Land use, worker heterogeneity and welfare benefits of public goods

Coen N. Teulings[#], Ioulia V. Ossokina[##*], Henri L.F. de Groot[###†]

[#]University of Cambridge, University of Amsterdam and Tinbergen Institute

[##]CPB Netherlands Bureau for Economic Policy Analysis and Eindhoven University of Technology

[###]Vrije Universiteit Amsterdam and Tinbergen Institute

October 15, 2017

Abstract

We show that investments in public goods change the optimal land use in their vicinity, leading to additional welfare benefits. This occurs through two sorting mechanisms. First, availability of public goods leads to higher population densities. Second, population groups sort according to their preferences for public goods. We develop a structural spatial general equilibrium model that accounts for these effects. The model is estimated using data on transport infrastructure, commuting behavior, land use and land rents for some 3000 ZIP-codes in the Netherlands and for three levels of education. Welfare benefits of investments in public transport infrastructure are shown to differ sharply by workers' educational attainment. Welfare gains from changes in land use account for up to 30% of the total benefits of a transport investment.

JEL Codes: H4, H54, R13, R23, R4

Keywords: local public goods, transport infrastructure, spatial equilibrium, residential sorting, land rents, educational attainment, commuting

*Corresponding author: i.v.ossokina@tue.nl.

†We thank Xichen Ji, Gerard Verweij, Stefan Groot en Arjen Deetman for excellent research assistance. We are grateful for comments and suggestions received from participants of workshops at CPB Netherlands Bureau for Economic Policy Analysis in the Hague, as well as participants of seminars and conferences in Amsterdam, Barcelona, Cambridge, Chicago, London, Manchester, Rotterdam, St Petersburg and Toulouse.

1 Introduction

House prices are routinely used to value welfare benefits from local public goods, like transport infrastructure, central business districts, and shopping malls. Although this is common practice, it raises two important issues. First, the supply and the use of housing are endogenous. Land close to valuable public goods is likely to be scarce and therefore commands a higher rent. This raises the density of construction per square meter of land and leads to a more intensive use of existing construction. Second, people are likely to differ in their preferences for living close to local public goods. This leads to spatial sorting, where people with a greater willingness to pay for the proximity to public goods will live closer to them. This will again affect the land rents and the land use. A proper model for the valuation of public goods should account for this 'double sorting' on density and heterogeneous preferences, and for the changes in optimal land use. Figs 1 and 2 illustrate the importance of the two mechanisms for the Netherlands. Fig. 1 shows that land rents vary by almost a factor 400, ranging from 3,800 euro per square meter in Amsterdam's Canal Zone to some 10 euro along the North-Eastern border with Germany. Fig. 2 documents that places with high land rents have high population densities (left panel). It also shows the spatial segregation in residential location between high and low educated workers (right panel). Roughly speaking, the high educated live in the cities where land rents peak, while low educated live in the countryside.

This paper develops a spatial general equilibrium model that allows the valuation of local public goods accounting for both mechanisms: changes in population density and heterogeneous preferences for local public goods. Changes in the supply of public goods lead in the model to a new equilibrium on the market for residential land. The underlying mechanisms include endogenous responses in home and work locations and modal choice of individuals, in land prices and population density. We estimate the parameters using detailed microdata for the Netherlands.

The model is applied to a hypothetical policy experiment. We consider the welfare implications of a better railway connection between an economic centre and its periphery. This improved connection leads to a shift in the modal split from car to train and to a relocation of jobs from the periphery to the more productive center. Indeed, as more people are willing to commute to the center, the labour supply in the center increases and so does the number of jobs. The periphery loses jobs, but becomes a more attractive residential location due to a better accessibility of the center. Land use intensity and land prices rise, especially in locations close to railway stations. Population increases there, and the population composition shifts towards high educated workers as they attach a relatively high value to railway connections and the accessibility of jobs in the economic center.

Our model extends the classical McFadden's random utility framework. This framework is frequently applied to recover heterogeneous preferences for local public goods and to value welfare benefits from discrete changes in these goods (see Epple and Sieg, 1999, Bayer et al., 2004, 2007, 2009 and Klaiber and Kuminoff, 2014). Our main extension is to model individual housing and land consumption explicitly. In our model, consumers optimally choose their residential location and the amounts of housing and other consumption of goods. We show that the choice of the consumption bundle reduces to the choice between land and other consumption given the local land rent and supply of public goods. Changes in the supply of local public goods shift the residential demand in a location outwards. This leads to higher rents, more intensive land use, and hence to higher housing and population densities. One can think of this process either as a more intensive use of the existing supply of housing, e.g. by splitting or merging apartments, or as a gradual reconstruction of a neighborhood after the land rent has changed. Most existing equilibrium sorting models (Klaiber and Phaneuf, 2009; Tra, 2010; Sieg et al., 2002, 2004) work with an equilibrium in housing services, holding housing supply constant. In these models changes in the supply of public goods affect housing prices and population composition, but not population density. Walsh (2007) includes the effects of public goods on land consumption in a spatial equilibrium setting. His way of modelling the preference heterogeneity is, however, based on a single amenity index. We apply a more flexible approach of modelling the preference for each public good explicitly. Furthermore, we allow for a non-constant elasticity of substitution between land and other consumption and estimate from the data the share of land in consumption.

As another extension we incorporate a separate home location, job location and commuting mode choice decisions in the model. This allows us to carry out a welfare analysis of concrete transportation improvements such as for example a better railway connection between the centre and its periphery.

To estimate the heterogeneity of preferences for home location, job location and commuting mode we use microdata on some 60000 Dutch employees of three education levels. We know the home and job location of these employees on the level of a four digit zip code, an area that in cities covers approximately one square kilometer. We also know which transportation mode the employees use for commuting. This dataset is enriched with information about land prices and amenities in the home zip codes, and travel time and cost characteristics of trips for all possible combinations of home and job locations. Variation in the location characteristics between zip codes and variation in trip characteristics between commuting modes are used to estimate the parameters of the consumer choice model. To estimate the share of land in consumption we use variation in land prices, residential land use and total wage income by zip code.

Our structural estimates provide a number of interesting insights. First, we find the elasticity of substitution between land and other consumption to be around 0.7. This result is roughly in line with Albouy and Ehrlich (2012) who report the elasticity of substitution between land and the value of construction to be about a half for the US. Since the elasticity is smaller than unity, the land share in consumption is increasing in the degree of agglomeration in the economy. Landlords are therefore the main beneficiaries of agglomeration. The land share varies from 6% in peripheral areas to almost 50% in the most expensive locations. Second, the preferences for local public goods differ widely across levels of education. Third, population groups significantly differ in their land consumption with high educated having a larger willingness to pay for residential space than low educated.[1] This has interesting policy implications. Investing in local public goods increases residential demand for a location, especially of high educated. To maximize the welfare benefits of new public goods, it is efficient that the high educated move to their vicinity. This requires, however, adjustments in the housing supply, as high educated have different land consumption preferences than other groups. In other words, investments in local public goods should optimally go together with redevelopment of the housing stock.

Our hypothetical policy experiment - closing down the railway connection between the city of Amsterdam and the area North of the city - illustrates the effect of spatial sorting and shifts in the intensity of land use. We show that 30% of the welfare benefits are due to relocation of people to other home and job locations; the other 70% are time savings of consumers who do not relocate. High skilled relocate relatively more and they get the

[1] Though our utility function is homothetic, we allow for differences in the utility function across levels of education. Since education is correlated to income, this yields outcomes that are comparable to a non-homothetic utility function. However, actually applying a non-homothetic utility function would greatly complicate the analysis.

major part (70%) of the total benefits. These results have important political economy implications. Investments in long-distance transportation benefit especially the high skilled. This result arises in our model through two channels: (i) high educated can gain relatively more in terms of wages by commuting longer distances; (ii) high educated are less sensitive than other groups to changes in land rents. Finally, the welfare gains from the new railway connection are divided between land owners and workers. Due to the impossibility of price discrimination, land owners cannot capture the whole gain (see also Kuminoff and Pope, 2013).

Fig. 1 Land rents (log) in the Netherlands[2]

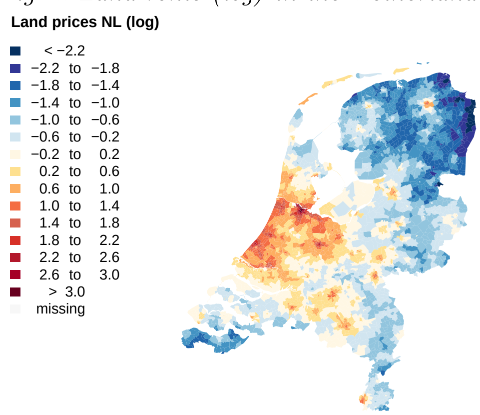
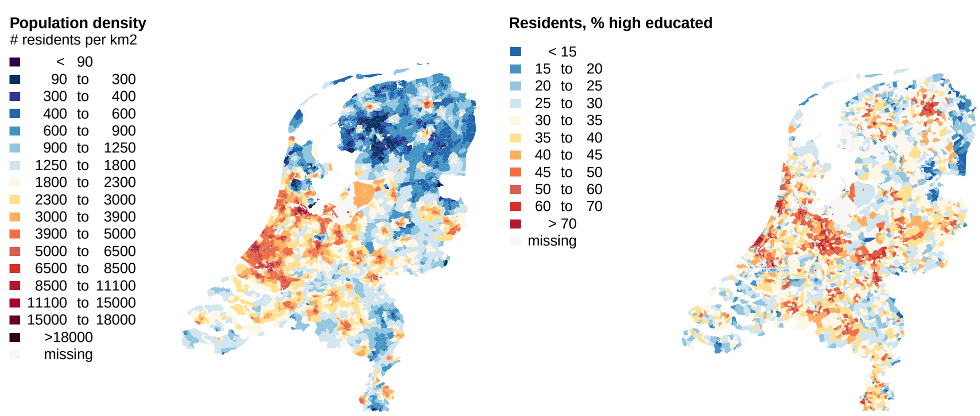


Fig.2 Population density (left) and share of high educated residents by zip code (right)



This paper links to two main strands of papers in the literature. The first strand studies how population size and density in different residential locations are affected by the public

[2]In Figure 1, log rents are normalized at 0 in Enschede with an average land price of 109 euro/m^2.

goods supply. Albouy (2009) shows that existing federal taxes induce population shifts from urban to rural areas. Desmet and Rossi-Hansberg (2013) use a general equilibrium model to explain the city size distribution from differences in congestion costs, amenities and productivity. Ahlfeldt et al. (2016) model the changes in residential density and land prices in Berlin. Diamond (2016) simulates sorting of high and low skilled between cities. The second strand focuses on the effect of investments in transportation. Haughwout (2002) develops a spatial model with aggregate investment in regional transport infrastructure. Anas and Liu (2007) model the interaction between land use and transportation infrastructure. Baum-Snow (2007) and Allen and Arkolakis (2014) analyze the effect of the highway system in the USA. Duranton and Turner (2012) estimate a model explaining the joint evolution of highways and employment. Bowes and Ihlandfeldt (2001), Gibbons and Macchin (2005), and Klaiber and Smith (2010) examine the effects of infrastructure investments on house prices. Baum-Snow and Kahn (2000) address changes in the modal shift.

The structure of this paper is as follows. Section 2 derives the indirect utility function of consumers and defines the equilibrium on the land market. Section 3 deals with identification issues in the structural estimation. Section 4 discusses the data and Section 5 the estimation results. Section 6 reports the results of the policy experiment and Section 7 concludes.

2 The model

2.1 Assumptions

We consider an economy with I individuals i. Each individual is endowed with an education level s: low, middle or high. The economy is made up of H locations, that are either indexed h for the home location of an individual, or j for her job location. Each individual has exactly one job. Within each location h, there are K_h houses. Individuals choose to live in one of the houses $k \in K_h \times H$; hence, by choosing a house k, an individual implicitly chooses to live in the home location h where that house is located. Individuals also choose job location j. Finally, they must choose a mode of transport m: car, train, other public transport, or walking/cycling. It is convenient to define the combination of these three choices for a house, a job location, and a mode of transport as one vector: $x \equiv \{k, j, m\}$. All three choices are discrete choices from a finite set of alternatives. While the set of regions H in the economy is exogenously given, the number of houses K_h at location h is determined endogenously.

The take-home pay of individual i is given by

$$W_i(x) = W_s e^{w_s(j) - c_i(x)},$$

where W_s is the nationwide mean wage for education level s, $w_s(j)$ is the relative deviation from the nationwide mean for education level s at job location j, and $c_i(x)$ is the generalized commuting cost relative to labour income. Hence, the wage for education level s varies between job locations. The commuting cost $c_i(x)$ depends on the house, the job location and mode of transport for commuting, that is, it depends on all elements of x. Commuting cost is modelled as an iceberg technology where commuting depletes a fraction of the wage. The cost $c_i(x)$ differs between individuals due to individual-specific factors.

