CESifo WORKING Papers

Welfare Benefits of Agglomeration and Worker Heterogeneity

Coen N. Teulings
Ioulia V. Ossokina
Henri L.F. de Groot

CESifo
Center for Economic Studies & Ifo Institute

Welfare Benefits of Agglomeration and Worker Heterogeneity

Abstract

The direct impact of local public goods on welfare is relatively easy to measure from land rents. However, the indirect effects on home and job location, on land use, and on agglomeration benefits are hard to pin down. We develop a spatial general equilibrium model for the valuation of these effects. The model is estimated using data on transport infrastructure, commuting behavior, wages, land use and land rents for 3000 ZIP-codes in the Netherlands and for three levels of education. Welfare benefits are shown to differ sharply by workers' educational attainment.

Coen N. Teulings
Amsterdam School of Economics
VU University of Amsterdam
Roetersstraat 11
The Netherlands - 1018 WB Amsterdam
c.n.teulings@uva.nl

Ioulia V. Ossokina
CPB Netherlands Bureau for
Economic Policy Analysis
P.O. Box 80510
The Netherlands - 2508 GM The Hague
ioulia.ossokina@cpb.nl

Henri L.F. de Groot
Faculty of Economics and Business
Administration
VU University of Amsterdam
De Boelelaan 1105
The Netherlands - 1081 HV Amsterdam
h.l.f.de.groot@vu.nl

1   Introduction

Since the work of Rosen (1974), Arnott and MacKinnon (1977), and Roback (1982), a large literature has developed on the impact of local public goods like transport infrastructure on aggregate welfare. Estimates of this impact are usually derived from the direct effect of these public goods on local factor prices, like wages and land rents. Since the economies of agglomeration play a crucial role in the spatial distribution of economic activity, there remains a suspicion that these studies underestimate the true impact because they do not properly account for the additional benefits from agglomeration. Moreover, locations differ by their supply of local public goods. Since people are heterogeneous in their preferences for these public goods, new public goods will invoke reallocation of people accross locations. On the one hand, this reallocation allows a country to take greater advantage of its supply of public goods, which should be included in the assessment of the impact of public goods. On the other hand, the heterogeneity in preferences implies that the benefits of new public goods are not evenly distributed among the population, raising political economy issues regarding decision making and funding of new projects. This paper contributes to this literature by constructing a detailed spatial general equilibrium model that is computationally tractable and at the same time allows the valuation of a wide range of indirect effects of concrete investment projects. The data requirements for this exercise are huge. The model is estimated for the Netherlands around 2004, using a refined regional disaggregation into some 3000 ZIP codes, for which we have information on the transport infrastructure between every combination of ZIP codes for four modes of transport, on local amenities, on wages, each for three levels of education, and on residential land rents and land use. Within bigger cities, each ZIP code covers about one square kilometre.

Land rents play a crucial role in the valuation of local public goods. Any direct benefit translates into higher land rents in the neighbourhood of the public good. The land rent gradient is a measure of the cost of travelling to the public good. The total land rent differential is therefore a first approximation of the net benefit, that is, the gross benefit of all people living sufficiently close to take advantage of the public good minus their cost of travelling to its location. Since land is more expensive at locations close to public goods, the average lot size is lower and population density is higher at these locations. This furthers an efficient use of the public goods, since the high population density allows many people to benefit from their proximity.

Figure 1 shows the land rents in the Netherlands for every ZIP code. Even in a rather homogeneous country as the Netherlands, land rents vary by almost a factor 400, ranging from 3800 euro per square meter in Amsterdam's Canal Zone to some 10 euro along the North-East border with Germany. Land rents have local peaks in cities, where there is large supply of local public goods, such as train stations, job availability, museums and cultural performances, the availability of shops and restaurants, and a nice scenery with historical buildings as exemplified by Amsterdam's Canal Zone. Figure 2 shows that there is also a wide dispersion in the composition of the population by level of education. The

maps show the shares of higher and lower educated workers living at each ZIP code, leaving the middle education level as the omitted category. In some ZIP codes, more than 70% of the working population is higher educated, while in others this is less than 15%. The same applies for the lower educated. Hence, there is a strong spatial segregation by education. Roughly speaking: the higher educated live in the city, the lower educated on the countryside. This outcome is consistent with the observation that the valuation of public goods differs between subgroups in the population.
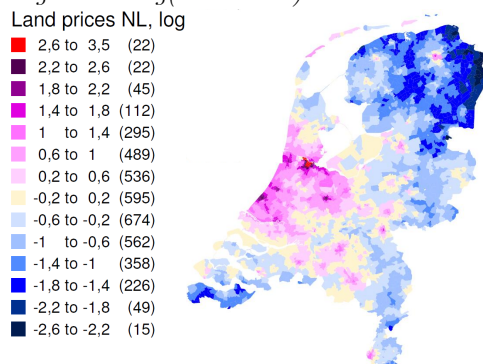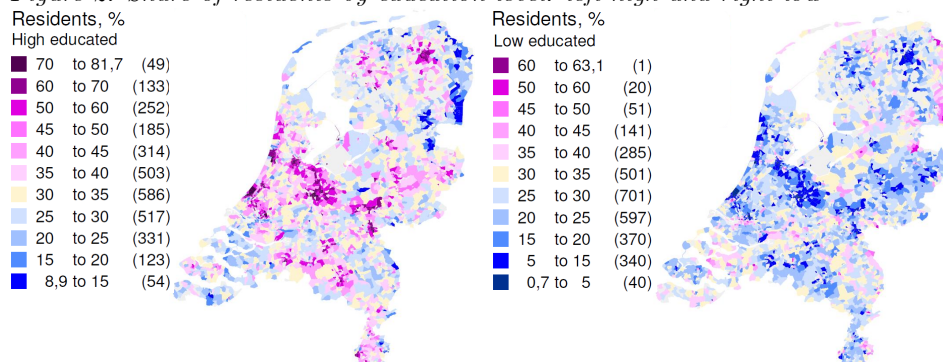
Figure 1. Log(land rents) in the Netherlands[1]

Land prices NL, log

2,6 to 3,5	(22)
2,2 to 2,6	(22)
1,8 to 2,2	(45)
1,4 to 1,8	(112)
1 to 1,4	(295)
0,6 to 1	(489)
0,2 to 0,6	(536)
-0,2 to 0,2	(595)
-0,6 to -0,2	(674)
-1 to -0,6	(562)
-1,4 to -1	(358)
-1,8 to -1,4	(226)
-2,2 to -1,8	(49)
-2,6 to -2,2	(15)

Figure 2. Share of residents by education level: left high and right low

Residents, %
High educated

70 to 81,7	(49)
60 to 70	(133)
50 to 60	(252)
45 to 50	(185)
40 to 45	(314)
35 to 40	(503)
30 to 35	(586)
25 to 30	(517)
20 to 25	(331)
15 to 20	(123)
8,9 to 15	(54)

Residents, %
Low educated

60 to 63,1	(1)
50 to 60	(20)
45 to 50	(51)
40 to 45	(141)
35 to 40	(285)
30 to 35	(501)
25 to 30	(701)
20 to 25	(597)
15 to 20	(370)
5 to 15	(340)
0,7 to 5	(40)

Our model is designed to capture these features. We model a spatial equilibrium where workers choose their job and home location and decide how to commute between these locations. Our basic tool is McFadden's well known utility maximizing interpretation of the logit model. For the purpose of our exercise, we add three extensions to this framework. First, we model individual land use. Where most researchers take the supply of housing and hence lot size as given, we model the choice between the consumption of land and other consumption. People decide how much land to consume depending on the local land rent. 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 by renting

---

[1] In Figure 1, 0 corresponds to a location in Enschede with an average land price of 109 euro/m^2.

out part of an apartment to students, or as a gradual reconstruction of a neigh-
bourhood after the land rent has changed. This extension of the model allows us
to estimate the elasticity of substitution between land and other consumption,
which is estimated to be around 0.75. Hence, the land share is increasing in the
level of the land rent.

