\mu}, \tag{14}$$

where $0 < \mu < 1$ and $\overline{h} > 0$ is the minimum lot size, which is supposed to be sufficiently low for the equilibrium consumption of the numéraire to be positive.[8] Maximizing (14) with respect to q and h subject to (2) leads to the linear expenditure system:

$$q^*(x, \omega) = (1 - \mu)[\omega t(x) - R(x)\overline{h}], \tag{15}$$

$$h^*(x, \omega) = (1 - \mu)\overline{h} + \mu \frac{\omega t(x)}{R(x)}. \tag{16}$$

The housing demand at any location x increases less than proportionally with income, which is in line with Albouy *et al.* (2016). This seems to oppose Davis and Ortalo-Magné (2011) who provide evidence that the expenditure share on housing is constant over time and across U.S. metropolitan areas. However, this does not mean that households having different incomes spend the same share of their incomes on housing within a city.

We show in Appendix B.4 that

$$\Psi_{x\omega} = \frac{t(x)\overline{h}}{H^2} \left[A(x) - (1 - \mu)T(x)\right], \tag{17}$$

which depends on the intensity of preferences for housing through the parameter μ. Set

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

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

Finally, consider the following utility:

$$u(q, h) = q^{\rho_1} + h^{\rho_2} \tag{19}$$

with $0 < \rho_i < 1$ and $\rho_1 \neq \rho_2$. The elasticity of substitution between housing and the numéraire is variable and equal to $1/(1 - s_1\rho_1 - s_2\rho_2)$ where s_i is the expenditure share on good $i = 1, 2$.[9] When $\rho_1 > \rho_2$, i.e., the composite good matters more than housing, it can be shown that the above preferences generate the index $\Delta(x) \equiv t(x) \left[a(x)\right]^{1/\rho_1}$, which is similar to (18).

Intuitively, (14) may be interpreted as a non-homothetic Cobb-Douglas utility and (19) as a non-homothetic CES. In what follows, we work with Stone-Geary preferences because they can be brought to the data in a direct way. However, our results hold true whenever the location-quality index $\Delta(x)$ is a function of $a(x)$ and $t(x)$ and is independent of income.

[8] Note that our results hold true if h is a congestible good such that the utility u decreases when the size of the household increases.

[9] An expression similar to (19) is used by Eeckhout *et al.* (2014) as a production function.

4.1 The spatial income distribution

Our objective is now to determine the equilibrium income mapping that specifies which ω-households are located at x under Stone-Geary preferences. 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 across the city 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 $G(\Delta)$ be the corresponding c.d.f. defined over $\mathbb{R}_+$.

Proposition 2. *Under Stone-Geary preferences, the residential equilibrium is unique. Furthermore, the income and location-quality index vary in the same way according to the relationship* $\omega^*(x) = F^{-1}[G(\Delta(x))]$.

Proof. The argument involves four steps.

(i) We show in Appendix B.4 that the bid-max lot size is unique and such that

$$H(\omega t(x), U/a(x)) \equiv H(\Delta(x), \omega, U), \tag{20}$$

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

(ii) Furthermore, as shown in Appendix B.5, the equilibrium condition $\Psi_\omega = 0$ is equivalent to the following differential equation:

$$\frac{\mathrm{d}U^*}{\mathrm{d}\omega} = \Delta^{\frac{1}{1-\mu}}(1-\mu)(H-\overline{h})^{\frac{\mu}{1-\mu}}\left(U^*(\omega)\right)^{-\frac{\mu}{1-\mu}}. \tag{21}$$

Since H depends on the location-quality index $\Delta(x)$, $U^*(\omega)$ also depends on the sole location-quality index.

(iii) It follows from Appendix B.5 that $U^*_\omega(\omega)$ is an increasing function of Δ:

$$\frac{\partial U^*_\omega(\omega)}{\partial \Delta} > 0. \tag{22}$$

In other words, the Spence-Mirrlees condition holds, which implies that the existence of a 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 $\omega(x)$ and $\Delta(x)$ (Chiappori, 2017). Since a given location x is associated with a unique value of the location-quality index, there exists an income value corresponding to x. Therefore, all the households located at x share the same income $\omega(x)$.

(iv) The equilibrium income mapping is unique and given by:

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

It then follows from (20) that the residential equilibrium exists and is unique.■

Thus, at the residential equilibrium, households ordered by increasing incomes are assigned to locations endowed with rising values of the location-quality index. As a result, it is sufficient to study how $\Delta(x)$ varies to determine the properties of the residential equilibrium. In particular, the functions $\omega^*(x)$ and $\Delta(x)$ have the same extrema. What is more, Proposition 2 has another intuitive implication: the willingness to pay for an additional unit of the location-quality index rises with the household's income. Indeed, as shown in Appendix B.5, the slope of the bid rent function increases with Δ if and only if the marginal utility of income also increases with Δ, that is, (22) is equivalent to

$$\Psi_{\omega\Delta}(x, \omega, U^*(\omega))|_{\Psi_\omega(.)=0} > 0,$$

which generalizes to the continuum the condition obtained by Fujita (1989) in the case of a finite number of income classes.

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

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

Therefore, income sorting is not mapped into spatial sorting: *the income gradient need not be a monotonic function of the distance from the CBD.* As a consequence, the affluent (or the poor) do not necessarily locate at the CBD or the city limit. Rather, the richest locate where the location-quality index is maximized, whereas the poorest reside in locations with the lowest location-quality index.[10]

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

Since $\Delta(0) > \Delta(x_2) > \Delta(B^*) > \Delta(x_1)$, the intermediate value theorem implies that z_1 in $[0, x_1)$, z_2 in (x_1, x_2) and z_3 in $(x_2, B^*]$ exist such that $\Delta(z_1) = \Delta(z_2) = \Delta(z_3)$. Proposition 2 implies that the households residing at these three locations have the same income. In other words, 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 3 depicts a spatial configuration where the middle class is split into two spatially separated neighborhoods with

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

the poor in between, while the affluent live near the city center. Such a pattern describes more accurately the spatial distribution of incomes in "old" US cities and in many European cities, than the homogeneous monocentric city model (Glaeser *et al.*, 2008).

[Figure 3 about here]

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

To determine the residential distribution of households, it remains to find the equilibrium values of the shares $\sigma(x, \omega(x))$. If $z_1 \neq z_2 ... \neq z_n$ exist such that $\Delta(z_1) = \Delta(z_j)$ for $j = 2, ..., n$, it follows from Proposition 2 that $\omega^*(z_1) = \omega^*(z_j)$ for $j = 2, ..., n$. Using (5) and (20), we also have:

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

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

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

Furthermore, it must be that

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

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

4.2 Land rent

It remains to characterize the equilibrium land rent. We show in Appendix B.6 that the equilibrium land rent is given by the following expression:

$$R^*(x) = \frac{\omega^*(x) t(x)}{H[\Delta(x), \omega^*(x), U^*(\omega^*(x))]} \left[1 - \frac{1 - \mu}{\varepsilon_{U,\omega}(x)}\right], \qquad (23)$$

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 (23) and rearranging terms, we obtain:

$$R^*(x) = \frac{\mu\omega^*(x)t(x)}{H[\Delta(x), \omega^*(x), U^*(\omega^*(x))] - (1-\mu)\overline{h}} > 0, \tag{24}$$

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

By totally differentiating (23) with respect to x, we obtain (see Appendix B.6):

$$R^*_x(x) = \frac{\omega^*(x)t(x)}{H[\Delta(x), \omega^*(x), U^*(\omega^*(x))]}\left[\frac{1}{\varepsilon_{U,\omega}(x)}A(x) - T(x)\right]. \tag{25}$$

Since $\varepsilon_{U,\omega}(x) > 1$, the above expression implies that the land rent gradient is always negative if $A(x)-T(x) < 0$ for all x. As x rises, the decreasing land rent compensates the $\omega^*(x)$-households for bearing higher commuting costs and being farther away from places endowed with more amenities. For example, in the standard monocentric city model in which $A(x) = 0$ and $T(x) > 0$ the land rent gradient is always negative. When $A(x) - T(x) > 0$ over some interval $[x_1, x_2]$, the land rent gradient can be positive or negative according to the value of $\varepsilon_{U,\omega}(x)$. Since household income increases over $[x_1, x_2]$, commuting costs also increase over this interval. Therefore, the land rent is a priori neither monotonic nor the mirror image of the spatial income distribution. However, $R^*(x)$ is upward sloping when $A(x) - \varepsilon_{U,\omega}(x)T(x) > 0$. In this case, moving toward locations with more amenities ($A(x) > 0$) is sufficient for the land rent to increase. In short, the interaction between amenities, commuting and income sorting may give rise to a variety of land rent profiles, which are not driven by the location-quality index alone. Therefore, what the land rent gradient looks like is essentially an empirical issue.

5 The equilibrium income mapping and land rent under Fréchet distributions

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 $\omega^*(x) = F^{-1}[G(\Delta)]$. 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 $\mathbb{R}_+$. 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 $\mathbb{R}_+$; the density is denoted $g(\Delta)$. The location-quality index covers a wider range of values when γ_Δ decreases.[11]

In the empirical analysis, we use the following utility function:

$$U = a^\beta q^{1-\mu}(h - \overline{h})^\mu \tag{26}$$

where we redefine the amenity to be equal to a^β, where $\beta > 0$ is the preference parameter for amenities. It follows from Proposition 2 that households ranked by decreasing incomes are assigned to locations having a decreasing location-quality index. Hence, the mapping $\omega^*(\Delta)$ can be retrieved from the condition:

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

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

$$\omega^*(x) = s_\omega \left[\frac{\Delta(x)}{s_\Delta}\right]^{\frac{\gamma_\Delta}{\gamma_\omega}}. \tag{27}$$

This expression shows that the spatial income distribution is a power of the location-quality index. Observe that what matters in the equilibrium mapping (27) is the ratio $\gamma \equiv \gamma_\Delta/\gamma_\omega$ and not the values of the two shape parameters. From now on, we work with γ.

Last, we show in Appendix B.8 that the equilibrium land rent at x is given by

$$R^*(x) = \mu(1-\mu)^{\frac{1-\mu}{\mu}} k^{-\frac{1}{\mu}} t(x) \left[\Delta(x)\right]^{\frac{1}{\mu}} \left[\frac{\mu t(x)}{R^*(x)} + \frac{(1-\mu)\overline{h}}{\omega^*(x)}\right]^{\frac{1}{(1-\mu)\mu\gamma}}, \tag{28}$$

where k is a positive constant determined in section 7.

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 X. In doing so, we run the risk of attributing different values of amenities and commuting costs to the same location $x \in X$. 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 (27) to quantify the sorting consequences of the spatial distribution of amenities.

Assume that labor time is given by $t(x) \equiv [\tau(x)]^{-\theta}$, where $\tau(x)$ is the commuting time, $\theta > 0$ is the elasticity of labor time with respect to commuting time. In this case, the location-quality index becomes:

$$\Delta(x) = [a(x)]^\beta [\tau(x)]^{-\theta(1-\mu)}. \tag{29}$$

[11] 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.

It follows from (27) that the income mapping is given by the following expression:

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

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

$$\tilde{\omega}(x) = \omega^*(x)\xi(x) \tag{30}$$

where the $\xi(x)$ are hourly labor income shocks that are independently and identically distributed according to a given distribution defined on $[0, \infty)$. Taking the log of (30), we obtain:

$$\log \tilde{\omega}(x) = \alpha_0 + \alpha_1 \log a(x) + \alpha_2 \log \tau(x) + \tilde{\xi}(x), \tag{31}$$

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

6 Data

6.1 Datasets

We have gained access to various nationwide non-public microdata from *Statistics Netherlands* between 2010 and 2015. In contrast to countries like 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 and gender. We only keep people that could be part of the working population, that is, who are between 18 and 65 years. We aggregate these data to the household level. Furthermore, we use information on household characteristic, such as household size, whether there are children in the household, as well as the marital status of the adults. Importantly, the *SSB* data enable us to determine where households exactly reside. More specifically, we know the location of a household up to the postcode level.[12] Hence, space is discrete in the plane (see Appendix B.8).

The data on income is 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. We focus on the gross yearly income of a household. The income data also provide information on whether households are homeowners or renters. In the Netherlands, about 90% of the properties in the rental sector is public housing. Public housing is rent controlled and there are often long waiting lists for public housing. So, households are not entirely free to choose their utility-maximizing location.

[12] A postcode is very small and contains on average $15-20$ addresses, implying that for cities, where most people live in apartments, the location is exactly known.

Therefore, we will focus on owner-occupied housing, which means that we keep about 70% of the data. 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 appears not to be an issue.

To estimate the commuting time for each household, we use again 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 elementary information on each establishment in the Netherlands, such as its exact location, the industrial sector, and the *estimated* number of employees in each establishment. To avoid miscoding and to exclude employment agencies (where people do not actually work), we exclude firms with more than 10 thousand employees. 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.[13] Overall, we are left with 95% of firms.

We first calculate the commuting time from each home location x to each job location e for each year. Then, we determine the commuting time $\tau(x)$ 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. More information on the road network and how we calculate the travel time between locations is provided in Appendix C.1.

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 *Dutch Association of Real Estate Agents* (*NVM*). The methodology used to calculate land values and lot sizes is described in Appendix C.2. The *NVM* data contains information on the large majority (about 75%) of owner-occupied house transactions between 2000 and 2015. We know the transaction price, the lot size, inside floor space size (both in 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.[14] We also know whether the house is a listed building.

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

[14] 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 $250m^2$ 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.

For ancillary analyses for which we do not need information on land (values) we exploit another dataset on housing transactions from the Land Registry (the *Kadaster*). Home buyers have to register transactions in the Land Registry, implying that we have information on the universe of owner-occupied housing transactions. We have information on the transaction price, inside floor space size (in m^2), the exact address, and a wide range of housing attributes such as house type, construction year, and whether a house is listed. The construction year controls for a range of house attributes are difficult to observe (e.g., building quality and architectural style).

We are interested in the impact of amenities and commuting time on income sorting and land prices. To this end, we gather data from Eric Fisher's *Geotagger's World Atlas*, which contain all geocoded pictures on the website *Flickr*. The idea is that locations with an abundant supply of aesthetic amenities will have a high picture density. We show in Appendix C.6 that there is a strong positive correlation 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.[15]

First, to avoid the possibility of inaccurate geocoding, we keep only one geocoded picture per location defined by its geographical coordinates.[16] 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 equal to 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 about 0.223. Fourth, one may argue that pictures are disproportionately clustered at public transport nodes because tourists usually arrive there. Therefore, we estimate specifications where we control for the distance to public transport nodes. Last, we recognize that people who take picture may belong to a specific socio-demographic group (e.g., young people with a smart phone) by including demographic controls and using instrumental variables.

Unlike Ahlfeldt (2014), we do not aggregate locations and use a continuous index of picture density because we know the exact locations of pictures. In line with (3), we measure the amenity

[15] Ahlfeldt (2014) 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.

[16] In 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.

level $a(x)$ as follows:

$$a(x) \equiv 1 + \delta \sum_{z=1}^{Z} e^{-\delta\tau(x,z)} n(z), \tag{32}$$

where z $(= 1, ..., Z)$ differs from x, $n(z)$ the number of pictures taken at z and $\delta > 0$ a decay parameter.[17] If pictures were uniformly distributed across space, the expected travel time to each picture location would be $E[\tau(x,z)] = 4\pi/\delta^2$. The burgeoning literature on the economic effects of amenities suggests that the effect of amenities is much more localized than the effects of employment (Rossi-Hansberg *et al.*, 2010; Koster *et al.*, 2016). We, therefore, assume that people take into account amenities within 15 minutes drive. Thus, if amenities were to be uniformly distributed, $E[\tau(x,z)] = 15$ yields $\delta = 0.915$. In the sensitivity analysis we show that our results are robust to this assumption.

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 an alternative hedonic amenity index in the spirit of Lee and Lin (2018) and an amenity index based on the augmented reality game *Pokémon Go* (see Appendix C.3 for more details). The hedonic index, based on housing transactions *outside* the Randstad, 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 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. We further use nationwide information on municipal population density in 1900 from *NLGIS* and the location of cinemas in 1910 from *SpinLab*. We gather additional data on the railway network in 1900 and the stations which by then existed (see Appendix C.4 for more information). To show robustness, similar instruments based on land use in 1832 obtained from *HISGIS* and *NLGIS* are constructed. *HISGIS* provides information on the exact space occupied by buildings and, importantly, the *cadastral income* per hectare. The cadastral income was used to determine the property tax and reflected the land value at that time. A disadvantage of the *HISGIS* is that it is only available for the inner cities of Amsterdam, Rotterdam and province of Utrecht, thereby reducing the number of observations by 75%. Additional information on the road network in 1821 is obtained from Levkovich *et al.* (2017).

[17] Note that we add a constant equal to one to the amenity index to avoid the issue that logged valued of the amenity index are highly negative. We have experimented with different constants leading to very similar results.

6.2 Descriptive statistics

We report descriptive statistics of the 4, 346, 889 observations of our sample in Table 1. ##The average (median) gross yearly income is €98, 514 (€86, 732)##. It appears that ##incomes are approximately Fréchet distributed## (see Appendix C.5). The average land price in the sample is €1, 714 , but there are stark spatial differences. For example, in Amsterdam, it is €3, 046, while in Rotterdam it is only €1, 715. Not surprisingly, in the rural Green Heart the land price is lower (€1, 507).[18] As expected, the correlation between the estimated land price and lot size is negative (the correlation ρ is -0.193).

[Table 1 about here]

The amenity index based on pictures range from approximately 1 to 843. The average amenity level in Amsterdam (23.3) is much higher than in Rotterdam (12.7), The Hague (12.4), and Utrecht (10.2). The underlying data on pictures is reported in Table C.3 (Appendix C.3). Recall that we only use pictures *outside* a building taken by *residents* in determining the amenity index. Table C.3 shows that 80% of the pictures are taken outside a building and about 60% of the pictures are taken by local residents. Going back to Table 1, we see that the average commuting time is 25 minutes. Given that the national average is about 30 minutes this seems a reasonable value (Department of Transport, Communications and Public Works, 2010). The unconditional correlation of the overall amenity index with the hourly income level is close to zero ($\rho = 0.0127$), but this is not very informative yet. The correlation of the amenity index with land prices is substantially higher ($\rho = 0.545$). Finally, households that have a shorter commute do not necessarily live in high amenity locations, as the correlation between the amenity level and commuting time is only -0.0454.

The average house size is 121m^2. However, in Amsterdam houses are only 90m^2, which corresponds to the higher land values in this city. About 24% of households occupy apartments and the correlation between occupying an apartment and the land price is indeed positive ($\rho = 0.145$).

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

7 Reduced-form estimation

In this section, we provide reduced-form evidence in support of the model's qualitative predictions made in Section 4. The analysis is complemented by a wide range of controls that provide evidence against alternative possible explanations.

[18] We report maps of the variables of interest, including land prices in Appendix A.1 and histograms of the variables of interest in Appendix C.5.

7.1 Identification strategy

We are interested in causal estimates of the coefficients α_1 and α_2, capturing the effects of respectively amenities and commuting time, on the spatial income distribution. Hence, a naive estimation would yield:

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

There are several problems associated with identifying α_1 and α_2.

First, with respect to commuting times, 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 an upward bias is that labor markets may not be fully competitive in the sense that workers may bargain over getting compensation when living further away. Mulalic *et al.* (2013) observe that about 15% of the costs of a longer commute is paid for by the employer.

Second, a more general concern with a causal interpretation of α_1 and α_2 as 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, 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 $a(x)$ and $\tau(x)$ 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). Furthermore, $\tau(x)$ may also be correlated to agglomeration economies: people living in dense places are likely more productive and therefore receive a higher wage, and have shorter commutes, implying that we would find an overestimate of α_2 (see e.g., Combes *et al.*, 2008).

Third, there may also be reverse causality between $\tilde{\omega}(x)$ and $a(x)$ and between $\tilde{\omega}(x)$ and $\tau(x)$. The provision of amenities, for example, may be a direct result of the presence of high-income households. Anecdotal evidence indeed 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 highly educated labor. Finally, because we do not observe the 'exact' amenity level, there may be a measurement error in $a(x)$, which may lead to a downward bias of α_1 when the error is random.[19]

The first step to mitigate the biases associated with these concerns is to add control variables. Most importantly, we control for *household* characteristics, $D(x)$. For example, members of the households are full-time or part-time workers, the size of the household and the age of the adults. This reduces the probability that we measure sorting on basis of household characteristics other than incomes. We also control for *housing* attributes, $C(x)$, such as housing type and construction year.[20]

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

The inclusion of the above variables is unlikely to address the issue that wages may be higher due to agglomeration economies. For a given population, agglomeration economies may be correlated to commuting times that are expected to be lower in denser areas. We therefore include work-location fixed effects $\eta_1(e)$. More specifically, we focus on the two jobs of the household that generate the highest number of working hours and use a work-location fixed effect for each job location pair. Hence, we compare households that work at the same location(s) but may have different commutes, which is in line with our model that takes job locations as given. If wages are higher because of spatial productivity differences, this is absorbed by the fixed effect. We further include a province dummies ($\eta_2(x)$), to control for differences in provincial taxes and policies, and year fixed effects ($\eta_2(y)$).

