Who Lives Where in Cities?
Amenities, Commuting and Income Sorting

Carl Gaigné, Hans R.A. Koster, Fabien Moizeau, Jacques-François Thisse

June 4, 2019

Abstract

We study the sorting of skill-heterogeneous consumers within and between cities. We allow for non-homothetic preferences and locations that are differentiated by their accessibility to exogenous amenities and distance to employment centers, where production is subject to spatial externalities. The residential equilibrium is driven by the properties of an amenity-commuting aggregator obtained from the primitives of the model. Using the model’s structure and estimated parameters based on micro-data on the Netherlands, we predict that exogenous amenities are a key driver of social sorting. In the absence of amenities, the GDP increases by 10% because commutes are shorter. However, income segregation rises.

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


*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. Proost, 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), HSE University (Russian Federation), Tinbergen Institute, and CEPR.
§CREM (Condorcet Center), Université de Rennes 1.
¶CORE-UCLouvain (Belgium), HSE University (Russian Federation) and CEPR.
1 Introduction

The purpose of this paper is to study the sorting of households endowed with different skills across space and to determine the corresponding aggregate output. Our approach combines the microspatial – the choice of a residential location in a city – and the macrospatial – the choice of a city – in a single framework. We tackle the problem from two complementary perspectives. First, we develop a new urban economics model that characterizes the residential choices made by heterogeneous households between and within cities. This is accomplished by considering a general equilibrium setting that accounts for the interplay between amenities, commuting, agglomeration economies, and housing consumption. Second, we estimate the model by using various datasets from the Netherlands and undertake counterfactuals that highlight the impact of amenities and commuting on income sorting and aggregate output.

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). Equally important, Moretti (2012) has documented the fact that a small number of cities attract a disproportionate share of high-skilled workers. This in turn fosters divergence between regions because the skilled are those who benefit most from agglomeration economies and who enjoy most amenities (Bacolod et al., 2009; Faggio et al., 2017; Diamond, 2016; Couture et al., 2019). Since spatial sorting seems to threaten social cohesion as a whole (Chetty et al., 2016; Chetty and Hendren, 2018), we find it important to study the various forces that underpin the sorting of skill-heterogeneous households within and between cities. 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 because the canonical monocentric city model leads to a fairly extreme prediction: households are sorted by increasing income order as the distance to the central business district rises (Fujita, 1989). 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; Couture et al., 2019). One missing key explanation, at least for cities that have a long history, is the existence of exogenous amenities, such as scenic landscapes, river and sea proximity, historic buildings, and architecture. That such amenities matter in residential choices has been empirically well documented (Brueckner et al., 1999; Glaeser et al., 2001; Koster et al., 2016; Koster and Rouwendal, 2017; Lee and Lin, 2018). In addition, there are substantial differences in total factor productivity among employment centers (Hornbeck and Moretti, 2019). It is well known that residential and employment locations are the outcome of the interplay between several forces: amenities, agglomeration economies, commuting, and housing consumption. What we do not know is how the interaction between these forces determines the social stratification of the urban system. This paper attempts to fulfill this gap by proposing a new approach in which cities are “featureful” in that locations are distinguished...
by the distance to employment centers, which have a specific total factor productivity, and the accessibility to given amenities. Like most many recent contributions in urban economics, we treat the skill distribution as a given (see, e.g., Behrens et al., 2014; Davis and Dingel, 2019).

What are our main findings? Using the bid-rent approach to deal with the assignment of heterogeneous households to a continuum of locations, we can pin down the properties of the equilibrium skill mapping. We then show within a general framework involving several employment centers that the interaction between amenities, commuting and housing consumption gives rise to turning points in the endogenous spatial income distribution. Regardless of the nature of the functional form of the amenity, commuting cost, and skill distributions, households’ spatial sorting is imperfect as a greater geographical distance between two households no longer implies a wider income gap. To investigate further how amenities, commuting and housing consumption interact to shape the space-economy and to test empirically our conclusions, we need a full characterization of the equilibrium skill and income mappings defined over the location set. Homothetic preferences must be ruled out because they are associated with multiplicity of equilibria while failing to capture the fact that housing expenditure shares decline with income. Under Stone-Geary preferences, the spatial equilibrium can be predicted from the spatial distribution of a location-quality index, which has the nature of an amenity-commuting aggregator. This index is built endogenously from the primitives of the model. We further show the existence and uniqueness of a spatial 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, the location-quality index is defined over a set that includes locations belonging to different cities. Consequently, we account for the fact that the spatial structure of amenities and employment centers within and between cities is critical in choosing a residential place.

The properties of the location-quality index may be used to test the predictions about how amenities and employment centers affect the income residential pattern. The upshot is that the bliss point is the global maximizer of the location-quality index, thus implying that this location is occupied by the affluent because they have the highest bid. As one moves away from this location along all admissible directions, households are sorted by decreasing incomes until a local minimizer of the location-quality index is reached where low-income households, but not necessarily the poorest, are located. Around this minimizer, household income starts rising. As a result, households get more exposure to other income-groups when the number of turning points of the location-quality index rises. Yet, the tyranny of the bid rent remains implacable: the affluent still occupy the best locations. In addition, there is imperfect spatial sorting because of the spatial splitting of households belonging to the same income class.

To make the model amenable for use with real-life data, we discuss two main extensions. First, we allow for cross-commuting by introducing shocks on commuting. We furthermore introduce agglomeration economies. That is, we allow workers’ productivity to be dependent on the density of jobs in their vicinity, which is in line with the large empirical literature that shows that ag-
Glomeration economies are key (Combes and Gobillon, 2015). A major advantage of our model is that it is parsimonious and tractable as the main parameters of the model can be estimated recursively. To be precise, we first estimate a commuting gravity equation, then estimate the skill and income mappings. This provides sufficient information to estimate the elasticity of agglomeration economies.

We estimate our model using data from the Netherlands. We use rich microdata for more than 10 million households covering the years 2010 to 2015 on incomes, residential and job locations at the household level, employment accessibility, as well as land values and amenities at each location. The Netherlands is one of the countries with the highest population densities in the world (if we disregard city states). It is also one of the richest, with a GDP per capita higher than the UK, Germany and Japan. Dutch cities, which were established long ago, are well known for providing different types of high-quality amenities. Despite being a small country, the Netherlands hosts no less than 8 UNESCO world heritage sites, which is almost as much as London and Paris together, while it hosts 61,908 listed buildings, which is more than three times the number of listed buildings in Greater London. Note also that Amsterdam, the capital, is the city that receives the second highest number of tourists per capita (after Dubai) in the world. Last, the basic public services that underpin social cohesion (e.g., education) 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. This amenity index captures both the heterogeneity in aesthetic quality of buildings and residents' perceived quality of a certain location. This allows us to move beyond the approach of defining amenities implicitly, as in Ahlfeldt et al. (2015) and Albouy (2016). We further show the robustness of our results by using alternative proxies for amenities based on Lee and Lin (2018), who use variation in housing prices, and an amenity index based on the augmented reality game Pokémon Go. Admittedly, households also 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. Employment accessibility is also likely to be endogenous due to correlation with unobservable household characteristics and agglomeration economies – the latter being more prevalent in dense areas where commutes are shorter.

We address the endogenous nature of amenities and accessibility in our econometric analysis in several ways. 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. Second, using Oster’s (2019) approach, we use coefficient movements together with changes in the $R^2$ after inclusion of controls to investigate whether omitted variable bias is important. It appears that bias-corrected estimates are very close to the standard OLS estimates, suggesting
that omitted variable bias is not a major issue.

Since Oster’s methodology neither addresses measurement error in the proxy for amenities, nor addresses reverse causality, we further instrument amenities and accessibility with proxies for observed amenities, such as distance to historic districts, and a measure of employment accessibility. In addition, we digitize historic maps from 1832 and 1900 and use the historic land use, amenity and population patterns to instrument for current amenities and employment accessibility. 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 also consider an alternative identification strategy using large-scale land reclamations from the sea between 1930 and 1970. We calculate the share of planned built-up areas and green space as instruments for current amenities and re-calculate the employment accessibility based on the projected population distribution. As these reclaimed locations are otherwise identical, and as no one was living in those locations at that time, this alternative identification strategy addresses the main endogeneity concerns.

We first report results from reduced-form regressions. 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 2.5% higher, while when employment accessibility doubles, households’ incomes increase by 4%. Hence, the impacts of amenities and accessibility seem to be of a similar order of magnitude. We then estimate structural parameters of the model, which enables us to undertake counterfactual experiments. For example, in the absence of exogenous amenities, which mimics many U.S. cities, consumers focus only on commuting. As a result, commutes are shorter and the overall output increases by 10.6%. Land rents, on the other hand, decrease. However, the spatial distribution of skills (and hence of household income) changes considerably with the rich living in the most accessible locations. Alternatively, when commuting costs are halved, households’ labor supply is higher. In this case, the aggregate real income rises by 7.3%. However, the spatial distribution of skills is hardly affected. Other counterfactuals confirm that, while amenities are a main determinant of sorting in and between cities, changes in commuting costs do have less pronounced implications for sorting.

Related literature. Suggesting the complexity of the issue, only a handful of papers in urban economics have studied the social stratification of cities with heterogeneous households. Beckmann (1969) was the first attempt to take into account a continuum of heterogeneous households in the monocentric city. Unfortunately, the assignment approach used by Beckmann was flawed (Montesano, 1972). Fujita (1989) proposed a rigorous analysis of the residential pattern for a finite number of income classes. Kamecke (1993) extended this result to a continuum of heterogeneous households by showing that there is perfect sorting moving out from the central business district by increasing incomes. More generally, recent surveys, such as Duranton and Puga (2015) and Behrens and Robert-Nicoud (2015), highlight the various difficulties associated with the spatial
assignment of heterogeneous agents and express some skepticism about the ability of the bid-rent
approach to deal with heterogeneous households and a continuum of locations.

Using Roback-like spatial equilibrium models, Moretti (2011) and Diamond (2016) studied
how local wages, urban costs and employment respond to local labor shocks. However, these
papers use reduced forms to capture the costs generated by a larger population. To be precise,
these authors focus on workers’ locational choices between cities but disregard workers’ residential
choices within cities. In a dynamic setting, Lee and Lin (2018) 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. We differ from them in at least one fundamental aspect: in their setting people are
assumed to work where they live. In our setting, households are free to choose where to live and
where to work, while accounting explicitly for commuting costs between the residence and the
workplace that may belong to different municipalities. By contrast, Lee and Lin ignore the issue
of commuting, which amounts implicitly to assuming that households live and work in the same
municipalities.

In an important paper, Ahlfeldt et al. (2015) highlight the role of amenities, agglomeration
economies and commuting in residential location choices in their study of the internal structure
of Berlin. We use a similar strategy to model commuting costs (i.e., by a gravity equation)
and identify parameters of the model (i.e., by exploiting the recursive structure of the model).
However, we differ from them in the following two fundamental aspects. First, these authors
assume 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 model in which the utility
level is endogenous and varies across households heterogeneous in skill. Second, Ahlfeldt et al.
consider a finite location space and equalize the supply and demand for land at each location to
determine the land rent. Our approach is more in line with mainstream urban economics since
we assume a continuum of locations and the bid rent approach, which allow us to determine the
main properties of the spatial equilibrium.

Ahlfeldt et al. 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 resid-
uals’, meaning that it is unclear what these amenities actually capture (e.g., they may capture
housing characteristics or sorting on unobserved household characteristics). In our paper, we
strive to show that amenities and employment accessibility 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 spatial
equilibrium. Last, but not least, apart from the existence and uniqueness of the spatial equilib-
rium, Ahlfeldt et al. do not provide any characterization of this equilibrium. By contrast, we
show that even in the presence of agglomeration economies residential choices are guided by the
location-quality index that allows us to pin down households’ location choices.
Couture et al. (2019) undertake a rich analysis of the impact of income inequality on the internal structure of American cities with different types of amenities. However, by assuming a fixed-lot size, these authors cannot capture the well-documented fact that housing is a normal good whose consumption rises with income. We depart from them by assuming a variable lot-size, which allows us to determine the spatial income mapping in a finer way. Furthermore, Couture et al. work with a given income distribution, while our primitive is the skill distribution.

Last, note 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. For example, 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. Importantly, 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.

The remainder of the paper is organized as follows. We provide a detailed description of our model in Section 2 and show how the bid rent function may be used to determine the social stratification within and between cities. Section 3 determines the equilibrium skill mapping. 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 first determine analytically the market outcome when skills and the location-quality index are Fréchet-distributed and, then, extend our setting by taking into account household heterogeneity in commuting and agglomeration economies. Data are discussed in Section 6. In Section 7, we outline the procedure to identify the model’s parameters, while we present the results of counterfactual analyses in Section 8. In Section 9, we summarize our main results and discuss two possible extensions of our basic setup.

2 The model and preliminary results

2.1 The urban economy

Consider a network defined by a finite set of vertices and a finite set of topological arcs. A topological arc is the image of a compact interval by a continuous one-to-one relationship into the plane. It thus contains a continuum of locations. These arcs connect different pairs of vertices that correspond to locations endowed with some degree of centrality in the transportation system. The amount of land available at each location of the network is 1, but the population density is endogenous. For expositional simplicity, we focus in Sections 2-4 on a single topological arc, which can be represented by a compact interval $X \subset \mathbb{R}_+$. As usual, location and distance from the origin are identically denoted by $x \in X$. However, we show in Section 5 that our results can be extended to the case of a network formed by streets and highways, as well as by subway and
train links, that cross and connect cities. Formally, this network may be viewed as the union of several urban areas. Therefore, our results hold true for location set that are much more general than linear spaces or graphs. Examples of networks similar to ours include Allen and Arkolakis (2014) and Redding (2016).

The economy involves two normal consumption goods: (i) land \( (h) \), which is a proxy for housing, and (ii) a homogeneous final consumption good \( (q) \). Each location \( x \) is endowed with one unit of land and the opportunity cost of land is given by the constant \( R_0 \geq 0 \). Shipping the final good within the city is costless. Therefore, its price is the same across city locations. This good is used as the numéraire.

2.2 Production

There is a finite but large number \( n \) of given employment locations \( i = 1, \ldots n \). The employment density over \( \{1, ..., n\} \) is endogenous. The \( i \)-th employment location is characterized by a total factor productivity \( A_i > 0 \). In line with standard urban economics, we assume that employment centers are dimensionless.

The economy involves a unit mass of skill-heterogeneous households. A household is characterized by her skill \( s \in \mathbb{R}_+ \) and is endowed with one unit of \( s \)-labor. The skill c.d.f. \( F(s) \) is continuously differentiable on \( \mathbb{R}_+ \) and its density is denoted \( f(s) \). Labor is used in employment locations by the sector producing the numéraire. A household residing at \( x \) and commuting to \( i \) has \( 0 < \ell_i(x) \equiv L(|x - i|) \leq 1 \) working units, where \( L(\cdot) \) is continuously differentiable, decreasing and such that \( L(0) = 1 \). If a \( s \)-household resides at \( x \) and is employed at \( i \), her effective skill is given by \( A_i \ell_i(x)s \). The effective skills are thus vertically differentiated by \( s \) and horizontally differentiated by their residence and workplace through the shifter \( A_i \ell_i(x) \). The \( s \)-households may be distributed over several locations; we denote by \( \zeta(x, s) \in [0, 1] \) the share of \( s \)-households who reside at \( x \).

The final sector operates under constant returns and perfect competition. The production function is given by

\[
Y = \left[ \sum_{i=1}^{n} \int_{0}^{\infty} \int_{\mathcal{Y}_i} [A_i \ell_i(x)s]^{(\sigma-1)/\sigma} \zeta(x, s) f(s) dx ds \right]^{\sigma/(\sigma-1)},
\]

where \( \mathcal{Y}_i \) is the set of locations hosting individuals working at \( i \) and \( \sigma > 1 \) the elasticity of substitution across skills. For simplicity in exposition, we rule out agglomeration economies at \( i \) by assuming that \( A_i \) is a constant. We will show in Section 5 that our main results remain valid when we account for such effects.

Since the final sector is competitive, the income earned by a \( s \)-household residing at \( x \) and working at \( i \) is given by

\[
A_i \ell_i(x)s \frac{\partial Y}{\partial [A_i \ell_i(x)s]} = [A_i \ell_i(x)s]^{(\sigma-1)/\sigma} Y^{1/\sigma} \equiv \omega(s)t_i(x),
\]
where $\omega(s) \equiv s^{(\sigma-1)/\sigma} Y^{1/\sigma}$ is a $s$-household’s skill-specific component of the household’s income, which also depends on the overall productivity $Y$ of the city, while $t_i(x) \equiv [A_i \ell_i(x)]^{(\sigma-1)/\sigma}$ is the commuting component of the household’s income. For any given $Y$, the variables $s$ and $\omega(s)$ vary together, while $\omega(s)$ is independent of $x$ and $t_i(x)$ independent of $s$. A $s$-household bears a commuting cost equal to $[A_i^{(\sigma-1)/\sigma} - t_i(x)]\omega(s)$. This formulation is consistent with the empirical literature, which shows that such costs increase with the (skill-specific) income (Börjesson et al., 2012). Note that commuting costs are not proportional to $\omega(s)$ because the coefficient $A_i^{(\sigma-1)/\sigma} - t_i(x)$ varies with $x$.

Setting
\[ t(x) \equiv \max_{i=1,\ldots,n} t_i(x) \equiv t^*(x), \] (3)

it follows from (2) that a household established at $x$ trades her commuting time to $i$ against the total factor productivity of $i$ and chooses the place $i^*(x)$ that maximizes $t_i(x)$. Therefore, $t(x)$ is not monotone in $x$. Furthermore, (3) implies that $\Upsilon_i^* = \{x; i^*(x) = i\}$. A household is indifferent between $i$ and $j$ at any solution to $t_i(x) = t_j(x)$. Let $I$ be the location set such that $t_i(x) = t_j(x) = t(x)$ for $i, j = 1, \ldots, n$ and $i \neq j$. Without much loss of generality, we assume that the set $I$ is finite.

The following remarks are in order. First, the $s$-households located at $x$ earn the income $\omega(s)t(x)$, which is strictly increasing in $s$. Second, it follows from (2) that a household’s income varies with the residence-workplace choices made by all households through the value of $Y$. In other words, individual incomes are endogenous and determined at the spatial equilibrium, which determines the whole distribution of skills over $\{1, \ldots, n\}$.

Second, the employment density over $\{1, \ldots, n\}$ is determined together with households’ residential choices. For any skill mapping $s(x)$ which specifies the skill level of the households residing at $x$, the employment density at $i$ is given by
\[ \mathbb{I}_i \equiv \int_{\Upsilon_i^*} \zeta(x, s(x)) f(s(x))dx. \]

Note that the skill composition of the workforce at $i$ varies with the employment location.

2.3 Consumption

Households share the same utility function. Since households prefer more amenities than less, we consider a preference structure similar to the one used in models of vertical product differentiation:
\[ U(q, h; b) = b \cdot u(q, h), \] (4)
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 (4) imply that the utility level associated with the consumption of a given bundle $(q, h)$ increases with the amenity level $b$, while the utility derived from consuming amenities rises with income 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.

Let $\eta(y) > 0$ be a given function whose value expresses the amenity level (or, equivalently, an aggregator of distinct amenities) available at $y \in X$, which are exogenous and intrinsic to a location. The corresponding utility level is negatively affected by the distance between the household and the place where these amenities are available. As a result, a household at $x$ ascribes the value $\varphi(|x-z|)\eta(z)$ to the amenity provided at $z \neq x$, which decreases with the distance $|x-z|$ between $x$ and $z$, with $\varphi(0) = 1$. In other words, $\varphi(\cdot) \geq 0$ has the nature of a distance-decay function. As shown by (4), households’ well-being at $x$ depends on the amenity field defined by the following expression:

$$b(x) \equiv \int_X \varphi(|x-z|)\eta(z)dz.$$  \hspace{1cm} (5)

We do not impose any functional restriction on $\eta(\cdot)$ and $\varphi(\cdot)$. In the featureless city of urban economics, $b(x)$ is constant across locations because both $\varphi$ and $\eta$ are constant across space. In this paper, $b(x)$ varies with $x$. However, $b(x)$ is independent of the household density at $x$.

A $s$-household’s residing at $x$ has the following budget constraint:

$$\omega(s)t(x) = q + R(x)h,$$  \hspace{1cm} (6)

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

Maximizing (4) with respect to $q$ and $h$ subject to (6) yields the numéraire demand

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

and the land demand $h^*(x, s) \equiv h(R(x), \omega(s)t(x))$, which is the unique solution to the equation:\footnote{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 $z$).}

$$u_h [\omega(s)t(x) - R(x)h^*, h^*] - R(x)u_q [\omega(s)t(x) - R(x)h^*, h^*] = 0,$$  \hspace{1cm} (7)

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

### 2.4 The spatial equilibrium with exogenous amenities

The study of skill-income sorting when locations are differentiated by more than one attribute brings about new and difficult issues. At first sight, the determination of a spatial 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 land consumption varies with both its

\[
\text{For any function } f(y, z), \text{ let } f_y (\text{resp., } f_{yz}) \text{ be the partial (cross-) derivative of } f \text{ with respect to } y (\text{resp., } y \text{ and } z).\]
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 land 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 skills and locations is imperfect, the rule $x(s)$ which assigns a particular skill $s$ to locations $x$ must be a correspondence. For example, for the same given land consumption, a household can be indifferent between living close to a high-productivity employment center while having a low level of amenities, or living far from employment centers while enjoying a high level of amenities. Therefore, apart from special cases, there is no one-to-one correspondence between the skill and location sets. Since $x(s)$ 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 skill.