The second extension is the way in which we model the choice of home
location and the implications for the land rent equation. People choose to live in
that location that offers them the highest utility. Land rents adjust to equalize
supply and demand for land at each location. Hence, if all individuals were
homogeneous, then the estimation of a logit model for home location would not
make sense, since land rents would adjust to equalize utility across locations.
Instead, one could better estimate a land rent equation directly. However, our
model features three types of workers, low, middle, and high educated. Hence,
locations might vary in their attractivity for various levels of education. We show
that the problem is analyzed far more easily by transposing the conditionality of
the probability, applying Bayes' rule, see Ellickson (1981): we do not analyze the
probability that an individual chooses home location h conditional on her level
of education s, but we analyze the probability of education level s conditional on
the characteristics of location h. This greatly reduces the computational burden.
Then, the local land rent gradient for a particular type of public good depends
on the type of individual that prefers to live at that location. If a location is
predominantly chosen by higher educated workers, the local land rent gradient
for public goods will correspond to the preferences of higher educated. Since
the local land rent is endogenous, it has to be instrumented. Our model shows
job availability to be a natural instrument.

Applying this methodology, we show that preferences for public goods differ
widely between levels of education. Our estimation method allows for an easy
inclusion of this selection effect on the local land rent gradient in the land rent
equation. Hence, the pay off of creating new public goods depends on the local
composition of the population, leading to a further increase in the pattern of
spatial specialization. This might explain the pattern of segregation shown in
Figure 2. Spatial differences in job availability explain about one third of the
variation in land rents, observed and unobserved amenities explain the other two
thirds. Our methodology allows us to calculate the hypothetical land rents that
would apply if everybody had a lower or a higher level of education, respectively.
Where the standard deviation of actual log land rents among ZIP codes is 0.90,
this standard deviation would be 1.35 if everybody were lower educated and it
would be 2.30 if everybody were higher educated. Locations that are in low
demand from one level of education, are in high demand from other education
levels. Since people sort into the locations in which they enjoy the greatest
comparative benefit from the local attributes, this process compresses land rent
differentials.

The combination of both extensions, land use and the land rent regression,
allows an interesting test of the model by an application of Shephard's lemma.
The weight of the land rent in the home location choice model should be equal
to the share of land rent in total consumption. We offer a formal derivation of

4

this test and show that this equality holds approximately.

The third extension of the logit model is that we allow for agglomeration externalities in production, a positive effect of higher employment density at a particular location on productivity of firms there. Estimating this relation is complicated, since the number of workers at a location is endogenous. There is a vast literature on this issue, using all kinds of instruments to tackle this problem, see Combes et al. (2012), Glaeser and Gottlieb (2009) and Ahlfeldt et al. (2012). We use a number of instruments, none of them being really satisfactory. The problem is that the spatial scale of the agglomeration externalities is small, less than a kilometre, while our instruments can only discriminate between wider areas. We have little new to offer here compared to the literature. In the end, we trust our OLS results most, which suggest that agglomeration economies predominantly apply at a very short range, in particular for higher educated workers.

We apply our model to the analysis of a policy experiment. The city of Amsterdam is located just south of a major canal, connecting the Amsterdam harbour to the North Sea. Many people commute from the North to jobs in Amsterdam and the neighbouring municipality of Haarlemmermeer (the location of Schiphol airport). The main connections between Amsterdam and the area north of the canal are five highway tunnels and two train tunnels. As a policy experiment, we calculate a counterfactual in which the train tunnels are closed. The availability of the tunnels leads to a shift in the modal split from the car to the train. However, more important is the shift in job and home locations, leading to a massive change in land rents. Land rents north of the canal are substantially higher in the equilibrium with train tunnels, in particular close to railway stations. Land rents in Amsterdam are slightly higher, due to the gain in agglomeration benefits from workers commuting from the North. Land rents in the rest of the country are lower, since people move to the greater Amsterdam region. The effect on the welfare of higher educated is ten times bigger than for lower educated, since the former prefer travelling by train more than others and since their wage differs more between job locations. The standard direct effect estimate would underestimate the total effect by some 30% in this particular example, mainly by ignoring job relocation.

This paper contributes to a number of strands in the literature. Haughwout (2002) develops a spatial general equilibrium framework with aggregate investment in regional transport infrastructure. Others examine the effects of infrastructure investments on the house prices, see Klaiber and Smith (2010), Bowes and Ihlandfeldt (2001), and Gibbons and Macchin (2005). The advantage of our study is that we can disentangle the direct effect from the indirect effects due to changes in home and job location, and in local productivity due to agglomeration. Baum-Snow and Kahn (2000) address changes in the modal shift. Anas and Liu (2007) model the interaction between land use and transportation infrastructure, while Inoa et al. (2014) analyse the interdependence of home and job location. Our paper is also related to the recent literature on the effect of transport infrastructure on the spatial distribution of economic activity. The theory of land use by Alonso (1964) predicts that faster commuting increases

the demand for land in suburbs relative to central cities. Baum-Snow (2007, 2010) finds empirical evidence for this prediction for the USA in the second part of the 20th century (see also Glaeser and Kahn, 2004).[2] Duranton and Turner (2012) estimate a model explaining the joint evolution of highways and employment in American cities. Desmet and Rossi-Hansberg (2013) explain city size distribution from differences in congestion costs, amenities and productivity.

The structure of the paper is as follows. Section 2 discusses the model, Section 3 the estimation strategy, Section 4 deals with data, while Section 5 discusses the estimation results. Section 6 discusses the method for solving the model and the methodology of the welfare analysis. Section 7 discusses the policy experiment and Section 8 concludes.

2 The model

Our spatial general equilibrium model consists of H locations available both for living and working. The geography of these locations is described by a matrix of commuting times and costs for every pair of locations by four transport modes. Each location h has a fixed area of land A_h available for residential use. Commercial land use is not modelled. There are three levels of education s. The total supply N_s of workers for each education level is fixed. We ignore household formation. Each worker supplies a fixed amount of labour.

Workers have to take three decisions. First, they have to decide where to live and how much land to rent at that location. Second, they have to decide where to work. Finally, they have to decide how to commute between their home and job location. They take these decisions as to maximize their utility. Land rents R_h adjust to clear the market for residential land at each home location h.

2.1 Workers' utility

Worker i is characterized by a mixed utility function, partly of the direct and partly of the indirect form:

$$
\begin{aligned}
v_{ihjm} &= \mu_{Hs}^{-1}\left(\alpha_s' a_h + z_h + \ln K_h\right) + y_{sj} + \gamma_s' x_{hjm} \\
&\quad + \omega_s\left[\ln\left(W_{sj} - X_{shjm}^c\right) - \rho_s \ln R_h^*\right] + e_{ihjm}, \\
\ln R_h^* &\equiv \ln R_h + \psi R_h.
\end{aligned}
\tag{1}
$$

Most parameters are indexed by the education level s. The suffix i implies a value of s; hence, we omit s whenever there is an index i.

The first line is the direct part of the utility function. The vector a_h measures observed amenities at home location $h \in H$, z_h with $\mathrm{E}[z_h] = 0$ measures

unobserved amenities that are uncorrelated to a_h, and $\ln K_h$ measures the number of houses (we return to this below); y_{sj} is an education level specific fixed effect capturing unobserved characteristics of job location $j \in H$; x_{hjm} is a vector of physical characteristics of commuting from location h to j by mode $m \in M$.

The second line is the indirect part of the utility function, a translog cost function for consumption with land and other consumption as inputs. The price of other consumption is normalized to unity. W_{sj} is equal to the net wage rate for education level s at job location j and X^c_{hjm} is the cost of commuting between h and j by mode m. Since all workers supply a fixed amount of labour and since we ignore other sources of income, $W_{sj} - X^c_{hjm}$ is equal to disposable income net of commuting cost. Since there is a unified market for land at each location, a single land rent R_h clears the market at location h. Workers choose their home location h, their job location j, their transport mode m, their land use L_{shj} and their other consumption C_{shj} as to maximize their utility, subject to their budget constraint:

$$W_{sj} - X^c_{hjm} = R_h L_{shj} + C_{shj}. \tag{2}$$

Define $x^c_{shjm} \equiv X^c_{hjm}/W_{sj}$. Since transport costs are just a small fraction of income, we have

$$\ln\left(W_{sj} - X^c_{hjm}\right) \cong \ln W_{sj} - x^c_{shjm}.$$

By Shephard's lemma, the land share in consumption satisfies

$$\frac{R_h L_{shj}}{W_{sj} - X_{hjm}} = \frac{\partial \ln\left(W_{sj} - X^c_{hjm}\right)}{\partial \ln R_h} \cong \frac{\partial\left(\ln W_{sj} - x^c_{shjm}\right)}{\partial \ln R_h} = \rho_s \ln R_h^+ \tag{3}$$

$$\ln R_h^+ \equiv \frac{d \ln R_h^*}{d \ln R_h} = 1 + \psi R_h.$$

The elasticity of substitution η between land use and other consumption reads:[3]

$$\eta = 1 - \frac{\rho_s \psi R_h}{\rho_s \ln R_h^+ \left(1 - \rho_s \ln R_h^+\right)}. \tag{4}$$

When $\psi = 0$, the translog cost function reduces to the Cobb-Douglas case where $\eta = 1$ and where the land share in disposable income is constant, $R_h L_{shj} = \rho_s(W_{sj} - X^c_{hjm})$. For $\psi > 0$, the elasticity of substitution is smaller than unity and the land share is increasing in the land rent R_h.