In sum, we estimate:

$$\log\tilde{\omega}(x) = \alpha_0 + \alpha_1 \log a(x) + \alpha_2 \log\tau(x) + \alpha_3 D(x) + \alpha_4 L(x) + \alpha_5 C(x) + \eta_1(e) + \eta_2(x) + \eta_3(y) + \tilde{\xi}(x), \tag{33}$$

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

It is unlikely that working with an endless string of controls will fully address the endogene-

[19] 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 US implies that the quality of schools and other neighborhood characteristics are often determined by the average income in the neighborhood (Bayer *et al.*, 2007). This is to be contrasted with what we observe in many other countries where local public goods such as schools are provided by centralized bodies.

[20] We do not directly control for house size, because house size is related to housing consumption and a direct result of the trade-off between commuting costs and amenities. Controlling for house size will not change the results (see Appendix C.8).

ity concerns raised above. Our data do not allow us to exploit quasi-experimental or temporal variation in $a(x)$ and $\tau(x)$. Therefore, we will rely on an instrumental variables approach. Our first set of specifications use *contemporary* instruments, while our second set of specifications will appeal to *historic* instruments. As contemporary instruments for amenities we use a set of observed, arguably exogenous, proxies for amenities, such as the weighted number of listed buildings (using the same weighting function as in (32)), whether x is in a historic district as well as the share of built-up areas and water bodies within 500m. By using other proxies for amenities, the measurement error of $a(x)$ is likely mitigated. For commuting times we need an instrument that is unrelated to household-specific differences in commuting such as unobserved ability. We then use the weighted number of jobs $J(x)$ that are within a commuting distance from location x. More specifically, we weight the number of jobs in e by calculating

$$J(x) = \sum_{e=1}^{E} F[\tau(x,e)]n(e), \tag{34}$$

where $\tau(x,e)$ is the commuting time between location x and employment location $e = 1, ..., E$ and $F[\tau(x,e)]$ the share of people who commute at most $\tau(x,e)$ minutes in the sample. Note that $J(x)$ is unrelated to the educational level or type of job, overcoming the issue that higher educated people search in thinner labor markets or that firms compensate individual workers for longer commutes.

One may argue that the contemporary instruments do not convincingly address the issue of unobserved locational and household characteristics that may be correlated with $a(x)$ and $\tau(x)$. Alternatively, we exploit the fact that $a(x)$ and $\tau(x)$ are autocorrelated. First, we use land use and amenities in 1900 as instruments. We expect aesthetic amenities to be positively correlated to the share of built-up area in 1900. The historic city center of Amsterdam, for example, has many buildings that are from (or before) 1900 and are now listed buildings. Furthermore, we also expect water bodies available in 1900 to be correlated to current water bodies, which are often considered as an amenity. We also count the number of cinemas in 1910 using the same distance decay function as in equation (32). As an instrument for commuting time we count the total number of people within 90 minutes travelling using the railway network in 1900. Since people at that time often work very close to their residence, this is a good proxy for the spatial employment distribution in 1900. Because of temporal autocorrelation we expect that a better population accessibility in 1900 implies a lower commute today.

Historic instruments are often criticized because of the strong identifying assumption that past unobserved locational features are correlated to current unobserved locational endowments. These instruments are hard to defend when analyzing income sorting patterns *between* urban areas. However, *within* the Randstad, these instruments are more likely to be valid because the patterns of income sorting within each city have considerably changed throughout the last century. Around 1900, open water and densely built-up areas were not necessarily considered amenities.

For example, the canals in Amsterdam were open sewers (Geels, 2006). Therefore, locations near a canal often repelled high-income households who located in lush areas just outside the city. It was also before the time when cars became the dominant mode of transport. People around 1900 usually walked to their working place, and thus commuting distances were very short. However, the rich could afford to live outside the city and take the train to their working place. The cities in 1900 were not yet influenced by (endogenous) planning regulations, as the first comprehensive city plans date from the 1930s. Hence, unobserved reasons that may cause the clustering of high-income people in the past are unlikely to be correlated to current amenities.

The main threat to the validity of the instrument is that built-up areas in 1900 are correlated to current unobservable attributes of the housing stock. To address this issue we estimate specifications where we only use the share of water in 1900 as an instrument and directly control for the share of built-up area in 1900.[21]

Historic instruments cease to be valid when unobserved characteristics of a location or building in the past are correlated with those in present time, which is less likely with instruments based on land use further back in time. On the other hand, by going back further in time, we may end up with weak instruments because the correlation between historic land use and current amenities and job locations will also be lower. Nevertheless, we exploit land use data from the census in 1832 because the data are 175 years before the sample period. We have exact information on the land use of each parcel in 1832, as well as information on the cadastral income – a proxy for land values – of each parcel at that time. If some past attractive features (e.g., housing attributes) are correlated to current sorting patterns, we expect this to be reflected in a positive coefficient of the cadastral income. Note that the 1832 data are only available for the province of Utrecht, as well as the inner cities of Amsterdam and Rotterdam, so that we have fewer observations. We again impute population per building using the municipal populations in 1832 and calculate the travel time of population within 90 minutes travel time using information on the road network from 1821. We further use the share of buildings, the share of built-up area excluding buildings and the share of water bodies within 500m as instruments.

In line with (23), we also repeat the above specifications where income is replaced by land prices or housing values. It follows immediately from (24) that a simple double log equation does not identify the structural parameters of the model. By contrast, we show in Section 8 that combining the income and land price equations allows us to separately identify β, γ and θ and, therefore, to predict $\omega(x)$ and $R(x)$.

[21] We also show that when controlling for current land use the results do not materially change (see Appendix C.7 for more details).

7.2 Empirical results

OLS and IV with contemporary instruments. We analyze the effects of amenities and commuting time on income sorting. Table 2 reports the results. Column (1) reports a naive regression of log income on log amenities, log commuting time and year fixed effects.[22] The results show that more amenities and a low commuting time are associated with *lower* incomes – contrary to the expectations. However, as mentioned above, these results may be severely biased because we do not control for household characteristics and housing attributes.

[Table 2 about here]

In column (2), we control for jobs' characteristics, such as whether the share of members of the household that have a permanent contract, work full time and have a company car, as well as households' characteristics, such as age and composition. The coefficient related to amenities is now positive and statistically significant: doubling the amenity level attracts households whose average incomes are 2.6% higher. The impact of commuting time still has the opposite sign. Column (3) investigates the importance of an uneven spatial distribution of housing quality. This issue does not seem to be major as the coefficients with respect to amenities and commuting time hardly change once we control for housing attributes.

We have argued above that the observed commuting time is probably correlated to (unobserved) job characteristics that are related to ability and the educational level – and therefore income – of members of the household. Therefore, in column (4), we instrument commuting time with the employment density as per (34). In other words, we expect that households living in denser areas (i.e., live closer to employment) have shorter commutes. Indeed, the first-stage results reported in Appendix C.6 show that the elasticity of commuting time with respect to employment density is between -0.35 and -0.55. Hence, doubling employment density implies a decrease in commuting time of at least 25%. When we go back to the second-stage results in column (4), Table 2, it appears that both amenities and commuting time now have the expected signs. Doubling the amenity level attracts households with incomes that are 1.9% higher, while doubling commuting times imply the attraction of household with incomes that are 20.0% lower.[23] In column (5) we also use instrument for picture density with observed proxies for amenities (e.g., nearby historic buildings or share water bodies). The first-stage results in Appendix C.6 show the expected signs: there is a higher picture density in built-up areas, in areas with more water bodies (e.g., the Am-

[22] We cluster our standard errors at the street (*PC5*) level to mitigate the issue that standard errors may be spatially correlated. We also experimented with clustering at higher levels (such as the neighbourhood), but this does not have strong implications for the results. We also clustered the standard errors at the household level to address correlation of the error term between different years, but this approach leads to somewhat smaller standard errors.

[23] Note that doubling the amenity level is a much more likely event (given a coefficient of variation of 3.262) than doubling the expected commuting costs (given a coefficient of variation of 1.637).

sterdam canal district) and where there are many historic buildings. The second-stage coefficient related to amenities in column (5), Table 2, is now somewhat higher. We note that the instruments are very strong in all specifications. When we include location attributes and province fixed effects in column (6), the results hardly change.[24]

The above specifications do not address the issue that incomes may be higher in dense areas with shorter commutes due to agglomeration economies. In column (7), Table 2, we therefore include work-location fixed effects, based on the two jobs in the household that generate the highest number of working hours. The results indicate that the coefficient of amenities is virtually the same, while the elasticity of commuting time is now about one-third of the effect reported in the previous column. This confirms that part of the higher incomes in denser areas is due to productivity effects. Column (8) improves on these results by only including households that have one job at a single-plant firm, which ensures that we measure commuting time correctly. In column (9) we only focus on households that have a company car, to address the potential issue that households may travel by foot, train or bicycle. In both specifications the results do not materially change.

IV with historic instruments. Amenities and job locations may be endogenous due to omitted correlated variables or reverse causality. The contemporary instruments may only partly address this issue. This is why we instrument amenities and commuting with historic variables.

[Table 3 about here]

The first set of specifications in Table 3 relies on instruments that are constructed from land use in 1900. The instruments are the shares of water bodies and of built-up area within a distance of 500m, the number of cinemas in the vicinity using the same weighting function as in (32), and the number of people within 90 minutes travelling in 1900.[25] In Appendix C.6, we report the corresponding first-stage results. The share of built-up area, the share of water bodies in 1900 and the cinema density in 1910 are strongly and positively correlated to the current amenity level. The number of people reachable within 90 minutes in 1900 is negatively related to current commuting times, which is in line with expectations, although the elasticity is much lower than the current employment density. 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.

The second-stage results reported in column (1), Table 3, reveal that when we instrument amenities and commuting times there is a positive effect of pictures on incomes and a negative

[24]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. So, we refrain from reporting them in the paper.

[25]We have experimented with other thresholds (e.g. 250 or 1000m; 60 or 120 minutes), but this leaves the results essentially unchanged. These results are available upon request.

impact of commuting times. However, we do not include work-location fixed effects, which may lead to an overestimate of the elasticity of incomes with respect to commuting times.[26] In Column (2), which is our preferred specification, we therefore include work-location fixed effects. Again, this leads to somewhat lower effect for commuting times. The results indicate that doubling the commuting time attracts households whose incomes are 16.7% lower. Doubling amenities leads to households whose incomes are 0.9% higher, so the impact is somewhat lower compared to the results using contemporary instruments. Column (3) improves on identification by directly controlling for built-up areas in 1900, as unobserved housing quality may be correlated to the share of built-up area in 1900. We also control for the *current* distribution of cinemas to control for the potential issue that the (past) pattern of cinemas is correlated to current unobservables (e.g., cinemas were often located in city centers). The results indicate that these issues are not very important, as the effect of amenities and commuting time are not significantly different from the estimates in the previous column, although the estimate of amenities is now only marginally significant.

Columns (4)-(6) repeat the same set of specifications but instead use instruments from the 1832 census. The census also provides information on the cadastral income, which was a proxy for the land value at that time. Column (4), which does not include work-location fixed effects seems to suggest that the effects of amenities and commuting are somewhat stronger than in our preferred specification. When the amenity level doubles, this will attract households whose incomes are 2.4% higher. A 100% increase in commuting times implies that households with a 16.4% lower income are attracted. The results are similar once we control for work-location fixed effects in column (5). When adding the share of buildings within 500m as a control in column (6), the effect of amenities becomes substantially stronger and, somewhat surprisingly, the share of land dedicated to buildings within 500m in 1832 seems to be associated with lower income households. In line with this, we note that it seems that more attractive locations in terms of cadastral income are now inhabited by poorer households. On the other hand, when the cadastral income was zero (e.g., in case of undeveloped land or when there were water bodies in 1832), current income seems to be lower. This is likely because these areas are still relatively remote. In any case, when we do not control for cadastral income in 1832, the results suggest that the elasticity of amenities is a bit lower and around 0.015, while the effect of commuting is essentially unaffected.

To sum up, the results unequivocally indicate that *the impact of amenities on income sorting is positive and 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. A standard deviation increase in log amenities attracts households whose incomes are about 1.8% higher (see column (2), Table 3). On the other hand, a standard deviation increase in log commuting time attracts households whose

[26] A large literature using historic instruments (e.g., Combes *et al.*, 2008) shows that agglomeration economies are also correlated to historic densities.

incomes are about 15% 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. We provide more evidence for this conclusion in Section 8.

A hedonic approach to amenities. 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 C.4 and the regression results are reported in Table 4. 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. In column (1), we show that this alternative index also has a strong impact on incomes. When we instrument commuting times with employment density, the impact of amenities becomes somewhat stronger. Columns (3) and (4) exclude and include work-location fixed effects respectively, while using contemporary instruments to instrument amenities and commuting times. Columns (5) and (6) rely on historic instruments with and without work-location fixed effects respectively. It appears that all elasticities with respect to amenities are between 0.01 and 0.04, which are in the same ballpark as the estimates using the picture index. The effect of commuting time is comparable to previous specifications. For other robustness checks with respect to the definition of the amenity index, we refer to Appendix C.7.

[Table 4 about here]

Effects on land prices. Land prices play a crucial role in households' locational decisions. Furthermore, we postulated that the signs of the effects of amenities and expected commuting time on land prices and incomes are the same. Therefore, we now estimate the effects of amenities and expected commuting time on land prices. We start in column (1), Table 5, with a simple specification including amenities and commuting time. This leads to a strong positive effect of amenities on land prices: doubling amenities implies a land price increase of 8.0%, while doubling the commuting time to the city center implies a land price decrease of 2.0%. When we include housing attributes and instrument for commuting time in column (2), the impact of commuting becomes much stronger, while the impact of amenities is somewhat lower.

[Table 5 about here]

In column (3), we instrument both for amenities and commuting time with contemporary instruments and include locational control variables. This is followed by the inclusion of work-location fixed effects in column (4). The latter results indicate that a 100% increase in the amenity level implies an increase in land prices of 11.4%. Doubling the commuting time decreases land prices by 20.1%. The results with historic instruments in column (5) and (6) are surprisingly similar to the ones using contemporary instruments. The preferred specification in column (6), including work-location fixed effects, imply that doubling the amenity level leads to a 9.0% increase in land prices, while doubling the commuting time is associated with an increase in land prices of 20.6%.

Largely, these results are in line with our previous findings that more attractive locations will attract higher income households. In particular, as shown by (28), *a higher amenity level and lower commuting time will lead to higher land prices.* To the extent one may be worried that these results are driven by the specific method used to determine land prices, we repeat the same set of regressions where land prices are replaced by the house prices from the Land Registry dataset. The results are very much similar to the ones presented in Table 5 (see Appendix C.7 for more details).

Sensitivity checks. Appendix C.7 shows that our results still hold for a wide range of robustness checks. We consider some alternative instruments for commuting time and alternative proxies for amenities, such as the amenity index based on listed buildings, and on locations provided by the augmented reality game *Pokémon Go*. We also investigate whether the choice of transport mode (car vs. train) matters for our results and whether controlling for disamenities (e.g., air and noise pollution) impacts the estimated coefficients.

We improve on identification by the inclusion of current land use as additional controls and the inclusion of more detailed fixed effects to make sure that our results are robust. We also consider other decay parameters with respect to the amenity index and the historic instruments and consider to use education level as dependent variable rather than income. Furthermore, we estimate separate regressions for each of the four largest cities in the Randstad. These regressions show that the exact definition of the city (i.e., a polycentric urban region as the Randstad, or separating the analysis into distinct cities) does not change the results. Overall, *the impact of amenities and commuting time on households' location choice is robust.*

8 Counterfactual analysis

8.1 Structural estimation

What happens to the spatial income distribution and land rents when the distribution of the location-quality index within the Randstad changes? When we aim to predict the spatial income distribution and land rents for alternative scenarios, (27) and (28) imply that we need to identify all the parameters $\{\beta, \theta, \gamma_\Delta, s_\Delta, \gamma_\omega, s_\omega, \mu, k\}$. Unfortunately, we cannot separately identify $\{\beta, \theta\}$ and $\{\gamma_\Delta, \gamma_\omega\}$ by solely using the income equation. Therefore, we also use information on land prices and lot sizes to estimate the necessary parameters, which we then use to determine the value of $\Delta(x)$, $\omega(x)$ and $R(x)$ at each location for each counterfactual scenario. More specifically, using $t(x) = [\tau(x)]^{-\theta}$ and $\tilde{h}(x) \equiv h^*(x) - (1-\mu)\overline{h}$, we rewrite (24) as follows:

$$r(x) \equiv \frac{R^*(x)}{\mu\omega^*(x)} = \frac{\tau(x)^{-\theta}}{\tilde{h}(x)}. \tag{35}$$

Let $\tilde{r}(x) \equiv r(x)\upsilon(x)$ where $\upsilon(x)$ are shocks that are independently and identically distributed according to a distribution defined on $[0, \infty)$. Taking the log of (35), we obtain:

$$\log \tilde{r}(x) = -\theta \log \tau(x) - \log \tilde{h}(x) + \tilde{\upsilon}(x),$$

where $\tilde{\upsilon}(x) \equiv \log \upsilon(x)$. In keeping with (33), we will use the following more general expression:

$$\log \tilde{r}(x) = \zeta_1 \log \tau(x) + \zeta_2 \log \tilde{h}(x) + \zeta_3 D(x) + \zeta_4 L(x) + \zeta_5 C(x) + \chi_1(e) + \chi_2(x) + \chi_3(y) + \tilde{\upsilon}(x), \tag{36}$$

where we instrument $\tau(x)$ while $\chi_1(e)$, $\chi_2(x)$ and $\chi_3(y)$ are work-location, province and year fixed effects.

Our estimation procedure encompasses four steps:

1. We estimate (36) to identify $\hat{\theta}$.
2. We estimate (33) and use the estimate $\hat{\theta}$ to identify $\{\hat{\beta}, \hat{\gamma}\}$.
3. We use $\{\hat{\beta}, \hat{\theta}, \hat{\gamma}\}$ and observations on amenities and commuting time to obtain $\hat{\Delta}(x)$ for all x. By fitting a Fréchet distribution on $\hat{\Delta}(x)$ and using $\hat{\gamma}$, we obtain $\{\hat{\gamma}_\Delta, \hat{\gamma}_\omega, \hat{s}_\Delta\}$.
4. We use the estimated parameters to obtain $\hat{\omega}(x)$ and the Newton-Raphson numerical iteration method to determine $\hat{R}(x)$ using (28) where k, which is defined in Appendix B.8, varies with the different counterfactual scenarios.

It is hard to estimate μ. Dutch households spend about one-third of their income on housing. According to Albouy *et al.* (2016), the income elasticity of housing demand is near two-thirds (from their preferred estimates). Therefore, we can determine the value of μ by using the income elasticity of (16):

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

Furthermore, our identification strategy is only able to identify $\hat{\omega}(x)$ and $\hat{R}(x)$ up to a constant because s_ω is not identified. We have chosen to proxy s_ω by fitting a Fréchet distribution to the observed distribution of hourly incomes and k in such a way that the minimum value of predicted land rents is equal to R_A, which we set to the average value of the land price of vacant land in municipalities bordering the Randstad.[27] Moreover, we set $\overline{h} = 25\text{m}^2$, which is the minimum value observed in the data, and use a discount rate of 2.5% to go from land prices to land rents. Standard errors are obtained by bootstrapping the whole estimation procedure.

[27] We obtain the information on vacant land from *Funda.nl*.

We report in Table 6 the results of the structural estimation. In column (1), using contemporary instruments, we find that $\hat{\beta}$ is 0.0540 while $\hat{\gamma}$ is about 0.5. Thus, the estimated preference parameter is somewhat higher than the reduced-form elasticities obtained in Tables 2 and 3. Likewise, $\hat{\theta}$ is equal to 0.190, which is ##also## somewhat higher than the reduced form estimate. The results are pretty robust to assumptions made on the fixed parameters. First, when setting μ to 0 in column (2), which implies a fixed lot size, the effect of commuting remains the same, but the impact of amenities become somewhat stronger. In column (3) we reduce the minimum lot size to 10m^2. This has hardly any impact on our results. In column (4), we use historic instruments and find that $\hat{\beta}$ is equal to 0.023. On the other hand, $\hat{\theta} = 0.611$, which is substantially higher compared to the specifications that use contemporary instruments, but in line with the reduced form estimates, which also show that the impact of commuting is stronger when we use historic instruments. Again, $\hat{\gamma}$ is about 0.5.