Given an amenity function $b(x)$, a given mass of heterogeneous households choose where to live and where to work in the city, how much land and how much of the composite good to consume. Consider a skill mapping $s(x)$. The $s$-households may be distributed over several locations. The land market clearing condition holds if $s(x)$ satisfies the following condition:

$$|\zeta(x, s) f(s) h^*(x, s) ds| = dx. \quad (8)$$

In other words, the amount of land available between any $x$ and $x + dx > x$ and the area occupied by the households whose skill varies from $s$ to $s + |ds|$ are the same. Since $s(x)$ need not be monotone, the land market clearing condition is expressed in absolute value. The corresponding city limit $L$ is, therefore, given by the solution to the equation:

$$\int_0^L \zeta(x, s(x)) f(s(x)) h^*(x, s(x)) dx = L. \quad (9)$$

The spatial equilibrium is such that no household has an incentive to move and to choose another workplace, and markets clear. Formally, a spatial equilibrium is defined by the following vector:

$$(s^*(x), \zeta^*(x, s^*(x)), i^*(x), Y^*, R^*(x), h^*(x, s^*(x)), q^*(x, s^*(x)), Y_i^*)$$

such that

$$b(x) \cdot u[q^*(x, s^*(x)), h^*(x, s^*(x))] \geq b(y) \cdot u[q^*(y, s^*(x)), h^*(y, s^*(x))] \quad 0 \leq y \leq L^* \quad (10)$$

holds under the constraints (6), (8), and (9). If the inequality is strict in (10) for all $y \neq x$, then all $s^*(x)$-households are located at $x (\zeta^*(x, s^*(x)) = 1)$. Otherwise, there exist at least two locations $x$ and $y$ such that the $s^*(x)$-households are indifferent between the locations $x$ and $y$. Thus, we have $0 < \zeta^*(\cdot, s^*(x)) < 1$ 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.\(^2\) Last, we have \(Y^*_i = \{x; i^*(x) = i\}\). Solving for the spatial equilibrium, we may characterize the equilibrium skill mapping \(s^*(x)\) from the location set \([0, L]\) to \(R_+\) that specifies which skill is available at \(x\), while the equilibrium income mapping is given by \(\omega(s^*(x))t(x)\).

In our setting, heterogeneous households reach different equilibrium utility levels, while households sharing the same skill enjoy the same equilibrium utility level. This is to be contrasted with Ahlfeldt *et al.* (2015) who assume that households share the same expected utility level, which is the exogenous reservation utility that prevails in the rest of the economy. Furthermore, here all the households residing at \(x\) work in the same employment location regardless of their skill level. However, when \(x_1\) and \(x_2 > x_1\) are such that \(t_i(x_k) = t_j(x_k) = t(x_k)\) for \(k = 1, 2\) and \(t(x) = t_i(x)\) for \(x < x_1\) and \(x > x_2\), we have \(t(x) = t_j(x)\) for \(x_1 < x < x_2\).

### 3 The equilibrium skill mapping

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

\[
\Psi(x, \omega(s), U) \equiv \max_{q, h} \left\{ \frac{\omega(s)t(x) - q}{h} \right\}_{\text{s.t.} \ b(x) \cdot u(q, h) = U}
\]

\[
= \max_{h} \frac{\omega(s)t(x) - Q(h, U/b(x))}{h},
\tag{11}
\]

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

The bid rent \(\Psi(x, \omega(s), U)\) is such that a \(s\)-household is indifferent across locations because she reaches the same utility level \(U\), (7) implies that the Alonso-Muth condition for a \(s\)-household under a differentiated space becomes:

\[
h^*(x, s)R_x(x) - \omega(s)t_x(x) = \frac{b_z(x)}{b(x)} \frac{u(q^*(x, s), h^*(x, s))}{u(q^*(x, s), h^*(x, s))},
\]

which reduces to the standard equation when \(b_z(x) = 0\) (Fujita, 1989).

Since each household treats the utility level as given, applying the first-order condition to (11) yields the equation:

\[
Q(h, U/b(x)) - hQ_h(h, U/b(x)) - \omega(s)t(x) = 0
\tag{12}
\]

whose solution, denoted \(H(\omega(s)t(x), U/b(x))\), is the quantity of land consumed by a \(s\)-household at \(x\) if her bid rent is equal to the land rent; the solution \(H(\cdot)\) is called the *bid-max lot size*\(^2\)For simplicity, we assume that the parameters of the economy are such that there is no agricultural land within the city. By implication, \(R^*(x) > R_0\) for \(x < \mathcal{L}^*\) and \(R^*(x) = R_0\) for \(x \geq \mathcal{L}^*\).
The equation (12) 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 utility-maximizer, hence to each equilibrium.

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

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

This expression shows that a household’s bid rent at $x$ depends separately on both $b(x)$ and $t(x)$ while its land 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:

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

where $U^*$ denotes the maximum utility reached by the households who set up at $x$ at the spatial equilibrium.

Land at $x$ is allocated to the highest bidder whose skill $s^*(x)$ solves the utility-maximizing condition:

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

while the second-order condition implies $\partial^2 \Psi/\partial \omega^2 < 0$. So far, we have implicitly assumed that (13) has a single maximizer in $\omega$. However, (14) may have several solutions in $\omega$. 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 is natural to assume that there is a single highest bidder at $x$. Characterizing the functions $u(q, h)$ and $t(x)$ for which this condition holds would require technical developments which are beyond the scope of this paper. Furthermore, we will show that this assumption holds for the Stone-Geary preferences used in the next sections.

Tottally differentiating (14) with respect to $s$ yields:

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

Therefore, since $d\omega/ds > 0$, $\Psi_{\omega s}(x, \omega(s^*(x)), U^*)$ and $ds^*(x)/dx$ have the same sign.

Set

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

and

$$
\varepsilon_{U,\omega} \equiv \frac{\omega}{U^*} \quad \varepsilon_{H,\omega} \equiv \frac{\omega}{H} (H_{\omega} + H_t U^*_\omega) \quad \varepsilon_{u,\omega} \equiv \frac{\omega}{u} \frac{\partial u}{\partial \omega}.
$$

We are now equipped to characterize $s^*(x)$.

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

$$
\Psi_{\omega s}(x, \omega(s^*(x)), U^*) \equiv \frac{t(x)}{H} \left[ \left( 1 - \frac{\varepsilon_{H,\omega} + \varepsilon_{u,\omega}}{\varepsilon_{U,\omega}} \right) B(x) - (1 - \varepsilon_{H,\omega}) T(x) \right]
$$

(16)
is positive (negative) at this location.

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

The expression (16) shows that, for any given utility $u(q, h)$, the interaction between amenities, commuting, and land consumption determines the social stratification of the city through the behavior of the function $\Psi_{wx}$. Furthermore, inspecting (16) shows that the impact of the amenity and commuting cost functions on the sign of $\Psi_{wx}$ is hard to predict because it depends on the values of $B$ and $T$. It also depends on how the elasticity of the land demand varies with $\omega$. Last, the sign of $\Psi_{wx}$, hence the slope of the skill mapping, changes at any solution of the equation $\Psi_{wx} = 0$.

To illustrate, consider first the benchmark case of a monocentric and featureless city, that is, $B(x) = 0$ 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 (16) that the sign of $\Psi_{wx}$ 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 land declines as income rises, the income elasticity of land is smaller than 1 (Albouy et al., 2016). Therefore, when skill, hence $\omega(s)$, 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. This effect is not offset by the higher land consumption because the income elasticity of land is smaller than 1. By implication, households are sorted by decreasing order of skill as the distance to the CBD increases. In this case, since $t(x) > t(y)$ when $\omega(s^*(x)) > \omega(s^*(y))$, skill-specific income gaps reflect the spatial sorting of households arising at the spatial equilibrium.

Consider now the case of a featureful monocentric city ($B(x) \neq 0$). Owing to the existence of amenities, even when the bid rent functions are downward sloping, the equation $\Psi_{wx} = 0$ may have several solutions in $x$. In this case, there is imperfect sorting, that is, greater skill differences are not mapped into more spatial separation.

The following three cases may arise.

(i) Assume that $\Psi_{wx} > 0$ for all $x$. As $s$, hence $\omega(s)$, 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 skills. In other words, the more skilled the household is, the closer she is 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 $\Psi_{wx} < 0$ for all $x$, the bid rent curve becomes steeper as $\omega$ rises. Therefore, the bid rent curves associated with any two different skills intersect once. In this case, $x = 0$ is the most-preferred city location and individuals are sorted out by decreasing skills.

(iii) The most interesting case arises when $\Psi_{wx}$ changes its sign because the slope of the skill gradient changes. In this case, there is imperfect sorting: household skill rises over some range of
sites and falls over others. We develop this argument in more detail in Section 4.

When the utility \( u(q, h) \) is specified, the condition (16) may be used to determine how households are distributed according to the behavior of \( B(x) \) and \( T(x) \) by calculating those elasticities. Note that \( T(x) = 0 \) when commuting is not accounted for, like in most models of local public finance. In this case, the sign of \( \Psi_{wx} \) is determined by the sign of \( (\varepsilon_{u_i,\omega} - \varepsilon_{H,\omega} - \varepsilon_{u_q,\omega})B(x) \) only.

Finally, when the city is featureless but polycentric, the function \( T(x) \) displays several extrema because \( t(x) \) is not monotone. Hence, even when \( b(x) \) is constant across locations, the decentralization of jobs favors income mixing.

4 The urban social structure under Stone-Geary preferences

Proposition 1 shows that the equilibrium income mapping is a linear combination of \( B(x) \) 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 A.2 that \( \varepsilon_{U,\omega} = \varepsilon_{H,\omega} = 1 \) and \( \varepsilon_{u_q,\omega} = 0 \). As a result, (16) can be reduced to \( \Psi_{wx} = 0 \) for all \( x \). In other words, there is a continuum of residential equilibria under homothetic preferences. Therefore, if the aim is 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 (16) reduces to \( B(x) - T(x) \). This expression suggests that quasi-linear preferences are a good candidate to study the spatial equilibrium. Unfortunately, assuming quasi-linear preferences is counterfactual as land is a normal good \( (\varepsilon_{H,\omega} > 0) \).

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

\[
u(q, h) = q^{1-\mu}(h - \overline{h})^\mu, \quad (17)\]

where \( 0 < \mu < 1 \) and \( \overline{h} > 0 \) is the minimum lot size, which is supposed to be sufficiently low for the equilibrium consumption of the numéraire to be positive. Maximizing (17) with respect to \( q \) and \( h \) subject to the budget constraint leads to the linear expenditure system:

\[
q^*(x, s) = (1 - \mu)\omega(s)t(x) - R(x)\overline{h}, \quad (18)
\]

\[
h^*(x, s) = (1 - \mu)\overline{h} + \mu\frac{\omega(s)t(x)}{R(x)}. \quad (19)
\]
The land 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 land is constant over time and across U.S. metropolitan areas. However, this result does not mean that households having different incomes spend the same share of their incomes on land within a given city.

We show in Appendix A.3 that

$$\Psi_{\omega x} = \frac{t(x)H}{H^2} [B(x) - (1 - \mu)T(x)], \quad (20)$$

which depends on the intensity of preferences for land through the parameter $\mu$. Set

$$\Delta(x) \equiv b(x)[t(x)]^{1-\mu}, \quad (21)$$

which subsumes the amenity and commuting levels 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 the income $\omega^*(s)$. The higher $\mu$, the stronger the preference for land. Therefore, as the intensity of preference for land increases, commuting matters less than the accessibility to amenities. Moreover, differentiating (21) shows that $\Delta_x(x)$ and $\Psi_{\omega x}$ have the same sign. Hence, $\Psi_{\omega x}$ changes sign at any extrema of the location-quality index.

Finally, consider the following utility:

$$u(q, h) = q^{\rho_1} + h^{\rho_2} \quad (22)$$

with $0 < \rho_i < 1$ and $\rho_1 \neq \rho_2$. The elasticity of substitution between land and the numéraire is variable and equal to $1/(1 - \delta_1 \rho_1 - \delta_2 \rho_2)$ where $\delta_i$ is the expenditure share on good $i = 1, 2$.\footnote{An expression similar to (22) is used by Eeckhout et al. (2014) as a production function.} When $\rho_1 > \rho_2$, i.e., the composite good matters more than land, it can be shown that the above preferences generate the index $\Delta(x) \equiv t(x)[b(x)]^{1/\rho_1}$, which is similar to (21).

Intuitively, (17) may be interpreted as a non-homothetic Cobb-Douglas utility and (22) 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 $b(x)$ and $t(x)$ and is independent of income.

### 4.1 The skill mapping

Our objective is now to determine the equilibrium income mapping that specifies which $\omega$-households are located at $x$ under Stone-Geary preferences. Since land 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 spatial equilibrium is unique. Furthermore, the equilibrium skill mapping \( s^*(x) \) and the location-quality index \( \Delta(x) \) vary together with \( x \).

**Proof.** The argument involves four steps.

(i) We show in Appendix A.3 that the bid-max lot size is unique and such that

\[
H(\omega(s)t(x), U/b(x)) \equiv H(\Delta(x), \omega(s), U),
\]

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

(ii) Furthermore, as shown in Appendix A.4, the equilibrium condition \( \Psi_\omega = 0 \) is equivalent to the following differential equation:

\[
\frac{dU^*}{ds} = \frac{dU^*}{d\omega} \frac{d\omega}{ds} = \left[ \Delta^{1-\rho}(1-\mu)(H-H)^{\frac{\mu}{1-\rho}}(U^*(\omega(s)))^{-\frac{\mu}{1-\rho}} \right] \frac{d\omega}{ds}.
\]

Since \( H \) depends on the location-quality index \( \Delta(x) \), the solution \( U^*(\omega(s)) \) to this differential equation also depends on \( b \) and \( t \) through the location-quality index \( \Delta \).

(iii) It follows from Appendix A.4 that \( U^*_\omega(\omega) \) is an increasing function of \( \Delta \):

\[
\frac{\partial}{\partial \Delta} \frac{dU^*}{ds} = \frac{\partial}{\partial \Delta} \frac{dU^*}{d\omega} \frac{d\omega}{ds} > 0.
\]

The Spence-Mirrlees condition thus holds, which implies the existence of a positive assortative matching between skills and the values of the location-quality index. In other words, there is a unique one-to-one and increasing relationship between \( s \) and \( \Delta \) (Chiappori, 2017). Therefore, regardless of the value of \( Y > 0 \), households ordered by increasing skills are always assigned to locations endowed with rising values of the location-quality index. To put it differently, the residential choices made by households are independent of the value of \( Y \). Since all households residing at \( x \) work at the same employment location \( i^*(x) \), plugging the corresponding workplace choices in (1) determines the unique equilibrium value \( Y^* \).

(iv) Conditional to \( Y^* \), the equilibrium skill mapping is unique and given by:

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

Since the c.d.f. of \( \omega(s) = (s^{\sigma-1}Y^*)^{1/\sigma} \) is

\[
\Lambda(\omega) \equiv F \left[ \left( \frac{\omega}{Y^*} \right)^{1/(\sigma-1)} \right]
\]

whose density is given by

\[
\lambda(\omega) = \frac{\sigma}{\sigma-1} \left( \frac{\omega}{Y^*} \right)^{1/(\sigma-1)} f \left[ \left( \frac{\omega}{Y^*} \right)^{1/(\sigma-1)} \right],
\]

the equilibrium skill-specific income mapping is thus unique and given by:

\[
\omega^*(x) = \Lambda^{-1}[G(\Delta(x))].
\]
It then follows from (23) that the spatial equilibrium exists and is unique.

This proposition implies that it is sufficient to study how $\Delta(x)$ varies to determine the properties of the spatial 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 A.4, the slope of the bid rent function increases with $\Delta$ if and only if the marginal utility of income also increases with $\Delta$, that is, (25) is equivalent to

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

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

Since the function (21) is in general not monotonic in $x$, we have:

$$\frac{\partial}{\partial x} \frac{dU^*}{d\omega} \geq 0.$$

Therefore, income sorting is not mapped into spatial sorting: the income gradient is not a monotone function of the distance from employment locations. As a consequence, the affluent (or the poor) do not necessarily locate at $x = 0$ or $x = L^*$. Rather, the richest locate where the location-quality index is maximized, whereas the poorest reside in locations with the lowest location-quality index.$^4$

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

Since $\Delta(0) > \Delta(x_2) > \Delta(L^*) > \Delta(x_1)$, the intermediate value theorem implies that $z_1$ in $[0, x_1)$, $z_2$ in $(x_1, x_2)$ and $z_3$ in $(x_2, L^*)$ exist such that $\Delta(z_1) = \Delta(z_2) = \Delta(z_3)$. Proposition 2 implies that the households residing at these three locations have the same income. In other

$^4$As for the housing demand, we have $\partial H/\partial \Delta < 0$, for otherwise the utility level $U$ of the $\omega$-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 $dH/dx$ 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.
words, there is spatial splitting because the households sharing the income $\omega^*(z_i)t(z_i)$ do not live in the same neighborhood. On the contrary, they are spatially separated by households having lower incomes in $(z_1, z_2)$ and higher incomes in $(z_2, z_3)$. Roughly speaking, Figure 3 depicts a spatial configuration where the middle class is split into two spatially separated neighborhoods with the poor in between, while the affluent live near the city center. Such a pattern describes more accurately the spatial distribution of incomes in “old” US cities and in many European cities, than the homogeneous monocentric city model (Glaeser et al., 2008).

[Figure 3 about here]

4.2 Land rent

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

$$R^*(x) = \frac{\omega^*(x)t(x)}{H[\Delta(x), \omega^*(x)t(x), U^*(\omega^*(x))]} \left[ 1 - \frac{1 - \mu}{\varepsilon_{U,\omega}(x)} \right], \quad (28)$$

where

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

Substituting $\varepsilon_{U,\omega}(x)$ in (28) and rearranging terms, we obtain:

$$R^*(x) = \frac{\mu\omega^*(x)t(x)}{H[\Delta(x), \omega^*(x)t(x), U^*(\omega^*(x))] - (1 - \mu)h} > 0, \quad (29)$$

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

By totally differentiating (28) with respect to $x$, we obtain (see Appendix A.5):

$$R^*_x(x) = \frac{\omega^*(x)t(x)}{H[\Delta(x), \omega^*(x)t(x), U^*(\omega^*(x))]} \left[ \frac{1}{\varepsilon_{U,\omega}(x)}B(x) - T(x) \right]. \quad (30)$$

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

To estimate the model, we need an explicit form of the skill-specific mapping \( s(x) = F^{-1}[G(\Delta(x))] \). For this, we must specify the distributions \( F \) and \( G \). Moreover, we introduce shocks in commuting costs enabling us to allow for cross-commuting. Finally, we introduce agglomeration economies based on skill density. In other words, we allow for the fact that households are more productive if the endogenous supply of skills in their vicinity is higher.

5.1 The skill mapping

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 what follows, we assume that the variable \( s \) is drawn from a Fréchet distribution to the power \( (\sigma - 1)/\sigma \) with the shape parameter \( \gamma_s > 0 \) and the scale parameter \( K_s > 0 \):

\[
F(z) = \exp(-K_sz^{-\gamma_s(\sigma-1)/\sigma}) \text{ over } \mathbb{R}_+ \text{ with density } f(s) = K_s\gamma_s \frac{1}{\sigma-1} \left[\exp(-K_sz^{-\gamma_s(\sigma-1)/\sigma})\right]^{1/(\gamma_s(\sigma-1) + 1)/\sigma}.
\]

An increase in \( \gamma_s \) leads to less income inequality. It is analytically convenient to assume the values of \( \Delta \) are also drawn from a Fréchet distribution with the c.d.f. \( G(z) = \exp(-K_\Delta z^{-\gamma_\Delta}) \text{ over } \mathbb{R}_+ \) and density \( g(z) \). The location-quality index covers a wider range of values when \( \gamma_\Delta \) decreases.

Using (27), the mapping \( s^*(x) \) can then be retrieved from the condition:

\[
\int_s^\infty f(z)dz = 1 - \exp(-K_sz^{-\gamma_s(\sigma-1)/\sigma}) = \int_\Delta^\infty g(\zeta)d\zeta = 1 - \exp(-K_\Delta \Delta^{-\gamma_\Delta}),
\]

which is the counterpart in the \( \Delta \)-space of the land market clearing condition (8). It follows from Proposition 2 that households ranked by decreasing incomes are assigned to locations having a decreasing location-quality index.

Set \( \gamma \equiv \gamma_\Delta/\gamma_s \) and \( K \equiv K_s/K_\Delta \). Solving the above equation yields the equilibrium skill mapping \( s^*(x) \) defined over \([0, L^*] \):

\[
s^*(x) = \left\{K^{1/\gamma_s} [\Delta(x)]^{\gamma} \right\}^{(\sigma-1)/\sigma}.
\]

Assuming that \( \zeta(x, s) = 1 \) for any \( x \), we obtain the equilibrium city output:

\[
(Y^*)(\sigma-1)/\sigma = \sum_{i=1}^n \int_{Y^*_i} [A_i\ell_i(x)]^{(\sigma-1)/\sigma} K_i^{1/\gamma_i} [\Delta(x)]^{\gamma} f \left\{ \left( \frac{K_i}{K_\Delta} \right)^{1/\gamma_i} [\Delta(x)]^{\gamma} \right\}^{(\sigma-1)/\sigma} \text{ d}x.
\]

Since \( \omega = s^{(\sigma-1)/\sigma} Y^{1/\sigma} \), the equilibrium income mapping is thus given by

\[
t(x)\omega^*(x) = K^{1/\gamma_s} [\Delta(x)]^{\gamma} t(x)(Y^*)^{1/\sigma}.
\]
This expression shows that this distribution is a power of the location-quality index and of the city output.

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

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

where $k$ is a positive constant.

5.2 Incomplete information on commuting

A $s$-household is characterized by an intrinsic income $\omega(s)$ that depends on her skill $s$ and the city output $Y$, while her actual income also depends on her residential location $x$ and her workplace $i$. In the data, households living at the same place do earn the same actual income but do not necessarily work in the same place. In line with discrete choice theory, we consider statistically identical and independent households. A $s$-household’s actual income is random and given by $\omega(s)t_i(x)\nu_{k,xi}$, where the $\nu_{k,xi}$ are i.i.d. shocks on commuting costs which are specific to the individual $k$ and locations $x$ and $i$. These shocks capture the idea that households located at the same place have idiosyncratic reasons for working in different parts of a city.

The effect of uncertainty on location and consumption decisions depends on the timing of uncertainty resolution and on the flexibility that allows a household to revise her decision in response to information. We assume here the timing that endows households with the possibility to adjust their workplace and total consumption conditional upon their residential choices. Before observing their actual income, households choose their residential locations $x$ at the spatial equilibrium associated with the distribution of expected incomes. Since households anticipate they will choose the best workplace after the resolution of uncertainty, residential choices follow the distribution of the location-quality index $(x) = b(x) \{\mathbb{E}[\max_i t_i(x)\nu_i]\}^{1-\mu}$. This index is the same for all households because these ones are statistically identical and independent.