Individuals are characterized by homothetic constant returns to scale utility function[3] with the consumption of housing services F and other consumption C as its arguments:

$$U_i(F, C; x) = U(F, C) e^{\chi_i(x)},$$

where the function $U(\cdot)$ is twice differentiable and satisfies the standard slope and curvature assumptions. Note that the function $\chi_i(x)$ is allowed to differ between individuals; on the contrary, the function $U(F, C)$ is the same for all i.

Let $P_h(F)$ be the price of a house that offers F housing services at residential location h. House prices differ across residential locations h, while the price of other consumption is the same across all locations in the economy. Without loss of generality, the price of other consumption is normalized to unity. Individuals choose F, C, and x as to maximize their utility subject to their budget constraint:

$$W_i(x) = P_h(F) + C.$$

Housing services are produced by real estate developers by means of a constant returns to scale production function $F = F(L, B)$ with the lot size L and the units of building B installed upon that lot as its inputs. Like the utility function, the production function is twice differentiable and satisfies standard assumptions. Each location h has an exogenously fixed supply of land A_h available for residential use. This land is most easily thought of

[3] Since a utility function is invariant to an increasing transformation, the concept of constant returns to scale has little meaning in this context. However, when using the indirect utility function, see equation (1) below, this assumption implies that the level of utility is linear in income. This simplifies notation.

as being owned by a class of absentee landlords, who maximize their income. Developers can choose how much land to buy from landowners and how much building to construct on that land. They make these choices as to maximize their profits. Perfect competition on the market for real estate development drives their profits down to zero. We assume that the land within a location h is homogeneous. Perfect competition on the land market yields a market clearing price R_h the sets the demand for that land equal to its exogenous supply A_h. The price of one unit of building B is the same across locations, which is again normalized to unity without loss of generality. These assumptions guarantee that the price for a house charged by developers at location h is equal to the cost of production of that house (see also Muth, 1969):

$$P_h\left[F(L,B)\right] = R_h L + B.$$

This formulation allows us to integrate the profit maximization problem of real estate developers and the utility maximization problem of individuals into a single unified problem. The individual chooses her optimal land use L^*, the units of building B^*, her expenditure on other consumption C^*, and her choice set x as to maximize her utility subject to her budget constraint;

$$\begin{aligned} \{L_i^*, B_i^*, C_i^*, x_i^*\} &= \arg\max_{L,B,C,x} U\left[F\left(L,B\right),C\right] e^{\chi_i(x)}, \\ \text{subject to} &: W_i(x) = R_h L + B + C. \end{aligned} \tag{1}$$

This reduced-form specification of the utility function encompasses the production function of housing services. It takes into account that in a perfectly competitive housing market, real estate developers will only develop houses that provide services to individuals in the most cost-effective way. This reduced form utility function does not have housing services F as its argument, but the inputs L and B that are required for the production of F.

2.2 Indirect utility function

The optimal demand for land L^*, for units of construction B^*, and other consumption C^* for an individual who chooses to live in a house in location h are all a function of the land price R_h in that location. For instance, for L^* we obtain (see Appendix A):

$$L_i^*\left(R_h;x\right) = -v'\left(R_h\right) W_i(x), \tag{2}$$

where $v(R_h) \equiv \ln U\left[F\left[\widehat{L}^*(R_h), \widehat{B}^*(R_h)\right], \widehat{C}^*(R_h)\right]$. The prices for B^* and C^* drop out, since they are equal for all regions (and normalized to unity). Substituting these individual demand functions back into the direct utility function yields the indirect utility function. This indirect utility function $V_i(R_h; x)$ takes the following convenient log additive form, see Appendix A for the derivation:

$$\begin{aligned} V_i(R_h; x) &\equiv U[F(L^*, B^*), C^*] \exp[\chi_i(x) + w_s(j) - c_i(x)] W_s \\ &= \exp[v(R_h) + \chi_i(x) + w_s(j) - c_i(x)] W_s, \end{aligned} \tag{3}$$

The factor W_s is just an exogenous constant that is irrelevant for the actual choices of the individual. The term in square brackets is the driving force. It consists of four components, those measuring the effect of respectively: (i) the land rents $v(R_h)$; (ii) home and job location specific factors $\chi_i(x)$; (iii) the wage surplus $w_s(j)$; (iv) commuting cost $c_i(x)$. These components are specified as follows:

$$\begin{aligned} v(R_h) &= -\rho(\ln R_h + \psi R_h), \\ \chi_i(x) &= \mu_{Ks}(\alpha_s' a_h + z_h + \widetilde{\varepsilon}_{ik} + \mu_{Js}\widetilde{\varepsilon}_{ikj}), \\ w_s(j) &= \mu_{Ks}\mu_{Js} y_{sj}, \\ c_i(x) &= \mu_{Ks}\mu_{Js}\mu_{Ms}(\gamma_s' c_{shjm} - \varepsilon_{ikjm}), \end{aligned} \tag{4}$$

where $\mu_{Xs} \in [0,1]$ for $X = K, J, M$. Some rearrangement of terms yields

$$\ln V_i(R_h; x) = v(R_h) + \mu_{Ks}[\alpha_s' a_h + z_h + \widetilde{\varepsilon}_{ik} + \underbrace{\mu_{Js}(y_{sj} + \widetilde{\varepsilon}_{ikj} + \underbrace{\mu_{Ms}(-\gamma_s' c_{shjm} + \varepsilon_{ikjm})}_{\text{commuting}})}_{\text{job location}}], \tag{5}$$

where we omit the irrelevant additive constant $\ln W_s$. This recursive specification of the indirect utility function, where the term capturing the effect of the commuting mode m is embedded in a broader term capturing the effect of job location j, which is in turn embedded in the overall utility of the house k in residential location h, implies a nested logit structure for the choice of the finite set $\{x\}$ (see Ben-Akiva and Lerman, 1985). Each term in this utility function is explained below, where we work our way back from the final term on commuting, to the term on job location, to the overall utility for a house k.

The term $c_i(x) = \mu_{Ks}\mu_{Js}\mu_{Ms}(\gamma_s' c_{shjm} - \varepsilon_{ikjm})$ reflects the commuting cost between h and j by mode m. It includes both a deterministic part $\gamma_s' c_{shjm}$ and a stochastic part ε_{ikjm}, capturing heterogeneity in individual preferences unobservable to the researcher and that is specific to each combination $x = \{k, j, m\}$; ε_{ikjm} follows a type I extreme value distribution, see Appendix B for details. The vector c_{shjm} covers the observable characteristics of the trip, like financial cost as share of income, travel time, convenience, and post- and pre-transport for commuting by train.

Commuting cost relative to income $c_i(x)$ enters the log indirect utility function (3) linearly with a coefficient equal to unity. Since the vector c_{shjm} includes the financial cost of commuting relative to income, our specification yields a restriction on the parameter for the corresponding element of γ_s. Let γ_{s0} denote this element. Hence, for the financial cost of commuting relative to income to enter linearly with coefficient unity, the following restriction must hold:

$$\mu_{Ks}\mu_{Js}\mu_{Ms}\gamma_{s0} = 1. \tag{6}$$

We refer to equation (6) as the transport cost identity. The economic intuition for this constraint is as follows. Commuting cost enters the budget constraint through $W_i(x) = W_s e^{w_s(j)-c_i(x)}$. An individual must be indifferent between losing one percent of income either via a lower wage $w_s(j)$ or via higher commuting cost. Since both $w_s(j)$ and c_{shjm0} measure the effect on the take home pay relative to the average wage W_s, the coefficient on c_{shjm0} must be equal to unity. Equation (6) achieves just that.

Working our way back, the commuting cost term is embedded in a more general term for utility derived from the job location. It consists of the job location fixed effect y_{sj} that captures, among other things, the relative wage surplus $w_s(j)$ at work location j for education level s. One could extend the model by allowing for other job location fixed effects. Since neither relative wages, nor other job location fixed effects are observed, we are unable to disentangle both. Hence, adding these effects would not change the empirical content of the model. Again, $\widetilde{\varepsilon}_{ikj}$ is a random individual component specific to each combination $\{k, j\}$ reflecting unobserved differences in preferences for job locations, see Appendix B for its distribution.

Working our way one step further back, the specification of the function $\chi_i(x)$ also includes both a deterministic and a stochastic part. Apart from the effect of the local land rent in the first term to be discussed below, there are two deterministic terms. First, the vector a_h measures the observed amenities at the location h, like the scenery in the neighborhood, the number of monuments, and the availability of restaurants and shops.

Second, the term z_h is a fixed location effect, which is assumed to be uncorrelated to a_h. These variables are constant across all houses k at location h. Finally, the term $\widetilde{\varepsilon}_{ik}$ is a random effect to each house k at location h, reflecting individual differences in preferences for a specific house k, see again Appendix B for its distribution.

Finally, we have the land rent function $v(R_h)$. For $\psi = 0$, we obtain $v(R_h) = -\rho \ln R_h$, which is the Cobb Douglas specification with the parameter ρ measuring the land share in total expenditure. The choice of the functional form of the second term of $v(R_h)$ is just a matter of convenience. For example, the translog cost function applies a second order polynomial in $\ln R_h$, hence, the second term reads $\frac{1}{2}\psi(\ln R_h)^2$ in that case. In principle, one can apply any functional form for $v(\cdot)$ that satisfies the curvature assumptions, $v'(R_h) < 0$ and $v''(R_h) > v'(R_h)^2$. For example, one can fit a non-parametric function. The specification proposed here brings the advantage of its simplicity and it fits the data on land use well, as we show in the empirical part of the paper. The absolute value of the elasticity of substitution implied by this specification is

$$\eta = 1 - \frac{\psi R_h}{(1+\psi R_h)(1-\rho(1+\psi R_h))}, \tag{7}$$

see Appendix A for the derivation. For $\psi > 0$, the elasticity of substitution is less than unity and hence the share of land is increasing in its price R_h.

Note that all parameters in the specification of equation (5) are allowed to vary between levels of education, except for the home location fixed effect z_h and the parameters ρ and ψ of the land rent function $v(R_h)$. The observed amenities $\alpha_s' a_h$ are likely to absorb most of the differences in preferences for various residential locations between levels of education. Hence, assuming z_h to be equal across levels of education is a justifiable restriction. Regarding ρ and ψ: since we do not have data on land use by level of education, ρ and ψ cannot be separately identified for each level of education. Hence, we constrain them to be equal across education levels.

2.3 Optimal choice of home and job location, commuting mode

The individual chooses her house, her work location and the commuting mode, $x = \{k, j, m\}$, as to maximize her utility. The recursive structure of the indirect utility function $\ln V_i(R_h; x)$ in equation (5) allows us to solve this utility maximization problem in three sequential steps, by backward induction. First, we solve for the optimal choice of commuting mode m_i^*

conditional on the residential and job location $\{k, j\}$. Next, we use these results to solve for the optimal job location j_i^* conditional on housing choice k. Then we use the expression for the optimal job location choice to analyze the optimal choice of the house k_i^*, and thereby the residential location h_i^*.

2.3.1 Choice of commuting mode

Since only the last term of utility function (5) depends on m, the optimal choice of m minimizes the commuting cost, taking k and j as given. Hence, the individual chooses m as to maximize

$$m_i^* = \arg\max_m \left[-\gamma_s' c_{shjm} + \varepsilon_{ikjm}\right].$$

Since ε_{ikjm} takes a standard type I extreme value distribution, the probability that individual i chooses commuting mode m conditional on h and j is described by a logit model.

$$\begin{aligned} \Pr\left[m_i^* = m | k, j, s\right] &= \frac{\exp\left(-\gamma_s' c_{shjm}\right)}{\exp\left(-c_{shj}\right)}, \\ c_{shj} &\equiv -\ln\left[\sum_{m \in M} \exp\left(-\gamma_s' c_{shjm}\right)\right]. \end{aligned} \tag{8}$$

The variable $-c_{shj}$ is the standard logsum in a logit model: a measure of the expected generalized commuting cost between the residential location h and the job location j for an individual with education level s.