[Table 6 about here]

8.2 Counterfactual scenarios

We undertake three counterfactuals in which the income distribution is kept constant. The objective is to assess the impacts of changes in amenities and commuting time on spatial sorting of households for a given income distribution. ##In the first scenario, we aim to mimic a typical U.S. monocentric city by putting all employment in each of the centers of the four largest cities in the Randstad and assuming that households commute to the nearest center. This implies that average commuting time is essentially the same (from 24.6 to 25 minutes), although the spatial distribution of employment considerably changes. In addition, we set historic amenities to zero, implying that the amenity level decreases on average by 30%. In the second scenario, we assume that cars are prohibited within the Randstad and that people have to take the train, a bike or both. This implies strong increases in commuting times of 300%, while amenities only decrease by 1.8% on average. In the third scenario, we consider the effects of a hypothetical increase in telecommuting. More specifically, we assume that highly-skilled workers can telecommute and do so 3 days a week. This drastically reduces the commuting time of these workers by 60%.##

We use the above-estimated parameters to construct $\hat{\omega}(x)$, $\hat{R}(x)$ and $h(x)$ for each of the three scenarios. Furthermore, we construct a measure of income mixing, i.e., the standard deviation of income $\sigma_\omega(x)$ in every postcode area to see how the counterfactual scenarios affect income mixing within the Randstad. ##In Table 7, we report the estimates.## The average income hardly changes across different scenarios because the aggregate income distribution is kept constant. However, our three scenarios have different implications for the amount of income mixing.

##The first scenario implies substantially less income mixing as $\bar{\sigma}_\omega$ is much lower than the baseline estimate, especially when using historic instruments. Income mixing is similar to the baseline scenario in the scenarios where people are not using cars or when telecommuting is allowed. The results reported in Table 7 show that the different scenarios have strong implications for the aggregate land rent. Rents decrease by about 3.3% (contemporary instruments) to 8.6% (historic instruments) when we move from a mixed city with amenities and employment spread across the urban area to monocentric cities without historic amenities. In line with anecdotal evidence, households will then consume more land: the average lot size increases by about 25%. The aggregate land rent is even $23-37\%$ lower when we consider the scenario in which people can only use mass transit or the bicycle.## As shown by (28), the average rise in commuting time by train reduces the land rent because households earn lower real incomes. In the final scenario, where telecommuting is implemented, rents are about $9-21\%$ higher than in the baseline scenario. Indeed, e-work makes the high-skilled workers wealthier, which intensifies competition for housing in the rich segment of the population. However, on average, land consumption is hardly affected. The final row of Table 7 includes the average location quality index $\bar{\Delta}$. The results for different scenarios are in line with the results obtained for the aggregate land rent.##

[Table 7 about here]

In Appendix C.8 we also plot the predicted income mapping $\hat{\omega}(x)$ and land rents $\hat{R}(x)$ over space. Predicted incomes are generally the highest in city centers, where the amenity level is high and commuting costs are low, but income patterns are non-monotonic and non-symetric iu distance to the (nearest) center. Scenario 1 – where we consider monocentric cities without historic amenities – considerably simplifies the spatial income pattern. Richer households are now residing in or near city centers while income gradients are more or less monotonically decreasing in distance to the city center, which is in line with the predictions by the monocentric city model. In scenario 2, we restrict households to travel by train or bicycle instead of travelling by car. Given the current railway network, this will have strong repercussions for the income distributions. High income households will now sort close to railway stations, where accessibility to jobs and amenities is higher. Hence, *changes in travel modes* (e.g., a switch to self-driving cars) *may have significant impacts on the social structure of cities.* This effect should not be overlooked when assessing the costs and benefits of policies that favor a particular travel mode. In counterfactual scenario 3, we assume that highly educated people will commute two days a week. This seems to imply that amenities are now more important for the rich, rather than accessibility. Hence, high amenity locations in and near Amsterdam and the city center of Rotterdam seem to disproportionally attract the rich.

9 Concluding remarks

In this paper, we used a new setup in which any city location is differentiated by two attributes, i.e., the benefit generated by the amenity field at this location and its distance to the nearest employment center. The bid rent function of urban economics may be used to show that the uneven provision of exogenous amenities is sufficient to break down the perfect sorting of households across the city. In other words, the equilibrium outcome now involves residential patterns in which households sharing the same income may live in spatially separated neighborhoods. As homothetic preferences generate a continuum of equilibria, we cannot assume a Cobb-Douglas or CES utility, i.e., the most preferred specifications used in the literature. Rather, we assume a Stone-Geary utility and go one step further by showing that there exists a location-quality index, which 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 consumers who choose to locate there. The relevance of exogenous amenities and commuting costs to explain the residential choices of consumers heterogeneous in income is confirmed by the empirical analysis of the Randstad, one of the main polycentric urban areas in Europe, where both effects are found to be significant. Moreover, given the polycentric nature of many cities in terms of amenities and employment accessibility, our results suggest that the classical monocentric model without amenities is a fairly poor predictor of the social structure of cities.

Two extensions of our model are worth mentioning. First, although considering a given income distribution is reasonable as a first step in the study of the effect of income inequality on the city's social structure, our model is tractable enough to start with individuals heterogeneous in skills s whose distribution is given. The gross income of an individual working in the employment center e_i is then given by $\omega_i = sp_i$, where $p_i \in [0, 1]$ is a premium that reflects the productivity of the center e_i. In this case, an individual earns different incomes in different employment centers. As a result, households may trade longer commutes against higher incomes earned in more productive employment centers. Our analysis holds true if we replace $t_i(x)$ by $p_i t_i(x)$.The next step is then to endogenize the premia p_i. For this, we must endow each center with a production function, while the labor market pins down the equilibrium premia p_i, in the spirit of Allen *et al.* (2015) and Lucas and Rossi-Hansberg (2002).

Second, we can account for the endogenous choice of local public goods (LPGs). Consumers at x choose their consumption level $b(x)$ 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 $b(x) = \gamma(x)R(x)$. Assuming that amenities and LPGs are bundled into a Cobb-Douglas aggregate, preferences become $U = a^\alpha b^{1-\alpha} \cdot q$, with $0 < \alpha < 1$. Solving the utility-maximizing condition for the equilibrium tax rate for a ω-consumer at x yields $b^*(x) = (1-\alpha)\left[\omega t(x) - (1+\gamma)R(x)\right]$. Using $b^*(x)$ and applying the same reasoning as in Section 5, it is readily verified that

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

In this case, the location-quality index becomes $\Delta(x) = [a(x)]^{\alpha/(2-\alpha)}\, t(x)$. Comparing this condition to (17) where $\mu = 0$ shows that, *the decentralized provision of LPGs weakens the impact* *of exogenous amenities in individual residential choices because they are substitutes.* In particular, when consumers do not value much the exogenous amenities (α is small) or when the level of exogenous amenities is almost constant across space, residential choices are mainly driven by commuting costs. More work is called for to deal with the case of endogenous housing consumption and neighborhood externalities (Calabrese *et al.*, 2012).

References

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

[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., C. Arkolakis and X. Li (2015) Optimal city structure. Yale University, memo.

[6] 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.

[7] Becker, S.O., T. Fetzer and D. Novy (2017) Who voted for Brexit? A comprehensive district-level analysis. *Economic Policy* 32: 601 – 51.

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

[9] Bénabou, R. (1996) Equity and efficiency in human capital investment: the local connection. *Review of Economic Studies* 63: 237 – 64.

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

[11] Börjesson, M., M. Fosgerau and S. Algers (2012) On the income elasticity of the value of travel time. *Transportation Research Part A. Policy and Practice* 46: 368 – 77.

[12] 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.

[13] Calabrese, S. D. Epple and R. Romano (2012) Inefficiencies from metropolitan political and fiscal decentralization: Failures of Tiebout competition. *Review of Economic Studies* 79: 1081 – 111.

[14] Carlino, G.A. and A. Saiz (2008) Beautiful city: Leisure amenities and urban growth. Federal Reserve Bank of Philadelphia. Working Paper No.0822.

[15] Chetty, R. and N. Hendren (2018) The effects of neighborhoods on intergenerational mobility I: Childhood exposure effects. *Quarterly Journal of Economics* 133: 1107 – 62.

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

[17] Chiappori, P.-A. (2017) *Matching with Transfers. The Economics of Love and Marriage.* Princeton: Princeton University Press.

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

[19] Davis, D.R. and J. Dingel (2018) A spatial knowledge economy. *American Economic Review*, forthcoming.

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

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

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

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

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

[25] Epple, D. and H. Sieg (1999) Estimating equilibrium models of local jurisdictions. *Journal of Political Economy* 107: 645 – 81.

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

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

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

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

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

[31] Glaeser, E.L., M.G. Ressger and K. Tobio (2008) Urban inequality. NBER Working Paper 14419.

[32] Gutiérrez-i-Puigarnau, E. and J.N. van Ommeren (2016) Do rich households live farther away from their workplaces. *Journal of Economic Geography* 16: 177 – 201.

[33] Inglehart, R.F. and P. Norris (2016) Trump, Brexit, and the rise of Populism: Economic have-nots and cultural backlash. Harvard Kenneny School WP 16-026.

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

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

[36] 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.

[37] 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.

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

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

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

[41] 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.

[42] Levkovich, O., J. Rouwendal, and J.N. van Ommeren (2017). The impact of highways on population redistribution: The role of land development restrictions. *Tinbergen Institute Discussion Paper 17-109/VIII.*

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

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

[45] Manning, A. (2003) The real thin theory: Monopsony in modern labour markets. *Labour Economics* 10: 105 – 131.

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

[47] Moretti, E. (2012) *The New Geography of Jobs*. Boston: Houghton Mifflin Harcourt.

[48] 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.

[49] Musterd, S., S. Marcińczak, M. van Ham and T. Tammaru (2017) Socio-Economic Segregation in European capital cities: Increasing separation between poor and rich. *Urban Geography* 38: 1062 – 83.

[50] Ong, C-b (2017) Tipping points in Dutch big city neighbourhoods. *Urban Studies* 54: 1016 – 37.

[51] O'Sullivan, A (2005) Gentrification and crime. *Journal of Urban Economics* 57: 73–85.

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

[53] 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.

[54] Roback, J. (1982) Wages, rents, and the quality of life. *Journal of Political Economy* 90: 1257 – 78.

[55] 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.

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

[57] Small, K.A. (2012) Valuation of travel time. *Economics of Transportation* 1: 2 – 14.

[58] Susilo, Y.O. and K. Maat (2007) The influence of built environment to the trends in commuting journeys in the Netherlands. *Transportation* 34: 589 – 609.

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

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

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

Tables

Table 1 – Descriptive statistics

	(1) mean	(2) sd	(3) min	(4) max
Gross income *(in €)*	98,514	59,806	7,337	912,093
Land price *(in € per m^2)*	1,714	828.0	10.32	14,962
Lot size *(in m^2)*	255.7	597.2	27	20,768
Amenities, $\delta = 0.915$	13.93	45.40	1.000	843.1
Hedonic amenity index	4,187	110.5	4,105	4,833
Commuting time *(in minutes)*	24.55	15.00	0.403	118.7
Employment accessibility	490,700	151,869	87,242	948,485
Total hours worked in household	2,178	904.5	420.5	6,207
Share full-time contracts	0.612	0.444	0	1
Number of different jobs in household	1.478	0.919	1	14
Household has company car	0.186	0.389	0	1
Share male adults	0.517	0.231	0	1
Average age of adults	41.77	8.945	21	64
Household size	2.959	1.291	1	11
Household type – single	0.142	0.349	0	1
Household type – couple	0.244	0.429	0	1
Household type – single parent with kids	0.0473	0.212	0	1
Household type – couple with kids	0.567	0.496	0	1
Share foreigners in household	0.104	0.260	0	1
Property size *(in m^2)*	121.0	44.00	26	350
Property type – apartment	0.244	0.430	0	1
Property type – terraced	0.444	0.497	0	1
Property type – semi-detached	0.238	0.426	0	1
Property type – detached	0.0737	0.261	0	1
Share open space <500m	0.716	0.199	0.00749	1
Share of open space <500m	0.229	0.189	0	0.989
Share of water bodies <500m	0.0555	0.0794	0	0.833
Listed buildings, $\delta = 0.915$	4.395	19.14	1	370.5
In historic district	0.0650	0.247	0	1
Share of historic districts <500m	0.0674	0.195	0	1

Notes: The number of observations is 4,346,889. For land price and lot size the number of observations is 964,314. Because of confidentiality restrictions the minimum and maximum values refer to the 0.0001$^{\text{th}}$ and 99.9999$^{\text{th}}$ percentile.

Table 2 – Regression results: OLS and IV with contemporary instruments

(Dependent variable: the log of hourly income)

	No instruments			*Contemporary instruments*					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	OLS	OLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Amenities, $\delta = 0.915$ *(log)*	-0.0036***	0.0371***	0.0425***	0.0281***	**0.0471*****	**0.0360*****	**0.0320*****	**0.0302*****	**0.0229*****
	(0.0010)	(0.0008)	(0.0007)	(0.0008)	(**0.0013**)	(**0.0017**)	(**0.0014**)	(**0.0016**)	(**0.0020**)
Commuting time in minutes (*log*)	0.0626***	0.0575***	0.0586***	**-0.2868*****	**-0.2485*****	**-0.2364*****	**-0.0863*****	**-0.0938*****	**-0.0996*****
	(0.0007)	(0.0006)	(0.0006)	(**0.0067**)	(**0.0073**)	(**0.0077**)	(**0.0043**)	(**0.0052**)	(**0.0059**)
Household characteristics	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Housing attributes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location attributes	No	No	No	No	No	Yes	Yes	Yes	Yes
Province fixed effects	No	No	No	No	No	Yes	Yes	Yes	Yes
Work location fixed effects	No	No	No	No	No	No	Yes	Yes	Yes
Number of observations	4,346,889	4,346,889	4,346,889	4,346,889	4,346,889	4,346,889	4,162,962	1,558,535	724,691
R^2	0.0585	0.2761	0.3027						
Kleibergen-Paap F-statistic				15,444	1,937	1,626	2,705	1,743	1,621

Notes: **Bold** indicates instrumented. We refer to Table C.7 in Appendix C for the corresponding first-stage results. Standard errors are clustered at the *PC5* level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 3 – Regression results: IV with historic instruments

(Dependent variable: the log of hourly income)

	Instruments from 1900			*Instruments from 1832*		
	(1) 2SLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
Amenities, $\delta = 0.915$ *(log)*	**0.0153*****	**0.0129*****	**0.0113***	**0.0351*****	**0.0210*****	**0.0732*****
	(**0.0030**)	(**0.0024**)	(**0.0060**)	(**0.0031**)	(**0.0032**)	(**0.0093**)
Commuting time in minutes (*log*)	**-0.4124*****	**-0.2415*****	**-0.2445*****	**-0.2376*****	**-0.2527*****	**-0.2220*****
	(**0.0396**)	(**0.0283**)	(**0.0308**)	(**0.0576**)	(**0.0410**)	(**0.0433**)
Share built-up land in 1900 <500m			0.0142			
			(0.0163)			
Cinemas in 2010, $\delta = 0.915$ (*log*)			-0.0619***			
			(0.0187)			
Cadastral income in 1832 per ha (*log*)				-0.0049***	-0.0053***	-0.0097***
				(0.0008)	(0.0007)	(0.0010)
Cadastral income in 1832 is zero				-0.1583***	-0.1569***	-0.2867***
				(0.0203)	(0.0183)	(0.0278)
Share of buildings in 1832 <500m						-0.3830***
						(0.0685)
Household characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Housing attributes	Yes	Yes	Yes	Yes	Yes	Yes
Location attributes	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Work location fixed effects	No	Yes	Yes	No	Yes	Yes
Number of observations	4,346,889	4,162,962	4,162,962	1,340,991	1,265,990	1,265,990
Kleibergen-Paap F-statistic	128	116.9	152.3	48.11	78.16	99.33

Notes: **Bold** indicates instrumented. We refer to Table C.8 in Appendix C for the corresponding first-stage results. Standard errors are clustered at the *PC5* level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 4 – Regression results: a hedonic amenity index

(Dependent variable: the log of hourly income)

		Contemporary instruments			Historic instruments	
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	2SLS	2SLS	2SLS	2SLS	2SLS
Hedonic amenity index (*log, std*)	0.0336***	0.0391***	**0.0313*****	**0.0283*****	**0.0113*****	**0.0101*****
	(0.0010)	(0.0008)	(**0.0009**)	(**0.0008**)	(**0.0025**)	(**0.0021**)
Commuting time in minutes (*log*)	0.0696***	**-0.3147*****	**-0.2755*****	**-0.1095*****	**-0.4783*****	**-0.2899*****
	(0.0007)	(**0.0069**)	(**0.0067**)	(**0.0039**)	(**0.0318**)	(**0.0222**)
Household characteristics	No	Yes	Yes	Yes	Yes	Yes
Housing attributes	No	Yes	Yes	Yes	Yes	Yes
Location attributes	No	No	Yes	Yes	Yes	Yes
Province fixed effects	No	No	Yes	Yes	Yes	Yes
Work location fixed effects	No	No	No	Yes	No	Yes
Number of observations	4,346,889	4,346,889	4,346,889	4,162,769	4,346,889	4,162,769
R^2	0.0542					
Kleibergen-Paap F-statistic		16,779	3303	6,828	156.2	141.9

Notes: **Bold** indicates instrumented. We refer to Appendix C.4 for the construction of the hedonic amenity index. Standard errors are clustered at the *PC5* level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 5 – Regression results: effects on land prices

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

		Contemporary instruments			Historic instruments	
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	2SLS	2SLS	2SLS	2SLS	2SLS
Amenities, $\delta = 0.915$ *(log)*	0.1149***	0.0693***	**0.1532*****	**0.1640*****	**0.1230*****	**0.1293*****
	(0.0018)	(0.0020)	(**0.0039**)	(**0.0033**)	(**0.0052**)	(**0.0059**)
Commuting time in minutes (*log*)	-0.0293***	**-0.8057*****	**-0.5742*****	**-0.2906*****	**-0.5258*****	**-0.2971*****
	(0.0009)	(**0.0150**)	(**0.0161**)	(**0.0088**)	(**0.0662**)	(**0.0634**)
Lot size *(log)*	-0.3727***	-0.3579***	-0.3281***	-0.3202***	-0.3367***	-0.3305***
	(0.0022)	(0.0027)	(0.0028)	(0.0026)	(0.0032)	(0.0033)
Household characteristics	No	Yes	Yes	Yes	Yes	Yes
Housing attributes	No	Yes	Yes	Yes	Yes	Yes
Location attributes	No	No	Yes	Yes	Yes	Yes
Province fixed effects	No	No	Yes	Yes	Yes	Yes
Work location fixed effects	No	No	No	Yes	No	Yes
Number of observations	964,314	964,314	964,314	886,857	964,314	886,857
R^2	0.4301					
Kleibergen-Paap F-statistic		8,834	1,122	1,796	96.47	75.85

Notes: **Bold** indicates instrumented. We refer to Appendix C.3 for the calculation of land prices. Standard errors are clustered at the *PC5* level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 6 – Structural estimation

	Contemporary instruments			*Historic instruments*		
	(1)	(2)	(3)	(4)	(5)	(6)
β	0.0540***	0.0696***	0.0537***	0.0231***	0.0297***	0.0232***
	(0.00711)	(0.00924)	(0.00738)	(0.00806)	(0.01033)	(0.00805)
θ	0.191***	0.192***	0.190***	0.611***	0.611***	0.612***
	(0.0104)	(0.0084)	(0.0155)	(0.0411)	(0.0353)	(0.0354)
$\gamma = \gamma^{\Delta}/\gamma^{\omega}$	0.598***	0.463***	0.602***	0.532***	0.414***	0.531***
	(0.0571)	(0.048)	(0.064)	(0.087)	(0.051)	(0.065)
γ^{Δ}	9.729***	7.537***	9.783***	3.651***	2.84***	3.644***
	(0.562)	(0.499)	(0.607)	(0.35)	(0.165)	(0.212)
s^{Δ}	257.041***	1252.067***	257.651***	84.085***	298.365***	83.85***
	(16.9279)	(84.7704)	(16.976)	(9.3262)	(35.1495)	(7.6714)
γ^{ω}	16.264***	16.26***	16.264***	6.866***	6.866***	6.866***
	(0.931)	(0.769)	(1.038)	(0.807)	(0.804)	(0.804)
s^{ω}	32.21***	32.21***	32.21***	32.21***	32.21***	32.21***
	(0.0466)	(0.0466)	(0.0466)	(0.0466)	(0.0467)	(0.0467)
Fixed parameters:						
μ	0.2222	0.0000	0.2222	0.2222	0.0000	0.2222
$\bar{h}$	25	25	10	25	25	10
Household characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Housing attributes	Yes	Yes	Yes	Yes	Yes	Yes
Location attributes	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Work location fixed effects	No	Yes	No	Yes	Yes	Yes
Number of observations	4,346,889	4,346,889	4,346,889	4,346,889	4,346,889	4,346,889