Under Stone-Geary preferences, we have seen that the utility maximization program of a household located at $x$ leads to the following demands for the numéraire:

$$q^*(x) = (1 - \mu) \left[ \omega(s)t_i(x)\nu_i - R(x)\bar{h} \right],$$

and for land:

$$h^*(x) = (1 - \mu)\bar{h} + \mu \frac{\omega(s)t_i(x)\nu_i}{R(x)}.$$

The corresponding indirect utility $V(R(x), \omega(s)t_i(x)\nu_i)$ is given by the expression:

$$V(R(x), \omega(s)t_i(x)\nu_i) = (1 - \mu)^{1-\mu} \mu^\mu b(x) \left[ \omega(s)t_i(x)\nu_i - R(x)\bar{h} \right] R(x)^{-\mu}.$$

The expected utility of a household at $x$ is as follows:

$$\mathbb{E}[V(R(x), \omega(s)t_i(x)\nu_i)] = \mathbb{E} \left\{ (1 - \mu)^{1-\mu} \mu^\mu b(x) \left[ \max_{i=1,...,n} \omega(s)t_i(x)\nu_i - R(x)\bar{h} \right] R(x)^{-\mu} \right\}.$$
When the $\nu_i$ are i.i.d. according to a Fréchet c.d.f. $I(z) = \exp(-K_iz^{-\varepsilon})$ with the shape parameter $\varepsilon > 1$ and the scale parameter $K_i$, the expected utility varies with the household’s expected income:

$$E \left[ \max_{i=1,\ldots,n} \omega(s)t_i(x)\nu_i \right] = \Gamma \left( \frac{\varepsilon - 1}{\varepsilon} \right) \omega(s) \left( \sum_{i=1}^{n} K_i t_i^\varepsilon(x) \right)^{1/\varepsilon},$$

where $\Gamma(\cdot)$ is the gamma function. Consequently, the argument developed in Section 4 holds true when the income is replaced by the expected income.

Since households anticipate they will choose the best workplace after the resolution of uncertainty, residential choices follow the distribution of the location-quality index $\Delta(x) = b(x) \{ E[\max_{i} t_i(x)\nu_i] \}^{1-\mu}$ as the right-hand side of (A.6.6) depends only on $\Delta(x)$ (see Appendix A.6). This distribution is the same for all consumers.

Once households are located, they are able to observe their actual incomes. The households then choose the workplaces that give them the highest incomes, as well as the corresponding consumption of land and numéraire. Since households are heterogeneous in commuting, those who choose the same residential location $x$ need not choose the same workplace, nor do they necessarily earn the same income and consume the same amount of land at the land rent $R(x)$. In particular, there is cross-commuting.

Since her residential location $x$ and the land rent $R(x)$ are given before the realizations, a household maximizes her indirect utility by maximizing her income. The probability that a household living at $x$ chooses to work at $i$ is given by the gravity equation:

$$\pi_{x|i} = \frac{K_i \left[ t_i(x) \right]^\varepsilon}{\sum_{j=1}^{n} K_j \left[ t_j(x) \right]^\varepsilon},$$

(35)

where the shape parameter $\varepsilon$ is an inverse measure of the dispersion of idiosyncratic tastes, which is assumed to be the same across employment locations, and $K_i$ is the scale parameter of the employment location $i$. As $\pi_{x|i} > 0$ for all $x \in X$, we have $\pi_{i} = [0, L_i^\delta]$ for $i = 1, \ldots, n$. Therefore, as in Ahlfeldt et al. (2015) and Diamond (2016), households residing at the same location work in different employment locations.

5.3 Agglomeration economies

In line with the literature on agglomeration economies (Duranton and Puga, 2004), we assume that the total factor productivity at $i$ is given by

$$A_i = A_i L_i^{\delta} \in [0, A_i], \quad i = 1, \ldots, n$$

(36)

where $A_i$ is a location-specific productivity shifter, $L_i$ the employment density at $i$, and $\delta > 0$ the elasticity of the agglomeration effect with respect to $L_i$ at any $i$.  

22
The equilibrium skill mapping (31) has been obtained conditional on a vector \( A = (A_1; \ldots; A_n) \). Since \( t(x) \) is the maximum of linear functions in \( A_i \), the function \( s^*(x; A) \) is continuous in \( A \) and piecewise continuously differentiable in each \( A_i \). Therefore, the function,

\[
F_i(A) \equiv A_i \left[ \int_0^{L^*} \frac{K_i[s(x; A)]^\varepsilon}{\sum_{j=1}^n K_j[t_j(x)]^\varepsilon} \, dx \right]^{\delta}, \quad i = 1, \ldots, n
\]

is also continuous in \( A \) and piecewise continuously differentiable in each \( A_i \).

Since \( F(A) \equiv (F_1(A), \ldots, F_n(A)) \) is a continuous mapping from \([0, A_1] \times \ldots \times [0, A_n]\) into \([0, A_1] \times \ldots \times [0, A_n]\), the Brouwer fixed point theorem implies that the system of equations (36) has at least one solution \( A^\ast \). Furthermore, \( A^\ast \) always belongs to \([0, A_1] \times \ldots \times [0, A_n]\) because the bracketed term in (37) is always positive and smaller than 1. Consequently, under Stone-Geary preferences, Fréchet distributions and the agglomeration economies (36), there exists a spatial equilibrium. We show in Appendix A.7 that this equilibrium is unique when \( \delta > 0 \) is small enough, that is, when the agglomeration effect in (36) is not too strong.\(^5\)

Plugging the corresponding values \( A_1^\ast, \ldots, A_n^\ast \) in (31), (33), and (36) yields the equilibrium skill and income mappings, as well as the equilibrium employment density in the presence of agglomeration economies.

To sum up, we have the following proposition.

**Proposition 3.** Under Stone-Geary preferences and Fréchet distributions, there exists a unique spatial equilibrium when the elasticity of (36) with respect to the employment density is not too large. Furthermore, the equilibrium income distribution is unique and behaves like the location-quality index.

### 5.4 The case of a general economic space

Since we do not make any specific assumption on the amenity and commuting time functions, our model is flexible enough to consider a general economic space such as that described in Section 2.1. Each location \( x \) belonging to the topological arc \( \ell \) is characterized by specific amenity and commuting time functions, \( b(x; \ell) \) and \( t(x; \ell) \). In this case, the location-quality index at \( x \in \ell \) is given by \( \Delta(x; \ell) = b(x; \ell)[t(x; \ell)]^{1-\mu} \). Households then choose the arc \( \ell \) and the location \( x \in \ell \) that maximize their utilities. It is straightforward that Propositions 2 and 3 hold true when the economy is formed by a network of cities.

\(^5\)We show in Section 8.2 that uniqueness holds for realistic values of \( \delta \).
6 Data

6.1 Datasets

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

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. 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. 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 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 B.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 B.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. We also know whether the house is a listed building. 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 on income sorting and land prices. We proxy the amenity level by the picture density in a neighborhood. More specifically, we gather data from Eric Fisher’s Geotagger’s World Atlas, which contain all geocoded pictures on the website Flickr. The idea is that locations with an abundant supply of aesthetic amenities will have a high picture density. We show in Appendix B.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.

First, to avoid the possibility of inaccurate geocoding, we keep only one geocoded picture per location defined by its geographical coordinates. This reduces the number of pictures by about 50%. Second, one may argue that the patterns of pictures taken by tourists and residents

---

6Alternatively, 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 leading to nearly identical results.

7We 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.

8Ahlfeldt (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.

9In 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.
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 0.223. 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.

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 B.3 for more details). The hedonic index aggregates the average impact of several proxies of amenities, such as the locations of historic buildings, proximity to open space and water bodies, by testing their joint impact on house prices. We also construct historic instruments. Knol et al. (2004) have scanned and digitized maps of land use in 1900 into 50 by 50 meter grids and classified each grid into 10 categories, including built-up areas, water, sand, and forest. We aggregate these 10 categories into 3 categories: built-up areas, open space, and water bodies and calculate the share of the area used for each type in each neighborhood. We further gather data from the 1909 census on occupations and employment in each municipality. Those ones were much smaller than current ones and about 4 times the size of the current neighborhoods. For each occupation we obtain the required skill level. This enables us to calculate the share of households who are medium and high-skilled. We gather additional data on the railway network in 1900 and the stations which by then existed (see Appendix B.4 for more information), enabling us to calculate employment accessibility in 1909. 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. 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, Leiden, Delft, Hoorn, and the 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).
6.2 Descriptive statistics

We show a map of the Netherlands, the study area, in Figure 2.A. We indicate the most important cities. The conurbation of the four largest cities Amsterdam, Rotterdam, The Hague, and Utrecht is referred to as the Randstad, which has a population of about 7.1 million. Figure 2.B shows the commuting intensity between different neighborhoods in the Netherlands. What it shows is that the Dutch urban structure is really polycentric with many commuting flows occurring between different cities. This underlines the need for a model that allows for location choices in and between cities.

We report descriptive statistics of the 10,213,524 households of our sample in Table 1. The average (median) gross yearly income is €91,535 (€86,732). It appears that incomes are approximately Fréchet distributed (see Appendix B.5). The average land price in the sample is €1,312, but there are stark spatial differences. For example, in the capital Amsterdam, it is €3,046, while in the rural province of Friesland it is only €716.10 As expected, the correlation between the estimated land price and lot size is negative (the correlation \( \rho = -0.245 \)). The average lot size is 364m². However, in Amsterdam lot sizes are only 253m², which corresponds to the higher land values in this city. About 15% of households occupy apartments and the correlation between occupying an apartment and the land price is indeed positive (\( \rho = 0.153 \)).

We use the neighbourhood definition following Statistics Netherlands, implying that we have 4,033 neighborhoods. The picture density, i.e., the proxy for amenities, range from 0 to 231 pictures per hectare. Only 0.2% of the households are in neighborhoods that do not have any pictures. We will disregard those households. The average picture density in Amsterdam (22.7) is much higher than in Rotterdam (9.63), The Hague (6.17), and Utrecht (7.66). Recall that we only use pictures outside a building taken by residents in determining the amenity index. It appears that 80% of the pictures are taken outside a building and about 60% of the pictures are taken by local residents. Going back to Table 1, we see that the average commuting time is 26 minutes, which is very close to statistics provided by other sources (Department of Transport, Communications and Public Works, 2010). The unconditional correlation of picture density with the gross income is close to zero (\( \rho = 0.0533 \)), but this is not very informative yet. The correlation of the amenity index with land prices is substantially higher (\( \rho = 0.431 \)). 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 low (\( \rho = -0.0454 \)).

10We report maps and histograms of income and land prices in Appendix C.5.
The descriptive of the historic instruments that we use are described in Table B.6 of Appendix B.4.

7 Structural estimation

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. Note that from here onwards we use subscripts to indicate households $k$, as well as neighborhoods of living $x$ and working $i$. Let us assume that amenities are given by $b_{kx} = e^{u_{kxi}} \tilde{b}_x$. Hence, amenities depend on a household-location specific shock and a location-specific observed amenity level proxied by picture density. Furthermore, we have labor supply, which is given by $t_{xi} \nu_{kxi} = (A_i \tau_{xi}^{-\kappa})(\sigma - 1)/\sigma \nu_{kxi}$, where $\nu_{kxi}$ are i.i.d. idiosyncratic shocks on commuting costs, as described in Section 5.2. The idiosyncratic component of $t \nu_{kxi}$ is drawn from a Fréchet distribution with a shape parameter $\varepsilon > 1$ and scale parameter $K_i > 0$.

We aim to estimate the parameters of the model $\{\beta, \kappa, \mu, \varepsilon, \delta, \gamma, \gamma_S, \gamma_\Delta \}$, while we fix the parameters $\{\sigma, \bar{h}\}$ to be in line with the literature. In this way we can calculate counterfactual income mappings and rents for cities in the Netherlands.

To obtain the parameters of the model, we first estimate a gravity equation of commuting flows to estimate the commuting time elasticity $\kappa = \kappa \varepsilon$. Second, using actual data on incomes, we can recover $\varepsilon$. Third, using information on land rents and lot sizes, which we observe for a subset of the data, we recover $\mu$. Fourth, we estimate the equilibrium income mapping to recover $\beta$ and $\gamma$. Fifth, by fitting Fréchet distributions to the location quality index, we recover $\gamma_\Delta$ and $\gamma_S$. Sixth, from the estimated income mapping we can recover the workplace productivity $A_i$ and skill density $L_i$, enabling us to back out the agglomeration elasticity $\delta$.

In what follows, we discuss the moment conditions and the identifying assumptions. We end this section discussing the choice of the fixed parameters $\{\sigma, \bar{h}\}$ and the estimation procedure.

7.1 Estimating the gravity equation

From equation (35), the probability that a household living in $x$ chooses to work in $i$ is equal to:

$$
\pi_{xi|x} = \frac{K_i(\omega_{x,t_{xi}})^\varepsilon}{\sum_{j=1}^{n} K_j(\omega_{x,t_{xj}})^\varepsilon} = \frac{K_i A_i^{\sigma(\sigma - 1)} \tau_{xi}^{-\sigma(\sigma - 1)} \tau_{xj}^{-\sigma(\sigma - 1)}}{\sum_{j=1}^{n} K_j A_j^{\sigma(\sigma - 1)} \tau_{xj}^{-\sigma(\sigma - 1)}}; \quad (38)
$$

11To fix a number of base parameters is common in spatial quantitative equilibrium models (see Ahlfeldt et al., 2015). Note that we identify everything up to a multiplication constant. Hence, the scale parameters of the Fréchet distribution are not strictly identified.
In line with Ahlfeldt et al. (2015), we first recover an estimate for $\tau \equiv -\varepsilon \kappa (\sigma - 1)/\sigma$ by estimating a log gravity model with residence and workplace fixed effects, which absorb $K_i$ and $A_i$. The first moment condition is then given by:

$$\mathbb{E}[\log \tau_{xi|x} - \varepsilon \kappa \log \tau_{xi} - \tilde{\Psi}_x - \tilde{\Omega}_i] = 0.$$  

(39)

By including residence fixed effects $\tilde{\Psi}_x$ and workplace fixed effects $\tilde{\Omega}_i$, we mitigate the endogeneity issues associated with $\tau_{xi}$. One remaining issue is the reverse causality between the flow and the travel time. Indeed, at the locations where there is more demand for travel, better transport infrastructures are likely to be provided, which in turn leads to a shorter travel time. We address this issue by instrumenting $\log \tau_{xi}$ with the log of Euclidian distance between two locations.

### 7.2 Commuting heterogeneity and preferences for land

The next step is to recover $\varepsilon$ from the data. Following Ahlfeldt et al. (2015), we choose to minimize the squared differences between variances within neighborhoods $x$ of adjusted labour supply and labour supply observed in the data. More specifically, let $w_{kxi} \equiv \omega_{kxi} t_{xi} \nu_{kxi}$ be the observed income in the data of a household $k$ located in neighborhood $x$ and working in neighborhood $i$. We observe income conditional on labor supply in the data. For example, if someone has a longer commute and therefore supplies less labour, we observe a lower gross income. More specifically, we use the observed income and control for household characteristics and detailed location pair fixed effects. We then recover $\log \tilde{w}_{xi}$ by taking the estimated values of the location pair fixed effects. Let $\tilde{t}_{xi} \equiv K_i t_{xi}$ be the transformed labour supply, obtained from (39). Note that $\sigma^2_{\log(\omega_{xi})|x} = \sigma^2_{\log(\tilde{t}_{xi})|x}$ because $\omega_x$ does not vary within the neighborhood. Hence, the relationship between the variance within neighborhood $x$ of log transformed incomes $\sigma^2_{\log(\tilde{t}_{xi})|x}$ and the variance of log observed incomes $\sigma^2_{\log \tilde{w}_{xi}|x}$ is given by $\sigma^2_{\log(\tilde{t}_{xi})|x} = \varepsilon^2 \sigma^2_{\log \tilde{w}_{xi}|x}$. This enables us to recover $\varepsilon$ from the second moment condition:

$$\mathbb{E}[\sigma^2_{\log \tilde{t}_{xi}|x} - \varepsilon^2 \sigma^2_{\log \tilde{w}_{xi}|x}] = 0.$$  

(40)

We also use information on land prices $R_x$ and lot sizes $h_x$ for a subset of the sample. Rewriting (29), we derive the third moment condition to determine $\mu$:

$$\mathbb{E} \left[ R_x - \frac{\mu w_{xi}}{(h_x - (1 - \mu)\bar{h})} \right] = 0.$$  

(41)

### 7.3 Estimating the income mapping

Recall that the income mapping (33) is derived from the skill mapping (31). The household $k$ who locates at $x$ is given by
\[ s_{kx} = \left[ K^{1/\gamma} \left( \Delta_x \right)^{\gamma} \right]^{\sigma/(\sigma-1)} \]

with \( K = K_S/K_\Delta \) and

\[ \Delta_x \equiv b_x \left[ \mathbb{E} \left( \max_{i=1,...,n} t_{xi}^{\nu_{kxi}} \right)^{1-\mu} \right] = b_x \cdot \left[ \Gamma \left( \frac{\varepsilon - 1}{\varepsilon} \right) \left( \sum_{i=1}^{n} \tilde{t}_{xi} \right)^{1/\varepsilon} \right]^{1-\mu}. \]

From here on, let us define accessibility as

\[ \tilde{a}_x = \sum_{i=1}^{n} \tilde{t}_{xi} = \sum_{i=1}^{n} K_i A_i^{\frac{\varepsilon(\sigma-1)}{\sigma}} \tau_{xi}^{-\frac{\varepsilon(\sigma-1)}{\sigma}}, \]

where \( \tilde{t}_{xi} \) is obtained from the gravity equation. Since the maximum of Fréchet variables is a Fréchet variable, we have

\[ \mathbb{E} \left[ \max_{i=1,...,n} t_{xi}^{\nu_{kxi}} \right] = \Gamma \left( \frac{\varepsilon - 1}{\varepsilon} \right) \tilde{a}_x^{1/\varepsilon} \]

where \( \Gamma (\cdot) \) is the gamma function. Recall that the observed income in the data \( \omega_{kxi} = (s_{kx})^{\frac{\varepsilon-1}{\varepsilon}} Y^{1/\sigma} \), the income mapping of a household \( k \) residing at \( x \) and working in neighborhood \( i \) can be rewritten as follows:

\[ w_{kxi} = K^{1/\gamma} \tilde{b}_x^{\beta_x} \left[ \Gamma \left( \frac{\varepsilon - 1}{\varepsilon} \right) \tilde{a}_x^{1/\varepsilon} \right]^{(1-\mu)\gamma} \left( t_{xi}^{\nu_{kxi}} \right) \left( 1^{1/\sigma} e^{\nu_{kxi}} \right). \]

Therefore,

\[ \log w_{kxi} + \frac{\sigma-1}{\sigma} \log \tau_{xi} = \Psi_x + \Omega_i + \nu_{kxi} \quad (42) \]

where \( \Psi_x \) are residence fixed effects and \( \Omega_i \) are again workplace fixed effects. Hence, we first estimate the location and workplace fixed effect. Then it should hold that:

\[ \Psi_x = \alpha_1 \log \tilde{b}_x + \alpha_2 \log \tilde{a}_x \quad \text{and} \quad \Omega_i = \frac{\sigma-1}{\sigma} \log A_i \]

where \( \alpha_1 \equiv \beta \gamma, \alpha_2 \equiv \left( 1 - \hat{\mu} \right) \gamma / \hat{\varepsilon}. \) Hence, \( \gamma = \hat{\alpha}_2 \hat{\varepsilon} / (1 - \hat{\mu}) \) and \( \beta = \hat{\alpha}_1 (1 - \hat{\mu}) / \hat{\alpha}_2 \hat{\varepsilon}. \)

Armed with estimated for \( \tilde{t}_{xi} \) (from the gravity equation), \( \tilde{\kappa}, \hat{\varepsilon}, \) and \( \hat{\mu}, \) we can infer \( \gamma \) and \( \beta \) from \( \Psi_x. \) Note further that \( \Omega_i \) are workplace fixed effects, so that wage differences associated with workplace productivity differences \( A_i \) (e.g., due to agglomeration economies) are absorbed by the fixed effects. More specifically, we focus on the job within the household that generates the highest number of working hours and use a work-location fixed effect for each location pair. Hence, we compare households that work at the same location(s), but have different commutes. Last, given \( \Omega_i, \) we recover the adjusted workplace productivities \( \tilde{A}_i \) (up to a constant):

\[ \tilde{A}_i = e^{\bar{\omega}_{xi} \Omega_i}. \]

There are several issues when using (42) to identify causal parameters of the model. First, with respect to accessibility \( \tilde{a}_x, \) since commuting times are on the left-hand side, there may be
reverse causality. Another reason for a bias is that labor markets may not be fully competitive in the sense that households may bargain over to get an income compensation for living further away. Hence, observed incomes $w_{kxi}$ may be higher when people live further away. Mulalic et al. (2013) observe that about 15% of the costs of a longer commute is paid by the employer.

Second, a more general concern with a causal interpretation of $\alpha_1$ and $\alpha_2$ as the impacts of amenities and accessibility 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 $\tilde{b}_x$ and $\tilde{a}_x$ could also be correlated with unobserved housing attributes because households with different incomes may have different preferences for housing quality, such as the age of the housing stock (Brueckner and Rosenthal, 2009). For example, a large share of the housing stock in the city center of Amsterdam takes the form of apartments. This may imply that the affluent are not willing to locate there because they eschew apartment living (Glaeser et al., 2008).