Note that z_h is uncorrelated to a_h but it is likely to be correlated to other variables, in particular the land R_h: when the unobserved amenities of a location

---

[3] The log cost function reads (leaving out the suffix h of R_h)

$$w\left(\ln R, \ln P\right) = \rho_s \ln R + (1 - \rho_s) \ln P + \rho_s \psi \exp\left(\ln R - \ln P\right)$$

where P is the price of other consumption, which is normalized to unity (note: this function is homogeneous of degree one: 1% increase in all prices increases $w\left(\cdot\right)$ by 1%). The elasticity of substitution is equal to $1 + w_{12}/\left(w_1 w_2\right)$, where the suffices refer to the partial derivatives. Using $P = 1$ and simplification yields the expression in the text.

are favourable, $z_h > 0$, this will lead to a higher demand for land at that location and hence a higher equilibrium land rent R_h. Similarly, y_{sj} is likely to be correlated to W_{sj}: if $y_{sj} > 0$, on the one hand workers like to work at that location anyway and hence firms can afford to pay lower wages, but on the other hand many workers want to work there, which leads to agglomeration externalities in production (to be discussed below), which allow firms to pay even higher wages.

Finally, the term e_{ihjm} is a stochast, which is independent of any other variable in the utility function. It can be decomposed as:

$$e_{ihjm} = \widetilde{e}_{ih} + \widetilde{e}_{ihj} + \widetilde{e}_{ihjm},$$

where components $\widetilde{e}_\chi$ are uncorrelated for all χ, see e.g. Ben-Akiva and Lerman (1985). The composite components $\widetilde{e}_{ihjm}$, $(\widetilde{e}_{ihj} + \widetilde{e}_{ihjm})$, and $(\widetilde{e}_{ih} + \widetilde{e}_{ihj} + \widetilde{e}_{ihjm})$ all follow a Gumbel or type I extreme value distribution with location parameter zero and scale parameters μ_{Ms}^{-2}, μ_{Js}^{-2}, and μ_{Hs}^{-2} respectively, $\mu_{Ms} \geqq \mu_{Js} \geqq \mu_{Hs}$.[4] Since utility functions are insensitive to increasing transformations, we normalize μ_{Ms} to unity without loss of generality. In equilibrium, every worker occupies one house. We assume that individual i draws K_h independent realisations of $\widetilde{e}_{ih}$, one for every house. She chooses that house that yields the highest utility. The term $\ln K_h$ captures the expected utility gains from this selection process.[5] Note that the number of houses K_h is endogenous, since it depends on the land rent R_h: the higher is the land rent, the lower is the land use per worker, and hence the higher is number of houses that fit within the area A_h of land available for residential use.

2.2 Production

Types of labour are the only input for production. Due to perfect competition at the labour market, wages W_{sj} are equal to productivity of education level s at job location j. Due to agglomeration externalities this productivity depends on the number of people N_{sj} of education level s working at location j:

$$\ln W_{sj} = \pi_{0s} + \pi_s \ln N_{sj} + u_{sj}, \tag{5}$$

where u_{sj} is a random productivity effect. The specification allows the agglomeration elasticity π_s to differ between education levels. In the presence of agglomeration economies, the random effect u_{sj} is correlated to both the right and the left hand side variables. The correlation can be positive or negative. Suppose that location j is more productive than other regions, $u_{sj} > 0$, allowing firms to pay higher wages, thereby increasing labour supply and hence employment N_{sj}; then, $\ln N_{sj}$ is positively correlated to u_{sj}. In contrast, suppose that jobs at location j are less attractive than in other regions, so that for

---

[4] A Gumbel distribution with scale parameter μ has a mean γ/μ (γ is Euler's constant), and variance $\mu^{-2}\pi^2/6$.

[5] The max of K Gumbel distributed stochasts, see previous note, has a mean $(\ln K + \gamma)/\mu$ and variance $\mu^{-2}\pi^2/6$.

a given $\ln W_{sj}$, $\ln N_{sj}$ is smaller. Then, $\ln W_{sj}$ must be higher to attract the same amount of workers and $\ln N_{sj}$ is negatively correlated to u_{sj}.

3   Estimation

3.1   Land use

Data for land use and disposable income per level of education are not available. Hence, we calculate total land use per home location from equation (3)

$$\frac{A_h R_h}{\Sigma_{s \in S} \overline{W}_{sh}} = (1 + \psi R_h) \frac{\Sigma_{s \in S} \rho_s \overline{W}_{sh}}{\Sigma_{s \in S} \overline{W}_{sh}} + \zeta_h, \tag{6}$$

where $\overline{W}_{sh}$ is total wage income of education level s living at location h minus the average financial cost of commuting,[6] and where ζ_h is an error term capturing unexplained variation in the land use. Since there is a multiplicative restriction on the coefficients ρ_s and ψ, this equation cannot be estimated by OLS. However, estimation by NLLS is straightforward, yielding estimates for ρ_s and ψ. The estimate for ψ is used to construct the variables $\ln R_h^*$ and $\ln R_h^+$.

3.2   The logits for transport mode and job location

Worker i chooses the transport mode, and the job and home location yielding the highest utility. The formal derivation of these models from utility maximization is relegated to Appendix A. The logit model for the choice of transport mode reads:

$$\begin{aligned}
\Pr[m|shj] &= \exp\left(c_{shjm} - c_{shj}\right), \tag{7}\\
c_{shjm} &\equiv \gamma'_s x_{hjm} - \omega_s x^c_{shjm},\\
c_{shj} &\equiv \ln\left[\sum_{m \in M} \exp c_{shjm}\right].
\end{aligned}$$

This model can be estimated from individual data on x_{hjm}, x^c_{shjm} and the actual choices of transport mode, yielding estimates for the parameters γ_s and ω_s. The logsum term c_{shj} can be interpreted as a transport cost indicator for the connection hj for a worker of education level s.

The logit model for the job location choice reads:

$$\begin{aligned}
\Pr[j|sh] &= \exp\left(g_{shj} - g_{sh}\right), \tag{8}\\
g_{shj} &\equiv \mu_{Js}\left(c_{shj} + y^*_{sj}\right),\\
g_{sh} &\equiv \ln\left[\sum_{j \in H} \exp g_{shj}\right],
\end{aligned}$$

---

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

where $y_{sj}^* \equiv y_{sj} + \omega_s \ln W_{sj}$ is the modified location fixed effect, that is, including the effect of wages. This logit model can be estimated using the calculated values for c_{shj} from the estimation results of model (7). The estimation of this model yields estimates for μ_{Js} and $y_{sj} = y_{sj}^* - \omega_s \ln W_{sj}$ for each job location $j \in H$. In fact, the coefficient μ_{Js} is identified from the between home location variation in the term c_{shj}. Note that ω_s cannot be estimated from this model, since $\ln W_{sj}$ is correlated to y_{sj}. However, ω_s can be estimated from equation (7).