Notes: Bootstrapped standard errors (250 replications) are clustered at the *PC5* level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 7 – Counterfactual scenarios

	Contemporary instruments				*Historic instruments*			
	Baseline scenario	*Scenario 1: Monocentric city*	*Scenario 2: No cars*	*Scenario 3: Telecommuting*	*Baseline scenario*	*Scenario 1: Monocentric city*	*Scenario 2: No cars*	*Scenario 3: Telecommuting*
$\bar{\omega}(x)$ *(in €)*	68,809	69,126	68,929	68,821	72,643	73,847	72,976	72,636
$\bar{\sigma}_{\omega}(x)$ *(in €)*	3,885	1,356	2,804	3,795	11,650	2,357	9,286	11,274
$\sum_x R(x)h(x)$ *(in million €))*	34,423	33,276	26,382	37,542	20,630	18,846	12,930	25,017
$\bar{h}(x)$ *(in m^2)*	400	499	313	398	117	147	75	115
$\bar{\Delta}(x)$	0.6317	0.6058	0.5078	0.6751	0.3265	0.3086	0.2065	0.3758

Figures

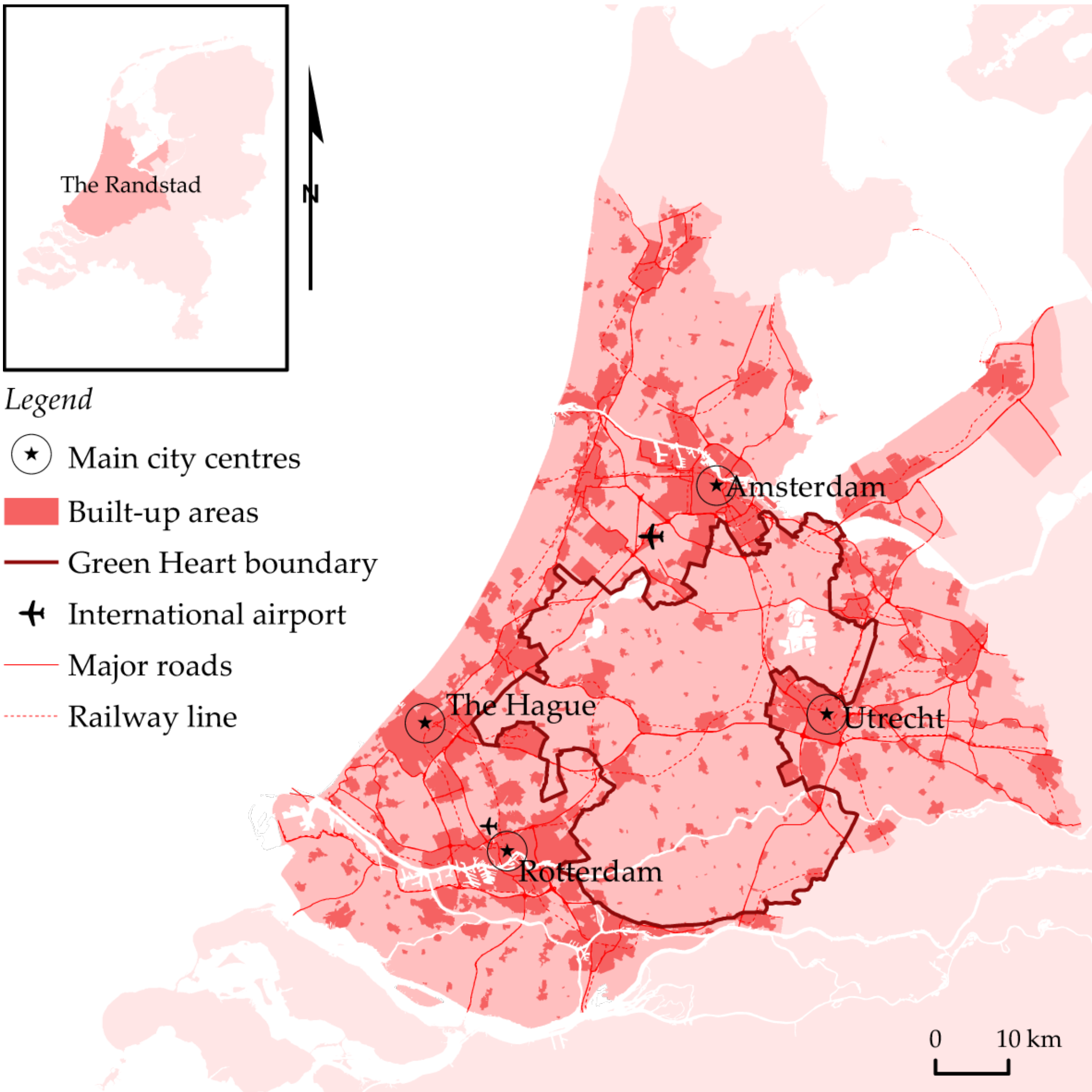
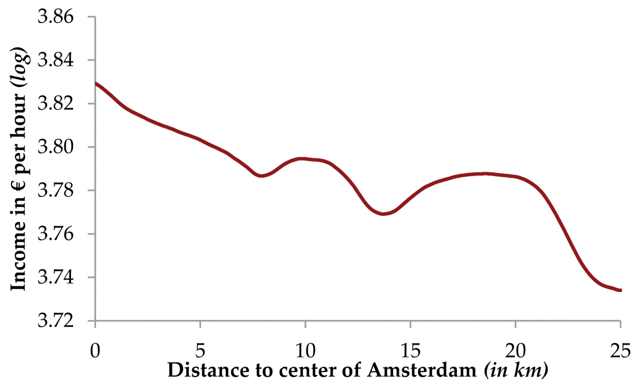
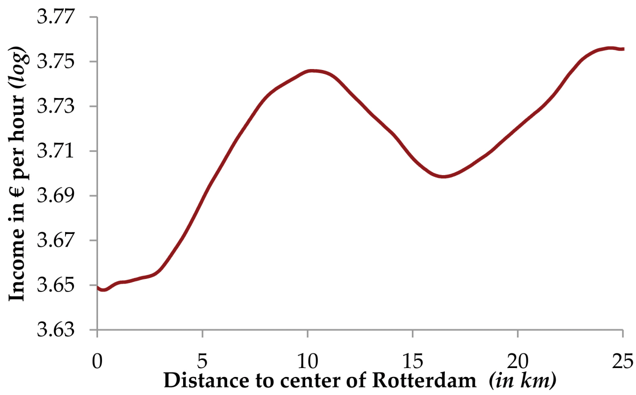


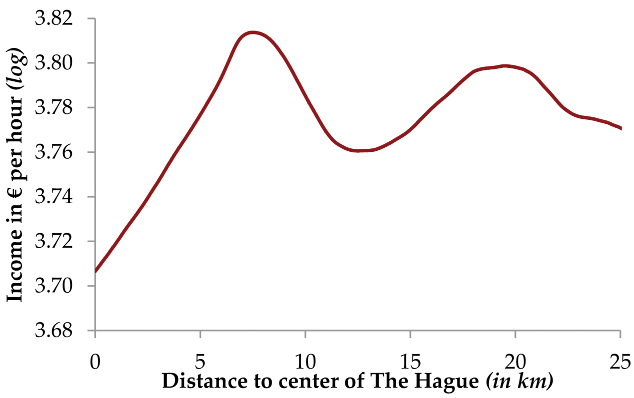
Figure 1 – The Randstad



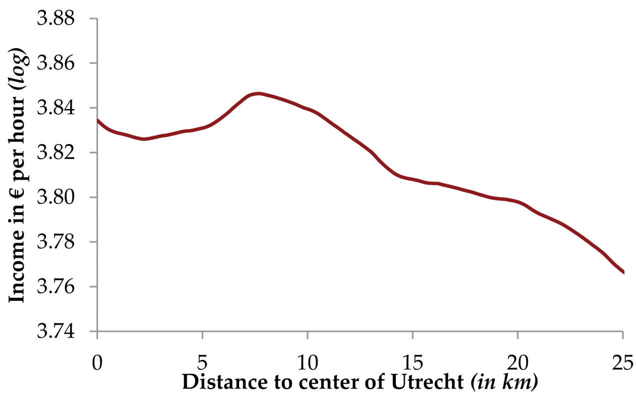
(A) Amsterdam



(B) Rotterdam



(C) The Hague



(D) Utrecht

Figure 2 – Income gradients for the Randstad

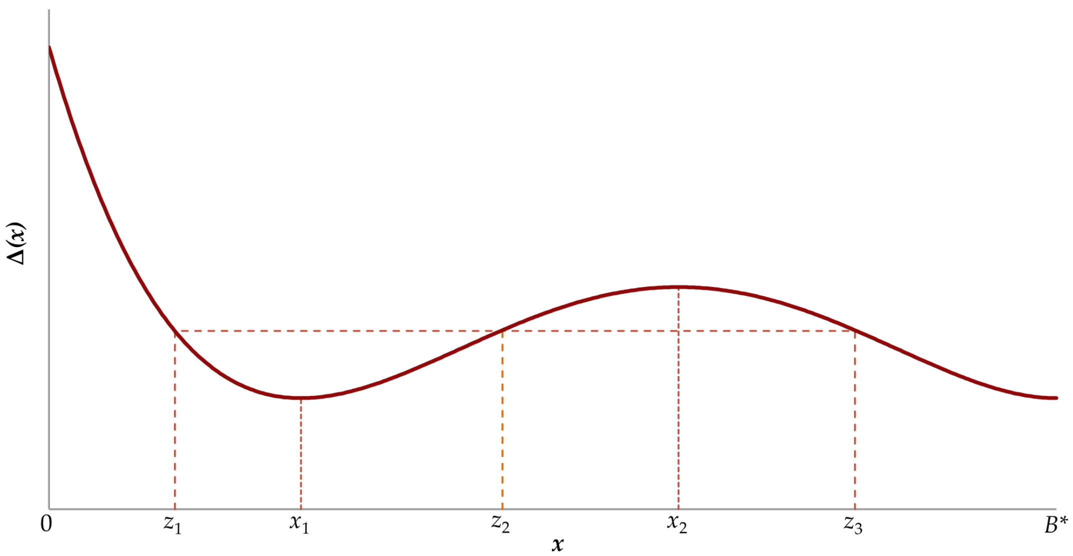


Figure 3 – Sorting and locational quality

Appendix A

A.1 Spatial patterns in the Randstad

In Figure A.1, we plot the incomes, land prices, amenities, and commuting times over space. In Panel A, we show the spatial income distribution for the Randstad. It can be seen that the income pattern is far from being symmetric and non-monotonic in distance to the center. In particular Rotterdam and The Hague have areas with high concentrations of poor households, in line with anecdotal evidence. We also observed that the corridor Utrecht–Amsterdam is generally inhabited by richer households. Land prices, which are displayed in Panel B of Figure A.1, are the highest in the centers of the bigger cities. It can also be seen that, although there seems to be a positive correlation between income and land prices, it is not always the case that the more expensive locations are inhabited by the rich.

[Figure A.1 about here]

By contrast, as mentioned earlier, the correlation between amenities and land values is high. Indeed, in Panel C of Figure A.1 we show that amenities are mainly concentrated in the areas with higher land prices. In Panel D of Figure A.1, we plot the average commuting times per location. In particular some locations near Rotterdam and The Hague seem to be inhabited by households with a short commute. The same holds for the city center of Amsterdam and an area to the north-west of Amsterdam.

A.2 Employment and commuting in the Randstad

One may question whether plotting the unconditional commuting time is informative, as we are predominantly interested in the variation in commuting times that is due to spatial variation in accessibility. In Panel A of Figure A.2, we show that there is indeed a negative correlation between commuting time and employment accessibility. This implies that areas with on average shorter commutes have a better accessibility to jobs (the unconditional correlation is -0.191). We provide more evidence for this relationship in Appendix C.6. In Panel B of Figure A.2, we investigate cross-commuting patterns in the Randstad. One could argue that the Randstad consists of many medium-sized cities that are not well connected to other places in the Randstad. Panel B shows the opposite: there is a lot of commuting between different zones in the Randstad, as well as *between* the four largest cities.

[Figure A.2 about here]

Appendix B

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

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

$$\Psi_x(x,\omega,U^*(\omega)) = \frac{\omega t}{H}\left(\frac{t_x}{t} - \frac{Q_a}{\omega t}a_x\right) \tag{B.1.1}$$

Differentiating (B.1.1) with respect to ω and rearranging terms yields the following expression:

$$\begin{aligned}\Psi_{x\omega}(x,\omega,U^*(\omega)) = \frac{t}{H}\Bigg\{&\frac{t_x}{t}\left[1-\frac{\omega}{H}(H_\omega + H_U U^*_\omega)\right]\\ &+\frac{a_x}{t}\left[\frac{H_\omega + H_U U^*_\omega}{H}Q_a - (Q_{aH}(H_\omega + H_U U^*_\omega) + Q_{aU}U^*_\omega)\right]\Bigg\}.\end{aligned} \tag{B.1.2}$$

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

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

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

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

Since

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

at the residential equilibrium, we have:

$$t = \frac{U^*_\omega}{a u_q}. \tag{B.1.3}$$

Plugging this expression, Q_a, Q_{aU} and Q_{aH} in (B.1.2), we get

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

which is equivalent to

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

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:

$$\Psi_{x\omega}(x,\omega,U^*(\omega)) = \frac{t}{H}\left[\left(1 - \frac{\varepsilon_{H,\omega} + \varepsilon_{u_q,\omega}}{\varepsilon_{U,\omega}}\right)A - \left(1 - \varepsilon_{H,\omega}\right)T\right], \tag{B.1.4}$$

which proves Proposition 1.

B.2 Modeling commuting costs

In the standard monocentric city model, the individual working time is supposed to be constant ($t_x = 0$), which implies commuting costs are given by an increasing function $c(x)$. In this case, if commuting costs are given by $\omega\tau(x) + c(x)$, the bid rent function $\Psi(x,\omega;U)$ is given by

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

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

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

Consequently,

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

and

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

Using (B.1.3), (B.2.1) may be rewritten as follows:

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

which reduces to (B.1.4) when $c_x = 0$. By contrast, when $t_x = 0$, we have $T(x) = 0$. Therefore, if $A(x) = 0$, we obtain:

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

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

Income-independent commuting costs under Stone-Geary preferences. When commuting costs are income-independent, the net income at x is given by $\omega - c(x)$. Using (B.6.1) in Appendix B.6 in which the net income $t(x)\omega$ is replaced by $\omega - c(x)$ and repeating the argument, we obtain the following aggregator:

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

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

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

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

B.3 Homothetic preferences

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

$$u_h = Ru_q.$$

It follows from Euler's theorem that

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

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

B.4 Stone-Geary preferences

It is readily verified from (14) that

$$Q(h, U/a(x)) = \left[\frac{1}{(h-\overline{h})^{\mu}}\frac{U}{a}\right]^{\frac{1}{1-\mu}}. \tag{B.4.1}$$

It follows from (B.4.1) that

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

Plugging Q_a, Q_{aH} and Q_{aU} into (B.1.2) and rearranging terms leads to

$$\begin{aligned}\Psi_{x\omega}(x,\omega,U^*(\omega)) = \frac{t}{H}\Bigg\{&\frac{t_x}{t}\left[1 - \frac{\omega}{H}(H_\omega + H_U U^*_\omega)\right] \\ &+ \frac{a_x}{a}\left[\frac{H_\omega + H_U U^*_\omega}{H}\left(-\frac{U}{t}Q_U\right)\left(\frac{h-(1-\mu)\overline{h}}{(1-\mu)(h-\overline{h})}\right)\right. \\ &\left.+\frac{Q_U}{(1-\mu)t}U^*_\omega\right]\Bigg\}. \end{aligned} \tag{B.4.2}$$

Plugging Q_h and Q in (9) and solving the corresponding equation yields

$$\frac{h-(1-\mu)\overline{h}}{(1-\mu)(h-\overline{h})} = \omega t\left[\frac{a}{U}(h-\overline{h})^{\mu}\right]^{\frac{1}{1-\mu}}. \tag{B.4.3}$$

Given the expression of Q_U, it turns out that

$$\left(-\frac{U}{t}Q_U\right)\left[\frac{h-(1-\mu)\overline{h}}{(1-\mu)(h-\overline{h})}\right] = -\frac{\omega}{1-\mu}. \tag{B.4.4}$$

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

$$\Psi_\omega(x,\omega,U^*(\omega)) = \frac{t}{H}\left(1 - \frac{Q_U}{t}U^*_\omega\right), \tag{B.4.5}$$

which is equal to 0 if and only if

$$U^*_\omega = \frac{t}{Q_U}. \tag{B.4.6}$$

Using (B.4.4) and (B.4.6), (B.4.2) can be rewritten as follows

$$\Psi_{x\omega}(x,\omega,U^*(\omega)) = \frac{t}{H}\cdot\left[1 - \frac{\omega}{H}(H_\omega + H_U U^*_\omega)\right]\cdot\frac{1}{1-\mu}\cdot\left[(1-\mu)\frac{t_x}{t} + \frac{a_x}{a}\right]. \tag{B.4.7}$$

Applying the implicit function theorem to (B.5.3) yields

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

and

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

Given Q_U, (B.4.6) can be expressed as the following differential equation:

$$U^*_\omega = t\cdot(1-\mu)\left[a\cdot(h-\overline{h})^\mu\right]^{\frac{1}{1-\mu}}(U^*(\omega))^{-\frac{\mu}{1-\mu}}. \tag{B.4.8}$$

We thus obtain

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

Therefore, by implication of (B.4.3), we have:

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

Substituting this expression into (B.4.7) yields:

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

Existence and uniqueness of the equilibrium housing consumption. The equilibrium housing demand satisfies (B.4.3). The LHS of (B.5.3) is decreasing and tends to $+\infty$ when $H \to \overline{h}$ and to $1/(1-\mu) > 0$ when $H \to +\infty$. The RHS of (B.4.3) is increasing in H. It tends to 0 when $H \to \overline{h}$ and to $+\infty$ when $H \to +\infty$. Therefore, (9) has a single solution $H(\omega t(x), U/a(x))$, which implies that there exists a unique equilibrium.

Using (9), we may rewrite (10) as follows:

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

Using Q_H leads to

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

B.5 Proof of Steps (i), (ii) and (iii) of Proposition 2

The bid-max lot size. From the definition of the location-quality index given by (18), (B.4.3) can be rewritten as follows

$$\frac{H-(1-\mu)\overline{h}}{(1-\mu)(H-\overline{h})}=\omega\Delta^{\frac{1}{1-\mu}}\left[\frac{(H-\overline{h})^{\mu}}{U}\right]^{\frac{1}{1-\mu}}, \tag{B.5.1}$$

which implies (20), so that the bid-max lot size depends on $a(x)$ and $t(x)$ through (18) only.

Equilibrium utility level. Applying the implicit function theorem to (B.5.1) yields

$$\frac{\partial H}{\partial \Delta}=-\left[U^{\frac{1}{1-\mu}}(H-\overline{h})^{-\frac{1}{1-\mu}-1}\frac{\mu H}{(1-\mu)}\right]^{-1}\omega\Delta^{\frac{\mu}{1-\mu}}<0. \tag{B.5.2}$$

Using the definition of the location-quality index, the differential equation (B.5.8) satisfied by the equilibrium utility level writes

$$U_{\omega}^{*}=\Delta^{\frac{1}{1-\mu}}(1-\mu)(h-\overline{h})^{\frac{\mu}{1-\mu}}(U^{*}(\omega))^{-\frac{\mu}{1-\mu}}. \tag{B.5.3}$$

Supermodularity of the equilibrium utility level. Differentiating (B.5.3) with respect to 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 (B.5.2), 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 (B.5.1), the expression in the bracketed term writes:

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

which implies

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

Since $\Psi_{\omega}(x,\omega,U^{*}(\omega))=0$, it follows from (B.4.5) that $U_{\omega}^{*}=t/Q_{U}$. Therefore, we obtain:

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

In other words, the supermodularity of $U^{*}(\omega)$ is equivalent to $\Psi_{\omega\Delta}>0$.

B.6 The land rent and land gradient

1.Rearranging (10) yields:

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

Using (B.4.6), 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_a into (B.1.1), we obtain:

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

and substituting t by its expression given in (B.4.6) we obtain:

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

B.7 City network

Since we do not make any specific assumption on the amenity and commuting time functions, our model is flexible enough to consider a city described by a transportation network defined by a finite set of nodes and a finite set of links that connect pairs of nodes. The network is supposed to be large enough to accommodate the whole population of households. Each link ℓ is characterized by specific amenity and commuting time functions, $a(x;\ell)$ and $t(x;\ell)$, which need not be specified in more detail. In this case, the location-quality index at $x \in \ell$ is specific to each link ℓ and given by $\Delta(x;\ell) = a(x;\ell)[t(x;\ell)]^{1-\mu}$.