Third, there may be reverse causality between $w_{kxi}$ and $\tilde{b}_x$ and between $w_{kxi}$ and $\tilde{a}_x$. For example, the provision of amenities may be a direct result of the presence of high-income households. Indeed, anecdotal evidence suggests that cultural and leisure services are often abundantly available in upscale neighborhoods (Glaeser et al., 2001). Similarly, high income neighborhoods may attract employers that are in need of specialized and highly educated labor. Finally, because we do not observe the ‘exact’ amenity level, there may be a measurement error in $\tilde{b}_x$, which may lead to a downward bias of $\alpha_1$ when the error is random.\footnote{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.}

The first step to mitigate the biases associated with these concerns is first to ‘purge’ household, job and housing characteristics, $C_k$, from neighborhood characteristics. Let us define: $v_{kxi} = \alpha_3 C_k + \tilde{v}_x + \tilde{v}_{kx}$, where $C_k$ captures, for example, the members of the households that are full-time or part-time households, the size of the household and the age of the adults, while housing attributes are, for example, housing type and construction year. Note that $\tilde{v}_x$ and $\tilde{v}_{kx}$ are shocks related to the neighborhood and the household, respectively. This approach reduces the probability that we measure sorting on basis of household characteristics other than incomes. Furthermore, as we include workplace fixed effects $\Omega_i$, we control for any productivity differences (e.g., due to agglomeration economies) at the workplace.
We estimate the income mapping in two stages. We then define the fourth moment condition:

$$E[\log w_{kxi} + \kappa \frac{\sigma-1}{\sigma} \log \tau_{xi} - \Psi_x - \Omega_i - \alpha_3 C_k] = 0. \quad (43)$$

The fifth moment condition is given by:

$$E[\Psi_x - \alpha_1 \log \tilde{b}_x - \alpha_2 \log \tilde{a}_x] = 0. \quad (44)$$

Working with an endless string of controls will not fully address the endogeneity concerns raised above. Unfortunately, our data do not allow us to exploit quasi-experimental or temporal variation in $\tilde{b}_x$ and $\tilde{a}_x$. Therefore, to investigate the importance of omitted variable bias we first analyze coefficient movements after including controls. Oster (2019) shows that coefficient movements together with changes in the $R^2$ can be used to estimate biased-corrected coefficients. We will outline this procedure in more detail in Appendix B.6.

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

Alternatively, we exploit the fact that $\tilde{b}_x$ and $\tilde{a}_x$ are autocorrelated. First, land use in 1900 is used as an instrument. We expect aesthetic amenities to be positively correlated to the share of built-up area in 1900. For example, the historic city center of Amsterdam has many buildings that have been built before 1900, which are now listed buildings. Furthermore, we also expect water bodies available in 1900 to be correlated to current water bodies, which are often considered as an amenity. As an instrument for commuting time, we count the total number of households $E_x; 1909$ in 1909 within a commuting distance by using the railway network in 1900:

$$E_x; 1909 = \sum_{i=1}^{n} F(\tau_{xi})n_{i,1909},$$

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

Historic instruments can be criticized because of the (strong) identifying assumption that past unobserved locational features are correlated to current unobserved locational endowments.
However, these instruments are more likely to be valid in the context of income sorting, because the patterns of income sorting within each city have considerably changed throughout the last century. Around 1900, open water and densely built-up areas were not necessarily considered as amenities. For example, the canals in Amsterdam were 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 often walked to their working place, and thus commuting distances were short. However, the rich could afford to live outside the city and take the train to their workplace. The cities in 1900 were not yet influenced by (endogenous) planning regulations, as the first comprehensive city plans date from the 1930s.

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

1. We go back further in time as it is less likely that unobserved characteristics of a location or building in the past are correlated with those in present 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. We exploit land use data from the census in 1832. We use municipal populations in 1832 and calculate the travel time of population within commuting distance using information on the road network of 1821. We further use the share of buildings, the share of built-up area, and the share of water bodies within neighborhood as instruments. Moreover, using data on the Cadastral Income, we can control for the value of land at that time. If the currently rich households sort themselves into the most attractive locations of the past, we expect to see a positive correlation with the Cadastral income in 1832.

2. We estimate specifications where we control for the current share of built-up areas and population density. Hence, the identifying assumption is that locations that were attractive in the past because of a share of built-up areas offer a higher amenity level because of the historic buildings they provide, and not because of a high population density today.

3. We gather data from the 1909 census on occupations and skills in each municipality. We then control in various ways for the average skill level of households in 1909 as a proxy for the income in the past. Controlling for the skill level should also address the issue that employment density in 1909 may be correlated to better matching opportunities. Since this
proxy may be considered imperfect, we also use the share of Protestants in 1899 as another proxy for income/skill. Indeed, at that time Protestantism was the dominant religion in the Netherlands while Protestants had a higher education level and were wealthier (Becker and Wößman, 2009).

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

5. Finally, we estimate specifications where we exclusively focus on areas of reclaimed land since 1900. These are areas that are reclaimed from the sea (about 5% of the land) just before and after World War II. We use the plans for development from the 1930s in which new built-up areas and green space are indicated. We calculate the share of planned built-up areas and green space as instruments for current amenities. We then re-calculate the employment accessibility by taking into account the projected population in each settlement. As these reclaimed locations are otherwise identical, and as no one was living in those locations at that time, we address reverse causality.

Given that all those checks lead to more or less the same estimates, we see this as a confirmation that our strategy to identify $\alpha_1$ and $\alpha_2$ is valid. When instrumenting for amenities and commuting time, we replace the fifth moment condition by the following alternative moment conditions:

$$\mathbb{E}[\log \tilde{b}_x - \tilde{\alpha}_0 - \tilde{\alpha}_1 \log z_x] = 0, \quad (45)$$

$$\mathbb{E}[\log \tilde{a}_x - \tilde{\alpha}_0 - \tilde{\alpha}_1 \log z_x] = 0, \quad (46)$$

where $z_x$ are the instruments described above. We then use the predicted values of adjusted amenities and commuting times in the following wage mapping:

$$\mathbb{E}[\tilde{\Psi}_x - \alpha_1 \log \tilde{b}_x - \alpha_2 \log \tilde{a}_x] = 0. \quad (47)$$

7.4 Recovering the parameters of the Fréchet distributions

Using estimates for $\{\beta, \kappa, \mu, \gamma\}$, we aim to obtain the shape parameters of the location-quality index and the income mapping. First, we calculate the expected labour supply at each location by using (38). Using observed amenities, commuting distances, and the adjusted workplace productivities, we can recover the adjusted location quality index (up to a multiplication constant)
at each location

$$\tilde{\Delta}_x = b_x \cdot \left[ \Gamma \left( \frac{\mu - 1}{\mu} \right) \left( \sum_{i=1}^{n} \tilde{t}_{xi} \right)^{\frac{1}{\mu}} \right].$$

Hence, the sixth moment condition may be written as follows:

$$E \left[ f_{\Delta} \left( \tilde{\Delta}_{x_1} \right) - \gamma_{\Delta} e^{-\left( \frac{\tilde{\Delta}_{x_1}}{K_{\Delta}} \right)^{-\gamma_{\Delta}}} \left( \frac{\tilde{\Delta}_{x_1}}{K_{\Delta}} \right)^{-(1+\gamma_{\Delta})} \right] = 0, \quad (48)$$

where \( f_{\Delta}(\tilde{\Delta}_{x_1}) \) is the p.d.f. of the adjusted location-quality index. From this, we obtain \( \gamma_S = \gamma_{\Delta}/\tilde{\gamma}. \)

### 7.5 Recovering the agglomeration elasticity

In the last step we estimate the parameter that indicates how important agglomeration economies are. We first determine the skill mapping for each location given the estimated parameters:

$$s_{kx} \equiv K^{1/\gamma_S} \tilde{b}_{kx}^{\tilde{\beta}_S} \left[ \Gamma \left( \frac{\mu - 1}{\mu} \right) a_x^{1/2} \right]^{(1-\tilde{\rho})\tilde{\gamma}}.$$

Note that we identify \( s_{kx} \) up to a multiplication constant. Hence, we set \( K_{\Delta} \) in such a way that the geometric mean of \( s_{kx} \) equals one and then fit a Fréchet distribution to \( s_{kx} \) to obtain \( K_S \).

From (36) we assumed that \( A_i = A_i L_i \), where \( L_i \) is given by:

$$L_i = \sum_{x=1}^{L^*} \frac{\tilde{t}_{xi}}{\sum_{j=1}^{n} \tilde{t}_{xj}} \left[ K_{S}^{\tilde{\gamma}_S} \sigma - 1 - e^{-K_{S}\tilde{\gamma}_S(s-1)/\sigma} - \tilde{\gamma}_S(s-1)+s/\sigma \right],$$

where the first term is the share of households living at \( x = 1, ..., L^* \) while commuting to \( i \) and the term between brackets is the skill density at location \( x \). Once again, one may argue that \( L_i \) is endogenous and correlated to unobserved endowments (Combes et al., 2010). We therefore use the same instruments as discussed above to mitigate endogeneity. This leads to the following moment conditions:

$$E \left[ L_i - \tilde{\delta}_0 - \tilde{\delta}_1 \log z_x \right] = 0, \quad (49)$$

and in the second stage:

$$E \left[ \log \tilde{A}_i - \log A_i - \delta \log \tilde{L}_i \right] = 0. \quad (50)$$

which identifies \( \delta \).

### 7.6 Estimation

Given the recursive structure of the model, we can just use standard regression techniques to estimate the model’s parameters. For the first moment condition (the gravity equation), we have
a high share of zero flows (about 93%). Therefore, we focus on area pairs that are at most within 60 minutes travelling time of each other. We will then use Poisson Pseudo-Maximum Likelihood methods to deal with the remaining zeroes. Since travel times are likely endogenous, we use a control function approach where the first stage residual is inserted as a control function in the second stage. Euclidian distances between two locations are used as an instrument. For the second and third moment conditions, we use ordinary least squares (OLS) and nonlinear least squares, respectively. For the income mapping (moment conditions 4 and 5), we also use linear regression techniques. When instrumenting for amenities and commuting time, we use two-stage least squares (2SLS) estimates. To obtain the Fréchet parameters (moment condition 6), we use Maximum Likelihood. For the final moment condition recovering agglomeration economies we again use OLS or 2SLS.

We obtain cluster-bootstrapped standard errors by first choosing a set of randomly drawn neighborhoods and then estimate the consecutive steps 250 times. Furthermore, we choose \( \sigma = 4 \) which is in line with the literature (Dustman et al., 2009). We set \( \bar{h} = 25 \text{m}^2 \), which corresponds to the minimum lot size in the sample, and we use a discount rate of 3.5% to go from land prices to land rents, which stems from Koster and Pinchbeck (2019).

8 Empirical results

In this section, we first report reduced-form results regarding the main equations to be estimated, that is, the gravity equation and the income mapping. We discuss several robustness checks regarding identification and consider other proxies for amenities. We then continue by reporting the structural parameters and consider several counterfactual scenarios.

8.1 Reduced-form results

Gravity equation. In Table 2 we report the results for the travel time elasticity. In column (1) we only include location pairs that are within 60 minutes drive from each other. Thus, we drop 77% of the data and we are left with 3.8 million residence-workplace pairs (note that many of those pairs have zero commuters so that more than 90% of the commutes are within 60 minutes). The estimated elasticity is \(-0.732\), implying that doubling the commuting time reduces the probability that someone commutes between \( x \) and \( i \) is reduced by about 50%. In column (2) we address the potential endogeneity of travel times. That is, locations that attract many commuters may invite transport investments, thus leading to lower travel times. We instrument travel times with the Euclidian distance. Unsurprisingly, this is a very strong instrument. We do include the first-stage residual in the second stage as a control function. As one may observe, the first-stage residual is highly statistically significant, strongly suggesting that endogeneity is an issue. The travel time elasticity is now somewhat lower \((-0.549)\), in line with the expectation that reverse causality would
lead to an overestimate. Given that endogeneity is quite important, we consider this specification as the preferred one.

[Table 2 about here]

In previous specifications we focus on commuting flows based on the job that generates the most hours in the household. In column (3), as a sensitivity check, we consider the two jobs that generate the most hours (if applicable). This hardly impacts the results. Column (4) investigates what happens if we use the railway travel time instead of travel time over the road. We show that this leads to similar estimates, although the elasticity is somewhat smaller. Rather than making a selection on maximum commuting time, we can also select locations with a sufficient number of commutes. In column (5) we include location pairs that have at least 25 commuters, including about 60% of the commutes. This leads to very similar results.

Reduced form income mapping. Let us now analyze the effects of amenities and commuting time on income sorting. We estimate regressions of the form:

$$\log w_{kxi} = \alpha_1 \log b_x + \alpha_2 \log a_x + \alpha_3 C_k + \xi_{kxi}$$

(51)

where $\xi_{kxi}$ is an error term. Because we do not yet obtain an estimate for $\tilde{t}_{xi}$, we proxy $a_x$ for now by:

$$E_x = \sum_{i=1}^{I} F(\tau_{xi}) n_i,$$

Hence, at location $x$ we weight the number of jobs $n_i$ by the share of people that commute maximally $\tau_{xi}$. Table 3 reports the reduced-form results.

Column (1) reports a simple regression of log income on log amenities and log accessibility, while we only control for demographic characteristics and year fixed effects. We show that more amenities and accessibility are associated with higher incomes. Doubling amenities implies an increase in income of $(\log 2 - \log 1) \times 0.0215 = 1.5\%$. Doubling of accessibility attracts households whose incomes are 6.9% higher. In column (2) we add a wider array of controls related to housing quality and job characteristics. Although the $R^2$ increases by almost 50%, the coefficients related to amenities and accessibility are hardly affected. This suggests that amenities are not so much correlated to building quality. In column (3) we include workplace fixed effects to control for agglomeration economies in the workplace and identify the ‘pure’ accessibility effect. We observe that the coefficient is somewhat lower. A 100% increase in amenities now attracts households whose incomes are 1.2% higher. The coefficient related to employment accessibility is hardly affected.

Despite the inclusion of $C_k$ and workplace fixed effects one may argue that we do not convincingly address the omitted variable bias. We deal with this issue by estimating bias-corrected
regressions following Oster (2019) in Appendix B.6. We show that when we choose the appropriate maximum attainable $R^2$ (as only part of the variation in incomes can be explained by variables varying at the neighborhood level), the estimates are very close to the OLS estimates. This strongly suggests that omitted variable bias is not a major issue.

In column (4) we aim to address potential measurement error in the picture density as a proxy for amenities by instrumenting for it with observed proxies for amenities (e.g., nearby historic buildings or share water bodies). The first-stage results in Appendix B.6 show the expected signs: there is a higher picture density in built-up areas, in areas with more water bodies (e.g., the Amsterdam canal district), and where there are many historic buildings.\textsuperscript{13} The contemporary instruments are strong instruments for amenities. The second-stage coefficient related to amenities in column (4), Table 3, is essentially identical, suggesting that measurement error is not a main concern.

[Table 3 about here]

Yet, amenities and accessibility may be endogenous due to reverse causality. The contemporary instruments may only partly address this issue. This is why we instrument amenities with historic variables in column (5). The instruments are the shares of water bodies and of built-up area in 1900 within a neighborhood $x$. In Appendix B.6, we report the corresponding first-stage results. The share of built-up area, the share of water bodies in 1900 are strongly and positively correlated to the current amenity level. Going back to Table 3, the coefficient of amenities is now somewhat higher: doubling amenities attracts households whose incomes are 2.3% higher. In column (6) we also instrument for employment accessibility with the number of households within commuting distance in 1909 using the railway network in 1900. The number of people reachable within commuting distance is positively correlated to current accessibility; the elasticity is 0.42. 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 (6), Table 3, reveal that when we instrument amenities and commuting times there is a positive effect of picture density and accessibility on incomes. We consider this specification to be the preferred specification. Doubling amenities attracts households whose incomes are 2.3% lower. Doubling accessibility leads to households whose incomes are 3.8% higher. The impacts of accessibility and amenities are thus very comparable to each other and very similar to the OLS results.

\textsuperscript{13}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.
Identification revisited. We consider additional robustness analyses in Table 4 that should increase confidence in the validity of our identification strategy. First, we show that our results are similar once we focus solely on urban areas. In column (1) we only include observations in the Randstad, i.e., the main polycentric metropolitan area in the Netherlands. This reduces the total number of observations by more than 50%. However, our results are similar, in particular for amenities. For employment accessibility we find that the coefficient is somewhat stronger, which may be due to traffic congestion in some parts of the Randstad (e.g. around Amsterdam and Rotterdam), which would imply that our estimates of commuting costs are underestimated. In column (2) we exclusively focus on observations close to city centers. That is, we only include locations within 15km of the center of an urban area with at least 100 thousand inhabitants. The coefficients are very similar, but for employment accessibility it becomes somewhat imprecise.

In column (3) we go back further in time and use instruments from 1832. This reduces the number of observations considerably, because the 1832 data is not available for whole of the Netherlands. The Kleibergen-Paap $F$-statistic in column (3) is lower, which is not too surprising as going back further in time implies that correlations between instruments and endogenous variables are becoming less strong. We find an effect for accessibility that is about twice as strong as when using instruments from 1900. In column (4) we use the information on the cadastral income, a proxy for the land value in 1832. This is missing in two thirds of the cases, so our number of observations drop further to about 1.8 million observations. Again, we find that the effect of amenities is very much comparable to the baseline specification. The effect of commuting time is even somewhat stronger. Interestingly, the effect of cadastral income is negative. A 10% decrease in the cadastral income in 1832 attracts households whose incomes are 0.03% higher, so the effect is economically small. This is in line with anecdotal evidence that amenities in the past are essentially uncorrelated, or even negatively correlated, to current amenities.

[Table 4 about here]

In column (5), Table 4, we estimate specifications where we again use instruments from 1900, but control for the current share of built-up areas and population density to make sure that our amenity proxy is not just capturing population density or built-up land. We find very similar effects for amenities and accessibility.

One may be more worried that concentrations of high income households are autocorrelated so that our instruments are correlated to concentrations of high income households in 1909. To investigate whether this is an issue, we calculate the share of medium and high skilled households in 1909. Municipalities then were much smaller, so this is a rather fine-grained measure of skill sorting across space. We also gather data on the share of Protestants in each municipality in 1899 and control for population accessibility in 1900. Including those measures does not impact our coefficients at all. Note that locations of high-skilled and medium-skilled households in 1909 are correlated to the locations of lower incomes nowadays, which is in line with anecdotal evidence
that the determinants of residential choices in the two periods are fairly different. This also confirms the negative association of Cadastral Incomes in 1832 to current incomes. Also, conditional on employment accessibility, population accessibility in 1900 is negatively correlated to current incomes. In column (7) we further study the sensitivity of our results by choosing another instrument for accessibility. We use the share of the population in 1909 born in the same municipality. If mobility of households is correlated over time, the share of locally born people should be negatively correlated to current accessibility, as the areas that host the most jobs (so have a better accessibility) are expected to attract workers from other places. We find that indeed the share of locally born people in 1909 is negatively associated with current employment accessibility. The Kleibergen-Paap $F$-statistic again indicates that these are strong instruments. We find a similar coefficient related to employment accessibility.

If one is still worried that household income sorting is autocorrelated, in column (8) we only include neighborhoods on reclaimed land. The Netherlands is well-known for its large-scale projects that reclaim land from the sea. Currently about 5% of the land is reclaimed. We consider three main projects (i.e. Wieringermeer, Noordoostpolder, Oostelijk, and Zuidelijk Flevoland) that occurred between 1930 and 1968, but permission by the government to reclaim those areas was already given in 1930. Most of the land was intended for agriculture, but a few small settlements were planned. For example, the Noordoostpolder was planned according to the Christaller’s central place theory. Also, there was an area Markerwaard, which was planned was in the end not reclaimed. Moreover, Lelystad was planned to be the largest city, but nowadays Almere is by far the largest city in the area. In other words, the plans differ considerably from the current spatial economic distribution. Because only a small share of the population lives in those areas, we only keep about 2.5% of the observations.

We then instrument for amenities with the share of planned built-up and green areas in column (8). We observe that the impact of amenities is slightly lower, but given its standard error it is not statistically significantly different from the baseline estimate. The coefficient of employment accessibility is very similar to the baseline estimate, albeit imprecise. When we also instrument for employment accessibility with the planned accessibility in column (9), the point estimates are again similar, but we have weak instruments leading to imprecise coefficients. In this way, we address reverse causality as no one was living in those locations at that time, and thus income was zero.

**Alternative proxies for amenities and effects on land prices.** One may worried that our results hinge on the particular choice of the amenity index. We therefore consider three alternative proxies for 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 B.3. 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), Panel A of Table 5, we re-estimate our preferred specification with historic instruments. This alternative index also has a strong impact on incomes. It appears that the amenity elasticity is essentially the same as the estimates obtained by using the picture index. We also gather data on ‘places of interest’ from the augmented reality game *Pokémon Go*, which was a hugely popular game in 2017.\textsuperscript{14} The game could be played at certain places of interest, the so-called ‘Pokéstops’.\textsuperscript{15} The locations of Pokéstops 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 Google. 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. In other words, these Pokéstops are not randomly located across space and signify locational attractiveness. We construct the Pokémon Go amenity index by using density of Pokéstops in a neighborhood. Our results show that the density of Pokéstops is positively associated with incomes: doubling the Pokéstop density attracts households whose incomes are 2.2% higher. The commuting time elasticity is very much the same compared to the baseline specification. In column (3) we just use the share of the land in a neighborhood that is part of an officially designated historic district. Using historic instruments we find a strong and statistically significant effect on household incomes: a 10% increase in the share of land part of a historic district attracts households whose incomes are 3% higher.

\textit{[Table 5 about here]}

In Panel B of Table 5, we investigate the impacts of amenities and commuting times on land prices. In our set up the signs of the effects of amenities and accessibility on land prices and incomes are the same (although magnitudes may differ). Therefore, we now estimate the effects of amenities and commuting time on land prices. We start in column (4), Table 5, with a simple OLS specification including amenities and accessibility, while controlling for households, job and housing characteristics. This leads to a strong positive effect of amenities on land prices: doubling amenities implies a land price increase of 8.7%, while doubling accessibility leads to land prices that are 22.4% higher. When we control for workplace fixed effects, the coefficients are hardly affected. In the final column we instrument for amenities and accessibility with historic instruments from around 1900. The effect of accessibility becomes somewhat lower, while the effect of picture density becomes about twice as strong. Hence, the reduced-form effects on land prices do indeed have the same signs, but are stronger in magnitude.

\textsuperscript{14}It was one of the most used and profitable mobile apps in 2016, having been downloaded more than 500 million times worldwide.