3.3 The home location logit

The estimation of the logit model for the home location is more involved. The model reads:

$$
\begin{aligned}
\Pr[h|s] &= \exp(z_h + q_{sh} + \ln K_h - q_s), & (9) \\
q_{sh} &\equiv \alpha_s' a_h + \mu_{Hs} g_{sh}^* - \rho_s^* \ln R_h^*, \\
q_s &\equiv \ln\left[\sum_{h \in H} \exp(z_h + q_{sh} + \ln K_h)\right],
\end{aligned}
$$

where $\rho_s^* \equiv \mu_{Hs} \omega_s \rho_s$ and $g_{sh}^* \equiv \mu_{Js}^{-1} g_{sh}$. The estimation of (9) is non-trivial due to two complications: (i) the high computational burden due to the large number of home locations H (several thousands in our data), each location with its own fixed effect z_h; (ii) land rents $\ln R_h^*$ are correlated with the unobserved location characteristics z_h. Indeed, we can expect that home locations with positive values of z_h will, ceteris paribus, command higher prices. The estimation can be simplified considerably by an estimation strategy similar to the one suggested by Berry et al. (1995) and applied later by Bayer et al. (2004). We extend that method by using a two step approach, similar to Heckman's method for estimating Tobit models. Where that method lost its appeal for labour supply models, it turns out to work well in the current context. Furthermore, we suggest a new instrument to solve for endogeneity of prices in a demand estimation. First, we transpose the conditionality of probabilities in the logit model using Bayes' rule. This transposed logit model can be estimated far more easily. The second step estimates the land market clearing condition, where we use a correction term to account for the selectivity in the error term. Both steps are explained below.

First, instead of estimating $Pr[h|s]$, the probability of home location h conditional on the worker's education level s, we estimate $Pr[s|h]$, the probability of worker's education level s conditional on the home location h. So, instead of analysing what home location is chosen by a particular worker-type, we analyze what worker-type chooses this particular home location. By Bayes' rule and equation (9) we have:

$$
\Pr[s|h] = \frac{\Pr[h|s]\Pr[s]}{\Sigma_{k \in S}\Pr[h|s=k]\Pr[s=k]} = \frac{\exp(q_{sh} - q_s + \ln N_s)}{\sum_{k \in S}\exp(q_{kh} - q_k + \ln N_k)}. \quad (10)
$$

The first equality applies Bayes' rule. The second equality substitutes the probability $\Pr[s]$ for the actual frequencies of that level of education in the data

$\Pr[s] = N_s/N$, where $N \equiv \Sigma_{s \in S} N_s$. Since $\rho_s^o \equiv \ln N_s - q_s$ is an education specific constant, this is again a logit model. However, this transposition of the original logit model (9) reduces the number of alternatives in the model from the number of home locations H to the number of education levels. Furthermore, the fixed location effects z_h drop out. Instead, we have to estimate only three additional intercepts ρ_s^o, one for each level of education. This reduces the computational burden from several days to several minutes. This transpose approach can be extended easily to a larger number of socioeconomic groups.

Estimation of this logit model yields estimates for α_s, ρ_s^*, and ρ_s^o up to a normalization for a reference alternative, as is standard in logit models. We can only estimate the difference of these parameters relative to the unknown value for the reference alternative. Such a normalization is not required for the parameter μ_{Hs}, since the corresponding variable g_{sh}^* varies between alternatives s. Hence, we have estimates for μ_{Hs} for every education level.

The second step uses the land market clearing condition. Bayes' rule implies $\Pr[h|s] = \Pr[s|h] \Pr[h] / \Pr[s]$. Taking logs and rearranging terms and using equation (9) to substitute for $\ln \Pr[h|s]$ yields:

$$
\begin{aligned}
\ln \Pr[s|h] &= -\ln \Pr[h] + \ln \Pr[s] + q_{sh} + z_h - q_s + \ln K_h \\
&= \ln N_s + q_{sh} + z_h - q_s \\
&= \rho_s^o + \alpha_s' a_h - \rho_s^* \ln R_h^* + \mu_{Hs} g_{sh}^* + z_h.
\end{aligned}
\tag{11}
$$

The second line substitutes the unconditional probabilities $\Pr[h]$ and $\Pr[s]$ by their actual frequencies in the data, K_h/N and N_s/N. Like the log probability $\ln \Pr[s|h]$, the transformed land rent $\ln R_h^*$ is correlated with the unobserved location characteristics z_h. As we know $\mu_{Hs} g_{sh}^*$ from the first stage, we can use this variable as an instrument for $\ln R_h^*$. This boils down to a 2SLS estimation of the following equation:[7]

$$
\ln \Pr[s|h] - \mu_{Hs} g_{sh}^* = \rho_s^o + \alpha_s' a_h - \rho_s^* \ln R_h^* + z_h.
\tag{12}
$$

One can follow two approaches for the calculation of $\ln \Pr[s|h]$ for the estimation of this equation. First, one can use the estimated coefficients from the logit model (10). The disadvantage of this approach is that calculation of the standard errors of the parameters becomes messy since the estimation errors of the first step enter the second step. Instead, one can substitute the log probability $\ln \Pr[s|h]$ for the log of the actual frequencies observed in the data, $\ln N_{sh} - \ln \sum_s N_{sh}$. This approach works if the number of observations for which s and h are known is sufficiently large. If not, the number of cells for which N_{sh} is (close to) zero becomes too large. The sample for which all of s, h, j, and m are known is relatively small, but the sample for which just s and h are known is large. Hence, we can use the second approach.

Estimation of equation (12) yields an estimate of $\rho_s^* \equiv \mu_{Hs} \omega_s \rho_s$, from which we can backout ρ_s. This value of ρ_s should be equal to the estimated value from the model for land use, equation (6). There are therefore two methods for

---

[7] We thank Guido Imbens for this idea.

establishing the value of ρ_s. The first method asks how much weight workers attach to the local land rent when choosing the optimal home location: the higher the land rent, the less attractive is the location. The second method analyzes how much land a worker uses: the higher the land rent, the less land he will use. By Sheppard's lemma, the weight of the land rent when choosing the home location should by equal to the share of land in total consumption. Even shorter: people should worry as much about high land rents when choosing their home location as when deciding on their consumption of land. Since there is nothing in the estimation procedure that guarantees this equality to be met, this offers a test to the model.

Equation (12) can be written as a land rent equation:

$$\ln R_h^* + \rho_s^{*-1} \ln \Pr[s|h] = \rho_s^{*-1} \left(\rho_s^o + \alpha_s' a_h + \mu_{Hs} g_{sh}^* + z_h \right). \tag{13}$$

The term $\rho_s^{*-1} \ln \Pr[s|h]$ on the left hand side deserves a further analysis. This term accounts for the degree of selectivity in stochastic term $\widetilde{e}_{ih} + \widetilde{e}_{ih}^*$ of those workers who choose to live in h. Suppose that for a particular location h, $\Pr[s|h] \to 1$: all workers living in this location have education level s. Hence, there is no selectivity and the term drops out and equation (13) simplifies to:

$$\ln R_h^* = \rho_s^{*-1} \left(\rho_s^o + \alpha_s' a_h + \mu_{Hs} g_{sh}^* + z_h \right).$$

Hence, the left hand side variable $\ln R_h^* + \rho_s^{*-1} \ln \Pr[s|h]$ can be interpreted as the log land rent that would have applied if all workers had education level s.

Now, suppose to the contrary that $\Pr[r|h] \to 1$ and hence $\Pr[s|h] \to 0$, implying $q_{rh} \to \infty$, see equation (10). Then, equation (13) converges to:[8]

$$\ln R_h^* = \rho_r^{*-1} \left(\rho_r^o + \alpha_r' a_h + \mu_{Hr} g_{rh}^* + z_h \right).$$

Which set of coefficients determines the marginal contribution of amenities a_h to $\ln R_h^*$, whether it is α_s/ρ_s^* or α_r/ρ_r^*, depends therefore on the composition of the group of workers who prefer to live in location h: the higher the share of education level s, the more the marginal contributions converge to coefficients α_s/ρ_s^*. The same applies to the job availability index g_{sh}^*.

Let v_s be defined as $v_s \equiv \mathrm{E}_i \left[\max_{h \in H, k \in h} (v_{ik}) \right]$, the expected utility for a worker with education level s. Similar to equation (19) in Appendix A, v_s satisfies:

$$v_s = \mu_{Hs}^{-1} q_s + \mathrm{E} \left[\widetilde{e}_i^* \right]. \tag{14}$$

Since v_s is an increasing transformation of q_s, see equation (9), and since utility is invariant to an increasing transformation, q_s can be used for the evaluation of welfare effects, as we do in Section 5.