Households choose the link ℓ and the location $x \in \ell$ that maximize their utilities. Households ordered by decreasing incomes are assigned to the link and the location endowed with falling values of the location-quality index. The equilibrium mapping on link ℓ, denoted by $\omega^*(x;\ell)$, solves (11). Households then determine the link which maximize their utilities. It is readily verified that Proposition 2 holds true along any link of the network. As a result, in equilibrium consumers sharing the same income may occupy separated locations along the same link or locations belonging to different links where the location-quality index takes the same value. By using sums rather than integrals, the above approach can be modified for x to vary discretely within the link.

9.1 B.8 The equilibrium land rent under Fréchet distributions

Rearranging terms in (16) yields:

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

and plugging the above expression into (B.5.10) leads to

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

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

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

Rearranging terms, this expression becomes:

$$\Phi = \mu(1-\mu)^{\frac{1-\mu}{\mu}}\left(\omega - \Phi\overline{h}\right)^{\frac{1}{\mu}}[U(\omega)]^{-\frac{1}{\mu}}\Delta^{\frac{1}{\mu}}. \qquad \text{(B.8.1)}$$

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

$$U^*(\omega) = \left(\omega - \Phi\overline{h}\right)X(\omega) \qquad \text{(B.8.2)}$$

be a solution to the above differential equation where $X(\omega)$ is determined below. Differentiating (B.8.2) 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

$$\Phi_\omega \equiv \frac{\mathrm{d}\Phi}{\mathrm{d}\omega} = \frac{\partial\Phi}{\partial\omega} + \Phi_\Delta\Delta_\omega = \Phi_\Delta\Delta_\omega. \qquad \text{(B.8.3)}$$

Differentiating (B.8.1) 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 (16) as follows:

$$H\Phi = \mu(\omega - \Phi\overline{h}) + \overline{h}\Phi. \qquad \text{(B.8.4)}$$

Plugging (B.8.4) into Φ_Δ leads to

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

Using Φ_ω and Δ_ω, (B.8.3) 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 (B.8.4) leads to the following differential equation in ω:

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

whose solution is

$$X(\omega) = k\left(\frac{\omega}{H}\right)^{\frac{\beta}{1-\mu}}, \tag{B.8.5}$$

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 (B.4.9), 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 (B.8.5) into (B.8.2) yields:

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

Plugging this expression into (B.8.1) and rearranging terms, we obtain the following implicit solution for the equilibrium land rent:

$$R^*(x) = \mu(1-\mu)^{\frac{1-\mu}{\mu}} k^{-\frac{1}{\mu}} t(x)\Delta^{\frac{1}{\mu}} \left[\frac{\mu t(x)}{R^*(x)} + \frac{(1-\mu)\overline{h}}{\omega^*(x)}\right]^{\frac{1}{(1-\mu)\mu\gamma}}. \tag{38}$$

Since the RHS of (B.8.5) is strictly decreasing and tends to 0 (∞) when $R(x) \to \infty$ (0), (B.R.8) has a unique solution in $R^*(x)$.

The lowest income in the sample, denoted by $\underline{\omega}$, is strictly positive. It follows from (27) that the lowest location-quality index associated with the poorest household is given by

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

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 (B.8.5) 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 C

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

The second part of this Appendix reports various additional econometric results. First, we report first-stage results in Appendix C.6. We undertake a wide range of robustness checks, including analyses using house prices, tests of identifying assumptions and other dependent variables and proxies for amenities, in Appendix C.7. In Appendix C.8 we report some additional results with respect to the counterfactual analyses.

C.1 Network distances

We obtain information on network distances from the *SpinLab* which enable us to calculate travel time τ between two locations. The dataset from *SpinLab* provides information on actual 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 postcode we calculate the straight-line distance to the nearest three 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 122m (the average is 153m). We also calculate the Euclidian distance from every job and photo location in the Randstad to the nearest three access points of the network. Then, 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 then choose the lowest of the network travel time and Euclidian travel time for observations that are within 2.5km of each other. This is because observations that are very close will not need to go via the network.

[Figure C.1 about here]

We illustrate this by plotting the relationship between the distance to the nearest center of the four largest cities and the travel time to one of these centers in Panel A of Figure C.1. The correlation between travel time and Euclidian distance is high ($\rho = 0.909$). For short distances (< 1km) we observe that it is often faster not to make use of the network so that the Euclidian travel speed is used. Beyond 10km the relationship between travel time and distance to the center is essentially linear. We also plot the relationship between speed and distance to the center of the Randstad in Panel B of Figure C.1. On average, the driving speed is 30km/h. However, it is shown that for short distances (< 5km), the speed is only 16km/h. In other words, speed is increasing in the distance travelled. This makes sense as for short distances people will likely have to use local streets on which speed limits are low; on the other hand, on longer trips it is more likely that people will make use of highways.

C.2 Land prices and lot sizes

Information on land values and lot sizes is not directly available but may be inferred from data on home sales. We 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 $1,478,871$ housing transactions.

Let $\mathcal{P}(x)$ denote the house price, $H(z)$ the observed lot size, $C(x)$ the housing characteristics at x. The log land rent $R(x)$ is equal to the fixed effects at the level of the postcode (about $15-20$ addresses), while $\eta_3(y)$ denote year y fixed effects. For each city, we estimate:

$$\log \frac{\mathcal{P}(x)}{H(z)} = \eta_1 C(x) + \log R(x) + \eta_3(y) + \epsilon(x), \tag{C.2.1}$$

where $\epsilon(x)$ is an identically and independently distributed error term that is assumed to be uncorrelated to land rents and housing characteristics, while η_1, η_2 and η_3 are parameters to be estimated. As $R(x)$ are given by the very local fixed effects, we do not impose any structure on how land rents $R(x)$ vary across locations.

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

Descriptive statistics for the housing sample are reported in Table C.1. Coefficients η_1 related to the housing attributes are reported in Table C.2. We see that the house price per square meter is generally a bit lower when the property is larger. However, the house price of properties that are (semi-)detached is generally lower. Furthermore, when the maintenance state of a property is good, prices are about 15% higher. When a property has central heating, the price per square meter is about 9% higher. The dummies related to the construction decades show the expected signs: in general, newer properties command higher prices. Properties constructed after World War II until 1970 generally have lower prices because this is a period associated with a lower building quality. In Appendix A.1 we already discussed the spatial distribution of land values. The lot sizes are inversely related to pattern of land prices ($\rho = -0.336$). In other words, more expensive locations generally have smaller lots, which makes sense.

C.3 Amenities

Picture index. We inspect the data underlying the picture index in Table C.3. We note that we already dropped all photos that have duplicate locations (55%). Recall that we only use pictures *outside* buildings taken by *residents* in the determination of the amenity index. More than 80% of the pictures are taken outside a building and about 60% of the pictures are taken by local residents.

[Tables C.3 about here]

In reporting the first-stage results in Appendix C.6, we show that there is a strong correlation with observed proxies of amenities, such as historic amenities (historic districts, number of listed buildings in the vicinity) and natural amenities (share open space, water within 500m).

Hedonic amenity index. We also test whether our results are robust to using an alternative hedonic amenity index, rather than relying on geocoded pictures. Following Lee and Lin (2018), we aim to construct an aggregate amenity index that describes the amenity level at every location

x.[28] We will make a distinction between *historic* amenities and *natural* amenities. We ignore consumption amenities (e.g., shops or restaurants) because they are potentially endogenous. However, it is worth noting that our results are similar when we also include consumption amenities, which suggests that historic and natural amenities are the main drivers for the results.

Let $\mathcal{A}(x)$ be a set of variables that describe amenities, $\mathcal{P}(x)$ the house price, while $C(x)$ are housing characteristics at x, and $\eta_2(y)$ are year y fixed effects. We also include street (*PC5*) fixed effects $\eta_3(x)$. *PC5* areas are rather small, so this should largely control for the accessibility to jobs and/or distance to the city center. To further reduce the bias of particular unobserved location attributes that may be correlated to $\mathcal{A}(x)$, we estimate the regression only using data from *outside the Randstad*, that is, data from outside our study area. This reduces the threat that local unobservable characteristics of a location are correlated with $\mathcal{A}(x)$. We then estimate:

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

where η_0, η_1, η_2 and η_3 are parameters to be estimated and $\epsilon(x)$ is an identically and independently distributed error term. We then use $\widehat{\eta}_0$ and $\mathcal{A}(x)$ to predict the amenity level in each location x in the Randstad:

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

where $a(x)$ is the (alternative) amenity value at x.

We use data on the universe of housing transactions in the Netherlands between 2008 and 2014 from the Land Registry. The descriptive statistics are reported in Table C.4. The average house price per square meter is about €2,000. Again, for the estimation of the hedonic amenity index, we only use data from outside the Randstad, which refers to a bit more than 50% of the data.

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

In Table C.5 we report the results of the regression of equation (C.3.1). For comparison, we first include data from the Randstad only. It can be seen that the share of built-up area in the vicinity has a negative effect on house prices. Given the mean of €2,000, the price is reduced by about 1% for a 10% increase in the share of built-up land. In other words, *households value open space*. Water bodies do not seem to have a statistically significant effect. The variables related to historic amenities are positive and statistically significant. For example, doubling the number of listed buildings located nearby implies a price increase of about €45-90 (2.2-4.4%). Being fully surrounded by historic district land implies a price increase of about 17%.

[28] 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 rather than *inter*-city variation. Using Albouy's proxy for amenities could capture the sorting of rich households in certain locations, but this is exactly the relationship we aim to test.

The results using data from outside the Randstad (see column (2)) show that the coefficients are a bit smaller, although the order of magnitude is similar. Most importantly, we find a negative effect of being in a historic district, but at the same time find a somewhat stronger effect of having a high share of historic district land around the property. This might be due to restrictions inside historic districts that prevent households from making substantial changes to their properties. In any case, the *net* effect of locating in a historic district is always positive.

Pokémon Go amenity index. To further show that our results are robust to other types of amenity indices, we also gather data on 'places of interest' from the augmented reality game *Pokémon Go*, which was a hugely popular game in 2017.[29] The game could be played at certain places of interest, the so-called 'Pokéstops'.[30] The locations of these Pokéstops often were determined in the geolocation game by *Ingress*. The developers then chose some of the first portals based on sites with historical or cultural significance, such as The Washington Monument, Big Ben, or museums. Other locations were chosen based on geotagged photos from Goggle. Many more portals were submitted as suggestions by *Ingress* players. There were approximately 15 million player-submitted portal locations, 5 million of which have been approved. Furthermore, *Ingress* player data have revealed the most popular of these portal locations; *Pokémon Go* has turned those into Pokéstops. In other words, these Pokéstops are not randomly located across space and signify locational attractiveness. So, very much like the picture index, Pokémon Go could be an alternative proxy for local amenities. We construct the Pokémon Go amenity index by using an expression similar to (32).

C.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

[29] It was one of the most used and profitable mobile apps in 2016, having been downloaded more than 500 million times worldwide.

[30] Another type of locations that are used in the game are so-called 'Gyms'. The latter type are unfortunately less useful, as these are almost uniformly distributed within urban areas in gardens, open spaces and public squares.

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, which implies a speed of about 70 km/h. We also gather data on the location of cinemas in 1910 from *SpinLab*.

We show a map of land use and the railway network for the Randstad in 1900 in Figure C.2. In Panel A it is shown that cities like Amsterdam, Rotterdam, The Hague, and Utrecht were already large by 1900. These cities host most of the cinemas at that time. It can also be seen that some areas that have been reclaimed from the sea (such as Flevoland in the upper-right part of the Randstad) did not exist in 1900. The Panel B of Figure C.2 shows the railway network and population accessibility. In particular, places around Amsterdam and Haarlem in the northern part of the Randstad have a high accessibility. The first railway line in the Netherlands was opened in 1839 between these two cities. The railway network was indeed the most developed in the northern part of the Randstad.[31]

[Figure C.2 about here]

We use data composed by *HISGIS*, which has compiled and digitized data from the first Dutch census in 1832. This dataset provides information on the land use of each parcel in the current inner cities of Amsterdam and Rotterdam, as well as for the province of Utrecht. The *HISGIS* data also provide information on the cadastral income, which was used to determine the tax at that time and is a proxy for land values. The Panel A of Figure C.3 shows that the study area is much smaller and excludes the city of The Hague. Hence, the results using data from 1832 is only based on a subsample of the population. We again rely on municipal population data from *NLGIS* to calculate the accessibility in 1832. When we do not have detailed information on buildings (such as in The Hague) we uniformly distribute the municipal population over space. Rail networks did not exist yet, so in order to calculate the population that could be reached within 90 minutes, we use information on the road network from 1832 obtained from Levkovich *et al.* (2017). Panel B of Figure C.3 shows that locations close to the roads in 1832 were generally much better accessible.

[Figure C.3 about here]

In Table C.6 we provide descriptives for all instruments. The average share of built-up area in 1900 was 7.3%, while it was $(0.0172 + 0.0487 =)$ 6.6% in 1832. However, this figure is a bit misleading because for 1832 we have mostly data near urban areas. On average about 404 thousand people could be reached within 90 minutes in 1900. Not surprisingly, this was much

[31] Note that in calculating the accessibility in each location in the Randstad, we take into account population and the railway network in whole of the Netherlands, thereby avoiding any problems related to boundaries.

lower (123 thousand) in 1832.

[Table C.6 about here]

C.5 Other descriptive statistics

In Figure C.4 we report the distributions of the log of income, land price, amenities and commuting. The distributions of land prices and amenities are positively skewed, while the distribution commuting time are somewhat negatively skewed.

[Figure C.4 about here]

C.6 First-stage results

Contemporary instruments. We first report first-stage estimates using contemporary instruments. The first five columns of Table C.7 focus on the effects on picture density. We find that *observable amenities have a strong effect on picture density.* For example, the amount of built-up area and water bodies in the vicinity imply an increase in the picture density. Historic amenities are generally positively related to the picture density, in line with expectations. For example, when the number of listed buildings in the neighborhood increases by 1%, the picture density increases by 0.6%. Historic districts also imply a positive increase in the picture density, although being *inside* historic districts seems to lower picture density compared to properties just outside historic districts. This may be due to the fact that transport nodes, like bus stops or railway stations, are often just outside historic districts. Furthermore, because historic districts are often small, it may be that people will take pictures of historic buildings located inside the historic district while standing outside the historic district. In any case, the net effect on amenities of being in a historic district is strongly positive: when a property is fully surrounded by historic districts within 500m, the picture density is $e^{0.80-0.27}-1=70\%$ higher. Employment accessibility seems to be positively related to picture density, which illustrates a potential endogeneity issue: it may be that jobs are located near high-amenity locations because of the presence of high income households so that there is a correlation with endogenous amenities (e.g., the presence of fancy shops). We will address this issue when using historic instruments. The coefficient related to amenities become somewhat smaller once we include location attributes and province fixed effects in column (2), but the first-stage coefficients remain more or less the same when including work-location fixed effects in column (3). Also, making selections (e.g., only focusing on households that have one job in a single-plant firm) in columns (4) and (5) does not change much.

[Table C.7 about here]

In columns (6)-(11) we report the first-stage results with respect to commuting time. The most important variable here is employment accessibility. We find that in all specifications a higher employment accessibility leads to lower commuting times, in line with expectations. When we do not include work-location fixed effects (columns (6)-(8)), the elasticity is about -0.35. The more believable specifications are in columns (9)-(11) where we keep the work location fixed. The elasticity increases to about -0.5, implying that the commuting time is reduced by 35% when the employment accessibility doubles. More specifically, given the work location, people who live in denser areas have on average a shorter commute. This is in line with the standard prediction of the monocentric city model.

Historic instruments. Because it is not unlikely that contemporary instruments are correlated to unobservable characteristics of a location, we also use historic instruments. The first-stage results are reported in Table C.8. These are the specifications corresponding to the estimates in Table 3. The instruments have the expected signs. In columns (1)-(5) we focus on the effects on amenities. The share of built-up area in 1900 and the share of water bodies in 1900 are strongly positively correlated to picture density. The former effect can be explained by the fact that the presence of historic amenities is strongly correlated to the number of historic buildings in the neighborhood. Cinemas in 1910 also have a positive effect on the current amenity level: a 1% increase in the number of cinemas in 1910 in the vicinity increases the picture density by 0.88% (see column (2)). The effect is about halved when we control for the current distribution of cinemas in column (3), Table C.8. In columns (4) and (5) we use instruments based on land use in 1832. The results are very similar. We find again positive effects of nearby buildings, built-up areas and weaker positive effects of nearby water bodies on the picture density. The impact of the instruments hardly changes when we control for work-location fixed effects in column (5).

[Table C.8 about here]

In columns (6)-(10) we investigate the effects of the instruments on commuting time. The most notable variable is the population that is accessible within 90 minutes. Since people at that time often reside relatively close to their working place, this is a proxy for the spatial employment distribution in 1900. Due to temporal autocorrelation we expect that a better population accessibility in 1900 implies a lower commute nowadays, which is indeed what we find. Doubling the population that is accessible within 90 minutes in 1900 leads to a reduction in commuting time of about 0.7%. The effect is highly statistically significant in all specifications, especially if we include work-location fixed effects in columns (7) and (8). The effect of the population in 1832 on current commuting time is somewhat stronger (see columns (9) and (10)). Hence, all our instruments have the expected signs and are statistically significant – and so seem to be strong instruments.

C.7 Sensitivity analysis

House prices. We commence the sensitivity analysis by testing the impact of amenities and commuting time on house prices. As shown in Sections 3 and 4, land prices play a major role in our model. They have been calculated by the procedure described in Appendix C.2. One may be worried that this may lead to measurement errors that are non-random. For example, the smoothing of data reduces the amount of variation in the data. In this subsection, we therefore repeat the same set of specifications as in Table 5, but use the log house price as the dependent variable. We use data from the Land Registry, which refer to the universe of housing transactions between 2008 and 2014. We already discussed descriptive statistics in Appendix C.4. Table C.9 reports the regression results.

[Table C.9 about here]

In column (1) we report a naive OLS specification of log house price per square meter on amenities and commuting time in minutes. We find a rather large effect of picture density on house prices, while the impact of commuting time is zero. This no longer holds when we instrument for commuting times in column (2) with current employment accessibility. We then find a strong negative effect of commuting times. The effect of commuting becomes smaller once we include location and province fixed effects, as well as work-location fixed effects in columns (3) and (4), respectively. The coefficients with respect to amenities become smaller once we instrument for the picture index with historic instruments. In our preferred specification in column (6) where we use historic instruments, doubling the amenity level implies a house price increase of 1.7%. House prices decrease by 19% when commuting time doubles. Compared to the results using land prices, the impacts of amenities and commuting times on land prices are slightly stronger, but otherwise very comparable.

Identification revisited. Since the validity of contemporary and historical instruments may be questioned, we estimate an additional set of specifications that should contribute to the belief that our identification strategy is valid. The results are reported in Table C.10. The first three columns focus on the use of contemporary instruments. In column (1), we instrument the commuting time with the distance to the nearest center of a city having at least 100 thousand inhabitants instead of the employment accessibility. This alternative instrument is unlikely to suffer from biases related to reverse causality. On the other hand, the definition of a city center is somewhat arbitrary. The results show that commuting times have a stronger effect, while the effect of amenities is somewhat lower. Nevertheless, both coefficients have the expected signs and are statistically significant. We may also try to mitigate biases associated with unobservable location characteristics or sorting by including more detailed fixed effects. In column (2), we include fixed effects for every city with at least 100 thousand inhabitants. The effect of amenities is essentially the same compared to the baseline estimate in column (7), Table 2. The effect of commuting is somewhat lower, but has the same order of magnitude. Detailed fixed effects are likely to absorb part of the variation in commuting times we are interested in, which may lead to an underestimation of the commuting effect. On the other hand, when we include detailed fixed effects we still may be able to identify the effects of amenities, for which there is much more local variation. In column (3), Table C.10, we include street (*PC5*) fixed effects, leading to a very similar effect for picture density. Identification of the coefficient with respect to commuting time is not possible because there is hardly any variation in employment accessibility within the street.

[Table C.10 about here]

Columns (4)-(7) rely on historic instruments. Instead of using population accessibility in 1900 as an instrument for commuting, we calculate the travel time using the railway network in 1900 to the nearest city center with at least 25 thousand inhabitants in 1900 and use that as an instrument. Column (4) shows that the results are virtually the same compared to the baseline estimate in column (2), Table 3. In column (5), Table C.10, we control for city fixed effects. This leads again to similar results. Column (6) includes *PC5* fixed effects. The impact of amenities becomes slightly stronger but is comparable to the baseline estimate. In the final specification, we improve on identification by *controlling* for the current share of built-up area and water bodies within 500m, as well as the cinemas in 2010 located nearby. Hence, we control for the fact that some places are currently more urban and, therefore, attract richer households. The results displayed in column (7) show that the impact of picture density is almost identical to the baseline estimate.