2.3.2 Choice of job location

Substituting the expression for the generalized cost of the optimal commuting mode $\max_m \left[-\gamma_s' c_{shjm} + \varepsilon_{ikjm}\right]$ in the utility function (5) and using equation (25) and (26) from Appendix B yields an expression for utility conditional on the optimal commuting mode:

$$\ln V_i\left(R_h; k, j, m_i^*\right) = v\left(R_h\right) + \mu_{Ks}\left[\alpha_s' a_h + z_h + \widetilde{\varepsilon}_{ik} + \mu_{Js}\left(y_{sj} - \mu_{Ms} c_{shj} + \varepsilon_{ikj}\right)\right]. \tag{9}$$

Since j enters only via the last term of this indirect utility function, the individual chooses j as to maximize that last term:

$$j_i^* = \arg\max_j \left[y_{sj} - \mu_{Ms} c_{shj} + \varepsilon_{ikj}\right].$$

Since ε_{ikj} follows a type I extreme value distribution, see Appendix B, this choice problem is again a logit model:

$$\begin{aligned}
\Pr\left[j_i^* = j|k, s\right] &= \frac{\exp\left(y_{sj} - \mu_{Ms}c_{shj}\right)}{\exp\left(g_{sh}\right)}, \qquad (10)\\
g_{sh} &\equiv \ln\left[\sum_{j\in H}\exp\left(y_{sj} - \mu_{Ms}c_{shj}\right)\right],\\
\ln V_i\left(R_h; k, j_i^*, m_i^*\right) &= v\left(R_h\right) + \mu_{Ks}\left(\alpha_s' a_h + z_h + \mu_{Js}g_{sh} + \varepsilon_{ik}\right).
\end{aligned}$$

where we use equation (25) and (26) from Appendix B in the final line. The variable g_{sh} measures the option value of finding a job for somebody living in location h. The logsum g_{sh} is therefore a generalized job attractivity measure for residential location h.

2.3.3 Choice of home location

Again, consider the utility function in equation (10). The individual chooses the house k as to maximize

$$\begin{aligned}
k_i^* &= \arg\max_k\left[v_{sh} + \varepsilon_{ik}\right], \qquad (11)\\
v_{sh} &\equiv \mu_{Ks}^{-1}v\left(R_h\right) + \alpha_s' a_h + z_h + \mu_{Js}g_{sh}.
\end{aligned}$$

The choice of a house k implies the choice for a residential location h. Hence, we obtain again a logit model:

$$\begin{aligned}
\Pr\left[h_i^* = h|s\right] &= \sum_{k\in K_h}\Pr\left[k_i^* = k\right] = \frac{\exp\left(v_{sh} + \ln N_h\right)}{\exp\left(v_s\right)}, \qquad (12)\\
v_s &\equiv \ln\left[\sum_{h\in H}\exp\left(v_{sh} + \ln N_h\right)\right] = \mathrm{E}\left[\max_k\left[v_{sh} + \varepsilon_{ik}\right]\right].
\end{aligned}$$

where N_h is the number of houses in location h. Equation (12) yields a simple expression for expected log utility:

$$\mathrm{E}\left[\ln V_i\left(R_h; x_i^*\right)|s\right] = \mu_{Ks}v_s. \qquad (13)$$

Why do we go through the complication of distinguishing between individual houses k within each residential location h? The reason is that if we fail to do so, the utility of an individual location would depend on the actual classification of the country as a whole into different locations (see also Lerman and Kern, 1983). To see this, consider the probability of choosing either of two neighboring locations, h_1 and h_2, with exactly the same amenities a_h and the same location fixed effect z_h in a model that ignores the variety of houses within a particular location.[4] It would satisfy equation (13), but now without the term $\ln N_h$:

$$\Pr\left[h_i^* = h_1 \vee h_2 | s\right] = \frac{\exp\left(v_{sh_1}\right) + \exp\left(v_{sh_2}\right)}{\exp\left(v_s\right)} = 2\Pr\left[h_i^* = h_1 | s\right].$$

Suppose the National Bureau of Statistics were to decide arbitrarily to merge both locations into one location h, $R_h = R_{h_1} = R_{h_2}$ and $v_{sh} = v_{sh_1} = v_{sh_2}$. Then, this specification implies that the probability $\Pr\left[h_i^* = h_1 \vee h_2 | s\right]$ would drop by a factor 2. By allowing individuals to choose between houses k instead of residential locations h, the term $\ln N_h$ enters the specification v_{sh}, which exactly offsets the effect of merging both locations. Hence, the probabilities become independent of the actual classification in locations. By acknowledging that each individual house is slightly different from the perspective of an individual buyer, we account for the fact that by doubling the number of houses in a particular subset K_h, the probability that an individual chooses a house in that subset also doubles, if we keep all observable differences constant.

2.4 Land rents and sorting

A simple transformation of the probability $\Pr[h|s]$ allows an insightful analysis of the land rents. First rewrite equation (12), using the definition of $v_{sh} \equiv \mu_{Ks}^{-1} v\left(R_h\right) + \alpha_s' a_h + z_h + \mu_{Js} g_{sh}$:

$$\Pr[h|s] = \exp\left[\mu_{Ks}^{-1} v\left(R_h\right) + \alpha_s' a_h + z_h + \mu_{Js} g_{sh} + \ln N_h - v_s\right], \tag{14}$$

Applying the Bayes' rule and a log transformation, substituting $\Pr[h]$ and $\Pr[s]$ for their observed values N_h/N and N_s/N and bringing $\mu_{Js} g_{sh}$ to the left hand side, we obtain:[5]

$$\ln \Pr[s|h] - \mu_{Js} g_{sh} = \overline{v}_s + \mu_{Ks}^{-1} v\left(R_h\right) + \alpha_s' a_h + z_h. \tag{15}$$

[4] Note that the argumentation below only holds if z_h in both locations are the same.

[5]
$$\begin{aligned} \Pr[h|s] &= \Pr[s|h]\Pr[h] / \Pr[s] \\ \ln \Pr[s|h] &= \ln \Pr[h|s] - \ln \Pr[h] + \ln \Pr[s] \\ &= -\ln \Pr[h] + \ln \Pr[s] + v_{sh} + \ln N_h - v_s \end{aligned}$$

Substituting $\Pr[h]$ and $\Pr[s]$ for their observed values N_h/N and N_s/N and cancelling common terms, obtain (15).

where $\overline{v}_s \equiv \ln N_s - v_s$. Since $-v(R_h)$ is an increasing function of R_h, equation (15) can be solved for $\ln R_h$. In general, the inverse function has no explicit analytical expression, except for the the Cobb Douglas case where $\psi = 0$ and where the inverse function simplifies to $-v(R_h) = \rho \ln R_h$. Though the subsequent argument applies for any admissible $\psi \neq 0$, we focus on this Cobb Douglas case for the sake of transparency. Then, the solution for $\ln R_h$ reads

$$\ln R_h = \frac{\overline{v}_s + \alpha_s' a_h + z_h + \mu_{Js} g_{sh} - \ln \Pr[s|h]}{\rho \mu_{Ks}^{-1}}. \tag{16}$$

Ignoring the term $-\ln \Pr[s|h]$ for the moment, this equation is a standard log linear land rent equation: land rents are increasing in observed and unobserved amenities, $\alpha_s' a_h + z_h$, and in the job availability g_{sh}. Note that the parameters of this equation and the job availability indicator g_{sh} differ between levels of education s. This seems to yield an inconsistency: though the left hand side does not depend on s, the right hand side does. We reframe this paradox in economic terms: how can a unified land market at each location h for all levels of education be consistent with education level specific returns to amenities? This paradox is resolved by the term $\ln \Pr[s|h]$. Take the higher education level $s = 3$ as a point of reference for our argument. When location h is predominantly inhabited by high educated workers, $\Pr[s = 3|h] \to 1$, it must be the case that the observed characteristics of location h are more attractive for higher than for middle or low-educated workers; in terms of our model: $v_{3h} \gg v_{sh,s=1,2}$, see equation (12). Hence, there is little sorting along unobserved preferences ε_{ik}. This is reflected by the term $\ln \Pr[s|h]$: since $\Pr[s = 3|h] \to 1$, $\ln \Pr[s = 3|h] \to 0$, and the log land rent equation simplifies to

$$\ln R_h = \frac{\overline{v}_3 + \alpha_3' a_h + z_h + \mu_{J3} g_{3h}}{\rho \mu_{K3}^{-1}}.$$

Land rent gradients correspond to the preferences of the higher educated.

Next, consider the case that location h is predominantly inhabited by low-educated workers, $s = 1$; hence, $\Pr[s = 1|h] \to 1$ and hence $\Pr[s = 3|h] \to 0$, or equivalently $v_{1h} \gg v_{sh,s=2,3}$. Hence, only higher educated with a strong unobserved preference ε_{ik} for a house at that location will live in location h. There is therefore strong positive sorting on unobservables. The lower $\Pr[s = 3|h]$, the stronger the positive sorting, which is captured by the term $\ln \Pr[s = 3|h]$. Since $\Pr[s|h] = \exp(v_{sh} + \overline{v}_s) / \Sigma_{l \in S} \exp(v_{lh} + \overline{v}_l)$, $\Pr[s = 3|h] \to$

$\exp(v_{3h} + \overline{v}_3 - v_{1h} - \overline{v}_1)$.[6] [7] Substitution of this expression for $\ln \Pr[s = 3|h]$ in equation (16) yields

$$\begin{aligned} \ln R_h &= \frac{\overline{v}_3 + \alpha_3' a_h + z_h + \mu_{J3} g_{3h} - (v_{3h} + \overline{v}_3 - v_{1h} - \overline{v}_1)}{\rho \mu_{K3}^{-1}} \\ &= \frac{\overline{v}_1 + \alpha_1' a_h + z_h + \mu_{J1} g_{1h}}{\rho \mu_{K1}^{-1}}. \end{aligned}$$

In this case, land rent gradients correspond to the preferences of the low educated.

Hence, what education level's s returns to amenities apply in a particular region depends on what level of education is predominant among the population there. If a location is mainly inhabited by high educated workers, then this education level's returns prevail. This conclusion has important implications for the cost benefit analysis of investments in transport infrastructure and other public goods. These investments have the highest return in those locations that are predominantly inhabited by people with a strong preference for these public goods. For example, a location close to a railway station attracts predominantly higher educated workers, who use the train more often, as we shall see in Section 6. This pushes the local land gradient towards the preferences of higher educated workers, making further specialization of amenities in favour of higher educated workers at that location more attractive.

2.5 Land market equilibrium

The model is closed by the constraint that the supply of residential land at location h, A_h, must be larger or equal to the demand. The demand can be calculated as the total number of workers with education level s, denoted N_s, times the probability that a person with education s chooses to live at location h times the expected land use among these

[6] Bayes' rule implies

$$\Pr[s|h] = \frac{\Pr[h|s]\Pr[s]}{\Sigma_{l \in S} \Pr[h|s = l]\Pr[s = l]} = \frac{\frac{N_s}{N}\exp(v_{sh} + \ln N_h - v_s)}{\Sigma_{l \in S}\frac{N_l}{N}\exp(v_{lh} + \ln N_h - v_l)} = \frac{\exp(v_{sh} + \overline{v}_s)}{\Sigma_{l \in S}\exp(v_{lh} + \overline{v}_k)},$$

where $N \equiv \sum_{s \in S} N_s$.

[7] This follows from

$$\lim_{v_{1h} \to \infty} [\Pr[s = 3|h]\exp(v_{1h} + \overline{v}_{1s} - v_{3h} - \overline{v}_3)] = \lim_{v_{1h} \to \infty} [1 + \exp(v_{2h} + \overline{v}_2 - v_{1h} - \overline{v}_1) + \exp(v_{3h} + \overline{v}_{3s} - v_{1h} - \overline{v}_1)] = 1,$$

where we substitute $\Pr[s = 3|h] = \exp(v_{3h} + \overline{v}_3)/\Sigma_{l \in S}\exp(v_{lh} + \overline{v}_l)$ in the second step.

individuals:

$$\begin{aligned}A_h &\geq \sum_{s\in S} N_s \Pr\left[h_i^* = h|s\right] \mathrm{E}\left[L_i^*|h,s\right] \\ &= \sum_{s\in S} -v'\left(R_h\right) N_s \Pr\left[h_i^* = h|s\right] \mathrm{E}\left[W_i(x)|h,s\right].\end{aligned} \tag{17}$$

where we use the expression for the optimal land consumption (2) in the second line.

In a market equilibrium, the above conditions are binding. Other things equal, the right hand side depends negatively on R_h by two mechanisms: first, the number of individuals that prefer a house at that location decreases, and second, the average lot size $\mathrm{E}[L_i^*|h,s]$ at that location becomes smaller. Land rents adjust till the supply and demand for land at each location are equal.

Finally, the number of houses at location h must be equal to the number of people who choose to locate there. Hence

$$N_h = \sum_{s\in S} \Pr\left[h_i^* = h|s\right] N_s. \tag{18}$$

where N_h is the number of houses that developers choose to construct at location h (the number of elements in the set K_h). The number of houses at location h adjusts such that conditional on the average lot size $\mathrm{E}[L_i^*|h,s]$, all available residential land A_h is used for residential construction, see equation (17).