---

[8] Equation (10) implies $\lim_{q_{rh} \to \infty} \exp(q_{rh} - q_{sh}) \Pr[s|h] = 1$. Taking logs yields $\lim_{q_{rh} \to \infty} (q_{rh} - q_{sh} + \ln \Pr[s|h]) = 0$. Hence, $\ln \Pr[s|h]$ converges to $q_{sh} - q_{rh}$. Substitution in equation (13) yields the result.

4 Data and descriptives

The data on commuting are based on 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 1 reports data selection. From the respondents for which home and job ZIP codes are available,[9] we select those aged between 18 and 65, not in full time education, working at least 12 hours per week. We drop respondents for whom education data are missing, with a home or work address outside the Netherlands or on one of the islands in the North Sea, 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 transportation 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 transportation 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 B.

Table 1 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

Data on living amenities originate 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 residential land use are derived from the digital map "Land use" by Statistics Netherlands, 2006.

Data on land prices for the years 2004–2006 have been calculated from microdata on 2.5 million 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 Groot et al. (2014). Land prices are converted into land rents per working day using capital cost of 4.2% per year and 228 working days per year, see Dijk and Romijn (2010). Data on wages for each level of education and ZIP code have been calculated from a Statistics Netherlands restricted-access micro dataset on individual earnings, which

---

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

combines data on earnings, level of education, and job location from various administrative databases. Wage data refer to 2010. Details of the calculation are discussed in Appendix B. Gross wages are converted to net wages using the gross-to-net wage calculator of CPB Netherlands Bureau for Economic Policy Analysis. The regional component of wages is estimated as ZIP code fixed effect from a Mincerian regression of log hourly wages on standard personal characteristics like gender, years of education, experience, and ethnicity. The regression includes dummies for 2-digit industries. Controlling for industry is likely to partly eliminate compensating differentials for job disamenities, like working in shifts, noise, hard work and the like. Summary statistics for the amenity variables, for land prices, and for ZIP code fixed effects in wages are reported in Appendix B.

5   Estimation results

5.1   Land use

Table 2 reports estimates of equation (6) under the condition $\rho_{high} = \rho_{middle} = \rho_{low}$.[10] The coefficient ψ is positive and highly statistically significant, showing that the substitution elasticity of land use and other consumption is less than one. Equation (4) implies an elasticity of 0.71 for the mean value of $\ln R_h$. This is reasonably in line with Albouy and Ehrlich (2012) who report the elasticity of substitution between land and other inputs in housing production to be about one-half using US data.

Table 2 Residential land use

parameter	coefficient	t-value
ψ	14.184	(20.9)
ρ	0.058	(38.9)
R^2		0.71
# observations		2812

Figure 3 shows that the model fits the general pattern in the data well, although it overestimates the land share for low levels of $\ln R_h$ substantially. This figure clarifies the reason for choosing a specification that combines a logarithmic term with a level term, $\ln R_h^* = \ln R_h + \psi R_h$, see equation (3): the second order polynomial in $\ln R_h$ usually applied in a translog cost function is unable to capture the curvature in the data. 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.

---

[10]Estimating separate values of ρ_s for each education level turns out to be problematic: $\rho_{middle} < 0$, which is theoretically infeasible. This is probably due to: i) the fact that we do not have data on land use by level of education, and: ii) the high negative correlation between the measured wage income across home locations for high and middle educated workers. Hence, we revert to estimating an average value for ρ_s for all education levels.

Figure 3. Actual vs. predicted income share of land
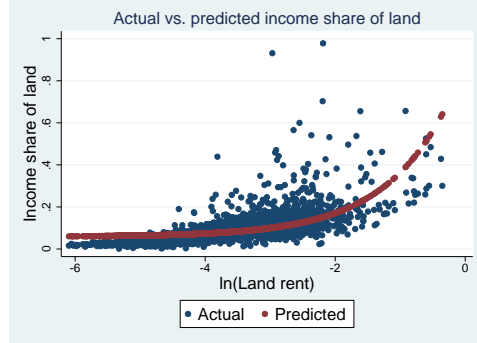
5.2 Modal split

Table 3 reports the estimation results for the modal split logit (7).[11] Most variables are highly statistically significant.[12]

The parameter ω_s on transport cost measures the marginal utility of money. Parking cost, defined as the local land rent per m^2 divided by disposable income, is a useful credibility check for the value of ω_s, since the ratio of the coefficient on parking cost to ω_s must be equal to the land use for parking. This calculation yields a land use of 34 m^2 at the home location and 21 m^2 at the job location, which are reasonable numbers. The car is a land intensive mode of transport. It is therefore a less popular mode of transport for consumers living in areas where land is expensive. The difference in land use at the home and the 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.

Higher educated have a strong preference for commuting by train or bike, keeping other factors constant. Since the rail infrastructure is better in cities, this contributes to an explanation why higher educated predominantly live and/or work in cities. Out-of-vehicle time is valued much more negatively than in-vehicle time for public transport. Distance to train station is valued strongly negatively also. 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.

---

[11]In the modal split logit we delete as irrelevant alternatives: (i) train if the total distance to transfer (home+job) is larger than 40 km; (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 km, all for a single trip. Furthermore, intra ZIP code commutes are excluded from the modal split estimation as we have no commuting data on these trips.

[12]We have tried a version where all coefficients varied by level of education. However, allowing the intercepts of each mode of transport to vary by level of education captures most of the variation.

Table 3 Estimation results for the modal split model

General variables	coef	t-val				
cost in % net wage	$-0.122^{a)}$	(8.6)				
time (minutes/10)	-0.261	(20.0)				
parking.costs at home[b)]	-4.184	(15.8)				
parking.costs at job[b)]	-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) This implies $\omega_s = 12.231$; b) in % wage/m^2

The implications of these estimation results are most easily judged from the implied values of time, see Table 4. The compensatory variation required to make people indifferent can be calculated from the cost variable. 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 middle, and 11 euro for low educated workers). Time spent in train is less costly, while time spent waiting for a bus is more costly.[13]

Table 4 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

[13] 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.

5.3   Job location

Table 5 reports the estimation results for the job location logit (8).[14] Some 10% of the observations work in the same ZIP code as where they live. We have no data on these commutes. Hence, the reported c_{shh} is equal to zero. 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 5. Estimation results for the job location logit

education	low		middle		high	
	coef	t-val	coef	t-val	coef	t-val
μ_J unrestricted						
job accessibility c_{shj}	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)
$\mu_J = 1$						
dummy home=job ZIP	0.567	(25.0)	0.233	(11.1)	-0.398	(13.2)
# observations	17151		26189		21999	

The estimation results show that $c_{shh} = 0$ is a good proxy for intra ZIP code commutes.[15] Theoretically, one would expect the coefficient μ_J by the transport cost indicator c_{shj} to be between zero and one. In that case, one minus the square of the estimated coefficient can be interpreted as the covariance between the error terms for the various transport modes within a ZIP code. For middle and low educated workers the estimates of the coefficient of generalized transport cost are larger than one. This result might reflect a non-linearity in the utility function, which does not show up in modal split model (since conditional on h and j, workers have to choose a mode of transport), but does show up in the job location choice (people dislike far away job locations more than proportional). 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 the coefficient on c_{shj} to 1, thus implicitly assuming a multinomial logit structure of the modal choice and job location choice.

---

[14]Theoretically, standard errors in the subsequent logit model are underestimated because part of the explanatory variables is constructed using estimates from previous logit model(s). In practice, this does make much of a difference. Hence, we ignore this complication.

[15]The c_{shj} varies from –34 to 0; the estimated dummy coefficients are thus reasonably close to zero.