Sensitivity checks with respect to the amenity index. In Table C.11, we investigate the robustness of our results with respect to the definition of the amenity index. In Panel A of Table C.11, we use contemporary instruments to address the endogeneity of amenities and commuting time; in Panel B we use historic instruments. In column (1), we use an alternative proxy for amenities: the number of listed buildings located nearby using the distance decay function of (32) and assuming $\delta = 0.915$. We then only instrument for commuting times with employment accessibility in Panel A, but instrument both variables in Panel B with historic land use and population accessibility. The results indicate that the effect of the number of listed buildings is positive and statistically significant, albeit somewhat lower than the impact of pictures when using contemporary instruments. This is not too surprising as listed buildings only capture historic amenities and do not include natural amenities. Hence, the picture density as a proxy for amenities is preferred. In column (2), we employ the amenity index based on the augmented reality game *Pokémon Go*. The results again confirm that the impact of amenities is positive and statistically significant. With historic instruments (Panel B), the coefficient implies that an amenity level that doubles will attract households whose incomes are 1.7% higher.

[Table C.11 about here]

In column (3), we test whether the choice of transport mode matters for the results. We calculate the travel time to work and to amenities by rail. This has limited impact on the coefficient related to amenities. The coefficient with respect to commuting time is lower, albeit negative and statistically significant, probably because we have measurement error. This holds in particular for locations that are not close to a station, where households are unlikely to take the train. Moreover, we do not have information on bus and tram routes, so this alternative way of calculating commuting time and the amenity index is imperfect. In column (4), Table C.11 we just calculate the Euclidian distance between photos instead of using car or train. Coefficients are very similar to the baseline estimates, although the coefficient with respect to commuting time is somewhat lower.

One may argue that part of the effect of amenities may be endogenous once the historic instruments are correlated to the locations of endogenous amenities. This might be the case when, for example, shops or bars are disproportionately located in historic buildings. In column (5), we therefore *control* for endogenous amenities by measuring the density of shops, bars, restaurants and other cultural amenities, using the distance decay function of (32) and assuming that $\delta = 0.915$. When we instrument with contemporary instruments in Panel A, we find *a strong negative impact of endogenous amenities*, potentially because shops and restaurants may imply noise pollution and lead to congestion and parking issues. The impact of "exogenous" amenities now becomes much stronger. The same holds, but to a lesser extent, when we instrument for the picture index with historic instruments.

The final column tests whether the presence of potential disamenities affects our conclusion that amenities attract richer households. We gather data on noise pollution (at a 25 by 25m grid), flood risk, as well as air pollution. We proxy air pollution by the intensity of Particulate Matter (PM_{10}), Carbon Monoxide (CO), Ammonia (NH_3) and Nitrogen Dioxide (NO_2). The results are essentially unaffected, which suggests that urban amenities are not much correlated to disamenities.

Other sensitivity checks. Table C.12 reports the results of additional robustness checks. In Panel A of Table C.12, we use contemporary instruments to address the endogeneity of amenities and commuting time; in Panel B we use historic instruments.

[Table C.12 about here]

Figure 1 shows that there is a vast enclave delimited by the cities of Amsterdam, Rotterdam, The Hague, and Utrecht where essentially no new construction is allowed, i.e., the Green Heart. Hence, the corresponding population and employment density is considerably lower. As we are mainly interested in the effects of amenities and commuting on the social structure in urban areas, we exclude observations in the Green Heart in column (1). The results are essentially unaffected, which is not too surprising as most of our observations stem from outside this area.

In the main specifications, we do not control for house size because the income mapping (30) already takes into account the endogenous determination of lot size. However, if we do control for (the log of) house size in column (2), Table C.12, the results are very similar.

In column (3), we replace the dependent variable – log of income per hour – by the share of adults in the households that have at least a bachelor's degree. Measuring household income is not always straightforward (e.g., when the income is based on dividends). In this case, it may be preferable to use educational level instead, i.e., a proxy that is highly correlated to income. We find that doubling amenities implies an increase in households with at least share bachelor's degree by $1.2 - 1.6\%$ higher. Doubling commuting time implies a reduction in households that have a bachelor's degree by $8 - 16\%$. In other words, the results are very robust when using a different proxy for household income.

In columns (4) and (5), we investigate whether our results are primarily driven by the arbitrary choice of the decay parameter δ. We have set $\delta = 0.915$, so that the expected travel time is 15 minutes when amenities are evenly distributed across space. We also run regressions when the expected travel time is, respectively, 5 and 30 minutes, which yield $\delta = 1.585$ and $\delta = 0.647$. Again, amenities have a positive and statistically significant effect, while the effect of commuting is negative (see columns (7) and (8), Table C.9). In general, for lower decay parameters, the impact of amenities becomes somewhat stronger and the impact of commuting somewhat lower.

When using historic instruments, we use population in 1900 within 90 minutes travelling as an instrument for current commuting time. In columns (6) and (7) in Panel B of Table C.12, we check that changing this assumption does not affect the results. In column (6), we take into account population within 60 minutes driving and in column (7) we take the current distribution of commuting times to calculate population accessibility in 1900, in line with equation (34). The coefficients are essentially unaffected.

City-specific results. It is also interesting to estimate regressions where the effects of amenities and commuting time on the spatial income distribution is city-specific. Hence, we estimate separate regressions for Amsterdam, Rotterdam, The Hague, and Utrecht. We report the results in Table C.13. Since we have fewer observations, there is less spatial variation in the instruments. For the regressions in columns (1)-(4), we therefore also use distance to the nearest center of a city with at least 100 thousand inhabitants and in columns (5)-(8) the travel time using the railway network in 1900 to the nearest city center with at least 25 thousand inhabitants in 1900 as additional instruments, respectively. Results without those instruments are, however, very comparable.

[Table C.13 about here]

In column (4), we use contemporary instruments. It appears that in *all cities amenities are* *positively related to incomes*. We find the strongest impact in The Hague. There a 100% increase in amenities attracts households whose incomes are 3.2% higher. The coefficients for amenities are very similar between Amsterdam, Rotterdam, and Utrecht. For commuting the impacts are a bit more different between the different cities. For Amsterdam we find the strongest impact, while for the Hague we find a small positive impact (although it is only marginally statistically significant) in column (2). We do not have a clear explanation for this, except that this may indicate that contemporary instruments are correlated to unobserved reasons why high income households sort into certain locations. Indeed, the more convincing identification strategy where we use historic instruments delivers the expected coefficients for all cities. The impact of amenities is remarkably similar between cities and ranges from 0.015 to 0.025. For commuting the effects are a bit more different across cities. With historic instruments, for Amsterdam and Utrecht, the commuting time elasticity is about 0.25, while for Rotterdam and The Hague it is about 0.08.

In short, the results for the different cities show that the definition of the urban area of reference, i.e., a polycentric urban region as the Randstad or focusing on individual cities instead, does not change the results.

C.8 Structural estimation and counterfactual analyses

**Presenting $\hat{\omega}(x)$ or $\hat{R}(x)$ as a function of distance to the center may not be the best way to illustrate the resulting distributions because the Randstad is asymmetric and polycentric, while the definition of the centers is somewhat arbitrary. We, therefore, compile maps with the predicted income pattern based on historic instruments and work location fixed effects. Because of the high correlation between $\hat{\omega}(x)$ and $\hat{\Delta}(x)$, we do not report $\hat{\Delta}(x)$.

We present the income mapping for the different scenarios in Figure C.5. In Panel A, we report results for the baseline scenario. Predicted incomes are generally the highest in city centers, where the amenity level is high and the commuting costs are low. However, there are some relatively low-dense areas (e.g., to the east of Rotterdam) with high predicted incomes because of superb accessibility to jobs. The map also highlights that predicted income patterns are often non-monotonic and non-symmetric in distance to the (nearest) city center. In Panel B, we show results when considering monocentric cities without historic amenities. We show that this considerably simplifies the spatial income pattern. High incomes are now residing in and near city centers and income gradients seem monotonically decreasing in distance to the city center, which is in line with the predictions by the monocentric city model. In Panel C, we restrict households to travel by train instead of travelling by car to work and to amenities. Given the current railway network, this will have strong repercussions for the income distributions. High income households will now sort close to railway stations, where accessibility to jobs and amenities is higher.

In counterfactual scenario 3 shown in Panel D, we assume that highly educated people will commute two days a week. This seems to imply that amenities are now more important for the rich, rather than accessibility. Hence, high amenity locations in and near Amsterdam and the city center of Rotterdam seem to disproportionally attract the rich.

[Figures C.5 and C.6 about here]

We repeat the same exercise for land rents, reported in Figure C.6. Because of the expected high correlation between $\hat{R}(x)$ and $\hat{\omega}(x)$, as the rich inhabit the most attractive locations, the spatial distribution of land rents looks very similar to the predicted income mapping.**

Appendix tables

TABLE C.1 – DESCRIPTIVES FOR *NVM* DATASET

	(1) mean	(2) sd	(3) min	(4) max
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 C.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 C.3 – Descriptives for pictures

	(1) mean	(2) sd	(3) min	(4) max
Year	2,011	2.201	2,004	2,014
Hour	13.91	4.452	0	23
Picture inside a building	0.195	0.396	0	1
Local resident	0.592	0.491	0	1

Notes: The number of observations is 665,105. The sample is taken from the Randstad (44% of all the pictures in the Netherlands). The data are from 2000-2014.

Table C.4 – Descriptive statistics for Land Registry data

	(1) mean	(2) sd	(3) min	(4) max
House price *(in € per m^2)*	2,050	771.8	201.7	5,984
Amenities, $\delta = 0.915$	11.38	38.45	1.000	806.3
Commuting time *(in minutes)*	27.39	17.48	0.656	119.5
Employment accessibility	523,462	158,887	91,769	966,672
Size of the property *(in m^2)*	118.8	44.42	26	350
Property type – apartment	0.233	0.423	0	1
Property type – terraced	0.381	0.486	0	1
Property type – semi-detached	0.279	0.448	0	1
Property type – detached	0.107	0.309	0	1
Share built-up area < 500m	0.712	0.213	0.00450	1
Share open space < 500m	0.251	0.211	0	0.993
Share water bodies < 500m	0.0368	0.0646	0	0.773
Number of listed buildings, $\delta = 0.915$	3.777	14.92	1	358.6
In historic district	0.0564	0.231	0	1
Share land in historic districts < 500m	0.0582	0.177	0	1
In Randstad	0.474	0.499	0	1

Notes: The number of observations is 451,760. Because of confidentiality restrictions the minimum and maximum values refer to the 0.0001^{th} and 99.9999^{th} percentile.

Table C.5 – Determining the hedonic amenity index

(Dependent variable: house price per m^2)

	Inside Randstad	*Outside Randstad*
	(1)	(2)
	OLS	OLS
Share built-up land < 500m	-284.5596***	-97.2366***
	(8.9879)	(4.9370)
Share water bodies < 500m	-49.0344**	-29.4911
	(22.3496)	(22.8423)
Listed buildings, $\delta = 0.915$	130.3866***	65.7576***
	(2.4862)	(2.3899)
In historic district	98.0307***	-59.7472***
	(9.8222)	(9.4094)
Share historic districts < 500m	284.9337***	389.1115***
	(15.2224)	(16.9976)
Housing attributes	Yes	Yes
PC5 fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Number of observations	214,001	237,759
R^2	0.6115	0.4374

Notes: Standard errors are clustered at the *PC5* level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table C.6 – Descriptive statistics for historic data

	(1) mean	(2) sd	(3) min	(4) max
Share built-up land in 1900 < 500m	0.0729	0.144	0	0.959
Share open space in 1900 < 500m	0.882	0.191	0.000264	1
Share water bodies in 1900 < 500m	0.0832	0.218	0	1
Cinemas in 1910, $\delta = 0.915$	1.001	0.0230	1	2.370
Population in 1900, < 90min	404,044	299,596	0	1.277e+06
Share buildings in 1832 < 500m	0.0172	0.0561	0	0.490
Share other built-up land in 1832 < 500m	0.0487	0.0799	0	0.801
Share open space in 1832 < 500m	0.864	0.198	0.00163	1
Share water bodies in 1832 < 500m	0.0742	0.165	0	0.997
Population in 1832, < 90min	122,761	87,115	1,949	318,553
Cadastral income in 1832 per ha	1,425	2,779	0.0235	122,205
Cadastral income in 1832 is zero	0.0530	0.224	0	1

Notes: The number of observations is 4,346,889 for the variables based on data from 1900 and 1,340,991 for the variables based on data from 1832. Because of confidentiality restrictions the minimum and maximum values refer to the 0.0001^{th} and 99.9999^{th} percentile.

Table C.7 – First-stage regression results: contemporary instruments

	(Dependent variable: the log of amenities, $\delta = 0.915$)					*(Dependent variable: the log of commuting time)*					
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS	(8) OLS	(9) OLS	(10) OLS	(11) OLS
Amenities, $\delta = 0.915$ (*log*)						-0.0160***					
						(0.0008)					
Employment accessibility	0.3945***	0.4109***	0.4559***	0.4873***	0.4441***	-0.3500***	-0.3442***	-0.3522***	-0.5140***	-0.5064***	-0.5404***
	(0.0124)	(0.0139)	(0.0133)	(0.0137)	(0.0147)	(0.0028)	(0.0028)	(0.0028)	(0.0027)	(0.0036)	(0.0040)
Share built-up land $<$ 500m	1.3591***	1.1625***	1.1246***	1.1548***	1.0926***		-0.0833***	-0.0875***	-0.0706***	-0.0512***	-0.0323***
	(0.0164)	(0.0183)	(0.0175)	(0.0183)	(0.0189)		(0.0039)	(0.0041)	(0.0034)	(0.0043)	(0.0063)
Share water bodies $<$ 500m	1.6538***	1.6733***	1.7639***	1.8607***	1.6345***		0.0444***	0.0939***	0.0307***	0.0088	-0.0102
	(0.0778)	(0.0830)	(0.0789)	(0.0799)	(0.0814)		(0.0106)	(0.0104)	(0.0082)	(0.0103)	(0.0135)
Listed buildings, $\delta = 0.915$ (*log*)	0.6109***	0.5365***	0.5350***	0.5345***	0.5495***		-0.0209***	-0.0222***	-0.0156***	-0.0180***	-0.0222***
	(0.0062)	(0.0066)	(0.0063)	(0.0064)	(0.0072)		(0.0013)	(0.0015)	(0.0014)	(0.0017)	(0.0023)
In historic district	-0.2695***	-0.2339***	-0.2235***	-0.2232***	-0.1933***		0.0207***	0.0127***	0.0134***	0.0157***	-0.0284***
	(0.0209)	(0.0199)	(0.0188)	(0.0192)	(0.0213)		(0.0048)	(0.0047)	(0.0042)	(0.0056)	(0.0085)
Share historic districts $<$ 500m	0.7997***	0.6553***	0.5939***	0.5662***	0.5657***		-0.0791***	-0.0449***	-0.0927***	-0.1256***	-0.0691***
	(0.0375)	(0.0366)	(0.0351)	(0.0355)	(0.0387)		(0.0082)	(0.0080)	(0.0075)	(0.0093)	(0.0127)
Household characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Housing attributes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location attributes	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Province fixed effects	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Work location fixed effects	No	No	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
Number of observations	4,346,889	4,346,889	4,162,962	1,558,535	724,691	4,346,889	4,346,889	4,346,889	4,162,962	1,558,535	724,691
R^2	0.6420	0.6718	0.7134	0.7095	0.7337	0.0618	0.0633	0.0669	0.3201	0.3144	0.4129

Notes: Standard errors are clustered at the *PC5* level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table C.8 – First-stage regression results: historic instruments

	(Dependent variable: the log of amenities, $\delta = 0.915$)					(Dependent variable: the log of commuting time)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Population in 1900, < 90min (*log*)	0.0447***	0.0442***	0.0437***			-0.0087***	-0.0099***	-0.0099***		
	(0.0021)	(0.0021)	(0.0020)			(0.0004)	(0.0004)	(0.0004)		
Share built-up land in 1900 < 500m	3.5053***	3.4473***	3.1109***			-0.1839***	-0.2058***	-0.2031***		
	(0.0386)	(0.0361)	(0.0364)			(0.0080)	(0.0083)	(0.0087)		
Share of water bodies in 1900 < 500m	1.0925***	1.2238***	1.2340***			0.0838***	0.0523***	0.0522***		
	(0.0841)	(0.0733)	(0.0724)			(0.0104)	(0.0111)	(0.0111)		
Cinemas in 1910, $\delta = 0.915$ (*log*)	0.7771***	0.8836***	0.4109***			0.1235***	0.0711**	0.0749**		
	(0.1201)	(0.1104)	(0.0980)			(0.0300)	(0.0296)	(0.0297)		
Population in 1832, < 90min (*log*)				0.1033***	0.0984***				-0.0164***	-0.0336***
				(0.0078)	(0.0083)				(0.0020)	(0.0020)
Share buildings in 1832 < 500m				5.1608***	5.0423***				-0.4139***	-0.3151***
				(0.1268)	(0.1239)				(0.0471)	(0.0424)
Share other built-up land in 1832 < 500m				1.5572***	1.5977***				0.2028***	0.1141***
				(0.0939)	(0.0922)				(0.0253)	(0.0230)
Share water bodies in 1832 < 500m				0.2361***	0.2408***				0.0314***	-0.0071
				(0.0443)	(0.0445)				(0.0090)	(0.0090)
Household characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Housing attributes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location attributes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Work location fixed effects	No	Yes	Yes	No	Yes	No	Yes	Yes	No	Yes
Number of observations	4,346,889	4,162,962	4,162,962	1,340,991	1,265,990	4,346,889	4,162,962	4,162,962	1,340,991	1,265,990
R^2	0.6445	0.6938	0.7016	0.7584	0.7892	0.0371	0.2774	0.2774	0.0450	0.3910

Notes: Standard errors are clustered at the *PC5* level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table C.9 – Sensitivity analysis: effects on housing prices

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

		Contemporary instruments			*Historic instruments*	
	(1) OLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
Amenities, $\delta = 0.915$ *(log)*	0.0794***	0.0462***	**0.0604***	**0.0630***	**0.0228***	**0.0250***
	(0.0019)	(0.0019)	(**0.0036**)	(**0.0031**)	(**0.0047**)	(**0.0047**)
Commuting time in minutes (*log*)	-0.0003	**-0.6260***	**-0.3845***	**-0.1933***	**-0.3586***	**-0.2740***
	(0.0014)	(**0.0183**)	(**0.0174**)	(**0.0114**)	(**0.0656**)	(**0.0648**)
Household characteristics	No	Yes	Yes	Yes	Yes	Yes
Housing attributes	No	Yes	Yes	Yes	Yes	Yes
Location attributes	No	No	Yes	Yes	Yes	Yes
Province fixed effects	No	No	Yes	Yes	Yes	Yes
Work location fixed effects	No	No	No	Yes	No	Yes
Number of observations	214,001	214,001	214,001	179,849	214,001	179,849
R^2	0.1284					
Kleibergen-Paap F-statistic		3,743	467.7	606.3	46.50	30.69

Notes: **Bold** indicates instrumented. Standard errors are clustered at the *PC5* level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table C.10 – Sensitivity analysis: identification revisited

(Dependent variable: the log of hourly income)

	Contemporary instruments			*Historic instruments*			
	(1) 2SLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS	(7) 2SLS
Amenities, $\delta = 0.915$ *(log)*	**0.0326*****	**0.0339*****	**0.0382*****	**0.0135*****	**0.0142*****	**0.0332*****	**0.0112*****
	(**0.0017**)	(**0.0013**)	(**0.0014**)	(**0.0020**)	(**0.0024**)	(**0.0017**)	(**0.0030**)
Commuting time in minutes	**-0.0918*****	**-0.0497*****		**-0.2253*****	**-0.2465*****		**-0.3877*****
	(**0.0093**)	(**0.0040**)		(**0.0150**)	(**0.0278**)		(**0.0349**)
Current land use and cinemas	No	No	No	No	No	No	Yes
Household characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Housing attributes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location attributes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City fixed effects	No	Yes	Yes	No	Yes	Yes	No
PC5 fixed effects	No	No	Yes	No	No	Yes	No
Work location fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	4,162,769	4,162,769	4,346,887	4,162,769	4,162,769	4,346,887	4,162,769
Kleibergen-Paap F-statistic	488.8	3677	3689	332.9	111.5	4122	73.18

Notes: **Bold** indicates instrumented. Standard errors are clustered at the *PC5* level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table C.11 – Sensitivity analysis: amenities

(Dependent variable: the log of hourly income)