\textsuperscript{15}Another type of locations that are used in the game are so-called ‘Gyms’. The latter types are unfortunately less useful, as these are almost uniformly distributed within urban areas in gardens, open spaces and public squares.
Other sensitivity checks. Appendix B.8 shows that our results still hold for a wide range of alternative robustness checks and sample selections. More specifically, to minimize any measurement error regarding accessibility and workplace productivities, we run specifications where we only keep households (i) with a single job, (ii) with a single job in a single-plant firm, and (iii) households with a company car (so that it is likely that those households actually use the car for commuting). We further test whether results change when using the share of highly educated adults in the household, which is a more direct way to estimate the (reduced-form) skills mapping. We find very similar effects, both in terms of sign and magnitude, which confirms that looking at income or skill level is equivalent. We also use commuting time by rail instead of commuting time over the road. Overall, the impact of amenities and commuting time on income sorting choice is robust.

8.2 Structural parameters

In Table 6 we report the results of the structural estimation. In column (1) we do not instrument for amenities $b_{x_i}$, employment accessibility $a_{x_i}$ and employment density $L_i$. We find that a commuting time elasticity equal to $\kappa = 0.22$, which is higher than in the literature. However, one should keep in mind that we use the log of commuting time, so that this represents an elasticity rather than the semi-elasticity usually reported in the literature. Commuting heterogeneity $\varepsilon$ is about 2.73, which is somewhat lower than Ahlfeldt et al. (2015), but close to the value picked by Brinkman and Lee (2019) and that estimated by Dericks and Koster (2019). The estimate $\mu$ indicates the preferences for land. We find that $\mu = 0.0955$, which confirms that richer households spend less of their income on land (Albouy et al., 2016). Note that $\mu$ may seem low, but we only include payments to land, not to housing itself.

[Table 6 about here]

So far, all the estimated parameters are identical for different specifications because the instruments are only used in the later steps to identify preferences for amenities, accessibility and agglomeration economies. The preference parameter $\beta$ that indicates how households value amenities in column (1) is similar to the baseline reduced-form result. However, when we use instruments based on data from 1900 $\beta$ is considerably larger. This is mainly because the relative location quality heterogeneity parameter $\gamma$ is about 50% smaller. The preference for amenities is not much affected if we use instruments based on data from 1832. The estimated elasticity of agglomeration economies is 0.0465 if we do not instrument, while it is higher when we use historic instruments (0.0745 and 0.0887 using instruments from 1900 and 1832, respectively). These estimates fall well within the range provided by the literature. For example, Rosenthal and Strange (2004) suggest a range of 0.03-0.08, while the mean estimate provided by the meta-analysis of Melo et al. (2009) is 0.058. Our estimates are higher than those reported by Combes and Gobillon (2015) who study
the elasticity of wages with respect to population density.

**Overidentification checks.** Our structural estimation procedure suggests natural overidentification checks that can be used to investigate whether our model is able to fit the data reasonably well. We do not expect to find a perfect fit because we consider only two determinants of location choices, that is, the level of exogenous amenities and the employment accessibility, while location choices may depend on many location attributes. First, our estimation procedure leads to an approximation for the employment level $L_i$ at each location $i$. If we compare the estimated $L_i$ to the observed employment level in each area, we find a correlation of 0.839, which is fairly high. One may be worried that this high correlation might be driven by a few locations that host many workers. This appears not to be an issue because the correlation between the log of estimated employment to the log of observed employment is equal to 0.907.

Another overidentification check involves the comparison of the ex-post estimated land rents (see equation (34)) to the observed land prices. Note that land prices are not a direct input in our model. Hence, there is no pre-determined mechanical correlation between estimated and actual land prices. We find a correlation between estimated and observed land prices of 0.643. When we correlate the log of estimated land prices to the log of observed land prices, we find a slightly higher correlation ($\rho = 0.718$). We believe that these correlations are quite high as we only include two determinants of locational choices. This in turn suggests that amenities and accessibility are very important and determinants of locational choices.

### 8.3 Counterfactual scenarios

Given the estimated parameters, our model allows for the undertaking of counterfactual analyses. We describe the exact procedure to solve for the counterfactual values and derive the aggregate land rent and real income in Appendix B.9. Let us consider three counterfactual scenarios. In the first one, we assume away amenities throughout the Netherlands. The idea is to mimic U.S. cities where amenity levels are considerably lower than in the Netherlands. In a second scenario, we reduce commuting time by 50%. This reflects a situation where a new transportation technology would reduce commuting times substantially (e.g., automated vehicles). In the third scenario, we consider driving restrictions in the main polycentric urban area in the Netherlands, i.e., the Randstad. In line with policies implemented in Mexico City, Bogotá and Beijing, we assume that cars are banned each other day (e.g., based on number plates). This implies that workers cannot use their cars each other day to get to work, thus implying an increase in commuting costs.

Before turning to the results, we estimate the outcomes for the baseline scenario. We report the results for the key variables in Table 7.

---

16This is except for the determination of $\mu$. However, the exact estimate of $\mu$ has essentially no implications for the correlation between actual and estimated land prices.
In the first scenario, we set the amenity level equal to the minimum value of amenities observed in the data. Since households do not care about amenities anymore, they live on average closer to their workplace and earn higher incomes. We find that the overall output increases by 10.6%, while the aggregate real income rise by 7.3%. The aggregate land rent decreases by 0.6% in the absence of amenities. We also construct a measure of income mixing, which is the standard deviation of skills in adjacent neighborhoods, to see how the counterfactual scenario affects income mixing within the Netherlands. A uniform amenity distribution implies substantially less income mixing as the standard deviation is much lower than the baseline estimate. This confirms the anecdotal evidence that European, especially Dutch, cities are more socially mixed than American cities.

Having a uniform distribution of amenities has strong repercussions for the spatial distribution of skills, hence of incomes. Indeed, the correlation between the values observed in the data and in the counterfactual is 0.556. Hence, amenities are a key-determinant of the skill-based sorting of households within and between cities. We report maps in Figure 3. Figure 3.A shows the relative change in skill at each location. In high amenity locations, such as the city center of Amsterdam or Utrecht, we observe a relatively large decrease in skills, thus confirming that high-skilled people value more amenities. Because of the strong positive correlation between land rents and the skill mapping, as shown in Figure 3.B, the highest amenity locations witness the highest decreases in land rents.

The second counterfactual considers a reduction in commuting time of 50%. This triggers a strong increase in output of 16.6%. Since commuting costs are much lower, households’ labor supply is higher, which implies a higher aggregate real income. To be precise, the aggregate real income rises by 29.8%. In line with this result, we observe that the aggregate land rent also increases by 3.8%. This increase is quite modest compared to the increase in labor income and output. The reason is that accessibility is now a much less important locational determinant. As a result, households substitute accessibility for the consumption of the composite good. Indeed, we observe that the consumption of the composite good increases by 33.2%. By contrast, the spatial implications of a strong reduction in commuting costs are limited (the correlation between the skill mapping in the baseline and counterfactual is 1). This is because we consider a proportional increase in commuting times, which leaves the commuting probabilities (essentially) unaffected (see Figure 4.A). Because of the non-linear relationship between rents and the skill mapping, we

17 The absolute amenity level makes no difference for the outcomes because we re-adjust the parameter $K_\Delta$ for the aggregate skill distribution to have a geometric mean equal to 1. Moreover, assuming an equal value for $\tilde{b}_x$ leads to the same result as when setting $\beta = 0$. 

44
find that rents do increase in some areas, which is particularly in (historic) centres of the large cities (see Figure 4.B). Because commuting costs become less important, households are willing to pay more for amenities. Hence, high amenity locations become relatively more expensive.

In the last counterfactual, we consider a spatially heterogeneous reduction in travel time by implementing driving restrictions in the largest metropolitan area in the Netherlands. The driving restrictions increases average commuting times from 24 to 32 minutes for households residing or working in the Randstad. The driving restrictions do not have a strong impact on total output, which decreases only by 0.3%. Net wages are not much affected either as the reduction is 1%. For the Randstad, the labor income decrease is somewhat higher (4.1%). Finally, the aggregate land rent decreases by 1%.

Despite the increase in the spatial heterogeneity of travel times, the spatial implications of the driving restrictions are limited. The correlation between the skill mapping in the baseline scenario and counterfactual is 0.996. Figure 5.A shows that the maximum difference in skills is only 2.7%. As for the land rents, we find a somewhat lower correlation ($\rho = 0.946$) and larger differences. The largest land rent difference occurs in the south of the Randstad and reaches up to 9% (see Figure 5.B). In sum, variations in commuting costs have a limited impact on the spatial sorting of skills, hence of incomes.

We confirm this conclusion in Appendix B.10 where we reduce the preference for commuting by 50%, which leads to a substantially higher output, but to essentially the same skill distribution.

So far, we did not discuss the implications of the different counterfactuals for the spatial employment distribution. This distribution is hardly affected by changes in commuting costs or amenities because agglomeration economies are relatively weak. Since the productivity constants $K_i$ and $A_i$ are unaffected by changes in commuting costs or amenities, production is, to a large extent, anchored in the same locations, thus reflecting the impact of history. We test this idea in Appendix B.10 where agglomeration economies are assumed to be much stronger. In this case, the spatial distribution of jobs change considerably, while the spatial distribution of skills is hardly affected. In sum, our last counterfactual reinforces the conclusion that the spatial distribution of skills is mainly determined by amenities, rather than by commuting costs or job locations.

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 its location and its distance to the nearest employment
location. 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 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.

The following extension is worth mentioning. We can account for the endogenous choice of local public goods (LPGs). Consumers at \( x \) choose their consumption level \( c(x) \) of LPGs, which, like in U.S. cities, are financed by a property tax \( \gamma(x) \). Hence, under the assumption of a fixed lot size (\( \mu = 0 \) and \( h = 1 \)), we have \( c(x) = \gamma(x)R(x) \). Assuming that amenities and LPGs are bundled into a Cobb-Douglas aggregate, preferences become \( U = b^\alpha c^{1-\alpha}q \), with \( 0 < \alpha < 1 \). Solving the utility-maximizing condition for the equilibrium tax rate for a \( \omega \)-consumer at \( x \) yields \( c^*(x) = (1 - \alpha) [\omega t(x) - (1 + \gamma)R(x)] \). Using \( c^*(x) \) and applying the same approach as in Section 4, it is readily verified that

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

In this case, the location-quality index becomes \( \Delta(x) = [b(x)]^{\alpha/(2-\alpha)} t(x) \). Comparing this condition to (20) where \( \mu = 0 \) shows that, the decentralized provision of LPGs weakens the impact of exogenous amenities in individual residential choices. In particular, when consumers do not value much the exogenous amenities (\( \alpha \) is small) or when the level of exogenous amenities is almost constant across space, residential choices will be mainly driven by commuting costs.

References


<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gross income (in €)</strong></td>
<td>91,535</td>
<td>53,683</td>
<td>3,589</td>
<td>999,897</td>
</tr>
<tr>
<td><strong>Land price (€ per m²)</strong></td>
<td>1,312</td>
<td>752.2</td>
<td>0.00753</td>
<td>22,418</td>
</tr>
<tr>
<td><strong>Lot size (m²)</strong></td>
<td>364.3</td>
<td>923.8</td>
<td>25</td>
<td>24,824</td>
</tr>
<tr>
<td><strong>Pictures per ha</strong></td>
<td>2.189</td>
<td>8.840</td>
<td>0</td>
<td>231.9</td>
</tr>
<tr>
<td><strong>Hedonic amenity index</strong></td>
<td>2.821</td>
<td>0.0915</td>
<td>2.723</td>
<td>3.885</td>
</tr>
<tr>
<td><strong>Share historic district</strong></td>
<td>0.0347</td>
<td>0.139</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Listed building</strong></td>
<td>0.0941</td>
<td>0.699</td>
<td>0</td>
<td>17.06</td>
</tr>
<tr>
<td><strong>Share built-up land</strong></td>
<td>0.449</td>
<td>0.298</td>
<td>0.000856</td>
<td>1</td>
</tr>
<tr>
<td><strong>Share water</strong></td>
<td>0.0496</td>
<td>0.0738</td>
<td>0</td>
<td>0.813</td>
</tr>
<tr>
<td><strong>Commuting time in minutes</strong></td>
<td>26.39</td>
<td>17.18</td>
<td>0</td>
<td>120.0</td>
</tr>
<tr>
<td><strong>Employment accessibility</strong></td>
<td>624,940</td>
<td>275,990</td>
<td>14,427</td>
<td>1.347e+06</td>
</tr>
<tr>
<td><strong>Total hours worked in household</strong></td>
<td>2,159</td>
<td>913.1</td>
<td>416.1</td>
<td>6,239</td>
</tr>
<tr>
<td><strong>Household has company car</strong></td>
<td>0.149</td>
<td>0.356</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Works at single-establishment firm</strong></td>
<td>0.443</td>
<td>0.497</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Number of jobs in household</strong></td>
<td>1.511</td>
<td>0.968</td>
<td>1</td>
<td>18</td>
</tr>
<tr>
<td><strong>Person is male</strong></td>
<td>0.521</td>
<td>0.215</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Person is foreigner</strong></td>
<td>0.0718</td>
<td>0.217</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Age of person</strong></td>
<td>41.99</td>
<td>9.008</td>
<td>18</td>
<td>64</td>
</tr>
<tr>
<td><strong>Apartment</strong></td>
<td>0.153</td>
<td>0.360</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>House built &lt;1945</strong></td>
<td>0.192</td>
<td>0.394</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

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

*Dependent variable: the number of commuters*

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Control</th>
<th>Flows based on two jobs</th>
<th>Travel time by train</th>
<th>Flow &gt;25</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Poisson</td>
<td>Poisson-CF</td>
<td>Poisson-CF</td>
<td>Poisson-CF</td>
<td>Poisson-CF</td>
</tr>
<tr>
<td>Commuting time elasticity, $\kappa$</td>
<td>(-0.7318^{***})</td>
<td>(-0.5485^{***})</td>
<td>(-0.5703^{***})</td>
<td>(-0.3393^{***})</td>
<td>(-0.5215^{***})</td>
</tr>
<tr>
<td></td>
<td>(0.0139)</td>
<td>(0.0122)</td>
<td>(0.0111)</td>
<td>(0.0080)</td>
<td>(0.0086)</td>
</tr>
<tr>
<td>First-stage error</td>
<td>(-0.2378^{***})</td>
<td>(-0.2079^{***})</td>
<td>0.3402***</td>
<td>0.4154***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0653)</td>
<td>(0.0475)</td>
<td>(0.0217)</td>
<td>(0.0207)</td>
<td></td>
</tr>
<tr>
<td>Residence location fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Workplace location fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Number of area pairs</td>
<td>3,904,262</td>
<td>3,904,262</td>
<td>3,904,262</td>
<td>3,904,262</td>
<td>66,147</td>
</tr>
</tbody>
</table>

*Notes:* We use commuting flows between neighborhoods based on the job that generates the most working hours. In columns (2)-(5) we use as instrument the euclidian distance between two neighbourhoods as instrument. In column (3) we derive the commuting flow based on the two jobs that generate the most working hours in the household. Standard errors are bootstrapped (250 replications) and in parentheses; *** $p < 0.01$, ** $p < 0.5$, * $p < 0.10$.  

### Table 3 – Baseline reduced-form regression results

*Dependent variable: the log of household gross income*

<table>
<thead>
<tr>
<th></th>
<th>+ Housing and job controls</th>
<th>+ Workplace fixed effects</th>
<th>Contemporary Instruments</th>
<th>Historic Instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>(1)</td>
<td>OLS</td>
<td>(2)</td>
</tr>
<tr>
<td>Pictures per ha (log)</td>
<td>0.0215***</td>
<td>(0.0016)</td>
<td>0.0285***</td>
<td>(0.0013)</td>
</tr>
<tr>
<td>Employment accessibility (log)</td>
<td>0.0999***</td>
<td>(0.0043)</td>
<td>0.0942***</td>
<td>(0.0040)</td>
</tr>
<tr>
<td>Household controls</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Housing and job controls</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Workplace fixed effects</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Number of observations</td>
<td>10,213,540</td>
<td>10,213,540</td>
<td>10,213,524</td>
<td>10,213,524</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.2041</td>
<td>0.2949</td>
<td>0.3316</td>
<td>386.2</td>
</tr>
<tr>
<td>Kleibergen-Paap F-statistic</td>
<td>936.2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Notes:* **Bold** indicates instrumented. Household controls include household size, mean age of adults, mean gender, household type (couple, single, kids), the share of the household that is foreign-born. Job controls are the total hours worked, whether the household has a company car, the share of full-time contracts, the share of permanent contracts. Housing controls include house type, height of the building, construction year dummies and whether a building is listed. Standard errors are clustered at the neighborhood level and in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.
Table 4 – Reduced form results: identification
(Dependent variable: the log of household gross income)

<table>
<thead>
<tr>
<th></th>
<th>Only</th>
<th>City center</th>
<th>1832</th>
<th>Control for</th>
<th>1909</th>
<th>Other</th>
<th>Only obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Randstad</td>
<td>&lt;15km</td>
<td>instruments</td>
<td>current land use</td>
<td>skills</td>
<td>instrument</td>
<td>on reclaimed land</td>
</tr>
<tr>
<td></td>
<td>2SLS</td>
<td>2SLS</td>
<td>2SLS</td>
<td>2SLS</td>
<td>2SLS</td>
<td>2SLS</td>
<td>2SLS</td>
</tr>
<tr>
<td><strong>Pictures per ha (log)</strong></td>
<td>0.0382***</td>
<td>0.0374***</td>
<td>0.0375***</td>
<td>0.0491***</td>
<td>0.0494***</td>
<td>0.0483***</td>
<td>0.0501***</td>
</tr>
<tr>
<td></td>
<td>(0.0048)</td>
<td>(0.0046)</td>
<td>(0.0048)</td>
<td>(0.0061)</td>
<td>(0.0050)</td>
<td>(0.0053)</td>
<td>(0.0044)</td>
</tr>
<tr>
<td><strong>Employment accessibility (log)</strong></td>
<td>0.1541***</td>
<td>0.1647***</td>
<td>0.1134***</td>
<td>0.1503***</td>
<td>0.0597***</td>
<td>0.1081***</td>
<td>0.2170***</td>
</tr>
<tr>
<td></td>
<td>(0.0023)</td>
<td>(0.0501)</td>
<td>(0.0160)</td>
<td>(0.0319)</td>
<td>(0.0100)</td>
<td>(0.0352)</td>
<td>(0.0645)</td>
</tr>
<tr>
<td><strong>Cadastral income in 1832 per ha (log)</strong></td>
<td>-0.0050**</td>
<td>(0.0023)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Share built-up land</strong></td>
<td>-0.0936***</td>
<td>(0.0138)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Population per ha (log)</strong></td>
<td>-0.0074**</td>
<td>(0.0033)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Population accessibility in 1900 (log)</strong></td>
<td>-0.0264*</td>
<td>-0.0681***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0150)</td>
<td>(0.0249)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Share of medium-skilled workers in 1909</strong></td>
<td>-0.1630***</td>
<td>-0.1860***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0211)</td>
<td>(0.0248)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Share of high-skilled workers in 1909</strong></td>
<td>-0.1378</td>
<td>-0.1645</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1081)</td>
<td>(0.1053)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Share protestants in 1899</strong></td>
<td>-0.0169***</td>
<td>-0.0033*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0061)</td>
<td>(0.0072)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Household controls</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Housing and job controls</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Year fixed effects</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Workplace fixed effects</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Number of observations</strong></td>
<td>4,340,639</td>
<td>6,023,886</td>
<td>5,549,488</td>
<td>1,782,784</td>
<td>10,213,325</td>
<td>9,778,046</td>
<td>9,778,046</td>
</tr>
<tr>
<td><strong>Kleibergen-Paap F-statistic</strong></td>
<td>70.51</td>
<td>34.43</td>
<td>22.73</td>
<td>33.87</td>
<td>61.92</td>
<td>21.16</td>
<td>15.79</td>
</tr>
</tbody>
</table>

Notes: **Bold** indicates instrumented. Household controls include household size, mean age of adults, mean gender, household type (couple, single, kids), the share of the household that is foreign-born. Job controls are the total hours worked, whether the household has a company car, the share of full-time contracts, the share of permanent contracts. Housing controls include house type, height of the building, construction year dummies and whether a building is listed. Standard errors are clustered at the neighborhood level and in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.10
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2SLS</td>
<td>2SLS</td>
<td>2SLS</td>
<td>OLS</td>
<td>OLS</td>
<td>2SLS</td>
</tr>
<tr>
<td>Pictures per ha (log)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hedonic amenity index (log, std)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pokéstops per ha (log)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share historic district</td>
<td>0.2914***</td>
<td>(0.0309)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employment accessibility (log)</td>
<td>0.0782***</td>
<td>(0.0086)</td>
<td>0.0580***</td>
<td>(0.0113)</td>
<td>0.0731***</td>
<td>(0.0088)</td>
</tr>
<tr>
<td>Number of observations</td>
<td>10,233,115</td>
<td>9,839,819</td>
<td>10,236,308</td>
<td>2,196,280</td>
<td>2,196,280</td>
<td>2,196,280</td>
</tr>
<tr>
<td>Household controls</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Housing and job controls</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Workplace fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>R²</td>
<td>0.5564</td>
<td>0.5891</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kleibergen-Paap F-statistic</td>
<td>29.02</td>
<td>68.45</td>
<td>29.14</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: **Bold** indicates instrumented. Household controls include household size, mean age of adults, mean gender, household type (couple, single, kids), the share of the household that is foreign-born. Job controls are the total hours worked, whether the household has a company car, the share of full-time contracts, the share of permanent contracts. Housing controls include house type, height of the building, construction year dummies and whether a building is listed. Standard errors are clustered at the neighborhood level and in parentheses. ***p < 0.01, **p < 0.05, *p < 0.10.
### Table 6 – Structural estimation

<table>
<thead>
<tr>
<th>Instruments</th>
<th>1900</th>
<th>1832</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Commuting time elasticity, $\kappa$: 0.2210*** (0.0048) (0.0048) (0.0048)
- Commuting heterogeneity, $\xi$: 2.7323*** (0.0144) (0.0144) (0.0144)
- Land preferences, $\mu$: 0.0955*** (0.0003) (0.0003) (0.0003)
- Amenity preferences, $\beta$: 0.0404** (0.0181) (0.0712) (0.0262)
- Relative location quality heterogeneity, $\gamma$: 0.3142*** (0.0124) (0.0370) (0.0408)
- Agglomeration elasticity, $\delta$: 0.0465*** (0.0016) (0.0040) (0.0034)
- Location quality heterogeneity, $\gamma_\Delta$: 6.0911*** (0.2162) (0.5683) (0.3531)
- Skills heterogeneity, $\gamma_s$: 23.0909*** (1.7681) (1.0818) (0.2162)