Table 6 Covariance/correlation matrix variables in job location logit[a)]

educ. level	μ_{Js} unrestricted					$\mu_{Js}=1$				
	c_{shj}	y^*_{sj}	y_{sj}	$\widetilde{e}$	share varian.	c_{shj}	y^*_{sj}	y_{sj}	$\widetilde{e}$	share varian.
low										
c_{shj}	15.45	0.83	0.60	0	0.795	15.45	0.55	0.31	0	0.793
y^*_{sj}	0.19	1.29	1.08	0	0.066	0.12	1.28	1.13	0	0.066
y_{sj}	0.14	0.91	1.10	0		0.07	0.91	1.20	0	
$\widetilde{e}$	0	0	0	1.04	0.053	0	0	0	1.64	0.084
middle										
c_{shj}	13.04	0.76	0.28	0	0.766	13.04	0.58	0.10	0	0.755
y^*_{sj}	0.18	1.32	0.95	0	0.077	0.13	1.43	1.13	0	0.083
y_{sj}	0.08	0.82	1.02	0		0.02	0.84	1.26	0	
$\widetilde{e}$	0	0	0	1.17	0.068	0	0	0	1.64	0.095
high										
c_{shj}	9.74	0.41	0.18	0	0.722	9.74	0.41	0.18	0	0.724
y^*_{sj}	0.12	1.25	1.05	0	0.093	0.12	1.24	1.04	0	0.092
y_{sj}	0.05	0.88	1.14	0		0.06	0.88	1.12	0	
$\widetilde{e}$	0	0	0	1.67	0.124	0	0	0	1.64	0.122

a) numbers below the main diagonal are correlations.

Table 6 presents the covariance/correlation matrix for the within home-location variance of the explanatory variables accessibility indicator c_{shj}, the location fixed effect y^*_{sj}, and the fixed effect corrected for wage differentials $y_{sj} = y^*_{sj} - \omega_s \ln W_{sj}$. The residual variance is calculated using the formula for the variance of a Type-I extreme value distribution, $\pi^2/6 = 1.64$. The generalized transport cost c_{shj} explains 70 to 80 % of the job location choice. The lower the education level, the larger the share of transport cost. Job location fixed effects play a minor role. The positive correlation between transport accessibility of a job location c_{shj} and its attractivity y^*_{sj} supports the notion that transport infrastructure is endogenous. More attractive job locations are, on average, better accessible. Apparently, infrastructure investment is steered towards improving the more valuable links.[16] The variance of y_{sj} is smaller than the variance of y^*_{sj}. Hence, wage differentials offer a partial explanation for the attractiveness of job locations, though other factors also play their role.

5.4 Land rents and home location choice

Table 7 reports the estimation results for the land rent model, equation (13). The estimation results of the first stage, home location logit model, equation (10) are reported in Appendix C. The parameters μ_{Hs}, which correspond to the parameter on the job availability measure g^*_{sh} are all well below one, which

---

[16] This emphasizes the importance of explicitly accounting for unobserved heterogeneity in attractivity of home and job locations, when estimating location choices of people (terms z_h and y_{sj} in utility function (1)). Not including these terms would lead to biased estimates of effect of commuting cost on location choice.

confirms the nesting structure. Higher educated are less sensitive to land rents. Since land rents are higher in the city, this adds to an explanation why higher 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 univerisities and the availability of restaurants.

Although the fit of the land rent model is quite good, the explanatory power of the home location logit is rather poor, see Figure 5, Panel A. As an experiment, we added the share of social housing to the model. This variable is highly statistically significant and improves in particular the predictive power of the home location logit, see Figure 5, panel B. Social housing is only available for low income families, so the supply of social housing is an important determinant of the share of low educated at a location. This suggests that supply side constraints imposed by public policy explain a large part of the regional segregation of workers by education. At first sight, this is remarkable, as Dutch policy makers are alleged to have a strong preference for socially mixed neighbourhoods. However, housing policy is endogenous: social housing is likely to be realized at locations where it is in high demand. Hence, the prevalence of social housing might just be a proxy for unobserved factors. For example, Schelling's (1969) finding that even mild individual preferences for living among people with similar education can produce full spatial segregation might explain why the demand for social housing is concentrated in particular neighbourhoods. This is an interesting topic for future research. For the moment, we just observe the empirical fact, but since supply side constraints do not fit well into the theoretical structure of our model, we ignore this correlation in our simulations.

Table 7. Estimation results for home location choice[a)]

education level	low		middle		high	
variable	coef	t-val	coef	t-val	coef	t-val
$\ln R_h^*$ (ρ_s^*)	0.645	(50.3)	0.464	(57.9)	0.311	(28.0)
job.availability[a)] g_{sh}^*	0.637	(3.6)	0.449	(2.3)	0.434	(2.4)
# monum.1km/1000	0.491	(5.8)	0.449	(8.5)	0.485	(6.6)
# monum.1-5km/1000	0.095	(4.1)	0.107	(7.5)	0.157	(7.9)
share nature within 5km	-0.079	(1.0)	0.052	(1.1)	0.484	(7.2)
dum. university in 10km	-0.074	(3.0)	-0.023	(1.5)	0.123	(5.8)
# restaurants 1km/100	0.338	(3.4)	0.581	(9.3)	0.625	(7.2)
# restaurants 1-5km/100	0.130	(9.0)	0.074	(8.2)	0.029	(2.3)
intercept	-3.238	(60.9)	-2.602	(78.3)	-2.300	(49.9)
implied ρ_s	0.083		0.085		0.059	
# observations		2758		2758		2758
addional variable:						
share social housing	1.298	(23.1)	-0.076	(1.8)	-1.109	(21.3)

[a)] Coefficients on job availability result from estimating equation (10); we use a clustered error correction, defining clusters to be home ZIP codes. Other coefficients come from estimating equation (12).

Figure 5. Actual vs. predicted shares of residents by education levels

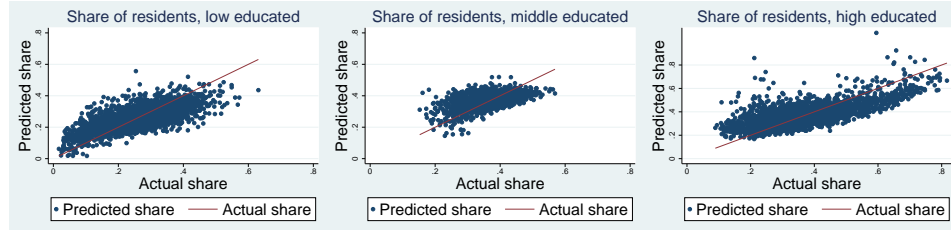*Reference model*
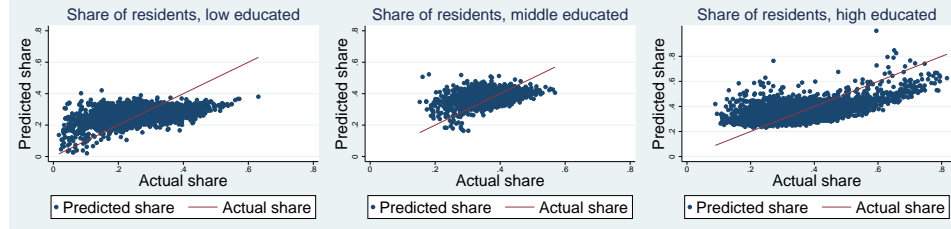
Model including social housing

Table 8 reports the contributions of job availability g_{sh}^* on the one hand and observed $\alpha_s' a_h$ and unobserved z_h amenities on the other hand to the total variance of $\ln R_h^* + \ln \Pr[s|h]/\rho_s^*$. The total variance is the highest for higher educated. The job availability explains 28 to 38% of the variance, where lower educated are on the high side. For the higher education levels amenities are more important, showing the importance of consumption externalities for location choice. Figure 6 provides a graphical documentation of the importance of job availability and amenities, for high and low educated. All panels depict the value of $\ln R_h^* + \ln \Pr[s|h]/\rho_s^*$ for each ZIP code, using intervals of 0.40. The upper panels show its level, which can be interpreted as the land rent different locations would command if their population consisted exclusively of lower and higher educated respectively. The middle panels show the contribution of job availability and the lower panels show the contribution of observed amenities. The relative contribution of job availability and amenities differs widely between levels of education. For lower educated, the contribution of amenities is rather flat, with the exception of the center of Amsterdam. For higher educated, the contribution of amenities is much more spread out. However, for both education levels, Amsterdam stands out relative to the rest of the country on the amenity side, while the contribution of job availability is spread out much more evenly among the central Western part of the country. In general, amenities contribute enormously to the popularity of cities as an area to live, in particular Amsterdam. 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 seggrageted location pattern of high and low educated documented in Figure 2.