	Listed buildings	*Pokémon index*	*By train*	*Euclidian*	*Shops*	*Disamenities*
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Contemporary instruments	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Amenities, $\delta = 0.915$ *(log)*			**0.0324*****	**0.0311*****	**0.0813*****	**0.0394*****
			(0.0014)	**(0.0013)**	**(0.0020)**	**(0.0014)**
Listed buildings, $\delta = 0.915$ (*log*)	0.0257***					
	(0.0009)					
Pokéstops, $\delta = 0.915$ (*log*)		**0.0232*****				
		(0.0023)				
Number of shops, $\delta = 0.915$ (*log*)					-0.0610***	
					(0.0012)	
Commuting time in minutes (*log*)	**-0.1164*****	**-0.1153*****	**-0.0930*****	**-0.0913*****	**-0.0754*****	**-0.0854*****
	(0.0038)	**(0.0047)**	**(0.0055)**	**(0.0042)**	**(0.0046)**	**(0.0042)**
Noise and air pollution	No	No	No	No	No	Yes
Household characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Housing attributes	Yes	Yes	Yes	Yes	Yes	Yes
Location attributes	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Work location fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	4,162,769	4,162,769	4,162,769	4,162,769	4,162,769	4,162,769
Kleibergen-Paap F-statistic	38,599	2,099	1,181	2,973	2,187	4,263
Panel B: Historic instruments						
Amenities, $\delta = 0.915$ (*log*)			**0.0200*****	**0.0124*****	**0.0447*****	**0.0171*****
			(0.0017)	**(0.0024)**	**(0.0022)**	**(0.0037)**
Listed buildings, $\delta = 0.915$ (*log*)	**0.0209*****					
	(0.0012)					
Pokéstops, $\delta = 0.915$ (*log*)		**0.0241*****				
		(0.0044)				
Number of shops, $\delta = 0.915$ (*log*)					-0.0502***	
					(0.0016)	
Commuting time in minutes (*log*)	**-0.2448*****	**-0.2662*****	**-0.0674*****	**-0.2467*****	**-0.2179*****	**-0.4201*****
	(0.0187)	**(0.0233)**	**(0.0070)**	**(0.0278)**	**(0.0256)**	**(0.0882)**
Noise and air pollution	No	No	No	No	No	Yes
Household characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Housing attributes	Yes	Yes	Yes	Yes	Yes	Yes
Location attributes	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Work location fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	4,162,769	4,162,769	4,162,769	4,162,769	4,162,769	4,162,769
Kleibergen-Paap F-statistic	178.8	137.3	162.7	116.5	117.8	29.98

Notes: **Bold** indicates instrumented.Standard errors are clustered at the *PC5* level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table C.12 – Sensitivity analysis: other sensitivity checks

	No Green Heart	*Include house size*	*Bachelor's degree*	*Decay of amenities*		*Population 1900*	
				$\delta = 1.585$	$\delta = 0.647$	< 60 *minutes*	*Current distr.*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Contemporary instruments	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Amenities, (*log*)	**0.0336*****	**0.0325*****	**0.0319*****	**0.0377*****	**0.0317*****		
	(**0.0014**)	(**0.0013**)	(**0.0013**)	(**0.0015**)	(**0.0014**)		
Commuting time in minutes (*log*)	**-0.0843*****	**-0.1155*****	**-0.1176*****	**-0.0946*****	**-0.0733*****		
	(**0.0047**)	(**0.0040**)	(**0.0046**)	(**0.0041**)	(**0.0047**)		
Household characteristics	Yes	Yes	Yes	Yes	Yes		
Housing attributes	Yes	Yes	Yes	Yes	Yes		
Location attributes	Yes	Yes	Yes	Yes	Yes		
Province fixed effects	Yes	Yes	Yes	Yes	Yes		
Work location fixed effects	Yes	Yes	Yes	Yes	Yes		
Number of observations	3,693,277	4,162,769	3,047,506	4,162,769	4,162,769		
Kleibergen-Paap F-statistic	2312	2828	2469	3344	1902		
Panel B: Historic instruments							
Amenities, (*log*)	**0.0114*****	**0.0120*****	**0.0254*****	**0.0124*****	**0.0141*****	**0.0135*****	**0.0124*****
	(**0.0027**)	(**0.0023**)	(**0.0028**)	(**0.0024**)	(**0.0028**)	(**0.0022**)	(**0.0022**)
Commuting time in minutes (*log*)	**-0.2709*****	**-0.2777*****	**-0.2240*****	**-0.2624*****	**-0.2242*****	**-0.2268*****	**-0.2502*****
	(**0.0347**)	(**0.0263**)	(**0.0338**)	(**0.0253**)	(**0.0320**)	(**0.0224**)	(**0.0214**)
Household characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Housing attributes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location attributes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Work location fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	3,693,277	4,162,769	3,047,506	4,162,769	4,162,769	4,162,769	4,162,769
Kleibergen-Paap F-statistic	87.11	116.7	100.9	133.3	103.6	103.3	220.6

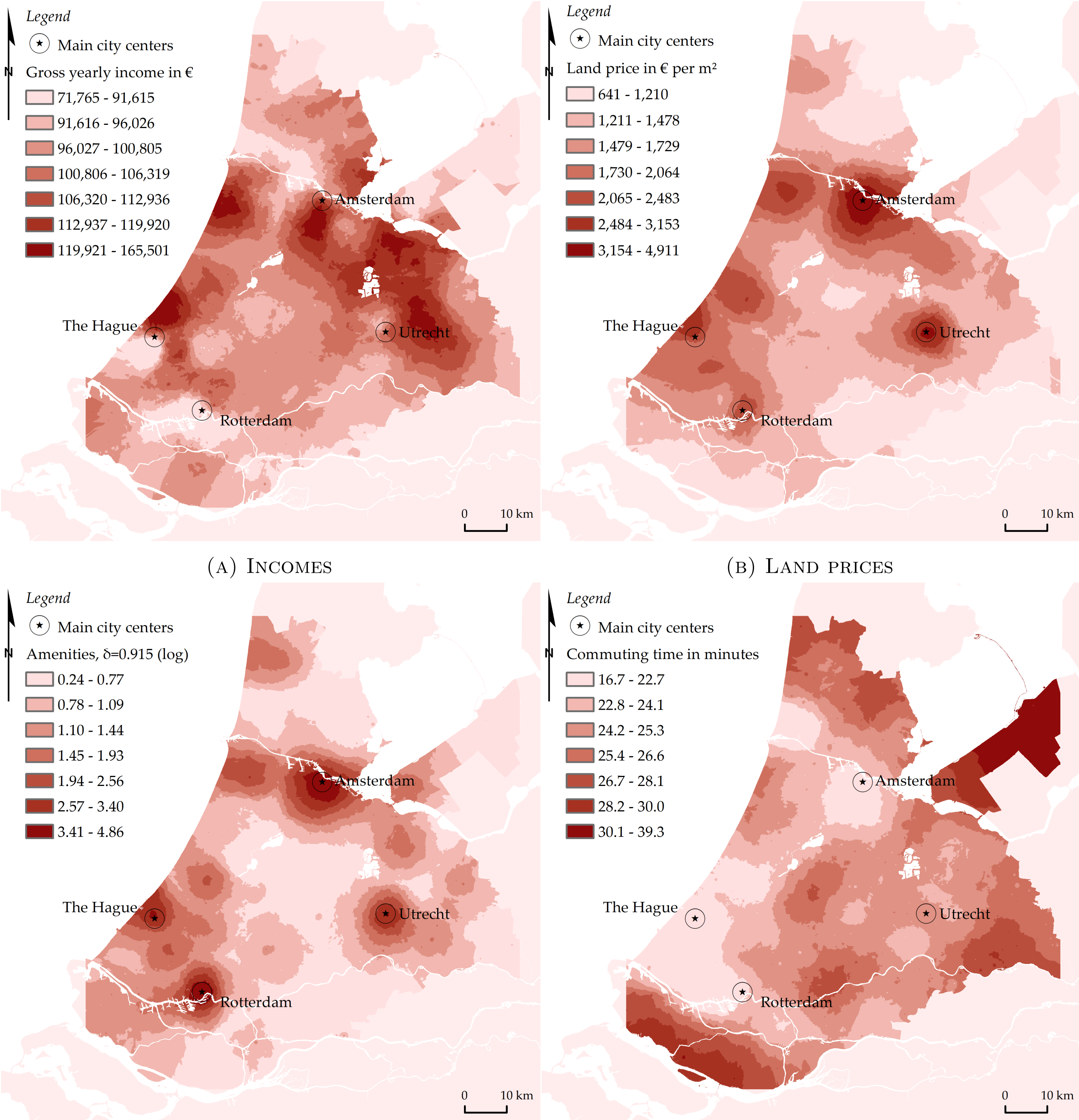
Notes: **Bold** indicates instrumented.In columns (1)-(4) and (5)-(7), we assume that $\delta = 0.915$ for amenities. In all columns but column (3), the dependent variable is the log of income per hour. In column (3) it is the share of the adult members of the household that have at least a Bachelor's degree. Standard errors are clustered at the *PC5* level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table C.13 – Sensitivity analysis: city-specific results

	Contemporary instruments				*Historic instruments*			
	Amsterdam	*Rotterdam*	*The Hague*	*Utrecht*	*Amsterdam*	*Rotterdam*	*The Hague*	*Utrecht*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Amenities, $\delta = 0.915$ (*log*)	**0.0279*****	**0.0237*****	**0.0466*****	**0.0254*****	**0.0145****	**0.0231*****	**0.0208*****	**0.0246*****
	(**0.0036**)	(**0.0017**)	(**0.0032**)	(**0.0025**)	(**0.0058**)	(**0.0018**)	(**0.0035**)	(**0.0035**)
Commuting time in minutes (*log*)	**-0.1013*****	**-0.0583*****	**0.0204***	**-0.0421*****	**-0.2600*****	**-0.0754*****	**-0.0945*****	**-0.2652*****
	(**0.0104**)	(**0.0060**)	(**0.0120**)	(**0.0068**)	(**0.0698**)	(**0.0096**)	(**0.0281**)	(**0.0239**)
Household characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Housing attributes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location attributes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Work-location fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	1,080,276	1,008,788	762,976	817,144	1,080,276	1,008,788	762,976	817,144
Kleibergen-Paap F-statistic	520.8	1904	436.1	866.2	20.45	362.8	160.6	92.77

Notes: **Bold** indicates instrumented.Standard errors are clustered at the *PC5* level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Appendix figures



(A) Incomes

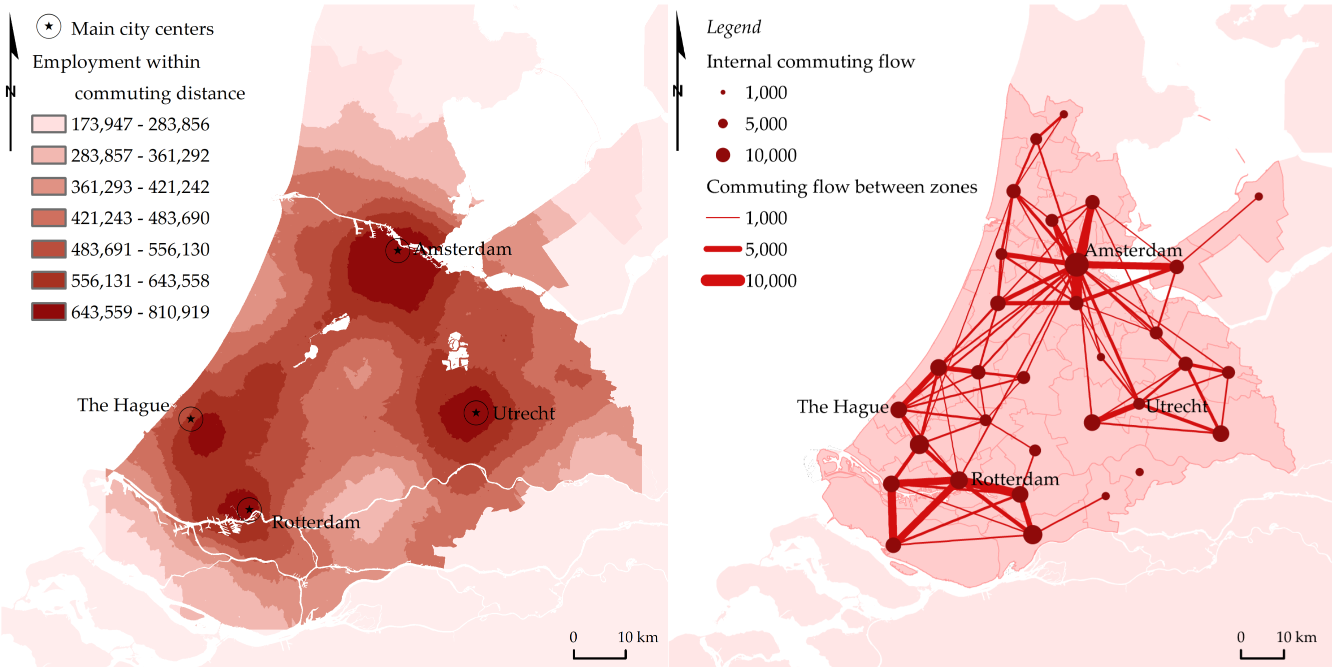
(B) Land prices

(C) Amenities

(D) Commuting times

Figure A.1 – Spatial gradients in the Randstad

Notes: The above figures report Kernel-smoothed values of the variables of interest.

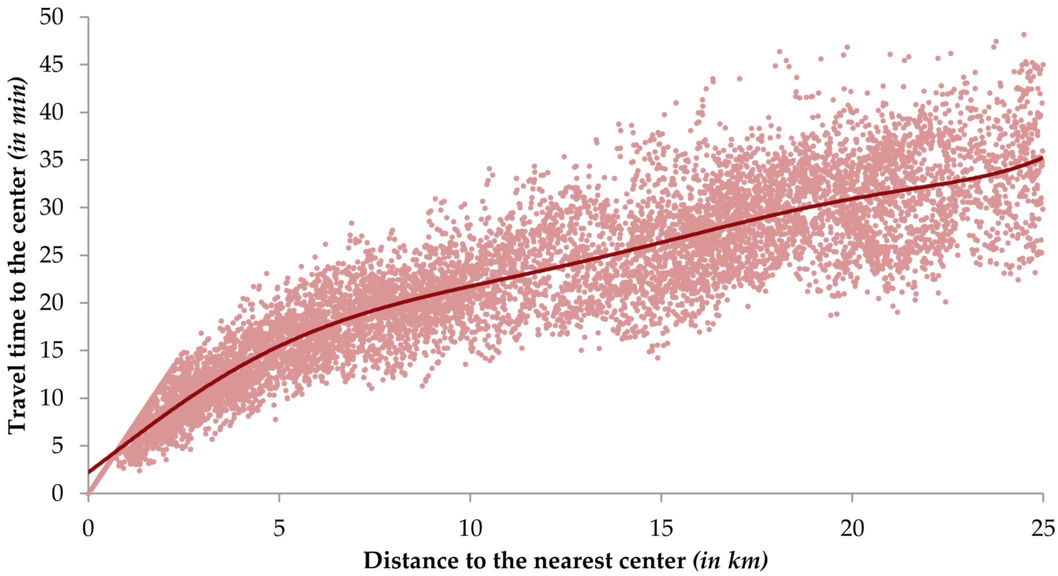


(A) Employment accessibility

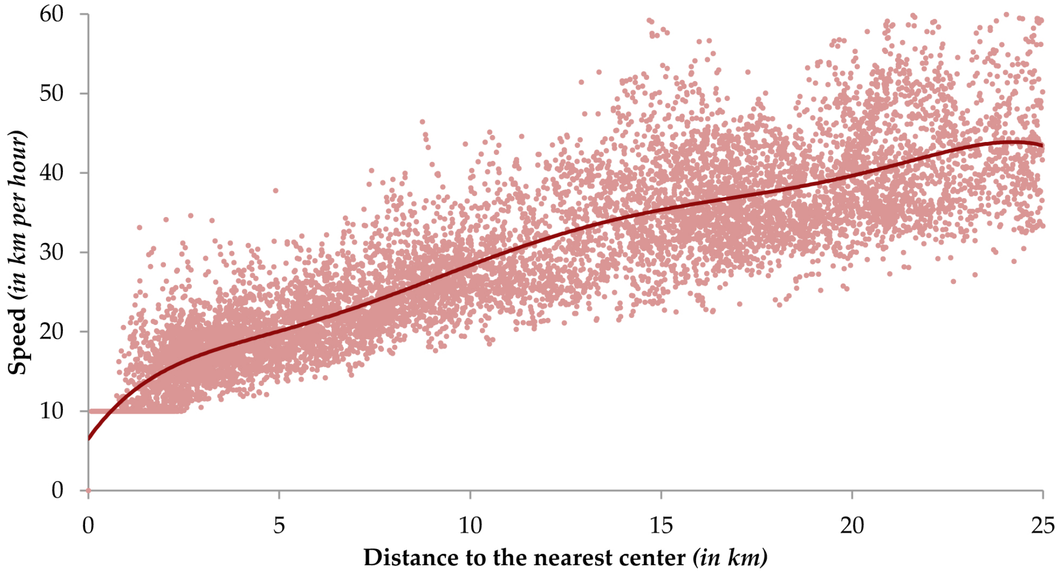
(B) Commuting flows in the Randstad

Figure A.2 – Commuting and accessibility

Notes: Panel A reports Kernel-smoothed values of employment accessibility.