**Fixed parameters:**
- Minimum lot size, $\bar{h}$: 25 25 25
- Elasticity of substitution, $\sigma$: 4 4 4

**Notes:** We estimate the parameters using data at neighborhood level. In column (2) we use as instruments the share of water bodies in 1900 in the neighborhood, the share of built-up land in 1900 in the neighborhood, the share of built-up land in 1900 <500m of the neighborhood, the share of built-up land in 1900 500-1000m, and employment accessibility in 1909. In column (3) we use as instruments the share of water bodies in 1832 in the neighborhood, the share of built-up land in 1832 in the neighborhood, the share of buildings in 1832 in the neighborhood, the share of buildings in 1832 <500m of the neighborhood, the share of buildings in 1832 500-1000m, and population accessibility in 1832. Standard errors are bootstrapped (250 replications) and clustered at the neighborhood level; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

### Table 7 – Counterfactual scenarios

<table>
<thead>
<tr>
<th>Scenario 1: no amenities</th>
<th>Scenario 2: 50% lower commuting costs in Randstad</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
</tr>
<tr>
<td>Total output</td>
<td>120,959</td>
</tr>
<tr>
<td>Aggregate land rents</td>
<td>437,039</td>
</tr>
<tr>
<td>Aggregate real income</td>
<td>11,821</td>
</tr>
<tr>
<td>Income mixing, $\sigma_x$</td>
<td>0.0472</td>
</tr>
</tbody>
</table>

**Notes:** We calculate aggregate land rents as: $\sum_{x=1}^L h_x^t R_x^t$ and aggregate net wages as: $\sum_{x=1}^L (1/h_x^t) \omega_x^t t_x^t$. Hence, we weight aggregate labor income by the density in each location.
Figures

Figure 1 – Sorting and location quality

(a) Overview map

(b) Commuting networks

Figure 2 – The Netherlands

Legend

- ▲ Airport
- ● City >100,000 inhabitants
- □ Cities

Legend

Number of commuters

<table>
<thead>
<tr>
<th>Range</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>110 - 208</td>
<td>6</td>
</tr>
<tr>
<td>209 - 369</td>
<td>12</td>
</tr>
<tr>
<td>370 - 639</td>
<td>18</td>
</tr>
<tr>
<td>640 - 1177</td>
<td>30</td>
</tr>
<tr>
<td>1178 - 3355</td>
<td>40</td>
</tr>
</tbody>
</table>
Figure 3 – Counterfactual 1: No amenities

Figure 4 – Counterfactual 2: 50% reduction in commuting costs
Figure 5 – Counterfactual 3: driving restrictions in the Randstad

(A) % change in $s_x$

(B) % change in $R_x$
Appendix A

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

Differentiating (13) with respect to \(x\) and using (12), we obtain:

\[
\Psi_x(x, \omega, U^*(\omega)) = \frac{\omega t}{H} \left( \frac{t_x}{t} - \frac{Q_b}{\omega t} b_x \right) \tag{A.1.1}
\]

Differentiating (A.1.1) with respect to \(\omega\) and rearranging terms yields the following expression:

\[
\Psi_{\omega x}(x, \omega, U^*(\omega)) = \frac{t}{H} \left\{ \frac{t_x}{t} \left[ 1 - \frac{\omega}{H}(H_\omega + H_UU^*_\omega) \right] b + \frac{b_x}{t} \left[ \frac{H_\omega + H_UU^*_\omega}{H} Q_b - (Q_{bH}(H_\omega + H_UU^*_\omega) + Q_{bU}U^*_\omega) \right] \right\} \tag{A.1.2}
\]

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

\[
Q_b = -\frac{U}{b^2 u_q}, \\
Q_{bU} = -\frac{1}{b^2 u_q} + \frac{U}{b^2 u_q^2} u_{qq} Q_U \\
Q_{bH} = \frac{U}{b^2 u_q^2} (u_{qq} Q_H + u_{qh}).
\]

Assume that the \(\omega\)-households are located at \(x\). Differentiating \(u = U^*(\omega)/b\) with respect to \(\omega\) and using the budget constraint \(Q = \omega t(x) - H\Psi\) and (14), we obtain:

\[
[t - (H_\omega + H_UU^*_\omega)\Psi] q + (H_\omega + H_UU^*_\omega) h = \frac{U^*_\omega}{b}. \\
\]

Since

\[
-u_q \Psi + u_h = 0
\]

at the spatial equilibrium, we have:

\[
t = \frac{U^*_\omega}{bu_q}. \tag{A.1.3}
\]

Plugging this expression, \(Q_b\), \(Q_{bU}\) and \(Q_{bH}\) in (A.1.2), we get

\[
\Psi_{\omega x}(x, \omega, U^*(\omega)) = \frac{t}{H} \left\{ \frac{t_x}{t} \left[ 1 - \frac{\omega}{H}(H_\omega + H_UU^*_\omega) \right] \right\} - \frac{b_x}{U^*_\omega} \left[ \frac{H_\omega + H_UU^*_\omega}{H} \right] b \left[ \frac{H_\omega + H_UU^*_\omega}{H} \right] Q_b - \left( Q_{bH}(H_\omega + H_UU^*_\omega) + Q_{bU}U^*_\omega \right) \right\},
\]

which is equivalent to

\[
\Psi_{\omega x}(x, \omega, U^*(\omega)) = \frac{t}{H} \left\{ -T(x) \left[ 1 - \frac{\omega}{H}(H_\omega + H_UU^*_\omega) \right] \right\} + B(x) \left[ 1 - \frac{H_\omega + H_UU^*_\omega}{H} \frac{\omega U}{U^*_\omega} - \frac{\omega U}{u_{qq} U^*_\omega} (u_{qq} Q_H + u_{qh}) (H_\omega + H_UU^*_\omega) - \frac{U^*_\omega}{u_{qq} Q_U} \right\}. \tag{A1}
\]
Using
\[ \frac{du_q}{d\omega} = u_{qq}QH (H_\omega + H_U U^*_\omega) + u_{qq}QU U^*_\omega + u_{qh} (H_\omega + H_U U^*_\omega), \]
we can rewrite \( \Psi_{x,\omega} \) as follows:

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

(A.1.4)

which proves Proposition 1.

**A.2 Homothetic preferences**

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

\[ u_h = Ru_q. \]

It follows from Euler’s theorem that

\[ hu_h + qu_q = u \]

\[ \Leftrightarrow h \frac{u_h}{u} + q \frac{u_q}{u} = 1, \]

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 \( \omega \) yields:

\[
\frac{du_q}{d\omega} = \frac{1}{t} \left( \frac{du}{d\omega} \right) _\omega - u \]

\[ = \frac{u}{\omega^2 t} (\varepsilon_{U,\omega} - 1) \]

\[ = u q / \omega (\varepsilon_{U,\omega} - 1) \]

so that

\[ \varepsilon_{u_q,\omega} = 0. \]

In short, we have \( \varepsilon_{U,\omega} = 1, \varepsilon_{H,\omega} = \varepsilon_{h,\omega} = 1 \) and \( \varepsilon_{u_q,\omega} = 0. \)
A.3 Stone-Geary preferences

It is readily verified from (17) that

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

(A.3.1)

It follows from (A.3.1) that

\[ Q_U = \frac{1}{1 - \mu} U^{-\mu - 1} \left[ \frac{1}{b(h - \bar{h})^\mu} \right]^{1/\mu} = \frac{1}{(1 - \mu) U^\mu}, \]

\[ Q_{Ub} = -\frac{1}{(1 - \mu) b} U, \]

\[ Q_b \equiv -\frac{1}{b} Q_U, \]

\[ Q_h = -\frac{\mu}{1 - \mu} \left[ \frac{1}{b(h - \bar{h})} \right]^{1/\mu}, \]

\[ Q_{bH} = \frac{U}{(1 - \mu) h(h - \bar{h})^{-1} Q_U}. \]

Plugging \( Q_b, Q_{bH} \) and \( Q_{bU} \) into (A.1.2) and rearranging terms leads to

\[ \Psi_{xx}(x, \omega, U^*(\omega)) = \frac{t_x}{H} \left\{ \frac{t_x}{t} \left[ 1 - \frac{\omega}{H} (H_\omega + H_U U^*_\omega) \right] \right. \]

\[ + \frac{b_x}{b} \left[ H_\omega + H_U U^*_\omega \left( -\frac{U}{t} Q_U \right) \left( \frac{h - (1 - \mu)\bar{h}}{(1 - \mu)(h - \bar{h})} \right) \right. \]

\[ \left. + \frac{Q_U}{(1 - \mu) t} U^*_\omega \right\}. \]  

(A.3.2)

Plugging \( Q_h \) and \( Q \) in (12) and solving the corresponding equation yields

\[ \frac{h - (1 - \mu)\bar{h}}{(1 - \mu)(h - \bar{h})} = \omega t \left[ \frac{b}{U} (h - \bar{h})^\mu \right]^{1/\mu}. \]  

(A.3.3)

Given the expression of \( Q_U \), it turns out that

\[ \left( -\frac{U}{t} Q_U \right) \left[ \frac{h - (1 - \mu)\bar{h}}{(1 - \mu)(h - \bar{h})} \right] = -\frac{\omega}{1 - \mu}. \]  

(A.3.4)

Differentiating (13) with respect to \( \omega \) and using (12), we obtain:

\[ \Psi_{\omega}(x, \omega, U^*(\omega)) = \frac{t}{H} \left( 1 - \frac{Q_U}{t} U^*_\omega \right), \]  

(A.3.5)

which is equal to 0 if and only if

\[ U^*_\omega = \frac{t}{Q_U}. \]  

(A.3.6)

Using (A.3.4) and (A.3.6), (A.3.2) can be rewritten as follows

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

(A.3.7)
Applying the implicit function theorem to (A.4.3) yields

\[ H_U = \frac{(h - (1 - \mu)\bar{h})(h - \bar{h})}{U\mu h} \]

and

\[ H_\omega = -\frac{t(1 - \mu)^2}{\mu h}U^{-\frac{1}{1 - \rho}}b^{\frac{1}{1 - \rho}}(h - \bar{h})^{1 + \frac{1}{1 - \rho}}. \]

Given \( Q_U \), (A.3.6) can be expressed as the following differential equation:

\[ U_\omega^* = t \cdot (1 - \mu) \left[ b \cdot (h - \bar{h})^\mu \right]^{\frac{1}{1 - \rho}} (U^*(\omega))^{-\frac{1}{1 - \rho}}. \]  
(A.3.8)

We thus obtain

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

Therefore, by implication of (A.3.3), we have:

\[ 1 - \frac{\omega}{H}(H_\omega + H_U U_\omega^*) = \frac{(1 - \mu)\bar{h}}{H}. \]  
(A.3.9)

Substituting this expression into (A.3.7) yields:

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

Existence and uniqueness of the equilibrium housing consumption. The equilibrium housing demand satisfies (A.3.3). The LHS of (A.3.3) is decreasing and tends to +\( \infty \) when \( H \to \bar{h} \) and to \( 1/(1 - \mu) > 0 \) when \( H \to +\infty \). The RHS of (A.3.3) is increasing in \( H \). It tends to 0 when \( H \to \bar{h} \) and to +\( \infty \) when \( H \to +\infty \). Therefore, (12) has a single solution \( H(\omega t(x), U/b(x)) \), which implies that there exists a unique equilibrium.

Using (12), we may rewrite (13) as follows:

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

Using \( Q_H \) leads to

\[ \Psi(x, \omega, U) = \frac{\mu}{1 - \mu} \left( H - \bar{h} \right)^{\frac{1}{1 - \rho}} \left[ \frac{U^H}{b} \right]^{\frac{1}{1 - \rho}}. \]  
(B.3.10)

A.4 Proof of Proposition 2

The bid-max lot size. From the definition of the location-quality index given by (21), (A.3.3) can be rewritten as follows

\[ \frac{H - (1 - \mu)\bar{h}}{(1 - \mu)(H - \bar{h})} = \omega \Delta^{\frac{1}{1 - \rho}} \left[ \frac{(H - \bar{h})^\mu}{U} \right]^{\frac{1}{1 - \rho}}, \]  
(A.4.1)

which implies (23), so that the bid-max lot size depends on \( b(x) \) and \( t(x) \) through (21) only.
**Equilibrium utility level.** Applying the implicit function theorem to (A.4.1) yields

$$\frac{\partial H}{\partial \Delta} = - \left[ U^{1-t} (H - \bar{h})^{-\frac{1}{1-t}} \frac{\mu H}{(1 - \mu)} \right]^{-1} \omega \Delta^{\frac{1}{1-t}} < 0. \quad \text{(A.4.2)}$$

Using the definition of the location-quality index, the differential equation (A.3.8) satisfied by the equilibrium utility level writes

$$U^*_\omega = \Delta^{\frac{1}{1-t}} (1 - \mu)(h - \bar{h})^{\frac{\mu}{1-t}} (U^*(\omega))^{-\frac{\mu}{1-t}}. \quad \text{(A.4.3)}$$

**Supermodularity of the equilibrium utility level.** Differentiating (A.4.3) with respect to $\Delta$, we obtain:

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

Using (A.4.2), this expression may be rewritten as follows:

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

From (A.4.1), the expression in the bracketed term writes:

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

which implies

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

Since $\Psi_\omega(x, \omega, U^*(\omega)) = 0$, it follows from (A.3.5) that $U^*_\omega = t/Q_U$. Therefore, we obtain:

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

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

**A.5 The land rent and land gradient**

1. Rearranging (13) yields:

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

Using (A.3.6), 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 $Q_b$ into (A.1.1), we obtain:

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

and substituting $t$ by its expression given in (A.3.6) we obtain:

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

### A.6 The equilibrium land rent under Fréchet distributions

Rearranging terms in (19) yields:

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

and plugging the above expression into (A.3.10) leads to

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

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

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

Rearranging terms, this expression becomes:

$$
\Phi = \mu (1 - \mu)^{\frac{1}{1-\mu}} \left( \omega - \Phi \bar{h} \right)^{\frac{1}{2}} \left[ U(\omega) \right]^{-\frac{1}{2}} \Delta^{\frac{1}{2}}.
$$

(A.6.1)

Applying the first-order condition to $\Phi$ yields the following differential equation in $\omega$:

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

Let

$$
U^*(\omega) = (\omega - \Phi \bar{h}) X(\omega)
$$

(A.6.2)

be a solution to the above differential equation where $X(\omega)$ is determined below. Differentiating (A.6.2) with respect to $\omega$, we obtain

$$
U^*_\omega(\omega) = \left[ \frac{1}{\omega - \Phi \bar{h}} - \frac{\bar{h}}{\omega - \Phi \bar{h}} \right] X(\omega).
$$

Totally differentiating $\Phi$ leads to

$$
\Phi_\omega = \frac{d\Phi}{d\omega} = \frac{\partial \Phi}{\partial \omega} + \Phi_\Delta \Delta_\omega = \Phi_\Delta \Delta_\omega.
$$

(A.6.3)

Differentiating (A.6.1) with respect to $\Delta$ yields:

$$
\Phi_\Delta = \Phi \left[ \frac{1}{\mu} \Delta^{-1} - \frac{1}{\mu} \Phi \bar{h} (\omega - \Phi \bar{h})^{-1} \right],
$$
whose solution in $\Phi_\Delta$ is
\[ \Phi_\Delta = \frac{1}{\Delta} \frac{\Phi}{\mu} \left[ \frac{\mu (\omega - \Phi \bar{h})}{\mu (\omega - \Phi \bar{h}) + \bar{h} \Phi} \right]. \]

Therefore, we may rewrite (19) as follows:
\[ H\Phi = \mu (\omega - \Phi \bar{h}) + \bar{h} \Phi. \]  

Plugging (A.6.4) into $\Phi_\Delta$ leads to
\[ \Phi_\Delta = \frac{\omega - \Phi \bar{h}}{\Delta H}. \]

Using $\Phi_\omega$ and $\Delta_\omega$, (A.6.3) becomes:
\[ \Phi_\omega = \Phi_\Delta \Delta_\omega = \frac{1}{\gamma} \frac{\omega - \Phi \bar{h}}{\omega H} = \frac{1}{\gamma} \frac{(H - \bar{h}) \Phi}{\omega H} > 0. \]

Since $U_\omega(\omega)/U(\omega)$ is equal to $1/(\omega - \Phi \bar{h})$ in equilibrium, it must be that
\[ \frac{X_\omega(\omega)}{X(\omega)} = \frac{\bar{h}}{\omega - \Phi \bar{h}} \Phi_\omega = \frac{\bar{h}}{\omega - \Phi \bar{h}} \frac{1}{\gamma} \frac{(H - \bar{h}) \Phi}{\omega H}. \]

Therefore, using (A.6.4) leads to the following differential equation in $\omega$:
\[ X_\omega(\omega) = \frac{1}{\gamma} \frac{\bar{h}}{\omega H} X(\omega), \]

whose solution is
\[ X(\omega) = k \left( \frac{\omega}{H} \right)^{\frac{1}{1 - \mu}}, \tag{A.6.5} \]
where $k > 0$ is the constant of integration. Indeed, differentiating the above equation with respect to $\omega$ leads to
\[ X_\omega(\omega) = \frac{1}{(1 - \mu)\gamma} \frac{H - \omega (H + H U_\omega)}{H^2} \frac{H}{\omega} X(\omega). \]

Using (A.3.9), we obtain:
\[ X_\omega(\omega) = \frac{1}{(1 - \mu)\gamma} \frac{1}{H} \frac{(1 - \mu)\bar{h}}{\omega} X(\omega) = \frac{1}{\gamma} \frac{\bar{h}}{\omega H} X(\omega). \]

Substituting (A.6.5) into (A.6.2) yields:
\[ U(\omega) = (\omega - \Phi \bar{h}) k \left( \frac{\omega}{H} \right)^{\frac{1}{1 - \mu}}. \]

Plugging this expression into (A.6.1) and rearranging terms, we obtain the following implicit solution for the equilibrium land rent:
\[ R^*(x) = \mu (1 - \mu) \frac{\omega}{K} k \frac{1}{\gamma} t(x) \Delta \frac{1}{\Delta} \left[ \frac{\mu t(x)}{R^*(x)} + \frac{(1 - \mu)\bar{h}}{K^{1/\gamma} \Delta(x)} + (Y^*)^{1/\sigma} \right]^{\frac{1}{1 - \mu} \frac{\sigma}{\gamma}}. \tag{A.6.6} \]
Since the RHS of (A.6.6) is strictly decreasing and tends to 0 (∞) when $R(x) \to \infty$ (0), (A.6.6) has a unique solution in $R^*(x)$.

The lowest income in the sample, denoted by $\omega$, is strictly positive. It follows from (33) that the lowest location-quality index associated with the poorest household is given by

$$\Delta = \left( \frac{K_S}{K_\Delta} \right)^{-\gamma/\gamma_S} (Y^*)^{-\gamma/\sigma} (\omega)^{1/\gamma} > 0.$$  

The constant $k$ may be obtained by evaluating $R^*(x)$ at the least enjoyable location $x$ where $\Delta(x)$ reached its minimum $\Delta$. We assume that $\Delta$ is unique. Furthermore, the land rent at $x$ is equal to the opportunity cost of land, $R_0$. Therefore, it is readily verified that $k$ is given by

$$k^{-1} \mu = R_0(1 - \mu)\frac{1 - \mu}{\mu} \left[ \frac{\mu t(x)}{R_0} \right] < \min, \quad \|J(F)\|_1$$

is smaller than 1. In other words, when $\rho > 0$ is small enough, $F(A)$ is a contraction.

A.7 The spatial equilibrium under agglomeration economies

We determine under which conditions the spatial equilibrium is unique. The function $F(A)$ is not differentiable over the the finite set $I$ of locations where $t_i(x) = t_j(x)$. It is well known that uniqueness holds when $F(A)$ is a contraction. This is so when the matrix norm $\|\cdot\|_\infty$ of the Jacobian $J(F)$ of $F(A)$, which is defined on the open set $]0, A_1[ \times \ldots \times ]0, A_n[ - \{I\}$, is smaller than 1.

Differentiating $F_i(A)$ with respect to $A_k$ yields the following expression over $]0, A_1[ \times \ldots \times ]0, A_n[ - \{I\}$:

$$\frac{\partial F_i(A)}{\partial A_k} = \delta_{ik} \cdot \left\{ \int_0^{\zeta} \frac{K_i[t_i(x)]^\epsilon}{\sum_{j=1}^n K_j[t_j(x)]^\epsilon} \frac{K_i[t_i(x)]^\epsilon}{\sum_{j=1}^n K_j[t_j(x)]^\epsilon} \right\}^{-\frac{\delta-1}{\delta}} \cdot \left\{ \int_0^{\zeta} \frac{K_i[t_i(x)]^\epsilon}{\sum_{j=1}^n K_j[t_j(x)]^\epsilon} \frac{\partial \ell^*(A)}{\partial A_k} \right\}$$

Since the two bracketed terms in the right-hand side of this expression are continuous, $\partial F_i(A)/\partial A_k$ has a supremum $C_{ik} \neq 0$. Therefore, we have:

$$\sum_{k=1}^n \left| \frac{\partial F_i(A)}{\partial A_k} \right| < \delta_{ik} \sum_{k=1}^n |C_{ik}| < 1,$$

where the second inequality holds for all $0 < \delta < \delta_k \equiv 1/(\delta_k \Sigma_k |C_{ik}|)$. Let $\delta_{\min}$ be the minimum of $\delta_k$ over $k = 1, \ldots, n$. If $\delta < \delta_{\min}$, $\|J(F)\|_\infty$ is smaller than 1. In other words, when $\delta > 0$ is small enough, $F(A)$ is a contraction.
A.8 The real wage under Stone-Geary preferences

With a Stone-Geary utility function, we have

\[ U = b \cdot u(q, h) \]

\[ u = (1 - \mu)^{(1-\mu)} \mu^{-\mu} q^{1-\mu} (h - \bar{h})^{\mu} \]  

(52)

and the budget constraint is given by \[ q + Rh = \omega t. \] The price index under Stone-Geary preferences is given by

\[ P = R^\mu \frac{\omega t}{\omega t - Rh}. \]

**Proof.** Inserting the equilibrium consumption of numéraire and housing in (52) yields the indirect utility of consumption:

\[ u^* = (\omega t - R\bar{h}) R^{-\mu} \]

\[ = \omega t R^{-\mu} \frac{\omega t - R\bar{h}}{\omega t}. \]

Hence, total expenditures are given by

\[ \omega t = u^* R^\mu \frac{\omega t}{\omega t - R\bar{h}} \]

so that the price index is

\[ P = R^\mu \frac{\omega t}{\omega t - R\bar{h}}. \]

Because the Stone-Geary utility function is non homogeneous, the price index \( P \) depends on income and varies across individuals.