An equilibrium is a set of land rents R_h and a set of number of houses N_h for each $h \in H$, satisfying equation (17) and (18).

3 Identification and estimation

Now that we have analyzed the individual behavior and the market equilibrium, the discussion of the identification of the model's parameters is relatively straightforward. Table 1 presents an overview. Each line of the table refers to one component of the estimation procedure. For each component, the table shows the relevant equation, the parameters that are estimated, the required inputs, and the outputs that are used in subsequent steps of the estimation. Lines 1 to 5 of the Table identify all parameters of the utility function (5).[8]

[8] Lines 2 to 5 estimate sequentially the nested logit model describing the consumer behaviour. This kind of estimation yields consistent estimates but is not efficient (Train, 2009). We sacrifice some efficiency for the sake of convenience.

Line 6 provides over-identifying restrictions.

Table 1 Identification of the parameters

	Model	equation	parameters	data/inputs	outputs
1.	OLS land use	(19)	ρ, ψ	$R_h, A_h, \overline{W}_h$	$v(R_h)$
2.	Logit modal split	(8)	γ_s	c_{shjm}	c_{shj}
3.	Logit job location	(10)	y_{sj}, μ_{Ms}	c_{shj}	g_{sh}
4.	Logit home location, first stage	(21)	μ_{Js}	g_{sh}, a_h	$\mu_{Js} g_{sh}$
5.	IV home location, 2nd stage	(22)	α_s, μ_{Ks}	$\mu_{Js} g_{sh}, a_h, v(R_h)$	
6.	Transport cost identity	(6)	μ_{Ks}	$\gamma_0, \mu_{Ms}, \mu_{Js}$	over-identified

Line 1 of Table 1 estimates the land share in total expenditure. By taking expectations over s in equation (2), substituting (4) for $v'(R_h)$, and multiplying the expected individual land use with the number of residents in h, N_h, we obtain

$$N_h \mathrm{E}\left[L_i^* | h, s\right] = \rho\left(1 + \psi R_h\right) R_h^{-1} \overline{W}_h.$$

where $\overline{W}_h$ is the total wage income in residential location h minus commuting cost.[9] Since the land availability constraint is binding in the market equilibrium, the total land use across all levels of education in location h must be equal to the endowment of land at location h, A_h:

$$A_h = N_h \mathrm{E}\left[L_i^* | h, s\right].$$

Combining these results yields:

$$\frac{A_h R_h}{\overline{W}_h} = \rho\left(1 + \psi R_h\right) + \zeta_h, \tag{19}$$

where ζ_h is an error term capturing unexplained variation in the land use. Equation (19) has an intuitive interpretation as it contains on both sides the average income share of land in location h.

Equation (19) can be estimated with OLS, yielding the parameter values for ρ and ψ. These values enter the land rent function $v(R_h)$, which will serve as an input for the estimation of the home location choice logit in line 4 of Table 1. However, equation (19) might suffer from the presence of measurement error in the data on R_h, as it includes R_h

[9] The average is taken over transport modes, where we use the probabilities $\Pr[m|shj] \Pr[j|sh] \Pr[s|h]$ to calculate the expectation.

on both sides. To study the importance of this problem we will also estimate an alternative specification:

$$A_h \overline{W}_h^{-1} = \rho \left(1 + \psi R_h\right) R_h^{-1} + \varsigma_h, \tag{20}$$

The logit for modal split (Table 1, line 2) can be estimated from individual data on trip characteristics c_{shjm} and the actual choices of commuting mode. This yields estimates for the modal split parameters γ_s. These parameters can be applied for the calculation of the transport logsum $-c_{shj}$, a generalized measure of commuting cost between job location j and residential location h. This measure is then used as an input in the logit of job location choice (Table 1, line 3). The job location logit yields estimates for the scaling parameter μ_{Ms} and the fixed effects y_{sj} for each job location j. The estimation results can be used for the calculation of the job availability measure g_{sh} that serves as an input in the logit for the residential location.

The estimation of the logit for residential location is more involved, since the land rent R_h is endogenous. The endogeneity problem can be seen easily from equation (14). Locations with high unobserved amenities z_h are more attractive than others. Since land rents R_h clear the market for residential land at each location, the unobserved amenities z_h and the land rent function $v\left(R_h\right)$ are positively correlated. Hence, the parameter estimates will be biased. We solve this problem by applying the two-step approach developed by Bayer et al. (2007). The first step (21) rewrites the logit (14) as follows:

$$\begin{aligned}
\Pr\left[h|s\right] &= \exp\left[\mu_{Ks}^{-1} v\left(R_h\right) + \alpha_s' a_h + z_h + \mu_{Js} g_{sh} + \ln N_h - v_s\right] && (21)\\
&= \exp\left[(\mu_{Ks}^{-1} - \mu_{K2}^{-1}) v\left(R_h\right) + (\alpha_s' - \alpha_2') a_h + \Theta_h + \mu_{Js} g_{sh} - v_s^0\right] \\
\text{where } \Theta_h &= \ln N_h + \mu_{K2}^{-1} v\left(R_h\right) + \alpha_2' a_h + z_h && (22)\\
v_s^0 &\equiv \ln\left[\sum_{l\in H} \exp\left(\left[(\mu_{Ks}^{-1} - \mu_{K2}^{-1}) v\left(R_l\right) + (\alpha_s' - \alpha_2') a_l + \Theta_l + \mu_{Js} g_{sl}\right]\right)\right]
\end{aligned}$$

The location-specific fixed effect Θ_h reflects the utility of location h for the reference education group $s = 2$. It encompasses the endogenous variable $v\left(R_h\right)$ and unobserved amenities z_h, thus allowing the other parameters to be estimated consistently. The fixed effects Θ_h can be estimated by contraction mapping. Note that while most of the preference parameters are estimated in deviations from the reference group, the valuation of job availability μ_{Js} can be estimated in absolute terms. This is due to the fact that g_{sh} varies by education group. The second step decomposes Θ_h by estimating (22); z_h is the error term of this regression model. We deal with the endogeneity of R_h by applying an instrumental variables technique. We instrument R_h with fixed characteristics (levels of amenities) of other locations that are close substitutes to h in geographical space.

Finally, the transportation cost identity (6), $\mu_{Ks}\mu_{Js}\mu_{Ms}\gamma_{s0} = 1$, yields an over-identifying restriction (Table 1, line 6). All parameters in this condition have been estimated in previous lines of Table 1. Hence, this condition provides three over-identifying constraints, one for each level of education. As explained in Section 2.2 this over-identification result has a straightforward economic interpretation. It ensures that an individual is indifferent between loosing one percent of income either via a lower wage or via higher commuting cost. This ratio is pinned down empirically by the estimated effects of the financial cost of commuting and the land rents on the preferences over various home locations, and hence by the parameter estimates of the home location logit derived from the observed location behavior, see equation (14). However, at the same time the estimated effect of land rent on the demand for residential land as derived from the observed demand for land must be equal to the estimated effect of land rent on utility by Shephard's lemma, see equation (19). Both estimates are fully independent but should be consistent. Since there is nothing in the estimation procedure that guarantees this consistency, this condition provides an over-identification test. The assymptotic distribution of this test statistic is discussed in Appendix C.

4 Data

In the estimation of the modal split and job location logit we exploit data on commuting from the 2004–2011 national travel survey for the Netherlands (Mobiliteitsonderzoek Nederland MON 2004–2009 and Onderzoek Verplaatsingen in Nederland OVIN 2010–2011). Respondents have been asked to report all their trips on a particular day. The response rate varies between 55 and 82%. Table 2 reports the data selection steps. From the respondents for whom home and job ZIP codes are available,[10] we select those aged between 18 and 65, not in full time education, working for at least 12 hours per week. We drop respondents for whom education level data are missing, with a home or work address outside the Netherlands

[10] A four-digit ZIP code contains on average 2000 houses. In urban areas, a ZIP code covers approximately a square kilometre.

or on one of the islands in the North Sea, those reporting a post-office box as work address, or having made more than eight trips on the day of survey. We restrict the set of commuting modes to four alternatives: car as a driver, train, bus/tram/metro, bike/walk, deleting respondents commuting by other modes. The remaining dataset is merged with data on travel times, costs and distances for each commuting mode provided by the Dutch Ministry of Transportation for every combination of home and job ZIP codes for 2004. Details of these travel data are discussed in Appendix D.

Table 2 Data selection

# persons	MON 2004–09	OVIN 2010–11
total respondents	310003	84339
working with known home and job ZIP code	75147	18463
selection on status and data availability (see text)	62130	14311
restriction to car, train, bus/tram/metro, bike/walk	56912	12964
travel data available and recorded correctly (see text)	53842	12003
land rents at home & job location available	53504	11835

For the home logit estimations we exploit restricted access microdata of Statistics Netherlands on the residential locations and education level of some 7.5 million Dutch workers. We also use data on amenities from three sources. Data on the area of nature are derived from the digital map "Land use" by Statistics Netherlands (2006). Data on the accessibility of amenities are derived from the dataset "Proximity of amenities" by Statistics Netherlands (2009). Data on the number of monuments are derived from the "Register of monuments" by Cultural Heritage Agency of the Netherlands.

Data on land prices have been calculated from microdata on housing transactions provided by the Dutch Association of Real Estate Brokers (NVM). The method for decomposing the value of the land and the value of the construction is discussed in Appendix C. Land prices are converted into land rents per working day using capital cost of 4.2% per year and 228 working days per year.

Finally, the OLS income share of land exploits data on residential land use from the digital map "Land use" by Statistics Netherlands (2006).

5 Estimation results

5.1 Land use

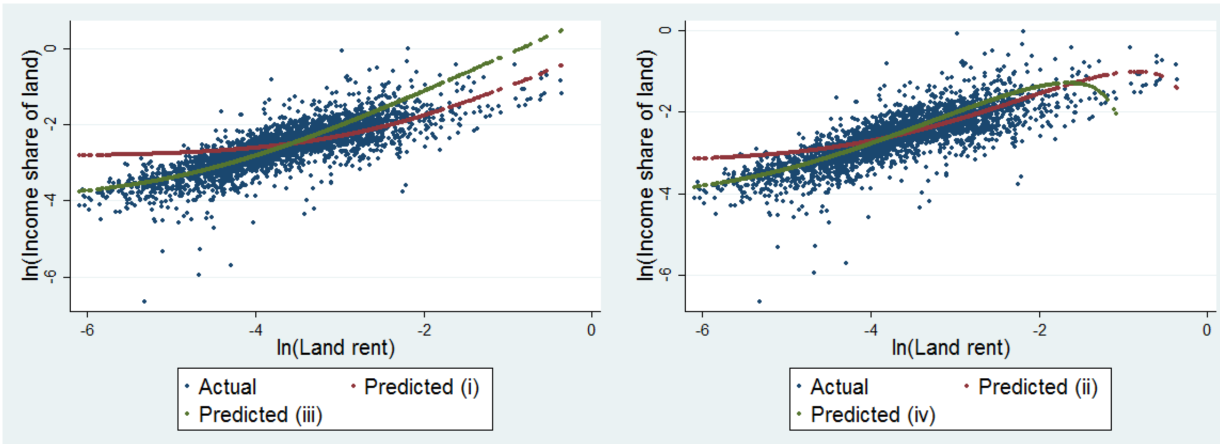
Table 3 reports estimates of equation (19), see line 1 of Table 1. We estimate four specifications: (i) the equation in its original form and (ii) with a quadratic term R_h^2 added; (iii) the equation in its alternative specification (20) and (iv) with a quadratic term added. In all specifications, the coefficients are highly significant. Fig. 3 reports the fit of these four specifications. All four specifications fit the data well in the intermediate segment of the land price distribution and in all four ψ is positive. Hence, the elasticity of substitution between land use and other consumption is less than one. When the quadratic term is added, the slope of the original and the alternative specifications becomes very similar. We conclude that the measurement error in R_h does not affect our results much. However, specification (i) (the original equation (19)) has a better fit in the right tail where land prices are high. In specification (iii), the land share is higher than unity in some extreme cases. Since these home locations in urban areas are important for our model, we shall use specification (i) in what follows.

The elasticity of substitution implied by our estimates (see equation (7)) is 0.71 for the mean value of $\ln R_h$. This is consistent with Albouy and Ehrlich (2012) who report the elasticity of substitution between land and the value of construction to be about one-half, using US data. One would expect the land rent elasticity of the population density to be higher than that of the intensity of construction, since people adjust both the intensity of construction per unit of land and the use of construction per person when the land rent is high. The predicted land share in consumption varies from about 6% for the ZIP codes with the lowest land rents to well above 50% for the most expensive ZIP codes.