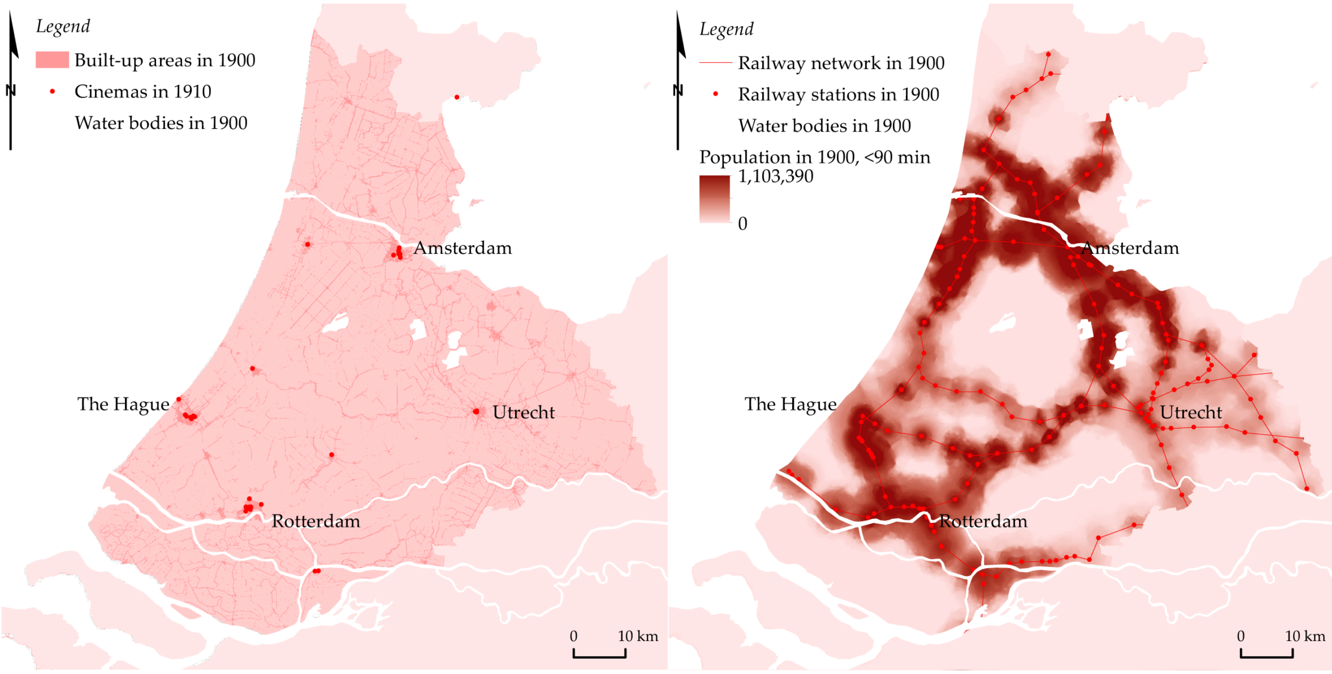


(A) Distance and travel time



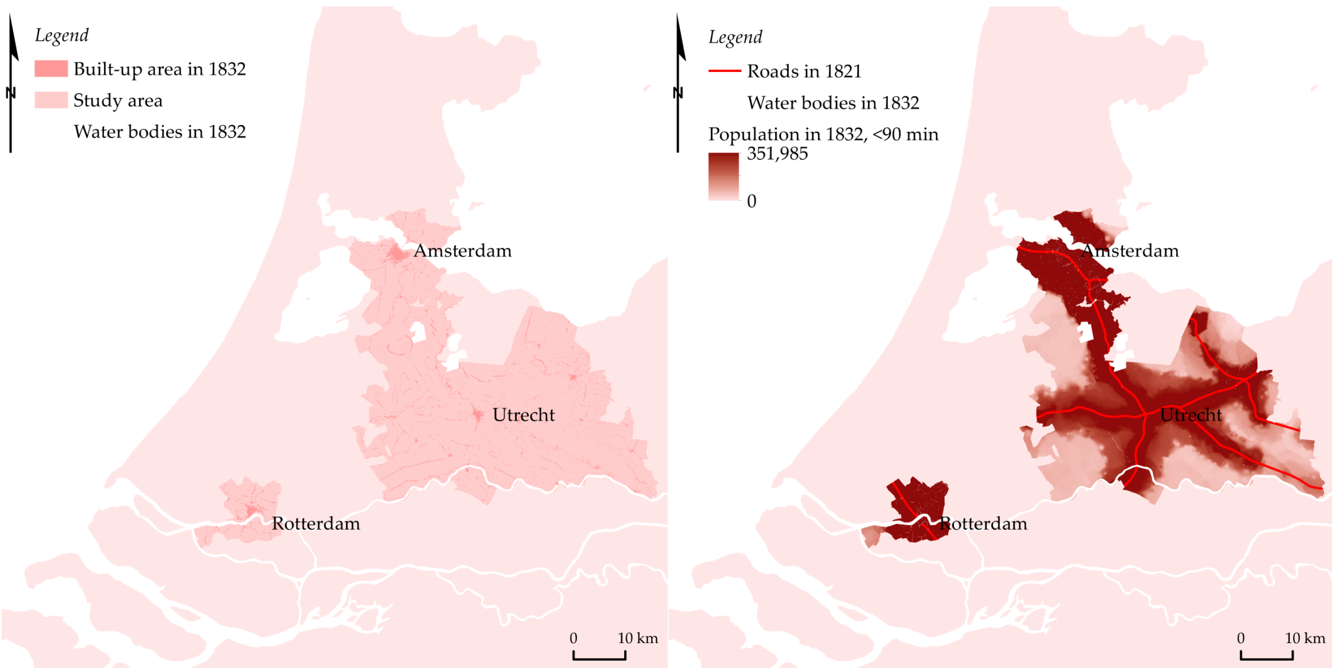
(B) Distance and speed

Figure C.1 – Calculation of travel time and speed



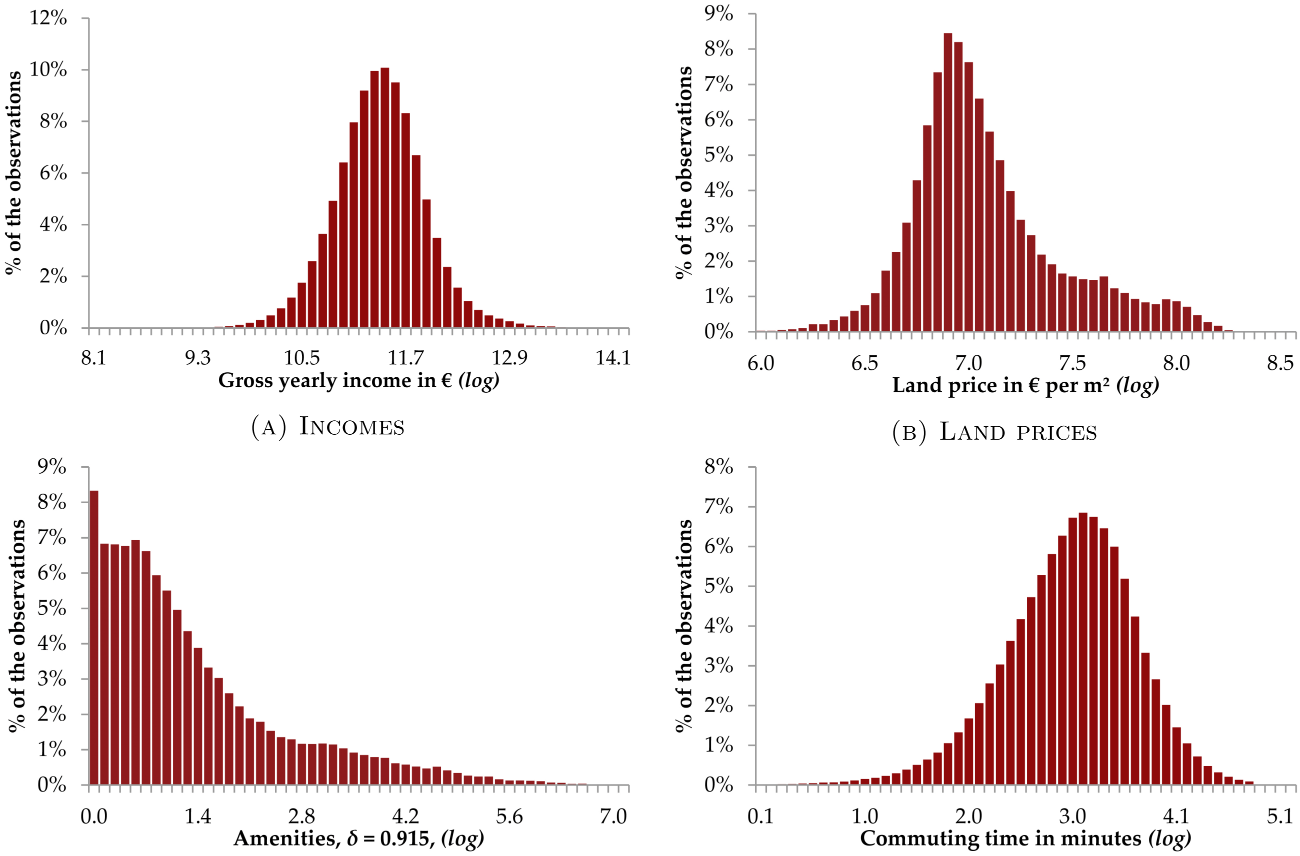
(A) Built-up land (B) The railway network and accessibility

Figure C.2 – Historic data from 1900



(A) Built-up land (B) The road network and accessibility

Figure C.3 – Historic data from 1832



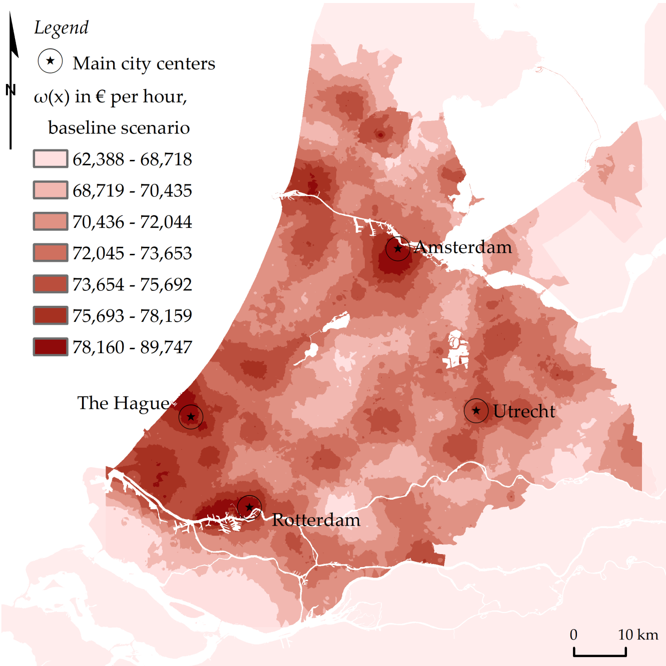
(A) Incomes

(B) Land prices

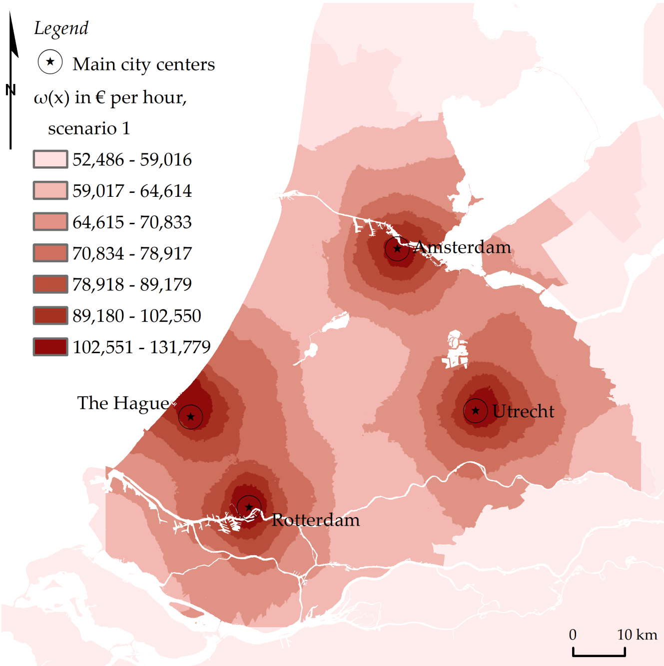
(C) Amenities

(D) Commuting times

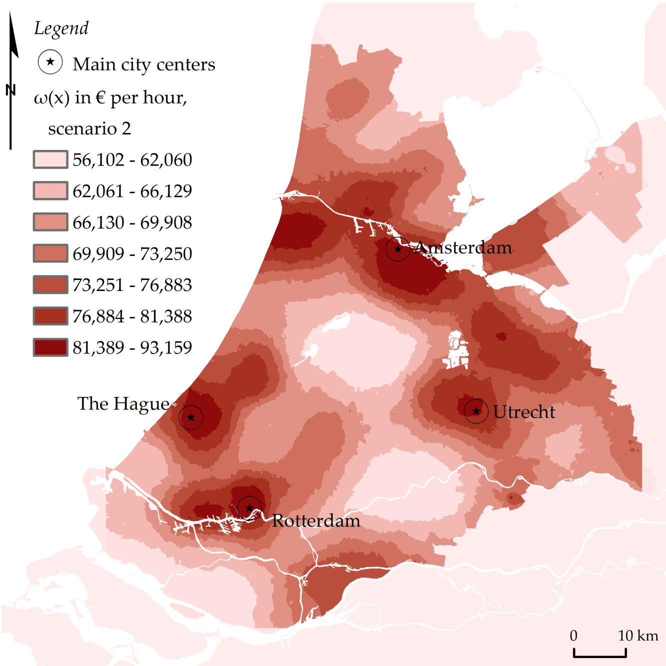
Figure C.4 – Histograms for the variables of interest



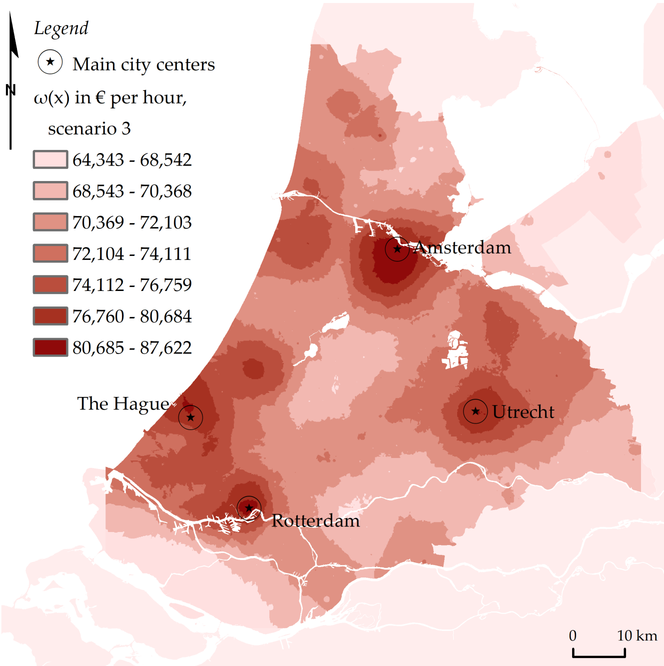
(A) Baseline scenario



(B) Scenario 1 – monocentric cities



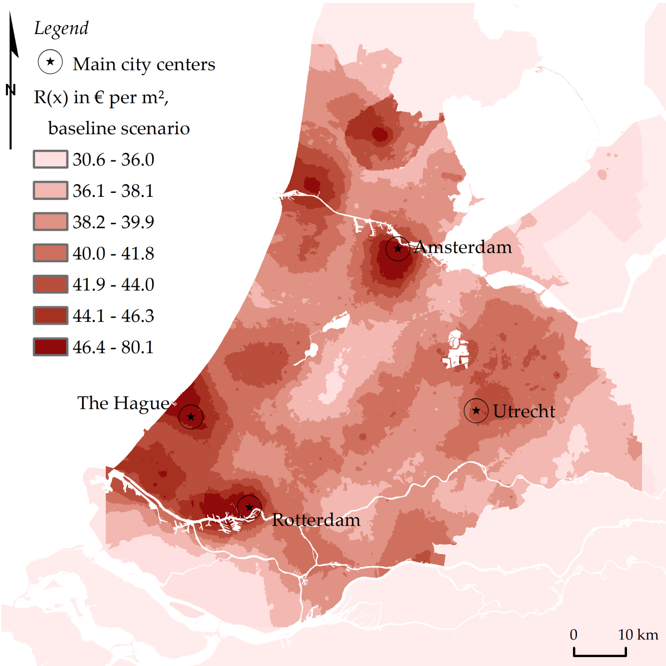
(C) Scenario 2 – no cars



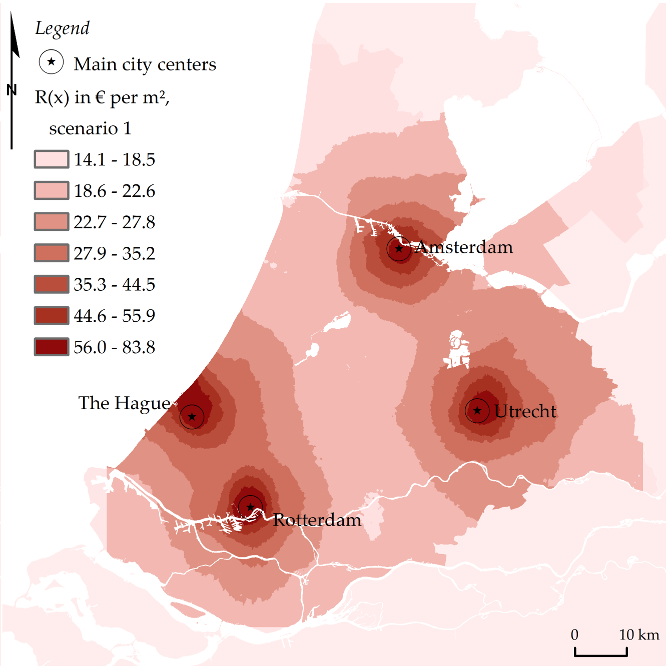
(D) Scenario 3 – telecommuting

Figure C.5 – Predicted income, $\omega(x)$

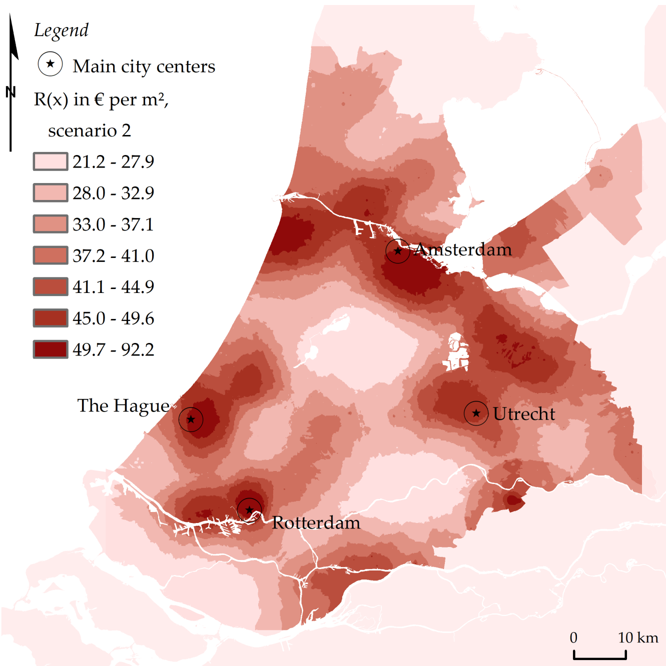
Notes: The above figures report Kernel-smoothed values of the variables of interest.



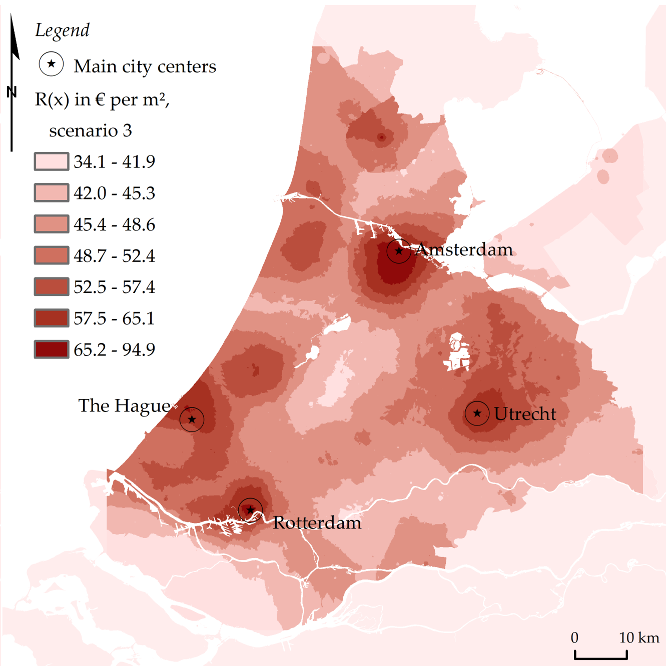
(A) Baseline scenario



(B) Scenario 1 – monocentric cities



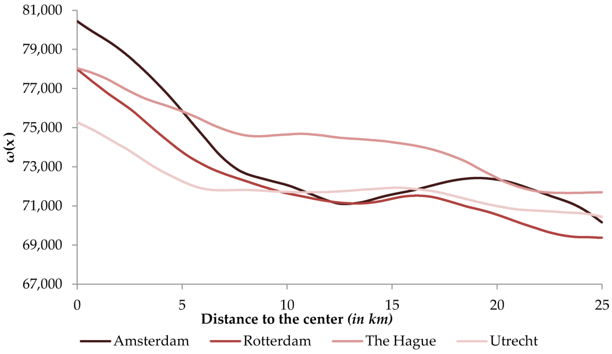
(C) Scenario 2 – no cars



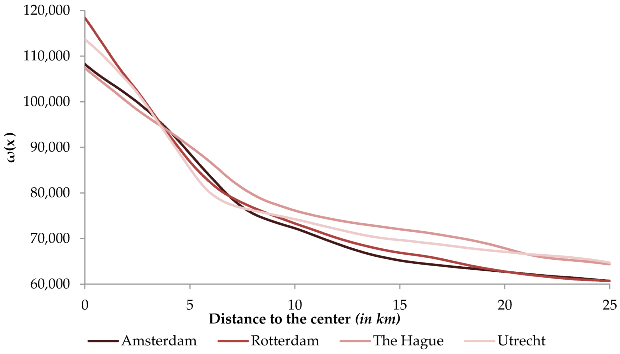
(D) Scenario 3 – telecommuting

Figure C.6 – Predicted land rents, $R(x)$

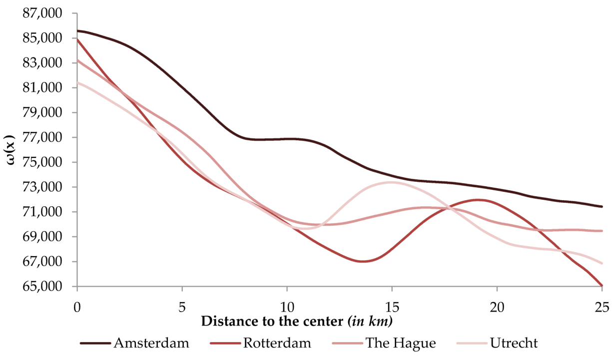
Notes: The above figures report Kernel-smoothed values of the variables of interest.



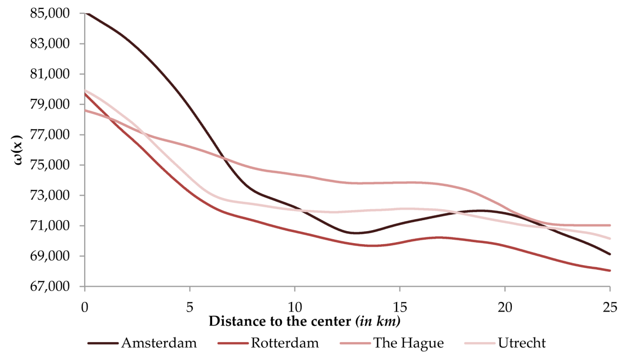
(A) Baseline scenario



(B) Scenario 1 – monocentric cities



(C) Scenario 2 – no cars



(D) Scenario 3 – telecommuting

Figure C.7 – Predicted income gradients for the Randstad