Hence, the real wage is

\[ \frac{\omega t}{P} = \frac{\omega t - R\bar{h}}{R^\mu} \]

**Remark.** Defining the variable \( \Pi = h - \bar{h} \) which represents consumption level in excess of "subsistence" and the variable \( m = \omega t - R\bar{h} \) which is the supernumerary income, \( u(q, h) \) becomes the standard Cobb-Douglas:

\[ u(q, h) = (1 - \mu)^{(1-\mu)} \mu^{-\mu} q^{1-\mu} \Pi^{\mu} \]

and the budget constraint is now expressed as \( q + R\Pi = m \). Hence, the real wage is

\[ \frac{m}{R^\mu} = \frac{\omega t - R\bar{h}}{R^\mu} \]

Appendix B

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

The second part of this appendix reports various additional econometric results. First, we report bias-corrected estimates using Oster’s (2019) methodology in Appendix B.6. Second, first-stage results in Appendix B.7. We undertake additional robustness checks, in Appendix B.8. In Appendix B.9 we outline the procedure to solve for counterfactual outcomes of the model. Appendix B.10 discusses the outcomes of a few additional counterfactual analyses.

**B.1 Network distances**

We obtain information on network distances from the *SpinLab* which enable us to calculate travel time \( \tau \) 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 neighborhood we calculate the straight-line distance to the nearest access points on the network and then calculate the network distance. The median distance from an observation in the dataset to the nearest access point of the network is 195m (the average is 326m). We assume that the average speed to get to the nearest access points is 10km/h. This is the speed based on the Euclidian distance; in reality the distance to get to the network will be higher because streets are usually curved. Hence, the assumption of 10km/h seems reasonable as the minimum speed on roads in the network is 20km/h. Furthermore, because of the dominance of the bicycle, this would be close to the average cycling speed. Using these information, we calculate the total driving time, which is the sum of the driving time to access the network, the network driving time and the time it takes from the network to arrive at the destination. Alternatively, we calculate for each location pair the Euclidian distance and assume again an average speed of 10km/h.

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

[Figure B.1 about here]

The correlation between travel time and Euclidian distance is modest \( (\rho = 0.643) \). For short distances \(< 10 \text{km}\) the correlation is, however, much higher \( (\rho = 0.862) \). We plot the relationship between distance and travel time in Figure B.1.A. This relationship is monotonic. Figure B.1.B shows the share of people commuting commute maximally \( \tau \) minutes, which we use to calculate employment accessibility in 1900.
B.2 Land prices and lot sizes

Information on land values and lot sizes is not directly available but may be inferred from data on home sales. We use information on home sales from NVM (The Dutch Association of Realtors), which comprises the large majority (about 75%) of owner-occupied house transactions between 2003 and 2017. We know the transaction price, the lot size, inside floor space size (both in m$^2$), the exact address, and a wide range of housing attributes such as house type, number of rooms, construction year, garden, state of maintenance, and whether a house is equipped with central heating. We make some selections to make sure that our results are not driven by outliers. First, we exclude transactions with prices that are above €1 million or below €25,000 and have a price per square meter which is above €5,000 or below €500. We also leave out transactions that refer to properties that are larger than 250m$^2$ or smaller than 25m$^2$, or have lot sizes above 5000m$^2$. These selections consist of less than 1% of the data and do not influence our results. We follow a similar procedure as Rossi-Hansberg et al. (2010), implying that we can only use information on residential properties with land. We are left with 1,337,445 housing transactions.

Let $P_{xy}$ denote the house price in year $y$, $H_{xy}$ the observed lot size and $C_{xy}$ the housing characteristics of property $x$. The log land rent $R_x$ is equal to the fixed effects at the level of the postcode (about 15 – 20 addresses), which we denote by $\varsigma_x$, while $\vartheta_y$ denote year $y$ fixed effects. For each city, we estimate:

$$\log \frac{P_{xy}}{H_{xy}} = \eta_1 C_{xy} + \varsigma_x + \vartheta_y + \epsilon_{xy},$$

(C.2.1)

where $\epsilon_{xy}$ is an identically and independently distributed error term that is assumed to be uncorrelated to land rents and housing characteristics, while $\eta_1$ are parameters to be estimated. As log $R_x$ are given by the very local fixed effects, we do not impose any structure on how land rents $R_x$ vary across locations. For about 80% of the data we do not observe land prices directly, because either there were no multiple sales in our study period or because there is no owner-occupied housing in the respective postcode. We therefore also estimate the above equation with neighborhood fixed effects instead.

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

Descriptive statistics for the housing sample are reported in Table B.1. Coefficients $\eta_1$ related to the housing attributes are reported in Table B.2. It appears that the house price per square meter of land is generally a bit lower when the property is larger. However, the house price per square meter of land of properties that are (semi-)detached is generally higher. Furthermore, when the maintenance state of a property is good, prices are about 502/1269 = 40% higher. When a property has central heating, the price per square meter is about 5.1% higher. The dummies related to the construction decades show the expected signs. Properties constructed after World War II until 1970 generally have lower prices because this is a period associated with a lower
building quality. Lot sizes are inversely related to pattern of land prices ($\rho = -0.245$). In other words, more expensive locations generally have smaller lots, which makes sense.

B.3 Amenities

Hedonic amenity index. We test whether our results are robust to using an alternative hedonic amenity index, rather than relying on geocoded pictures. Following Lee and Lin (2018), we aim to construct an aggregate amenity index that describes the amenity level in every neighborhood $x$.\(^{18}\) We will make a distinction between historic amenities and natural amenities.

Let $A_\tilde{x}$ be a set of variables that describe amenities of property $\tilde{x}$ (so the location is more detailed than the neighborhood $x$). For example, we calculate the share of historic districts, the number of listed buildings, water bodies and open space within 500m of each property. Let $P_\tilde{x}$ the house price, while $C_\tilde{x}$ are housing characteristics at location $\tilde{x}$, and $\vartheta_y$ are year $y$ fixed effects. We also include neighborhood fixed effects $\varsigma_x$, so we identify the effects of amenities on prices within neighborhoods. We then estimate:

$$\log P_\tilde{x} = \eta_0 A_\tilde{x} + \eta_1 C_\tilde{x} + \vartheta_y + \varsigma_x + \epsilon_{\tilde{x}y},$$

(B.3.1)

where $\eta_0$ and $\eta_1$ are parameters to be estimated and $\epsilon_{\tilde{x}y}$ is an identically and independently distributed error term. We then use $\hat{\eta}_0$ and $A_x$ to predict the amenity level in each location $x$ in the Randstad:

$$\tilde{b}_x = \frac{1}{N_x} \sum_{x=1}^{N_x} \hat{\eta}_0 A_\tilde{x},$$

(B.3.2)

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

We use data on the universe of housing transactions in the Netherlands between 2010 and 2015 from the NVM. Additional descriptive statistics of the NVM data are reported in Table B.3. We have 695,709 observations and the average house price is €229 thousand.

[Tables B.3 and B.4 about here]

In Table B.4 we report the results of the regression of equation (B.3.1). We first investigate the impact of listed buildings. It can be seen that the share of historic districts leads to higher price. A 10 percentage point increase in the share of land designated as historic district increases prices by 1.8%. Listed buildings do have a small additional effect of 0.5% per listed building. In column

\(^{18}\)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 also interested in intra-city variation in amenities. 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.
(2) we investigate the impact of water bodies and open space. For a 10 percentage point increase in water bodies, prices rise by 3%. Moreover, a 10 percentage point increase in open space implies a price increase of 0.6%, so this effect is considerably smaller. When we put historic amenities and natural amenities together, the coefficients are essentially unaffected. We consider this as the preferred specification. In the last specification we investigate whether the results change when we include endogenous amenities, such as shops, cafés, and leisure establishments. This appears not to be the case. Only hotels restaurants and cafés have a statistically significant impact on prices, which suggests that exogenous amenities related to the built environment and land use are more important than endogenous amenities.

B.4 Historic data

To instrument current amenity levels and commuting time we use information on land use, the railway network and amenities in 1900. For the 1900 land use maps, Knol et al. (2004) have scanned and digitized maps into 50 by 50 meter grids and classified these grids into 10 categories, including built-up areas, water, sand and forest. We aggregate these 10 categories into built-up, open space and water bodies. Knol et al. document large changes in land use across the Netherlands from 1900 to 2000. For example, the total land used for buildings has increased more than fivefold. On the other hand, the amount of open space has decreased by about 10%. We also use information on municipal population in 1900 from NLGIS. Municipalities were much smaller at that time and about the size of a large neighborhood nowadays. We impute the local population distribution using the location of buildings and assuming that the population per building is the same within each municipality. We further use information on railway stations from Koopmans et al. (2012). We enrich these data by adding missing stations from various sources on the internet and create a network with travel times. To approximate the speed, we fit a regression of the length of (current) railway segments between stations on current travel time on the railway network. Based on historic sources, it appears that the average speed is about 50% of what it is currently (about 70km/h).

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

[Figure B.2 about here]

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, Rotterdam, Leiden, Delft, Hoorn, as well as for the province of Utrecht, Drenthe, Groningen, Friesland, Overijssel, Gelderland, and parts of Noord-Brabant. The HISGIS data also provide information on the cadastral income for about one-third of the observations, which was used to determine the tax at that time and is a proxy for land values. In Panel A of Figure B.3 we show 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 rely on municipal population data from NLGIS to calculate the accessibility in 1832. We assume that population is uniformly distributed within the municipality. Rail networks did not exist yet, so in order to calculate the population that could be reached within commuting time, we use information on the road network from 1821 obtained from Levkovich et al. (2017). Panel B of Figure B.3 shows the network back then.

In Table B.5 we provide descriptives for all instruments. The average share of built-up area in 1900 was 4.3%, while it was 4.2% in 1832. However, this figure is a bit misleading because for 1832 we have more data near urban areas. On average about 38 thousand jobs and 89 thousand people could be reached within commuting distance in 1900. Not surprisingly, this was much lower (40 thousand) in 1832.

B.5 Other descriptive statistics

In Figure B.4 we report the distributions of the log of income and the log of land price. The distributions of land prices is somewhat positively skewed.

In Figure B.5 we show maps of income and land price distributions across the Netherlands. As expected, land prices are generally higher in cities. The pattern for incomes is less clear, but generally speaking we find that wealthier households locate close to or in cities.

B.6 Bias-corrected estimates

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

Oster (2019) then derives a GMM estimator to derive bias-corrected estimates of the impact of amenities and employment accessibility on gross incomes. There are two key input parameters that have to be determined. First, there is the maximum $R^2$ from a hypothetical regression of income on amenities, accessibility and controls, which we denote as $\tilde{R}$. Given that our variables are neighborhood-specific, rather than household-specific variables, $\tilde{R}$ is likely to be much smaller than 1. Second, a parameter must be chosen that determines the relative degree of selection on observed and unobserved variables, which we denote by $\varpi$. Although this parameter is fundamentally unknown, Altonji et al. (2005) and Oster (2019) show that $\varpi = 1$ is a reasonable (upper-bound) value. Oster (2019) then shows that

$$\alpha_1^* \approx \hat{\alpha}_1 - \varpi [\hat{\alpha}_1 - \hat{\alpha}_1] \frac{R^2 - \tilde{R}^2}{R^2 - \tilde{R}^2} \quad \text{and} \quad \alpha_2^* \approx \hat{\alpha}_2 - \varpi [\hat{\alpha}_2 - \hat{\alpha}_2] \frac{R^2 - \tilde{R}^2}{R^2 - \tilde{R}^2}, \quad \text{(B.3.2)}$$

where $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are parameter estimates obtained from a regression with controls (say household, job and housing controls, as well as workplace fixed effects), and $R^2$ is the corresponding $R^2$. $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$ are parameter estimates obtained from a regression without controls and $\tilde{R}_1^2$ and $\tilde{R}_2^2$ are the corresponding $R^2$s. Hence, this equation provides a simple way to evaluate robustness of the results. We then report bootstrapped bias-corrected estimates in Table B.6 of the coefficients of interest. We replicate the first three specifications reported in Table 3.

[Table B.6 about here]

In columns (1)-(3) of Table 6, we naively assume that in theory we can fully explain variation in wages, so that $\tilde{R} = 1$. Given this assumption, we find in column (1) – where we only include household controls and year fixed effects – that the effect of amenities is about 10 times as strong, and the effect of employment accessibility is about twice as strong as in the corresponding OLS specification. The effect of amenities become even stronger once we add housing and job controls in column (2) and is again comparable to column (1) once we add workplace fixed effects. This may lead to the conclusion that the OLS estimates are not robust and subject to omitted variable bias. However, the assumption that $\tilde{R} = 1$ is likely to be wrong because the dependent variable is a variable measured at the micro-level (the household), while amenities and employment accessibility are measured at the neighborhood level. Hence, the maximum attainable $\tilde{R}$ when omitted variables are important is likely substantially lower. To determine $\tilde{R}$ we therefore run a regression of income
on household, job and housing controls, as well as residence, workplace and year fixed effects. This leads to an \( R^2 \) of 0.357, which is considerably smaller than 1. Moreover, it is around the value of \( \bar{R} = 1.3\bar{R}^2 \), which is supported by experimental data (Oster, 2019).

Columns (4)-(6) then show that the effect of amenities and employment accessibility are very similar to the OLS estimates. In our preferred specification, we find an elasticity of 0.0239, which is close to 0.0166 found in the OLS specifications. For employment accessibility we find an elasticity of 0.0936, which is essentially the same as 0.0881 reported in the corresponding OLS specification.

In other words, these results strongly suggest that omitted variable bias is not a major issue. Having said this, Oster’s (2019) methodology does not account for measurement error in amenities or employment accessibility or reverse causality. It is therefore still important to apply our instrumental variables strategy.

**B.7 First-stage results**

We report first-stage estimates in Table B.7. In column (1) we use contemporary instruments for amenities. We show that current proxies for amenities are strongly positively correlated to picture density. For example, when the number of listed buildings per hectare increases by 1, picture density increases by 12.7%. Also the share of a neighborhood designated as historic district is positively correlated to the picture density. Furthermore, we find positive correlations with the share of built-up land and water bodies located in the neighborhood. Hence, picture density seems a meaningful proxy for amenities.

[Table B.7 about here]

In column (2) we use historic instruments. This means that we use the share of built-up land in 1900 and share of water in 1900 as instruments for picture density. We find strong positive effects of the share of built-up land in 1900 on picture density. This effect is about twice as strong as the share of contemporary built-up land, likely because the share of built-up land in 1900 is positively correlated to the current intensity of historic amenities.

Column (3) also includes the instruments for employment accessibility: the share of built-up land in 1900 within 500m, the share of built-up land in 1900 between 500 and 1000m and, most importantly, employment accessibility. This leaves the effects of the share of built-up land in 1900 in the own neighborhood almost unaffected.

In column (4) we take employment accessibility as dependent variable. The instruments for accessibility are relevant. We find a strong positive effect of the share of built-up land in 1900 between 500 and 1000m on accessibility, which makes sense. Also employment accessibility in 1909 has a strong positive effect on current employment accessibility. More specifically, doubling employment accessibility in 1909 is associated with an increase in current employment accessibility of 29%.
B.8 Sensitivity checks for the reduced-form income mapping

Table B.8 reports the results of additional robustness checks. Our dataset contains observations on households. When calculating the commuting elasticity and when including workplace fixed effects, we focus on the job that generates the most working hours. This may be problematic when more people are employed in the household that work in different location. In column (1) we therefore only include households that are associated with one job. This does not lead to material differences in outcomes. When calculating the commuting time, we calculate the commuting time to the nearest plant of a firm, if it has multiple establishments. We test whether this introduces error by only including households that are associated with one job in a single plant firm in column (2). In this way we address any measurement error in commuting time. Again, the estimates are very similar.

[Table B.8 about here]

Our measures of commuting time rely on the minimum of travel time on the road and rail. However, in almost all cases travel time over the road is shorter. To make sure that households actually consider this travel time, we only keep households having a company car in column (3). This does not materially change the results. Column (4) replaces the dependent variable income by the share of adults in the household that have a college degree or more. We find very similar effects. For example, when the picture density doubles this increases the share of highly educated households by 3.3 percentage points. Conversely, doubling commuting times decreases the share of highly educated households by 20.3 percentage points.

Column (5) tests whether the results are robust when using commuting time by rail instead of commuting time over the road or rail. The results are comparable. Overall, the impact of amenities and commuting time on income sorting choice is robust.

B.9 Counterfactual analyses

We outline the procedure for the three counterfactual analyses discussed in Section 8, which we denote $c = 1, 2, 3$, respectively.

1. The first step is to determine the location-specific scale parameters $K_i$, and productivity endowments $A_i$. We set $\kappa_{x_i} = 1$ and use the estimated $\bar{\Omega}_i$ and $\Omega_i$ to obtain

   $$ K_i = e^{\bar{\Omega}_i - \Omega_i}, \quad A_i = e^{\gamma_i \Omega_i} / \bar{L}_i^{\gamma_i}. $$

2. We build the values for counterfactual scenario $c = 1, 2, 3$ for commuting times $\tau_{x_i}^c$, exogenous amenities $\bar{b}_x^c$, and productivity endowments $A_i^c$. If values do not vary for the specific scenario under consideration, we take the values from the data. Moreover, we set the starting values for $L_i^c$ equal to the estimated value from the data and the initial value for the parameters $\gamma^c$. 
$\hat{\gamma}_\Delta^c$, $K_\Delta^c$, $K_S$ to the values obtained in the structural estimation. We treat the parameters $\hat{\kappa}$, $\hat{\varepsilon}$, $\hat{\mu}$, $\hat{\beta}$, and $\hat{\delta}$ as given and obtain them from the structural estimation results.

3. We calculate labor productivity $t_{xi}^c = \left[ K_i^c (L_i^c)^{\delta} (\tau_{xi}^c)^{\hat{\kappa}} \right]^{(\sigma-1)/\sigma}$ for each location pair $(xi)$, as well as the accessibility $\tilde{a}_{x}^c = \sum_{i=1}^{n} \tilde{t}_{xi}^c = \sum_{i=1}^{n} K_i (t_{xi}^c)^{\hat{\delta}}$ of location $i$.

4. We calculate the location-quality indices:

$$\Delta_x^c = (\tilde{a}_x^c)^{\hat{\gamma}} \Gamma \left( \frac{\hat{\varepsilon} - 1}{\hat{\varepsilon}} \right) \left( \tilde{a}_x^c \right)^{\frac{1}{\hat{\varepsilon}}}.$$  

5. We fit a Fréchet distribution to $\Delta_x^c$ to obtain the adjusted values of the shape parameter $\hat{\gamma}_\Delta^c$.

Since the aggregate skill distribution is given, it must be that $\hat{\gamma}^c = \hat{\gamma}_\Delta^c / \hat{\gamma}_S$.

6. We determine the skill mapping $s_x^c = \left[ (K_S^c/K_\Delta^c)^{1/\hat{\gamma}_S} (\Delta_x^c)^{\hat{\gamma}^c} \right]^{\sigma/(\sigma-1)}$ and re-adjust $K_\Delta$ for the geometric mean of $s_x^c$ to remain equal to 1. This step is important because the aggregate skill distribution does not change under our counterfactual scenarios. Hence, $K_\Delta^c$, $\hat{\gamma}_S$ and the geometric mean should not change in the different counterfactuals.

7. We calculate total counterfactual labor supply in each employment location $i$. We have:

$$L_x^c = \sum_{x=1}^{\ell_x^c} \frac{\tilde{t}_{xi}^c}{\sum_{j=1}^{t_{xi}^c}} f(\tilde{s}_x^c).$$

where

$$f(\tilde{s}_x^c) = \frac{\sigma - 1}{\sigma} \hat{K}_S \hat{\gamma}_S e^{-\hat{K}_S (\tilde{s}_x^c)^{-\hat{\gamma}_S} - \hat{K}_S^c (\tilde{s}_x^c)^{-\hat{\gamma}^c - 1}} \left( \tilde{s}_x^c \right)^{-\hat{\gamma}_S(\sigma-1)+\sigma} / \sigma,$$

is the skill density. Since $L_x^c$ is an input to Step 3, we repeat steps (3)-(7) until $L_x^c$ converges, which is usually within 10 iterations.

8. We now have all the information to solve for the total output in the city:

$$Y^c = \left[ \sum_{i=1}^{n} \sum_{x=1}^{\ell_x^c} K_i \left[ t_{i}^c (x) \right]^{\hat{\varepsilon}} t_{xi}^c (\tilde{s}_x^c)^{\frac{\sigma-1}{\sigma}} f(\tilde{s}_x^c) \right]^{\frac{\sigma-1}{\sigma}}.$$  

9. We also determine the income mapping $\omega_x^c t_{xi} = \left( K_S^c / K_\Delta^c \right)^{1/\hat{\gamma}_S} (\Delta_x^c)^{1/\hat{\gamma}^c} (Y^c)^{1/\sigma} (t_{xi})$, which enables us to determine the land rent at each location $x$:

$$R_x^c = R_0 \left[ \frac{\Delta_x^c}{\Delta_x} \right]^{\hat{\mu} / \hat{\mu}^c} \left[ \frac{\hat{\mu}^c R_x^c + (1-\hat{\mu}) x}{\hat{\mu} R_0^c + (1-\hat{\mu}) x} \right]^{1/(\hat{\mu}^c)}.$$

where $x$ is the location where the poorest household (with the lowest $\omega_x$) lives, while $R_0$ is the agricultural land rent. We do not have good data on agricultural land prices. In any case, these will be not very useful as agricultural land prices in the Netherlands are highly regulated. We therefore set $R_0$ equal to the 5th percentile value of the observed land rents in our data. We use a standard Newton-Raphson procedure to determine the solution $R_x^c$.  