Table 3 Land use equation

	(i)		(ii)		(iii)		(iv)	
parameter	coef.	t-value	coef.	t-value	coef.	t-value	coef.	t-value
ψ	14.195	(23.2)	37.837	(16.2)	123.293	(25.8)	172.694	(21.0)
ρ	0.058	(43.5)	0.039	(25.2)	0.018	(39.8)	0.016	(31.4)
2^{nd} order term			-1.716	(19.8)			-6.894	(12.8)
R^2		0.76		0.79		0.37		0.40
# observ.		2753		2753		2753		2753

Fig. 3 Actual vs. predicted income share of land



5.2 Modal split

Table 4 reports the estimation results for the modal split logit, described in line 2 of Table 1.[11] Most variables are highly statistically significant. Higher educated have a strong preference for commuting by train or bike, holding other factors constant. Since the rail infrastructure is better in cities, this contributes to an explanation of why higher educated predominantly live and/or work in cities. Out-of-vehicle time is valued more negatively than in-vehicle time for public transport. Distance to a train station is valued negatively. A high degree of urbanization leads to a higher preference for travelling by bus/tram/metro. This might be related to the higher network quality and the higher service frequency. The car is a land intensive mode of transport. It is therefore less popular for consumers living in locations where land is expensive. Finally, the parameter on the transport cost will be used in the transport cost identify.

The implications of these estimation results are most easily judged from the implied values of time, see Table 5. The compensating variation required to make people indifferent to a marginal increase in travel time can be calculated as $\gamma_{s.time}/\gamma_{s0} * W_s$, where $\gamma_{s0} \in \gamma_s$ is the estimated coefficient for the financial cost of commuting as a fraction of wage income.

[11] We delete irrelevant alternatives, i.e.: (i) train if the total distance to transfer (home+job) is larger than 40 kilometre; (ii) bus/tram/metro if in-vehicle time is larger than 2 hours or out-of-vehicle time is larger than 1.5 hour; (iii) bike if commuting distance is larger than 40 kilometre, all for a single trip.

The value of time is higher for higher educated workers, because they earn a higher wage. An hour spent riding a car or waiting for the train is valued at the average wage rate in our data (18 euro for high, 14 euro for medium, and 11 euro for low educated workers). Time spent in train is less costly, while time spent waiting for a bus is more costly.[12]

Parking costs are measured by including the land rent per square meter relative to the wage at both the home and job location, R_j/W_s. A higher parking cost lowers the probability of choosing a car. Let γ_{sP} be the estimated coefficient on R_j/W_s and let A_p be the land use for parking. Hence, the cost of parking relative to income is equal to A_pR_j/W_s. The effect of the financial cost of commuting relative to income is measured by the coefficient γ_{s0}, see equation (6). Hence $\gamma_{s0}A_pR_j/W_s = \gamma_{sP}R_j/W_s$ and therefore $A_p = \gamma_{sP}/\gamma_{s0}$: the ratio of both parameters is an estimate of the square meters land used for parking. This calculation yields a land use of 34 m^2 and 21 m^2 at the home and the job location respectively. Land use for parking might not always be adequately priced for the consumer, but one would expect land use to adjust to its shadow price one way or the other, e.g. by the employer not making parking space available. Since car use is land intensive due to parking space, it is less popular in locations where land is expensive. The difference in land use at the home and job location can either be due to more efficient land use at the job location (e.g. parking garages) or to the fact that most facilities at the job location are paid for from pre-tax income while facilities at home are paid for from after tax income.

[12] The values of time found for car and bus are somewhat higher, and the values of time for train somewhat lower than those reported in the recent stated preferences study for the Netherlands (Significance et al., 2013). The stated preferences values of time are (averaged over education levels, and over in- and out-of-vehicle time): 9 euro/hour car, 12 euro/hour train and 8 euro/hour bus.

Table 4 Estimation results for the modal split model

General variables	coef	t-val				
cost in % of net wage	-12.24	(8.6)				
time (minutes/10)	-0.261	(20.0)				
parking cost at home[a)]	-4.184	(15.8)				
parking cost at job[a)]	-2.543	(18.0)				
Alternative specific variables	train		bus		bike	
	coef	t-val	coef	t-val	coef	t-val
intercept	-1.574	(14.7)	0.029	(0.3)	0.622	(19.2)
high educated	0.578	(11.6)	-0.003	(0.1)	0.527	(16.9)
low educated	-0.480	(6.7)	0.006	(0.1)	-0.108	(3.4)
distance home-transfer (km/10)	-0.404	(9.7)				
distance job-transfer (km/10)	-0.518	(7.8)				
urbanization at home location			0.258	(15.2)		
urbanization at job location			0.289	(17.0)		
Δ time in vehicle (minutes/10)	0.177	(15.5)	0.071	(5.6)	-0.190	(13.7)
Δ time out vehicle (minutes/10)	0.000	(0.0)	-0.146	(4.8)		
# observations	58778					

[a)] Measured as: land rent divided by the wage income, in %.

Table 5 The values of time (euro/hour)

education	low	middle	high
car time	12	14	19
in-vehicle time train	4	5	6
out–of-vehicle time train	12	14	19
in-vehicle time bus	9	10	14
out–of-vehicle time bus	18	22	29

5.3 Job location

Table 6 reports the estimation results for the job location logit (described in line 3 in Table 1). Some 10% of the individuals work in the same ZIP code as where they live. We have no data on these commutes. We add a dummy for the average cost of intra ZIP code commuting. Since all parameters are education level specific, the model can be estimated for each education level separately.

Table 6 Estimation results for the job location logit

education	low		middle		high	
	coef	t-val	coef	t-val	coef	t-val
Free estimation of the coefficient on generalized commuting cost:						
generalized commuting cost	1.261	(225.2)	1.187	(296.8)	0.992	(247.5)
dummy home=job ZIP	-0.128	(4.7)	-0.282	(11.3)	-0.373	(11.3)
The coefficient on generalized commuting cost constrained to unity:						
dummy home=job ZIP	0.567	(25.0)	0.233	(11.1)	-0.398	(13.2)
# observations	17151		26189		21999	

For medium and low educated workers the coefficient on generalized commuting cost (coefficient μ_{Ms} in (10)) is larger than one, which is inconsistent with the assumptions of a nested logit. An explanation might be that commuting costs per mode are estimated with a fair amount of measurement error for short commuting distances, because within a ZIP code heterogeneity is ignored. This is consistent with the fact that the coefficient is larger than 1 for the lower educated. Higher educated commute longer distances and are therefore less vulnerable to measurement error in short run commutes. We have experimented with different specifications of the modal split model, but by and large this does not change this outcome much. In what follows, we restrict μ_{Ms} to 1, implicitly assuming a multinomial logit structure of the modal split and job location choice.

5.4 Home location

Table 7 reports the estimated coefficients from the first stage home logit (21). The standard errors are clustered. The coefficients by job availability indicate the weights people of education level s attach to job availability g_{sh} in their home location. These are the only coefficients estimated in levels. High educated are less sensitive to land rents. Since land rents are higher in the city, this adds to the explanation why high educated workers predominantly live in the city. Alternatively, this result can be interpreted as saying that higher educated are prepared to pay a higher premium for amenities of the city, such as an environment with many monuments, the proximity of universities and the availability of restaurants. Table 8 reports the estimation results of the second step (22) (see line 5 in Table 1).

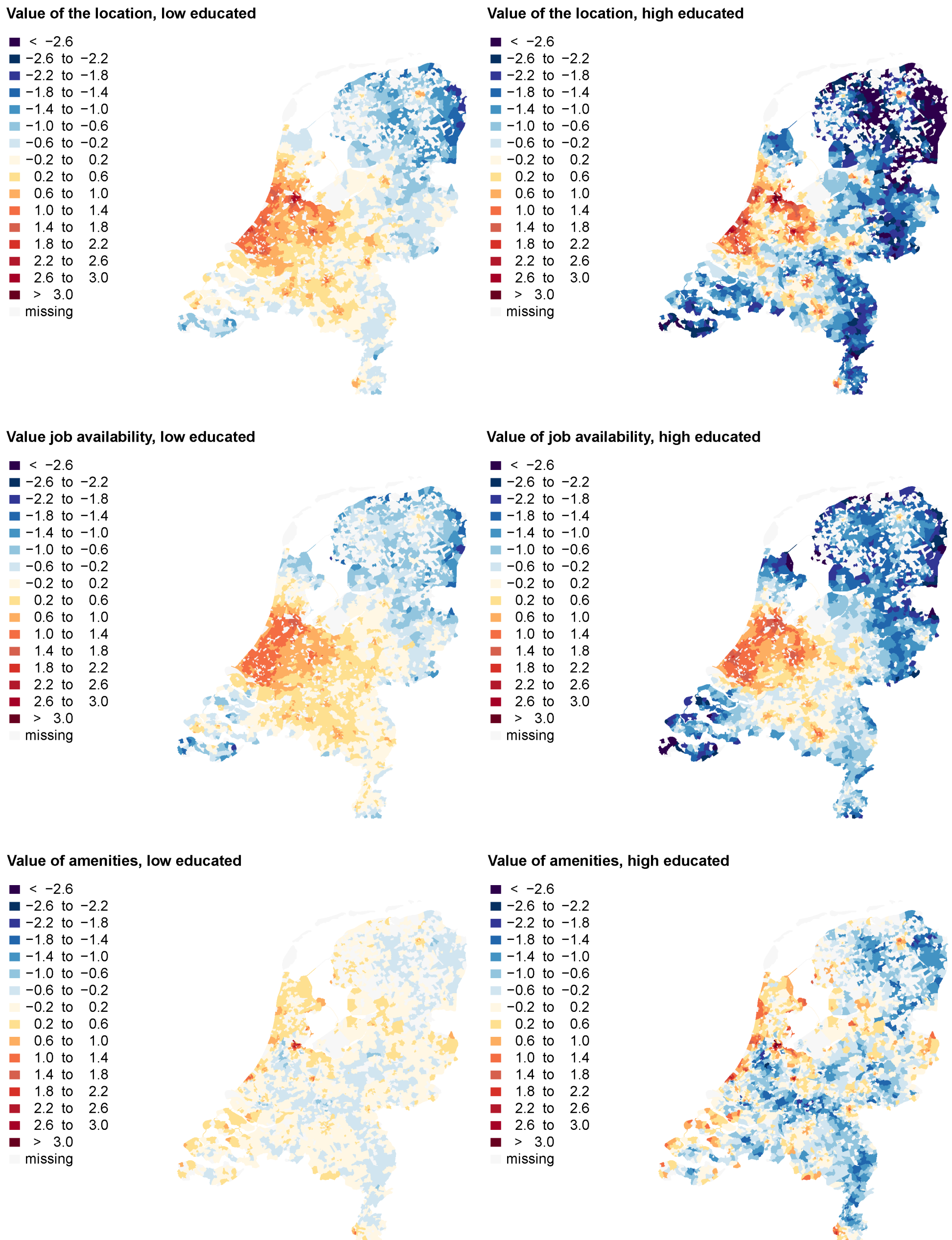
Table 7 Home location choice, first stage

education level	low		middle		high	
variable	coef	t-val	coef	t-val	coef	t-val
job availability (level)	0.627	(10.6)	0.438	(7.4)	0.423	(6.7)
transformed land rent	2.971	(11.7)			-2.422	(8.5)
# monum.1km/1000	-0.135	(1.3)			-0.071	(0.7)
# monum.1-5km/1000	-0.021	(0.7)			0.020	(0.8)
share nature within 5km	0.097	(1.6)			0.347	(4.9)
dum. university in 10km	-0.007	(0.4)			0.133	(6.0)
# restaurants 1km/100	-0.349	(3.7)			0.100	(1.0)
# restaurants 1-5km/100	0.063	(3.7)			-0.021	(1.4)
# observations		2753		2753		2753

Table 8 Home location choice, second stage, IV

education level	middle	
variable	coef	t-val
transformed land rent	6.795	(21.4)
# monum.1km/1000	0.399	(9.2)
# monum.1-5km/1000	0.094	(7.7)
share nature within 5km	-0.024	(0.6)
dum. university in 10km	-0.065	(3.3)
# restaurants 1km/100	0.437	(7.4)
# restaurants 1-5km/100	0.051	(5.2)
intercept	-8.264	(110.7)
# observations		2753
first stage F-value		198

Fig. 4 Value of the location, job availability and observed amenities, expressed in terms of log land price differentials. Low educated left, high educated right.[13]



[13] In all panels, by normalization, 0 corresponds to a location in Enschede with an average land price of 109 euro/m^2.