Figure 6. Value of job availability g_h^* and observed amenities a_h, expressed in terms of $\ln R_h^* + \ln \Pr[s|h]/\rho_s^*$. Low educated left, high educated right.[17]

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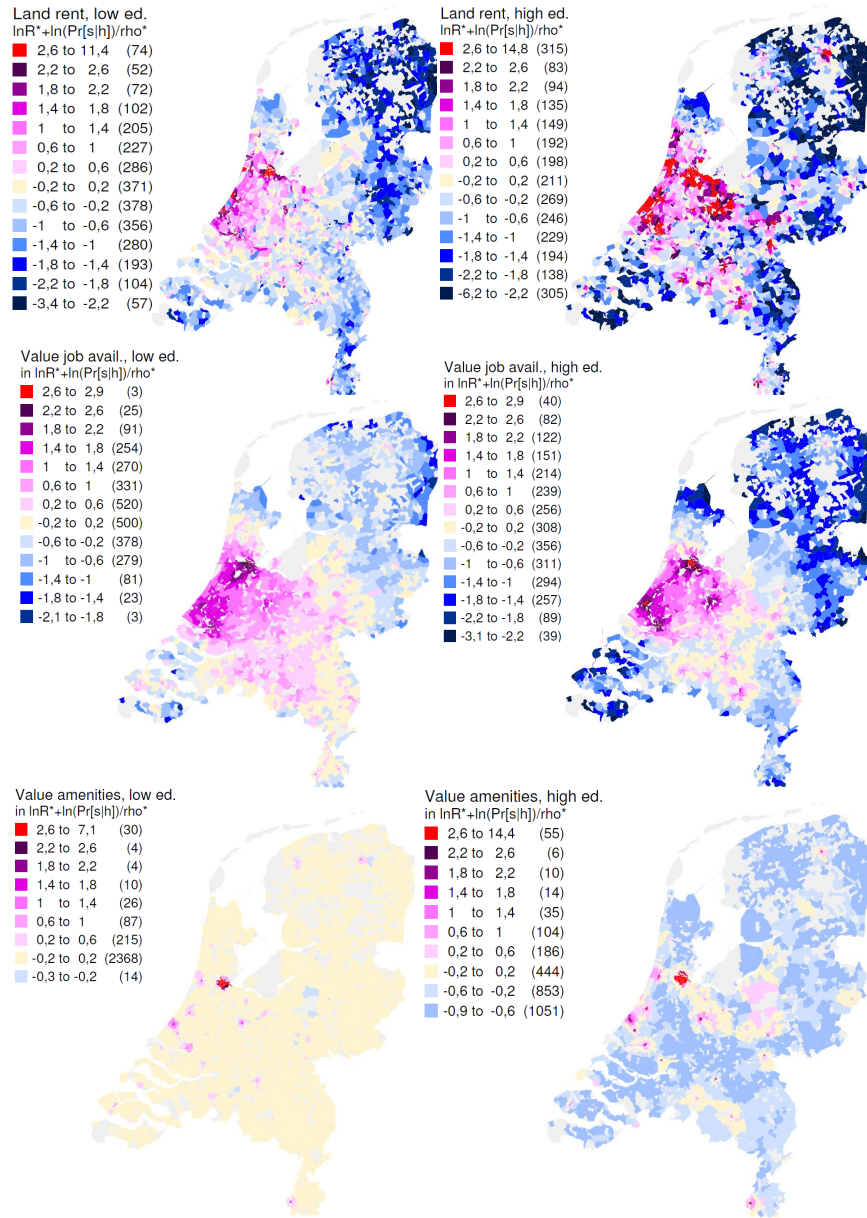
[17]In all panels, 0 corresponds to a location in Enschede with an average land price of 109 euro/m^2.

Table 8. Explained variance of $\ln R_h^* + \ln \Pr[s|h]/\rho_s^*$

level	low		middle		high	
	variance	share[18]	variance	share	variance	share
total	1.816	1.00	1.824	1.00	5.277	1.00
g_{sh}^*/ρ_s^*	0.696	0.38	0.622	0.34	1.472	0.28
$\alpha_s' a_h/\rho_s^*$	0.255	0.14	0.385	0.21	0.949	0.18
z_h/ρ_s^*	0.487	0.27	0.365	0.20	1.574	0.30

5.5 Effect of agglomeration on wages

Table 9 provides estimation results for the agglomeration elasticity, equation (5). We report both OLS and IV estimates, the latter to account for the potential endogeneity of employment. A crucial issue is the range of the effect. In their study of the benefits of spatial proximity to other firms for the advertising industry in Manhattan, Arzaghi and Henderson (2008) provide evidence that the range of these benefits is some 500 meter. This finding is supported by the study on the effect of the separation and subsequent reunification of Berlin by Ahlfeldt et al. (2012): they report a range of 1 km or less. Hence, we use employment at the lowest available aggregation level (1 km around the centre of a ZIP code) as an explanatory variable. We add the log employment for all ZIP codes within 5 km of the own ZIP code (including the own ZIP code) as an additional variable to test the spatial distance decay of agglomeration externalities in our case. Since ZIP codes differ widely in their surface and since employment density might be what matters most, we correct for the area of commercial land use in a ZIP code: the larger the area, the lower the density and hence the lower the agglomeration elasticity. We expect the coefficient for this variable to be halfway between zero and the coefficient on log employment, since the short range of agglomeration externalities imply that they are partly driven by density and partly by absolute numbers. The estimation results confirm this prior.

The OLS results suggest that agglomeration effects of about 3% at the short range are relevant for the higher education levels, but not at all for the lowest level, while the effects at the longer range are relevant for the lower levels, but not at all for the highest level. The longer range effect for lower educated can be interpreted as the spillover effect of the higher cost of living in urban areas to the wages of lower educated who produce the nontradables consumed by the higher educated, compare Autor et al. (2006). This interpretation is consistent with the concentration of higher educated in urban areas as documented in Figure 2 and with Gennaioli et al. (2013), who report regional clustering of higher educated within most countries.[19]

---

[18] The shares of explained variance are exclusive the contribution of covariances; for this reason the shares do not sum up to 1.

[19] This can explain the high agglomeration elasticity of 10% reported by Ahlfeldt et al.: if agglomeration yields a concentration of high skilled workers, Berlin's reunification might have led to an increase in the average skill level of the population. However, Combes et al. (2012) and De la Roca and Puga (2013) find that productivity gains from job relocation to denser areas are mainly due to agglomeration externalities and less due to selection.

	education	low		middle		high	
model	variable	coef	t-val	coef	t-val	coef	t-val
OLS 1	$\ln N_{s1km}$	0.002	(1.3)	0.033	(18.2)	0.026	(19.5)
(2034)	$\ln A_{1km}$	-0.001	(0.9)	-0.015	(10.9)	-0.008	(6.4)
	R^2	0.001		0.14		0.16	
OLS 2	$\ln N_{s1km}$	-0.006	(2.8)	0.019	(8.8)	0.025	(14.3)
(2034)	$\ln A_{1km}$	0.003	(1.7)	-0.010	(6.7)	-0.009	(6.4)
	$\ln N_{s5km}$	0.019	(5.5)	0.021	(7.0)	-0.003	(1.4)
	$\ln A_{5km}$	-0.010	(2.7)	-0.005	(1.5)	0.008	(3.0)
	R^2	0.03		0.20		0.16	
IV 1[a]	$\ln N_{1km}$	-0.007	(0.9)	-0.003	(0.5)	-0.005	(0.9)
(2034)	$\ln A_{1km}$	0.005	(1.6)	0.002	(0.6)	0.003	(1.2)
	$\ln N_{5km}$	0.053	(6.6)	0.061	(9.1)	0.020	(3.6)
	$\ln A_{5km}$	-0.044	(6.0)	-0.039	(6.4)	-0.005	(0.9)
	$Sargan(p-val)$	7	(0.17)	17	(0.00)	21	(0.00)
	$F_{\ln N_{1km}}(p-val)$	46	(0.00)	64	(0.00)	74	(0.00)
	$F_{\ln N_{5km}}(p-val)$	100	(0.00)	131	(0.00)	145	(0.00)
IV 2[b]	$\ln N_{5km}$	0.047	(7.2)	0.059	(11.6)	0.015	(4.4)
(2034)	$\ln A_{5km}$	-0.039	(-6.0)	-0.037	(-7.0)	-0.002	(-0.4)
	$Sargan(p-val)$	2	(0.31)	8	(0.02)	13	(0.00)
	$F_{\ln N_{5km}}(p-val)$	208	(0.00)	270	(0.00)	299	(0.00)

[a] Instruments: number of monuments, number of houses in 1930, and existence of a highway in 1930; all within 1 km respectively 5 km.