A18
10. We find consumption level of the composite good \( q^c_x = (1 - \hat{\mu}) [\omega^c_{x} t^c_{x} - R^c_x \hat{h}] \) and the housing consumption \( h^c_x = (1 - \hat{\mu}) \hat{h} + \hat{\mu} \omega^c_{x} t^c_{x} / R^c_x \), which is identified up to multiplication constant, so that the utility level is given by \( u^c_x = (q^c_x)^{1-\hat{\mu}} (h^c_x - \hat{h})^{\hat{\mu}} \). This enables us to determine the aggregate land rent and aggregate real income:

\[
ALR^c = \sum_{x=1}^{c^*} h^c_x R^c_x, \quad \text{and} \quad ARI^c = \sum_{x=1}^{c^*} \frac{1}{h^c_x} \frac{\omega^c_{x} t^c_{x} - R^c_x \hat{h}}{(R^c_x)^{\hat{\mu}}},
\]

where \( 1/h^c_x \) is the density of households in neighborhood \( x \). We derive the real income \((\omega^c_{x} t^c_{x} - R^c_x \hat{h}) (R^c_x)^{\hat{\mu}}\) in Appendix A.8.

### B.10 Other counterfactual analyses

We consider three more counterfactual analyses. First, we assume that households do not value amenities \((\beta^c = 0)\). Second, we assume that the commuting decay is less substantial \((\kappa^c = \hat{\kappa}/2)\). Last, we assume that agglomeration economies are much stronger \((\delta^c = \hat{\delta} \times 5)\). We report the main results in Table B.10.

![Table B.10 about here](image)

The first scenario \((\beta^c = 0)\) replicates the scenario with uniform amenities over space \((c = 1)\). Therefore, the overall output increases is still equal to 10.6\%\, while aggregate real incomes rise by 7.3\%. The aggregate land rent decreases by 0.6\%\, when there are no amenities. The spatial implications are the same as those reported in Figure 3.

In the second scenario \((\kappa^c = \hat{\kappa}/2)\), we observe a substantial increase in output of 50.9\%. Note that reducing \(\kappa\) by 50\% implies much bigger drop in commuting costs than a 50\% decrease. Because of the substantial reduction in commuting costs, we observe a high increase in the aggregate real income of 103\%. By contrast, we witness a modest increase in the aggregate land rent of 5.6\%. Since accessibility is now much less important in locational choices, households substitute accessibility for the consumption of the composite good.

Despite the stark decrease in commuting costs, the spatial implications of the reduction in \(\kappa\) are limited as we find a correlation between the baseline and counterfactual skills mapping of 0.997. We show in Figure B.6.A that the spatial differences are at most 5\%. Maximal rent differences are again larger (up to 19\%, see Figure B.6.B).

![Figure B.6 about here](image)

In the last scenario, we consider stronger agglomeration economies, that is, we multiply the estimated agglomeration elasticity by 5. Even then, the model converges to a unique equilibrium, which shows that Proposition 3 still holds. Stronger agglomeration economies have substantial
implications: the output increases more than twelvefold, which is quite unrealistic. Aggregate real incomes rise by 416%, while land rents decrease by 63%.

Despite the stronger intensity of agglomeration economies, we observe limited changes in the spatial skill distribution. The correlation with the baseline spatial skill distribution is 0.986. The correlation of land rents is smaller and land rent differences can be much more substantial. The spatial distribution of jobs is now considerably different from that of the baseline model as the correlation is only 0.479. Jobs are now more much concentrated than in the baseline scenario. This counterfactual reinforces our conclusion that the spatial distribution of skills is mainly driven by amenities, rather than commuting costs or agglomeration economies.
### Table B.1 – Descriptives for NVM dataset

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>sd</td>
<td>min</td>
<td>max</td>
</tr>
<tr>
<td>Price (in € per m²)</td>
<td>1,269</td>
<td>927.2</td>
<td>25</td>
<td>25,000</td>
</tr>
<tr>
<td>Lot size (in m²)</td>
<td>445.7</td>
<td>1,189</td>
<td>25</td>
<td>25,000</td>
</tr>
<tr>
<td>Size of property (in m²)</td>
<td>132.4</td>
<td>45.16</td>
<td>26</td>
<td>538</td>
</tr>
<tr>
<td>Number of rooms</td>
<td>4.944</td>
<td>1.363</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>Terraced property</td>
<td>0.417</td>
<td>0.493</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Semi-detached property</td>
<td>0.370</td>
<td>0.483</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Detached property</td>
<td>0.189</td>
<td>0.392</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Private parking space</td>
<td>0.454</td>
<td>0.498</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Garage</td>
<td>0.394</td>
<td>0.489</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Garden</td>
<td>0.966</td>
<td>0.182</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Number of bathrooms</td>
<td>0.929</td>
<td>0.483</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Number of kitchens</td>
<td>0.677</td>
<td>0.484</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Number of balconies</td>
<td>0.132</td>
<td>0.354</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Number of roof terraces</td>
<td>0.0674</td>
<td>0.257</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Number of floors</td>
<td>2.717</td>
<td>0.636</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>Internal office space</td>
<td>0.00444</td>
<td>0.0665</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Maintenance score of the outside</td>
<td>0.758</td>
<td>0.131</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Maintenance score of the inside</td>
<td>0.753</td>
<td>0.143</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Number of types of insulation</td>
<td>2.381</td>
<td>1.831</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Central heating</td>
<td>0.920</td>
<td>0.271</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Listed building</td>
<td>0.00652</td>
<td>0.0805</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Newly built property</td>
<td>0.0417</td>
<td>0.200</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Construction year</td>
<td>1.967</td>
<td>34.95</td>
<td>1,362</td>
<td>2,017</td>
</tr>
<tr>
<td>Year of observation</td>
<td>2,011</td>
<td>4,389</td>
<td>2,004</td>
<td>2,017</td>
</tr>
</tbody>
</table>

Notes: The number of observations is 1,337,495. Because of confidentiality restrictions the minimum and maximum values refer to the 0.01% and 99.99% percentile. This implies that we exclude the bottom and top observations.
Table B.2 – Estimating land prices and lot sizes  
(Independent variable: the log of land price per m²)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rooms</td>
<td>-6.1664***</td>
</tr>
<tr>
<td></td>
<td>(0.4506)</td>
</tr>
<tr>
<td>Terraced property</td>
<td>702.4875***</td>
</tr>
<tr>
<td></td>
<td>(6.5087)</td>
</tr>
<tr>
<td>Semi-detached property</td>
<td>510.0447***</td>
</tr>
<tr>
<td></td>
<td>(6.5516)</td>
</tr>
<tr>
<td>Detached property</td>
<td>360.7740***</td>
</tr>
<tr>
<td></td>
<td>(6.7580)</td>
</tr>
<tr>
<td>Private parking space</td>
<td>-56.3558***</td>
</tr>
<tr>
<td></td>
<td>(1.9988)</td>
</tr>
<tr>
<td>Garage</td>
<td>-42.8160***</td>
</tr>
<tr>
<td></td>
<td>(2.0556)</td>
</tr>
<tr>
<td>Garden</td>
<td>47.5907***</td>
</tr>
<tr>
<td></td>
<td>(2.8356)</td>
</tr>
<tr>
<td>Number of bathrooms</td>
<td>17.3274***</td>
</tr>
<tr>
<td></td>
<td>(0.9885)</td>
</tr>
<tr>
<td>Number of kitchens</td>
<td>-7.2575***</td>
</tr>
<tr>
<td></td>
<td>(1.0818)</td>
</tr>
<tr>
<td>Number of balconies</td>
<td>47.8147***</td>
</tr>
<tr>
<td></td>
<td>(1.5204)</td>
</tr>
<tr>
<td>Number of roof terraces</td>
<td>109.0801***</td>
</tr>
<tr>
<td></td>
<td>(1.8878)</td>
</tr>
<tr>
<td>Number of floors</td>
<td>94.9407***</td>
</tr>
<tr>
<td></td>
<td>(1.0148)</td>
</tr>
<tr>
<td>(Internal) office space</td>
<td>-55.3454***</td>
</tr>
<tr>
<td></td>
<td>(6.3595)</td>
</tr>
<tr>
<td>Maintenance score of the outside</td>
<td>29.5137***</td>
</tr>
<tr>
<td></td>
<td>(6.3366)</td>
</tr>
<tr>
<td>Maintenance score of the inside</td>
<td>501.7345***</td>
</tr>
<tr>
<td></td>
<td>(5.8143)</td>
</tr>
<tr>
<td>Number of types of insulation</td>
<td>8.3945***</td>
</tr>
<tr>
<td></td>
<td>(0.3138)</td>
</tr>
<tr>
<td>Central heating</td>
<td>65.8404***</td>
</tr>
<tr>
<td></td>
<td>(1.7719)</td>
</tr>
<tr>
<td>Listed building</td>
<td>27.9334***</td>
</tr>
<tr>
<td></td>
<td>(6.2691)</td>
</tr>
<tr>
<td>Newly built property</td>
<td>-13.3758***</td>
</tr>
<tr>
<td></td>
<td>(4.3108)</td>
</tr>
<tr>
<td>3rd-order polynomial of property size</td>
<td>Yes</td>
</tr>
<tr>
<td>Construction decade dummies</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>Yes</td>
</tr>
<tr>
<td>Postcode fixed effects</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1,280,031</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.8295</td>
</tr>
</tbody>
</table>

Notes: Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. 
### Table B.3 – Other descriptive statistics for NVM data

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>sd</td>
<td>min</td>
<td>max</td>
</tr>
<tr>
<td>Price of home (in €)</td>
<td>229,238</td>
<td>116,074</td>
<td>25,000</td>
<td>1,000,000</td>
</tr>
<tr>
<td>Share land in historic district &lt;500m</td>
<td>0.0695</td>
<td>0.192</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Listed buildings &lt;500m</td>
<td>0.179</td>
<td>0.894</td>
<td>0</td>
<td>19.53</td>
</tr>
<tr>
<td>Share water bodies &lt;500m</td>
<td>0.0411</td>
<td>0.0713</td>
<td>0</td>
<td>0.920</td>
</tr>
<tr>
<td>Share open space &lt;500m</td>
<td>0.244</td>
<td>0.217</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Shops, &lt;500m</td>
<td>0.254</td>
<td>0.394</td>
<td>0</td>
<td>4.711</td>
</tr>
<tr>
<td>Hotels, restaurants, cafés &lt;500m</td>
<td>0.159</td>
<td>0.364</td>
<td>0</td>
<td>7.983</td>
</tr>
<tr>
<td>Leisure establishments &lt;500m</td>
<td>0.0127</td>
<td>0.0215</td>
<td>0</td>
<td>0.318</td>
</tr>
</tbody>
</table>

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

### Table B.4 – Determining the hedonic amenity index

(Independent variable: the log of house price per m²)

<table>
<thead>
<tr>
<th></th>
<th>(1) OLS</th>
<th>(2) OLS</th>
<th>(3) OLS</th>
<th>(4) OLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share land in historic district &lt;500m</td>
<td>0.1796*** (0.0210)</td>
<td>0.1710*** (0.0204)</td>
<td>0.1695*** (0.0209)</td>
<td></td>
</tr>
<tr>
<td>Listed buildings &lt;500m</td>
<td>0.0047** (0.0024)</td>
<td>0.0052*** (0.0024)</td>
<td>-0.0043 (0.0029)</td>
<td></td>
</tr>
<tr>
<td>Share water bodies &lt;500m</td>
<td>0.3014*** (0.0255)</td>
<td>0.2824*** (0.0253)</td>
<td>0.2869*** (0.0251)</td>
<td></td>
</tr>
<tr>
<td>Share open space &lt;500m</td>
<td>0.0604*** (0.0084)</td>
<td>0.0636*** (0.0084)</td>
<td>0.0690*** (0.0085)</td>
<td></td>
</tr>
<tr>
<td>Shops, &lt;500m</td>
<td>-0.0084 (0.0074)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hotels, restaurants, cafés &lt;500m</td>
<td>0.0423*** (0.0118)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cultural establishments &lt;500m</td>
<td>0.0480 (0.0640)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leisure establishments &lt;500m</td>
<td>0.0232 (0.0730)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Housing controls: Yes
Neighborhood fixed effects: Yes
Year fixed effects: Yes

Number of observations: 695,709

Notes: Housing controls include house type, house size, whether the property has a garage, garden and/or central heating, the number of layers of insulation, the maintenance quality, the number of rooms, construction year dummies and whether a building is listed. Standard errors are clustered at the neighborhood level and in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.10
Table B.5 – Descriptive statistics for historic data

<table>
<thead>
<tr>
<th></th>
<th>(1) mean</th>
<th>(2) sd</th>
<th>(3) min</th>
<th>(4) max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employment accessibility in 1909</td>
<td>38,029</td>
<td>23,884</td>
<td>1,494</td>
<td>163,349</td>
</tr>
<tr>
<td>Share of high-skilled workers in 1909</td>
<td>0.0298</td>
<td>0.0285</td>
<td>0</td>
<td>0.197</td>
</tr>
<tr>
<td>Share of medium-skilled workers in 1909</td>
<td>0.216</td>
<td>0.128</td>
<td>0.00386</td>
<td>0.688</td>
</tr>
<tr>
<td>Population accessibility in 1900</td>
<td>89,184</td>
<td>62,641</td>
<td>3,008</td>
<td>422,544</td>
</tr>
<tr>
<td>Share built-up land in 1900</td>
<td>0.0432</td>
<td>0.103</td>
<td>0</td>
<td>0.930</td>
</tr>
<tr>
<td>Share water in 1900</td>
<td>0.0591</td>
<td>0.175</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Share locals in 1899</td>
<td>0.643</td>
<td>0.102</td>
<td>0.217</td>
<td>0.950</td>
</tr>
<tr>
<td>Share protestants in 1899</td>
<td>0.518</td>
<td>0.337</td>
<td>0</td>
<td>0.998</td>
</tr>
<tr>
<td>Population accessibility in 1832</td>
<td>40,389</td>
<td>20,970</td>
<td>1,986</td>
<td>135,168</td>
</tr>
<tr>
<td>Cadastral income in 1832 per ha</td>
<td>603.6</td>
<td>2.235</td>
<td>0</td>
<td>61,866</td>
</tr>
<tr>
<td>Share buildings in 1832</td>
<td>0.00726</td>
<td>0.0338</td>
<td>0</td>
<td>0.412</td>
</tr>
<tr>
<td>Share built-up land in 1832</td>
<td>0.0416</td>
<td>0.0890</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Share water in 1832</td>
<td>0.120</td>
<td>0.264</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

The number of observations is 10,213,524. For the 1832 data it is 5,556,498. Because of confidentiality restrictions the minimum and maximum values refer to the 0.01% and 99.99% percentile. This implies that we exclude the bottom and top 1,024 observations.

Table B.6 – Bias corrected estimates

<table>
<thead>
<tr>
<th></th>
<th>(1) GMM</th>
<th>(2) GMM</th>
<th>(3) GMM</th>
<th>(4) GMM</th>
<th>(5) GMM</th>
<th>(6) GMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pictures per ha (log)</td>
<td>0.2735***</td>
<td>0.5509***</td>
<td>0.2105***</td>
<td>0.0326***</td>
<td>0.0345***</td>
<td>0.0239***</td>
</tr>
<tr>
<td></td>
<td>(0.0495)</td>
<td>(0.0802)</td>
<td>(0.0162)</td>
<td>(0.0017)</td>
<td>(0.0011)</td>
<td>(0.0009)</td>
</tr>
<tr>
<td>Employment accessibility (log)</td>
<td>0.1900***</td>
<td>0.1900***</td>
<td>0.0775***</td>
<td>0.2377***</td>
<td>0.1041***</td>
<td>0.0936***</td>
</tr>
<tr>
<td></td>
<td>(0.0647)</td>
<td>(0.0647)</td>
<td>(0.0285)</td>
<td>(0.0154)</td>
<td>(0.0054)</td>
<td>(0.0045)</td>
</tr>
<tr>
<td>Household controls</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Housing and job controls</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Housing and job controls</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>$\bar{R}$</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.357</td>
<td>0.357</td>
<td>0.357</td>
</tr>
<tr>
<td>$\varpi$</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Number of observations   | 10,213,540  | 10,213,540  | 10,213,540  | 10,213,540  | 10,213,540  | 10,213,540  |

Notes: We use commuting flows between neighborhoods based on the job that generates the most working hours. In columns (2)-(5) we use as instrument the euclidian distance between two neighbourhoods as instrument. In column (3) we derive the commuting flow based on the two jobs that generate the most working hours in the household. Standard errors are bootstrapped (250 replications) and in parentheses; *** p < 0.01, ** p < 0.5, * p < 0.10.
### Table B.7 – First-stage regression results

<table>
<thead>
<tr>
<th></th>
<th>Dep.var. Pictures per ha (log)</th>
<th>Dep.var. Employment accessibility (log)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td></td>
<td>OLS</td>
<td>OLS</td>
</tr>
<tr>
<td>Listed buildings per ha</td>
<td>0.1269***</td>
<td>(0.0346)</td>
</tr>
<tr>
<td>Share historic district</td>
<td>2.2677***</td>
<td>(0.1894)</td>
</tr>
<tr>
<td>Share built-up land</td>
<td>2.4775***</td>
<td>(0.0847)</td>
</tr>
<tr>
<td>Share water</td>
<td>2.5560***</td>
<td>(0.3015)</td>
</tr>
<tr>
<td>Share built-up land in 1900</td>
<td>5.419***</td>
<td>(0.2521)</td>
</tr>
<tr>
<td>Share water in 1900</td>
<td>0.6163***</td>
<td>(0.1520)</td>
</tr>
<tr>
<td>Share built-up land in 1900, 0-500m</td>
<td>0.2646</td>
<td>(1.1085)</td>
</tr>
<tr>
<td>Share built-up land in 1900, 500-1000m</td>
<td>5.8759***</td>
<td>(1.3125)</td>
</tr>
<tr>
<td>Employment accessibility (log)</td>
<td>0.3448***</td>
<td>(0.0420)</td>
</tr>
<tr>
<td>Employment accessibility in 1909 (log)</td>
<td>0.3150***</td>
<td>(0.0458)</td>
</tr>
<tr>
<td>Household controls</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Housing and job controls</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Workplace fixed effects</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Observations</td>
<td>10,213,524</td>
<td>10,213,524</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.6046</td>
<td>0.5036</td>
</tr>
</tbody>
</table>

Notes: Household controls include household size, mean age of adults, mean gender, household type (couple, single, kids), the share of the household that is foreign-born. Job controls are the total hours worked, whether the household has a company car, the share of full-time contracts, the share of permanent contracts. Housing controls include house type, height of the building, construction year dummies and whether a building is listed. Standard errors are clustered at the neighborhood level and in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$
Table B.8 – Sensitivity analysis for reduced form regressions

<table>
<thead>
<tr>
<th></th>
<th>One job households</th>
<th>+ Single Company plant firm</th>
<th>Education level</th>
<th>Commuting by rail</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td>2SLS</td>
<td>2SLS</td>
<td>2SLS</td>
<td>2SLS</td>
</tr>
<tr>
<td>2SLS pictures per ha (log)</td>
<td>0.0401***</td>
<td>0.0367***</td>
<td>0.0325***</td>
<td>0.0540***</td>
</tr>
<tr>
<td></td>
<td>(0.0038)</td>
<td>(0.0038)</td>
<td>(0.0040)</td>
<td>(0.0032)</td>
</tr>
<tr>
<td>Employment accessibility (log)</td>
<td>0.0551***</td>
<td>0.0455***</td>
<td>0.0708***</td>
<td>0.0333***</td>
</tr>
<tr>
<td></td>
<td>(0.0111)</td>
<td>(0.0134)</td>
<td>(0.0112)</td>
<td>(0.0099)</td>
</tr>
<tr>
<td>Household controls</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Housing and job controls</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Workplace fixed effects</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Number of observations</td>
<td>6,706,524</td>
<td>3,532,906</td>
<td>1,523,567</td>
<td>7,626,355</td>
</tr>
<tr>
<td>Kleibergen-Paap F-statistic</td>
<td>85.60</td>
<td>88.36</td>
<td>77.53</td>
<td>82.87</td>
</tr>
</tbody>
</table>

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

Table B.9 – Additional counterfactual analyses

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Scenario 4: ( \beta^e = 0 )</th>
<th>Scenario 5: ( \kappa^e = \hat{\kappa}/2 )</th>
<th>Scenario 6: ( \delta^e = \hat{\delta} \times 5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Total output</td>
<td>120,959</td>
<td>133,805</td>
<td>182,472</td>
<td>1,492,015</td>
</tr>
<tr>
<td>Aggregate land rents</td>
<td>437,039</td>
<td>434,299</td>
<td>461,534</td>
<td>713,349</td>
</tr>
<tr>
<td>Aggregate real income</td>
<td>11,821</td>
<td>12,680</td>
<td>24,003</td>
<td>61,066</td>
</tr>
<tr>
<td>Income mixing, ( \bar{\sigma}_x )</td>
<td>0.0472</td>
<td>0.0148</td>
<td>0.0491</td>
<td>0.0473</td>
</tr>
</tbody>
</table>

Notes: We calculate aggregate land rents as: \( \sum_{z=1}^{Z} h_z R_{x}^e \) and aggregate real income as: \( \sum_{z=1}^{Z} (1/h_z)(\omega_x^e R_{x}^e - R_y h)R_{x}^e \). Hence, we weight aggregate real income by the density in each location. See Appendix A.8 for more information.
Appendix figures

Figure B.1 – Calculation of travel time and speed

(a) Distance and travel time
(b) Commuting time distribution

Figure B.2 – Historic data from 1900

(a) Built-up land
(b) The railway network and accessibility

Legend
- Water bodies in 1900
- Built-up areas in 1900

Legend
- Railway network in 1900
- Railway stations in 1900
- Water bodies in 1900

A27
Figure B.3 – Historic data from 1832

(a) Built-up land
(b) The road network and accessibility

Figure B.4 – Histograms for the variables of interest

(a) Incomes
(b) Land prices
Figure B.5 – Spatial distribution of variables of interest

Figure B.6 – Counterfactual 5: $\kappa^c = \hat{k}/2$