Our methodology allows to calculate the land rents that different locations would command if their population consisted exclusively of lower of high educated respectively, as well as the contribution of job availability and observed amenities to these rents. The calculation proceeds as follows. We calculate the value of $-v(R_h) + \mu_{Ks} \ln \Pr[s|h]$ for each ZIP code and then solve for lnR_h. Fig. 4 provides a graphical documentation of the results, using intervals of 0.40. The upper two panels depict the calculated log land rent differentials for high and low educated. The middle and lower panels show the valuation of job availability and amenities, respectively, expressed in log land rents. The relative contribution of job availability and amenities differs widely between levels of education. For high educated, there is a much higher variation in the attractivity of locations, both in terms of job availability and in terms of amenities. This adds to the explanation of the spatial segregation between high and low educated workers documented in Fig. 2. Amenities contribute substantially to the popularity of cities as an area to live in, in particular Amsterdam. The contribution of job availability is spread out much more evenly among the central Western part of the country.

5.5 Transport cost identity

The transport cost identity reads $\mu_{Ks}\mu_{Js}\mu_{Ms}\gamma_{s0} = 1$. All of its parameters have been estimated, so we can check whether this over-identifying restriction holds (see line 6 of Table 1). Table 9 reports the transport cost identity calculation for the three education levels. The over-identifying restrictions hold remarkably close.

Table 9. Transport cost identity

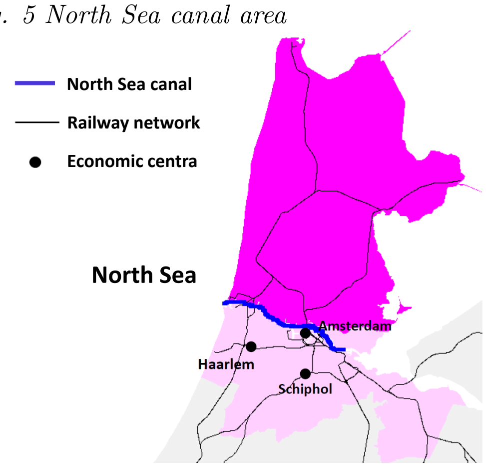
s	μ_{Ks}	t-val	μ_{Js}	t-val	μ_{Ms}	t-val	γ_{s0}	t-val	product	st.err.
1	0.102[a]	(17.1)[a]	0.627	(10.6)	1	∞	12.24	(8.6)	0.78	(0.27)
2	0.147[a]	(21.4)[a]	0.438	(7.4)	1	∞	12.24	(8.6)	0.79	(0.30)
3	0.228[a]	(7.3)[a]	0.423	(6.7)	0.992	(247.5)	12.24	(8.6)	1.17	(0.41)

a) Calculated from Table 7 and 8, using equation (21) and (22).

6 Policy experiment

The city of Amsterdam is located just south of a major canal, connecting the Amsterdam harbour to the North Sea. The main connections between Amsterdam and the area North of the canal consist of five highway tunnels and two train tunnels. Fig. 5 illustrates the location of the canal and the railway network in the region. The areas North and South are indicated in dark pink respectively light pink. Since many people commute from the North to jobs in Amsterdam and the neighboring municipality of Haarlemmermeer (the location of Schiphol airport), this connection is important for the Dutch economy. As a policy experiment, we consider what difference the availability of these rail tunnels makes. We calculate a counterfactual in which the rail tunnels are closed, so that no train connection is possible between North and South, and compare it with the current equilibrium.

Fig. 5 North Sea canal area



6.1 Framework for the welfare analysis

There are four types of agents in our model, three types of workers differing by their level of education s and the class of absentee landlords. Landlords might be further subdivided in local subgroups, as we will do in our empirical application. The effect of the change in transport accessibility on the wealth of landlords Q_l is equal to the sum of the effect on land rents across all locations:

$$Q_l = \sum_{h \in H} A_h \left(R_h^n - R_h^o \right),$$

where the superscripts n and o refer to the new and the old equilibrium, respectively. The effect on the utility of consumers with education level s is derived from their expected utility, see equation (13). This general equilibrium effect can be decomposed into four components:

(i) the effect of the change in the commuting cost for mode m for people who actually use that mode;

(ii) the effect of people changing modes because the relative cost have changed;

(iii) the effect of people changing jobs location because some jobs have become more easily accessible;

(iv) the effect of people changing their home location because locations change in their relative attractivity.

The expression for the calculation of these components are presented in Table 10. They follow from equations (8)-(13). For example, for the sum of effect (i), (ii), and (iii):

$$\frac{d\mathrm{E}\left[\ln V_i\left(R_h; x_i^*\right)|s\right]}{dg_{sh}} = \frac{d\mathrm{E}\left[\ln V_i\left(R_h; x_i^*\right)|s\right]}{dv_s}\frac{dv_s}{dv_{sh}}\frac{dv_{sh}}{dg_{sh}} = \mu_{Ks}\Pr[h|s]\mu_{Js}.$$

The first factor follows from equation (**??**), the second from equation (12) (since $\frac{dv_s}{dv_{sh}} = \exp\left(v_{sh} + \ln N_h - v_s\right) = \Pr[h|s]$), and the third from equation (10). The other effects are derived similarly. All effects are expressed in terms of money equivalents by multiplying them by the average wage W_s for education level s.

Note that the calculation of the first three components in Table 10 does not require the calculation of the new equilibrium. However, when considering the relocation of people between residential locations h we have to solve for changes in the number of houses N_h and land rents R_h for each location h. This requires finding a solution to a system of $2H$ simultaneous equations. This system is solved by starting with a vector of land rents for each h, calculating N_h from the system of equations (18), and then calculating the demand for land at each location from equation (17). For those locations where demand exceeds supply A_h, the land rent is increased and the other way around. This algorithm converges to an equilibrium.

Table 10 Decomposition of the general equilibrium effect per person[a)]

effect	equation	expression
users of mode m	(4)	$W_s\mu_{Ks}\mu_{Js}\mu_{Ms}\sum_{h\in H}\sum_{j\in H}\Pr\left[hjm|s\right]\gamma_s(c^n_{shjm}-c^o_{shjm})$
idem + modal shift	(9)	$W_s\mu_{Ks}\mu_{Js}\mu_{Ms}\sum_{h\in H}\sum_{j\in H}\Pr\left[hj|s\right]\left(c^n_{shj}-c^o_{shj}\right)$
idem + job relocation	(10)	$W_s\mu_{Ks}\mu_{Js}\sum_{h\in H}\Pr\left[h|s\right]\left(g^n_{sh}-g^o_{sh}\right)$
total	(12)	$W_s\mu_{Ks}\left(v^n_s-v^o_s\right)$

[a)] All probabilities are evaluated in the old equilibrium.

6.2 Results

The model and the expressions in Table 10 are applied to a policy experiment, using the parameter values estimated in Section 6. We compare the counterfactual equilibrium to the current equilibrium.

Table 11 describes the relocation of economic activity between the regions north and south of the canal due to the availability of the tunnels. The better connection of the less productive region north of the canal to the vibrant metropolitan area around Amsterdam leads to a relocation of jobs from the North to the South. The number of jobs in the North declines by 3%. Some 33.000 workers commute by train from the North to the South; 80% are additional commuters. Fig. 6 documents the job relocation process. However, the lower concentration of jobs in the North comes along with a higher quality of living, as can be seen from the increase in land prices in the North in particular along the railway corridors (see Fig. 7). Higher land prices lead to a lower land use per worker. Hence, the total population in the North goes up. Since Amsterdam is particularly attractive as a job location for higher educated and since higher educated prefer travelling by train, the main part of the population increase are higher educated, their population being 7% higher due to the availability of the tunnels. The analysis shows that a new commuting link may lead to a flight of jobs from the periphery, but also to an increase in the price of residential land by making the region a more attractive residential area, especially along the railways and in particular for higher educated.

In our framework transport infrastructure affects welfare through two channels: changes in population composition and changes in land use intensity. This policy experiment illustrates the interaction between the two mechanisms. Improved rail accessibility of the North attracts new, mostly high educated, population to the region (first mechanism). It is efficient that the newcomers live next to the rail stations as they value their proximity the most. This requires, however, adjustments in land use (second mechanism), as the newcomers also have other land consumption preferences than the incumbents. Stated differently, investments in local public goods may fail to generate the expected benefits if they do not go hand in hand with redevelopment of the housing stock.

Table 11 Residents, job, and commuting North and South, in thousands

tunnels:	no				yes			
	low	middle	high	total	low	middle	high	total
residents:								
North	126	183	169	478	128	190	181	499
South	171	270	378	819	171	270	377	818
jobs:								
North	109	147	135	391	106	143	131	380
South	206	341	443	989	210	350	453	1013
commuting:								
North-South	24	43	42	109	28	53	55	136
train	0	0	0	0	5	12	16	33
car	19	35	35	89	18	33	33	84
South-North	8	11	14	33	8	11	16	35
train	0	0	0	0	0.3	0.6	1	2
car	7	10	13	30	7	10	13	30
North-North	98	132	115	345	95	126	108	329
South-South	133	208	289	630	133	208	288	628

Table 12 reports the welfare gains from the tunnels. The benefits are distributed unevenly among education levels: high skilled individuals benefit more, since they have the highest preference for commuting by train and the most to gain from being able to commute to the vibrant Amsterdam economy with its wide availability of high paying jobs. Their benefits are three times as large as the benefits of middle educated and ten times larger than the gains of low educated individuals. The net welfare benefits for the landowners are relatively small: landowners in North and South gain, landowners elsewhere loose. This is due to the greater attractivity of the North for living, which reduces the demand for land elsewhere in the country. The table also shows that land owners cannot expropriate the total benefits from the public good. This is in line with Kuminoff and Pope (2014) and Bayer et al. (2007) who report a wedge between the capitalization effect and the total welfare effect of a policy measure. This effect arises due to changes in hedonic schedule caused by relocation of people and can be very substantial, as illustrated by our counterfactual example.

Fig. 6 Jobs flight to highly productive job locations south of the canal

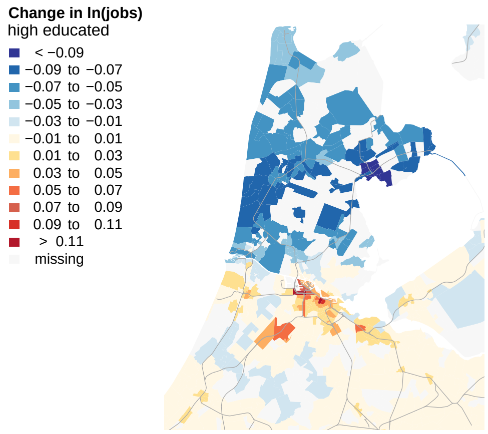
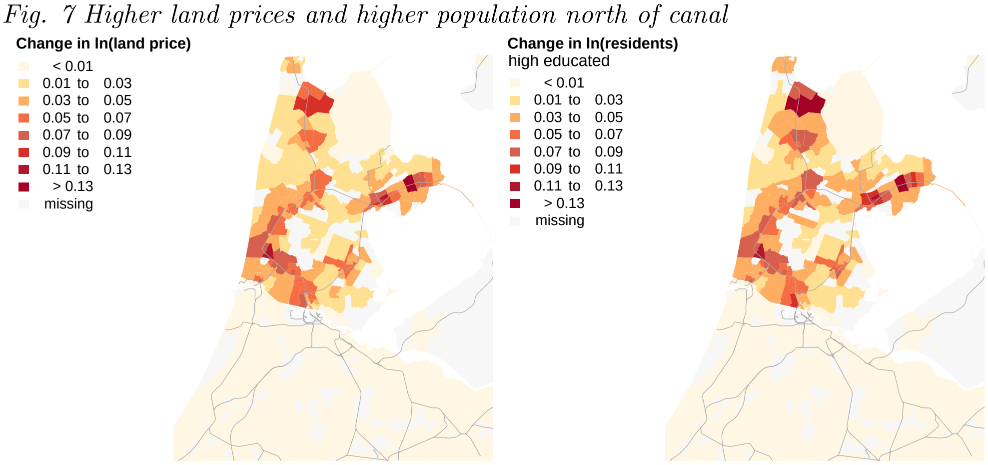


Fig. 7 Higher land prices and higher population north of canal



The direct gain via the modal split makes up for about two thirds of the total welfare gain, while job relocation accounts for a quarter. Although the home relocation effect and the effect for landowners are relatively small in total, they are very important for the distribution of the gains. The land owners transfer part of their benefits to consumers. This result arises because the new transport connection relaxes the tense land market in Amsterdam leading to lower land rents there. High educated benefit most from moving to the North, their benefits from home relocation are therefore the largest. Low and middle educated gain much less because they derive little benefit from the new infrastructure while they face higher land rents due to high educated workers driving up land rents at locations close to stations.

Table 12 Decomposition welfare effects, in mln euros

effect	education level			land owners			total
	low	middle	high	North	South	elsewhere	
modal split	151	470	1329				1950
job relocation	51	183	578				812
home relocation	13	43	183				239
land owners				1659	-188	-1570	-99
total	215	696	2090	1659	-188	-1570	2902

7 Conclusion

We have developed and estimated a structural spatial general equilibrium model for the valuation of the effects of investments in public goods on home and job location choice and land use. Our results suggest that ignoring the changes in the intensity of land use and relocation of people with different education levels misses important determinants of the size and distribution of welfare gains from infrastructural investments. As a policy experiment we calculated the welfare benefits of two railway tunnels connecting Amsterdam to the region North of the city. The direct effect of the tunnels on the travel times and modal split ignores up to 30% of the total general equilibrium effect. These wider gains come together with increases in the population density and the share of high educated in the North, due to better job market access. The benefits of the railway tunnels are distributed highly unequally across education levels, the gains for high educated being ten times larger than for low educated. This unequal distribution of benefits poses a challenge for the political economy of investments in public goods and specifically transport infrastructure. Considerable changes in land use intensity and population composition show that large investments in public goods should be accompanied by land redevelopment. Keeping housing supply fixed prohibits the efficient use of new infrastructure by population groups who value it most and leads to foregone benefits.

Our model focusses on some main mechanisms through which investments in public goods affect the intensity of land use and the composition of population. This allows to keep the analysis tractable. We do not account for the option of transferring land from agricultural to residential use. Similarly, we do not allow local wages or the local supply of amenities to be adjusted to changes in the structure of the economy. Ignoring these margins of adjustment leads to an underestimation of the benefits of public goods. Furthermore, there is no feedback of changes in modal split on travel times. For example, if the closure of the railway tunnels were to lead to a massive increase in car traffic, that would increase travel times for these trips. However, travel times are treated as exogenous in our application. This also leads to an underestimation of the benefits. Since travel by car did not massively increase in our policy experiment, this does not substantially affect our conclusions. Finally, we have studied preference heterogeneity between three education levels. Our framework can easily be extended to more socioeconomic groups, e.g. males versus females, singles versus couples, yielding new interesting insights.

References

Ahlfeldt, G.M., Redding, S.J., Sturm, D.M., Wolf, N., 2015. The economics of density: Evidence from the Berlin Wall, Econometrica 83: 2127-2189.

Albouy, D., Ehrlich, G. 2012. Metropolitan land values and housing productivity, NBER Working Paper 18110.

Albouy, D., 2009. The unequal geographic burden of federal taxation, Journal of Political Economy 117, 635–667.

Allen, T., Arkolakis, C., 2014. Trade and the topography of the spatial economy, Quarterly Journal of Economics 129 (3), 1085–1140.

Anas, A., Liu, Y., 2007. A regional economy, land use, and transportation model (RELUTRAN): Formulation, algorithm design, and testing, Journal of Regional Science 47, 415–455.

Anderson, M., 2014. Subways, strikes, and slowdowns: The impacts of public transit on traffic congestion, American Economic Review 104 (9): 2763–2796.

Arnott, R., MacKinnon, J. 1977. The effects of urban transportation changes: A general equilibrium simulation, Journal of Public Economics 8, 19–36.

Baum-Snow, N., 2007. Did highways cause suburbanization?, Quarterly Journal of Economics 122 (2), 775–805.

Baum-Snow, N., 2010. Changes in transportation infrastructure and commuting patterns in US metropolitan areas, 1960-2000, American Economic Review 100, 378–382.

Baum-Snow, N., Kahn, M.E., 2000. The effects of new public projects to expand urban rail transit, Journal of Public Economics 77 (2), 241–263.

Bayer, P., McMillan, R., Rueben, K., 2004. An equilibrium model of sorting in an urban housing market, NBER Working Paper 10865.

Bayer, P., Ferreira, F., McMillan, R., 2007. A unified framework for measuring preferences for schools and neighborhoods, Journal of Political Economy 115 (4), 588–638.

Bayer, P., Keohane, N., Timmins, Ch., 2009. Migration and hedonic valuation: The case of air quality, Journal of Environmental Economics and Management 58, 1–14.

Bayer, P., McMillan, R., 2012. Tiebout sorting and neighborhood stratification, Journal of Public Economics 96,1129–1143.

Bayer, P., Fang, H., McMillan, R., 2014. Separate when equal? Racial inequality and residential segregation, Journal of Urban Economics 82, 32–48.

Ben-Akiva, M., Lerman, S.R., 1985. Discrete Choice Analysis. The MIT Press, Cambridge, MA.

Berry, S., Levinsohn, J., Pakes, A. 1995. Automobile prices in market equilibrium. Econometrica 63 (4), 841–890.

Bowes, D.R., Ihlanfeldt, K.R., 2001. Identifying the impacts of rail transit stations on residential property values, Journal of Urban Economics 50 (1), 1–25.

Buchinsky, M., Gotlibovsky, C., Lifshitz, O., 2014. Residential location, work location, and labor market outcomes of immigrants in Israel, Econometrica 82, 995–1054.

Cardell, N.S., 1997. Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity, Econometric Theory 13, 185–213.

Charles, K.K., Hurst, E., Notowidigdo, M.J., 2015. Housing Booms and Busts, Labor Market Opportunities, and College Attendance, mimeo

Combes, P-Ph., Duranton, G., Gobillon, L., Puga, D., Roux, S., 2012. The productivity advantages of large cities: Distinguishing agglomeration from firm selection, Econometrica 80 (6), 2543–2594.

Davis, M.A., Heathcote, J., 2007. The price and quantity of residential land in the United States, Journal of Monetary Economics 54, 2595–2620.

Davis, M.A., Palumbo, M.G., 2008. The price of residential land in large US cities, Journal of Urban Economics, 63, 352–384.

Desmet, K., Rossi-Hansberg, E., 2013. Urban accounting and welfare, American Economic Review 103 (6), 2296–2327.

Diamond, R., 2016. The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000, American Economic Review, 106 (3), 479–524.

Dijk, M. van, Romijn, G., 2010. Reforming the Dutch housing policies (Hervorming van het Nederlandse woonbeleid, in Dutch), CPB Netherlands Bureau for Economic Policy Analysis, The Hague.

Duranton, G., Turner, M.A., 2012. Urban growth and transportation, Review of Economic Studies 79 (4), 1407–1440.

Duranton, G., Turner, M.A. 2011. The fundamental law of road congestion: Evidence from US cities, American Economic Review 101, 2616–2652.

Ellickson, B., 1981. An alternative test of the hedonic theory of housing markets, Journal of Urban Economics 9, 56–79.

Epple, D., Sieg, H., 1999. Estimating equilibrium models of local jurisdictions, Journal of Political Economy 107, 645–681.

Gamper-Rabindran, S., Timmins, Ch., 2011. Hazardous waste cleanup, neighborhood gentrification, and environmental justice: Evidence from restricted access census block data, American Economic Review: Papers and Proceedings 101 (3), 620–624.

Gautier, P.A., Svarer, M., Teulings, C.N., 2010. Marriage and the city: Search frictions and sorting of singles, Journal of Urban Economics 67, 206–218.

Gibbons, S., Machin, S., 2005. Valuing rail access using transport innovations, Journal of Urban Economics 57 (1), 148–169.

Glaeser, E.L., Gottlieb, J.D., 2009. The wealth of cities: Agglomeration economies and spatial equilibrium in the United States, Journal of Economic Literature 47 (4), 983–1028.

Glaeser, E.L., Gyourko, J., Saks, R., 2005. Why is Manhattan so expensive? Regulation and the rise in housing prices, Journal of Law and Economics 48, 331–369.

Glaeser, E.L., Kahn, M.E., 2004. Sprawl and urban growth, in: Henderson, J.V., Thisse, J.-F. (Eds), Handbook of Regional and Urban Economics, vol. 1, Elsevier, Amsterdam, pp. 2481–2527.

Groot, H.L.F. de, 2011. Determination of land rents: a simple approach.

https://www.cpb.nl/sites/default/files/Determination-of-Land-Rents-2011.pdf

Gyourko, J., Mayer, Ch., Sinai, T., 2013. Superstar cities, American Economic Journal: Economic Policy 5 (4), 167–199.

Haughwout, A.F., 2002. Public infrastructure investments, productivity and welfare in fixed geographic areas, Journal of Public Economics 83, 405–428.

Hornbeck, R., 2012. The enduring impact of the American dust bowl: Short- and long-run adjustments to environmental catastrophe, American Economic Review 102 (4), 1477–1507.

Horowitz, J.L., 1986. Bidding models of housing markets. Journal of Urban Economics 20, 168–190.

Inoa, I., Picard, N., Palma, A. de, 2014. Effect of an accessibility measure in a model for choice of residential location, workplace, and type of employment, mimeo.

Klaiber, H.A. and N.V.Kuminoff (2014). Equilibrium Sorting Models of Land Use and Residential Choice. In Duke J.M and J-J. Wu (eds.) The Oxford Handbook of Land Economics.

Klaiber, H.A., Smith, V.K., 2010. Valuing incremental highway capacity in a network, NBER Working Paper 15989.

Kline, P., Moretti, E., 2013. Place based policies with unemployment, American Economic Review 103, 238–243.

Kuminoff, N.V., Pope, J.C., 2014. Do “capitalization effects” for public goods reveal the public’s willingness to pay?, International Economic Review 55, 1227–1250.

Lau, L. 1976. A Note on Elasticity of Substitution Functions, The Review of Economic Studies 43(2), 353-358.

Lerman, S.R., Kern, C.R., 1983. Hedonic theory, bid rents, and willingness-to-pay: Some extensions of Ellickson’s results, Journal of Urban Economics 13, 358–363.

Lucas, R.E., Rossi-Hansberg, E., 2002. On the internal structure of cities, Econometrica 70, 1445–1476.

Roback, J., 1982. Wages, rents and the quality of life, Journal of Political Economy 90, 1257–1278.

Sieg, H., Smith, V.K., Banzhaf, H.S., Walsh, R., 2004. Estimating the general equilibrium benefits of large changes in spatially delineated public goods. International Economic Review 45 (4), 1047–1077.

Sieg, H., Smith, V.K., Banzhaf, H.S., Walsh, R., 2002. Interjurisdictional housing prices in locational equilibrium, Journal of Urban Economics 52, 131–153.

Significance, VU University Amsterdam, John Bates Services, in collaboration with TNO, NEA, TNS NIPO and PanelClix, 2013. Values of time and reliability in passenger and freight transport in The Netherlands. Report for the Ministry of Infrastructure and the Environment.

Tra, C., 2010. A discrete choice equilibrium approach to valuing large environmental changes, Journal of Public Economics 94, 183–196.

Train, K., 2009. Discrete choice methods with simulation. Cambridge University Press, Cambridge.

Waddell, P., 1993. Exogenous workplace choice in residential location models: is the assumption valid? Geographical Analysis 25: 65-82.

Appendix A Derivation of the indirect utility function

The first-order conditions of the maximization problem in equation (1) yield expressions for L_i^*, B_i^*, and C_i^* as a function of R_h conditional on x. For example, for L_i^* we obtain:

$$L_i^*\left(R_h; x\right) = \widehat{L}^*\left(R_h\right) W_i(x),$$

where $\widehat{L}^*\left(R_h\right)$ is the optimal land consumption per unit of income. The individual specific term $\chi_i\left(x\right)$ drops out due to the multiplicative specification of the utility function. The effects of the take home pay $W_i(x)$ can be factored out since the utility function is homothetic and the production function of housing services features constant returns to scale. We omit the prices of other consumption and building as formal arguments, since these are constant across locations. We can write similar functions for B_i^* and L_i^*. Substitution of these demand functions for L, B and C in the utility function $U_i\left(\cdot\right)$ yields an expression for the indirect utility function $V_i\left(R_h; x\right)$ of the form given in equation (3), where we use $W_i(x) = W_s e^{w_s(j) - c_i(x)}$ and where $v\left(R_h\right) \equiv \ln U\left[F\left[\widehat{L}^*\left(R_h\right), \widehat{B}^*\left(R_h\right)\right], \widehat{C}^*\left(R_h\right)\right]$ is a twice differentiable function. Since $dU\left(\cdot\right)/dR_h < 0$ and $d^2U\left(\cdot\right)/dR_h^2 > 0$, the function $v\left(R_h\right)$ must satisfy $v'\left(R_h\right) < 0, v''\left(R_h\right) > -v'\left(R_h\right)^2$. Since $W_i(x)$ is the cost to an individual of acquiring a utility level $V = V_i\left(R_h; x\right)$, the cost function $C\left(\cdot\right)$ that goes with this indirect utility function reads:

$$C\left(R_h; V\right) = V \exp\left[-v\left(R_h\right) - \chi_i\left(x\right)\right] = W_i(x).$$

By Shephard's lemma the demand for land is the partial derivative of the cost function with respect to the price of land:

$$L_i^*\left(R_h; x\right) = C_R\left(R_h; V\right) = -v'\left(R_h\right) W_i(x). \tag{23}$$

The full cost function includes as arguments the prices of both land and all other expenditure. Since a cost function is homogeneous of degree one in prices, we can write

$$C\left(R_h, P; V\right) = PC\left(P^{-1}R_h, 1; V\right) = PV \exp\left[-v\left(P^{-1}R_h\right) - \chi_i\left(x\right)\right],$$

which is equal to the expression in the text for the normalization $P = 1$. The derivatives reads

$$\begin{aligned}
C_R &= -v'\left(P^{-1}R_h\right) V \exp\left[-v\left(P^{-1}R_h\right) - \chi_i\left(x\right)\right] =_{P=1} -v'\left(R_h\right) C, \\
C_P &= \left[1 + P^{-2}R_h v'\left(P^{-1}R_h\right)\right] V \exp\left[-v\left(P^{-1}R_h\right) - \chi_i\left(x\right)\right] =_{P=1} \left[1 + R_h v'\left(R_h\right)\right] C, \\
C_{RP} &= -P^{-2}R_h\left[v'\left(P^{-1}R_h\right)^2 + v''\left(P^{-1}R_h\right)\right] V \exp\left[-v\left(P^{-1}R_h\right) - \chi_i\left(x\right)\right] =_{P=1} R_h\left[v''\left(R_h\right) - v'\left(R_h\right)^2\right] C.
\end{aligned}$$

The absolute value of the elasticity of substitution η between land and all other expenditure (building plus other consumption) can be calculated as (see Lau, 1976):

$$\eta = -\frac{C_{RP}C}{C_R C_P} = -R_h \frac{v'' - {v'}^2}{v'\left(1 + R_h v'\right)}, \tag{24}$$

leaving out the argument(s) of the functions C and v and where the subscripts R of C denote the relevant partial derivatives.

Appendix B The error structure of the model

The terms $\varepsilon_{ikjm}, \varepsilon_{ikjM}, \widetilde{\varepsilon}_{ikj}, \varepsilon_{ikW}$, and $\widetilde{\varepsilon}_{ik}$ are individual specific random effects which follow a type I extreme value distribution with zero mean and variance $\pi^2/6$. The error terms ε_{ikjM} and ε_{ikW} are defined by

$$\begin{aligned} \varepsilon_{ikjM} &\equiv \max_m \left[-\gamma'_s c_{shjm} + \varepsilon_{ikjm}\right] + c_{shj}, \\ \varepsilon_{ikW} &\equiv \max_j \left[y_{sj} - \mu_{Ms} c_{shj} + \varepsilon_{ikj}\right] - g_{sh}. \end{aligned} \tag{25}$$

The distributions of $\widetilde{\varepsilon}_{ikj}$ and of of $\widetilde{\varepsilon}_{ik}$ are such that

$$\begin{aligned} \varepsilon_{ikj} &\equiv \widetilde{\varepsilon}_{ikj} + \mu_{Ms}\varepsilon_{ikjM}, \\ \varepsilon_{ik} &\equiv \widetilde{\varepsilon}_{ik} + \mu_{Js}\varepsilon_{ikW}. \end{aligned} \tag{26}$$

where $\widetilde{\varepsilon}_{ikj}$ and ε_{ikjM} and where $\widetilde{\varepsilon}_{ik}$ and ε_{ikW} are uncorrelated. Usually, the parameter μ_{Ks} can be normalized to unity without loss of generality. This is not the case in our model due to the land rent function $v\left(R_h\right)$ and due to the interpretation of $\ln V_i\left(R_h; x\right)$ as a log cost function, such that $L_i^* = -v'\left(R_h\right) W_i(x)$, see equation (23). Since $\varepsilon_{ikjm}, \widetilde{\varepsilon}_{ikj}$, and $\widetilde{\varepsilon}_{ik}$ follow a type I extreme value distribution, the choice problems in equation (8) and (10) are described by a logit model, see Ben Akiva and Lerman (1985), Waddell (1993), Cardell (1997), and Train (2009).

Appendix C Standard error in the transport cost identity

The estimation error of the left hand side of equation (6) is obtained by observing that the error terms in the various submodels that identify each of these parameters are independent: ε_{ikjm} for the estimation of γ_{s0}, see equation (8); ε_{ikj} for μ_{Ms}, see equation (10); ε_{ik} for μ_{Js}, see equation (14); and z_h for μ_{Ks}, see equation (15). A first order expansion of the variance of a product of independent random variable satisfies

$$\mathrm{Var}\left[XY\right] \cong \mathrm{E}^2\left[Y\right]\mathrm{Var}\left[X\right] + \mathrm{E}^2\left[X\right]\mathrm{Var}\left[Y\right],$$

Using $\mathrm{E}[\mu_{Ks}\mu_{Js}\mu_{Ms}\gamma_{s0}] = 1$, we obtain

$$\begin{aligned}\mathrm{Var}\left[\mu_{Ks}\widehat{\mu_{Js}\mu_{Ms}}\gamma_{s0}\right] &\cong \mu_{Ks}^{-2}\mathrm{Var}\left[\widehat{\mu}_{Ks}\right] + \mu_{Js}^{-2}\mathrm{Var}\left[\widehat{\mu}_{Js}\right] + \mu_{Ms}^{-2}\mathrm{Var}\left[\widehat{\mu}_{Ms}\right] + \gamma_{s0}^{-2}\mathrm{Var}\left[\widehat{\gamma}_{s0}\right] \\ &= t\left(\widehat{\mu}_{Ks}\right)^{-2} + t\left(\widehat{\mu}_{Js}\right)^{-2} + t\left(\widehat{\mu}_{Ms}\right)^{-2} + t\left(\widehat{\gamma}_{s0}\right)^{-2},\end{aligned}$$

where we use $t\left(\widehat{\mu}_{Ks}\right) \equiv \mu_{Ks}/\sqrt{\mathrm{Var}\left[\widehat{\mu}_{Ks}\right]}$, where $t\left(\widehat{\mu}_{Ks}\right)$ is the t-statistique of $\widehat{\mu}_{Ks}$. For $\widehat{\mu}_{Ks}$, we estimate $\widehat{\mu}_{Ks}^{-1}$ and $t\left(\widehat{\mu}_{Ks}^{-1}\right)$. Since

$$\begin{aligned}\mathrm{Var}\left[f\left(X\right)\right] &\cong f'\left(\mathrm{E}\left[X\right]\right)^2\mathrm{Var}\left[X\right] \Rightarrow \\ \mathrm{Var}\left[X^{-1}\right] &\cong \mathrm{E}^{-4}\left[X\right]\mathrm{Var}\left[X\right] = \mathrm{E}^{-2}\left[X\right]t\left(X\right)^{-2} \Rightarrow t\left(X^{-1}\right) \cong t\left(X\right),\end{aligned}$$

where we use $\mathrm{E}\left[X^{-1}\right] \cong \mathrm{E}^{-1}\left[X\right]$ in the final step. The assymptotic t-statistique of X is equal to the assumptotic t-statistique of X^{-1}. Like the standard errors for the coefficients of the nested logit models, this expression for the standard error does not account for the estimation error of coefficients estimated in early stages. Hence, it is a lower bound of the true standard error.

Appendix D Data

Travel times by car are reported for the morning peak hour between 7 and 9 a.m. When multiple routes are possible, travel times, costs and distances are calculated as averages over all possible routes, weighted by the number of commuters using each route. The cost of car travel has been set at 0.3 euro for every kilometer traveled plus toll costs.[14] Travel times by train and bus/tram/metro are split up between in- and out-of-vehicle times. Travel costs for the train have been provided by the Ministry of Transportation; travel costs for bus/tram/metro are calculated from the number of urban transit zones traveled.[15] Biking and walking travel times are calculated by using the travel distances calculated for car trips, assuming an average speed of 16 kilometer per hour. The costs of these trips are set equal to zero. We deleted implausible observations, e.g. for which the actually chosen travel mode

[14] This includes fuel, amortization, insurance, maintenance, and taxes for a car in a medium-price range, using a gasoline price of €1.25 per litre or €0.10 per kilometre for 2005 and of €1.78 per litre or €0.15 per kilometre for 2012 (http://www.autoweek.nl/kostenberekening.php?id=35685&jaar=2005).

[15] Cost = €0.43 times the number of urban transit zones plus one.

is characterized by very large or very small travel times and/or distances (below the 2.5 or above the 97.5 percentile for the mode concerned), or home-work distances smaller than the home-work straight line.

Table 14 presents the descriptive statistics of the mode choice by education level. Table 15 presents the number of ZIP codes with a positive number of workers and people living there. Summary statistics for the amenity variables, for land prices, and for ZIP code fixed effects in wages are presented in Table 16.

Table 14 Descriptive statistics for commuting data

mode	car		train		bus		bike, walk	
	mean	st.dev.	mean	st.dev.	mean	st.dev.	mean	st.dev.
modal share	0.71		0.05		0.03		0.21	
distance, km	44.7	(36.9)	87.9	(47.4)	29.9	(19.2)	11.6	(7.2)
duration, min	48.9	(30.0)	143.6	(42.4)	85.9	(34.4)	43.4	(27.0)
cost, euro	6.8	(5.4)	6.8	(3.3)	3.7	(1.5)	0	

Table 15 Coverage of the population data commuters

	our dataset	The Netherlands
# home ZIP codes	3520	4019
# work ZIP codes	3247	4015
working population (mln)	7.40	7.50
mean # residents by ZIP code	2103	1867
mean # jobs by ZIP code	2384	1951
fraction males	0.56	0.56
fraction per education level:		
- low	0.25	0.27
- middle	0.36	0.44
- high	0.38	0.30

Table 16 Descriptive statistics on wages, land prices, and amenities by ZIP code

variable	unweighted		weighted by # residents/workers		# ZIP codes
	mean	st.dev.	mean	st.dev.	
ln net daily wage low	4.514	0.039	4.515	0.038	2729
ln net daily wage middle	4.668	0.054	4.678	0.053	2834
ln net daily wage high	4.961	0.044	4.976	0.039	2510
ln daily land rent	-3.689	0.901	-3.329	0.891	2758
# monum.within 1km/1000	0.032	0.161	0.051	0.238	2758
# monum.1 to 5km/1000	0.284	0.779	0.463	1.161	2758
share nature within 5km	0.127	0.114	0.132	0.109	2758
dummy uni.10km	0.269	0.444	0.380	0.485	2758
# rest. within 1km/100	0.054	0.159	0.086	0.224	2758
# rest.1 to 5km/100	0.634	1.507	1.055	2.242	2758
share social housing	0.256	0.157	0.298	0.155	2754

Appendix E Calculation of land rents

We calculate land rents for some 4,000 four digit zip codes in the Netherlands using the hedonic price methodology. We exploit unique geo-referenced microdata on more than 1 million housing sale transactions 1985-2007 provided by the Dutch Organization of Real Estate Brokers (NVM). The NVM covers 70–80% of all housing transactions on average, with urban regions being somewhat overrepresented and peripheral regions being somewhat underrepresented. For each house the NVM documents a range of structural characteristics including land lot size, living space, type of house (terraced, corner, semi-detached, etc.), presence of a garage or own parking space, presence of central heating, the year of construction, etc. Table 17 contains descriptive statistics of the main variables.

We estimate the following regression model (see Glaeser et al., 2005; Davis and Heathcote, 2007; and Davis and Palumbo, 2008, Groot, 2011):

$$\ln P_{ijt} = \alpha + \sum_{j=1}^{J} \beta_j \ln L_{ijt} + \sum_{k=1}^{K} \gamma_k X_{kijt} + \sum_{t=1985}^{T} \delta_t D_{ijt} + \varepsilon_{ijt},$$

where P_{ijt} is the price of house i in area j at time t, L_{ijt} stands for the lot size, and the X's are house characteristics that we control for. The key parameters of interest are the β_j's. These capture the share of land in the total transaction price. We allow β_j to vary over four-digit zip codes; in urban areas these zip codes cover one squared kilometer. Note that $\beta_j = dlnP_{ijt}/dlnL_{ijt} = L_{ijt}/P_{ijt} \cdot dP_{ijt}/dL_{ijt}$. Since dP_{ijt}/dL_{ijt} is the marginal effect of an additional square meter of land on the transaction price, it can be interpreted as the marginal price of land. Therefore β_j is the share of land in the housing price. Using this information, the price of land per square meter can be easily derived as $\beta_j \cdot P_{ijt}/L_{ijt}$. We correct for overall price increases by adding the time dummies.