[b] Same instruments as a), but only for the 5 km range.

We try to address the endogeneity of employment by instrumenting it. While agglomeration is likely to be persistent, the factors driving current productivity probably differ from those in the past. Hence, historical data are reasonable instruments. We use the number of monuments within 1 km and 5 km from the centre of a ZIP code, the number of houses in 1930 within 1 km and 5 km, and a dummy variable taking the value 1 if the nearest highway in 1950 is no more than 1 km respectively 5 km from the centre of a ZIP code.

We run two IV regressions. With the exception of the model for the low educated, none of our regressions passes the Sargan test. The short range agglomeration elasticities are all insignificant. We suspect that this is merely an artefact of the lack of discriminatory power of our instruments at this range. Our instruments can explain why there is more employment in ZIP codes in Amsterdam than in the low density areas in the North-East of the country; they cannot explain why this employment is concentrated in a particular ZIP code within the Amsterdam region. Hence, we exclude the short range effect in the second IV regression. In that case, we find agglomeration elasticities that are within the range reported elsewhere in the literature and which even

pass the Sargan test in the case of lower educated. However, based on the evidence from other studies, we tend to believe that agglomeration externalities might predominantly operate on the short range. Despite the endogeneity of employment, the OLS results might be the best reflection of the real world.

6 Comparative statics and welfare analysis

6.1 Equilibrium and comparative statics

An equilibrium to this economy is a set of land rents R_h, wages W_{sj}, land use L_{shj}, and number of houses K_h and jobs N_{sj} that satisfy equation (5) for wages, equation (3) for land use, and constraints on the number of houses and jobs, and land availability

$$N_{sj} = \sum_{h \in H} \Pr[h|s] \Pr[j|sh] N_s, \tag{15}$$

$$K_h = \sum_{s \in S} \Pr[h|s] N_s,$$

$$A_h = \sum_{s \in S} \sum_{j \in H} \Pr[h|s] \Pr[j|sh] L_{shj} N_s.$$

taking as given the supply of land A_h, labour supply N_s, the transport characteristics x_{hjm} and x^c_{shjm}, observed and unobserved amenities $\alpha'_s a_h + z_h$, and job location fixed effects y_{sj}. The calculation of an equilibrium is challenging as it involves solving a system of $2 \times H$ equations. The solution procedure is discussed in Appendix D.

Since this economy features agglomeration economies, there is no reason why an equilibrium would be unique. Our focus is therefore on a comparative static analysis of small changes in the transport infrastructure (shifts in c_{shjm}). These comparative statics take the current equilibrium as starting point and calculate how this equilibrium shifts by a change in c_{shjm}. Even this shift might be non-unique, in particular for larger changes in c_{shjm}. An analysis of the multiplicity of equilibriums falls outside the scope of this paper.

6.2 Welfare analysis

Our model consists of four types of agents, three types of workers differing by their level of education and a class of absentee landlords. It is most convenient to think of workers renting the land from these landlords. Landlords might be further subdivided in local subgroups, as we will do in our empirical application. The effect on the wealth of landlords Q_l is equal to the sum of the effect on land rent 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 wealth of consumers of education level s, Q_s is derived from their utility q_s, see equation (14). We are looking for the equivalent variation EV_s that makes consumers of that level of education equally happy before and after the change in c_{shjm} :

$$q_s^n [R^n, W_s^n] = q_s^o(R^o, W_s^o + EV_s),$$

where the dependence of q_s on R and W_s is explicitly addressed. The superfixes n and o for the utility function q_s indicate what value of c_{shjm} is applied, before of after the investment. A Taylor expansion of this equation implies

$$\frac{Q_s}{N_s} = EV_s \cong \frac{q_s^n (R^n, W_s^n) - q_s^o(R^o, W_s^o)}{\sum_{j\in H} \partial q_s^o(R^o, W_s^o)/\partial W_{sj}^o} \qquad (16)$$

This general equilibrium effect can be decomposed in a number of components, which are listed in Table 10. The final line is again the integral effect described in equation (16). The derivation of these expressions is presented in Appendix D.

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

users of mode m	$\omega_s^{-1} \sum_{h\in H} \sum_{j\in H} \Pr[hjm|s] \left(c_{shjm}^n - c_{shjm}^o \right) W_{sj}^o$
idem + modal shift	$\omega_s^{-1} \sum_{h\in H} \sum_{j\in H} \Pr[hj|s] \left(c_{shj}^n - c_{shj}^o \right) W_{sj}^o$
idem + job relocation	$\left[\omega_s \mu_{Js} \sum_{j\in H} \frac{\Pr[j|sh]}{W_{sj}^o} \right]^{-1} \sum_{h\in H} \Pr[h|s] [g_{sh}^n (W_s^o) - g_{sh}^o (W_s^o)]$
idem + effect shift W_s	$\left[\omega_s \mu_{Js} \sum_{j\in H} \frac{\Pr[j|sh]}{W_{sj}^o} \right]^{-1} \sum_{h\in H} \Pr[h|s] [g_{sh}^n (W_s^n) - g_{sh}^o (W_s^o)]$
idem + home relocation (= total)	equation (16)

[a] $\Pr[hj|s] = \Pr[h|s]\Pr[j|sh], \Pr[hjm|s] = \Pr[hj|s]\Pr[m|shj]$,
All probabilities are evaluated in the old equilibrium R^o, W_s^o.

7 Policy experiment

The model and the expression in Table 10 are applied to a policy experiment, using the parameter values reported in Table 11. 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. Since many people commute from the North to jobs in Amsterdam and the neighbouring municipality of Haarlemmermeer (the location of Schiphol airport), this connection is crucial 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, and compare it to the present equilibrium. Figure 7 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.
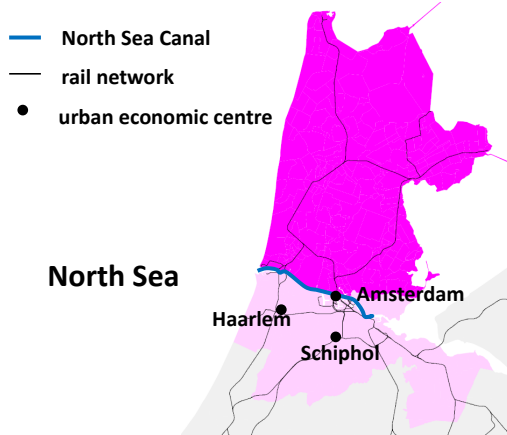
Figure 7 North Sea canal area

Table 11. Parameter values used in the experiment

parameter	all	low	middle	high
ω	12.24			
μ_J	1.00			
ψ	14.18			
μ_{Hs}		0.64	0.45	0.43
ρ_s [20]		0.078	0.080	0.055
# monuments within 1km/1000		0.49	0.45	0.49
# monuments 1 to 5km/1000		0.095	0.107	0.157
share nature within 5km		-0.079	0.052	0.484
dummy university 10km		-0.074	-0.023	0.123
# restaurants within 1km/100		0.338	0.581	0.625
# restaurants 1 to 5km/100		0.130	0.074	0.029
ρ_s^0		-3.24	-2.60	-2.30
agglomeration elasticity		0.00	0.03	0.03

Table 12 describes the relocation of economic activity between the regions north and south of the canal due to the availability the tunnels. The better connection of the less productive region north of the canal to 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 5%. Some 32.000 workers commute by train from the North to the South; 80% are additional commuters. The relief the rail tunnels offer for the congestion in the auto tunnels (these tunnels figured in the top of the Dutch traffic jam hit list for years) is therefore limited. Figure 8 documents the wage differentials between the North and the

---

[20]The values of ρ_s in the experiment are weighted averages of the ρ_s from the land use regression (3) and the ρ_s from the home logit (9).

South and 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 Figure 9. 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 5% higher due to the availability of the tunnels.

Figure 8. Jobs flight to highly productive job locations south of the canal
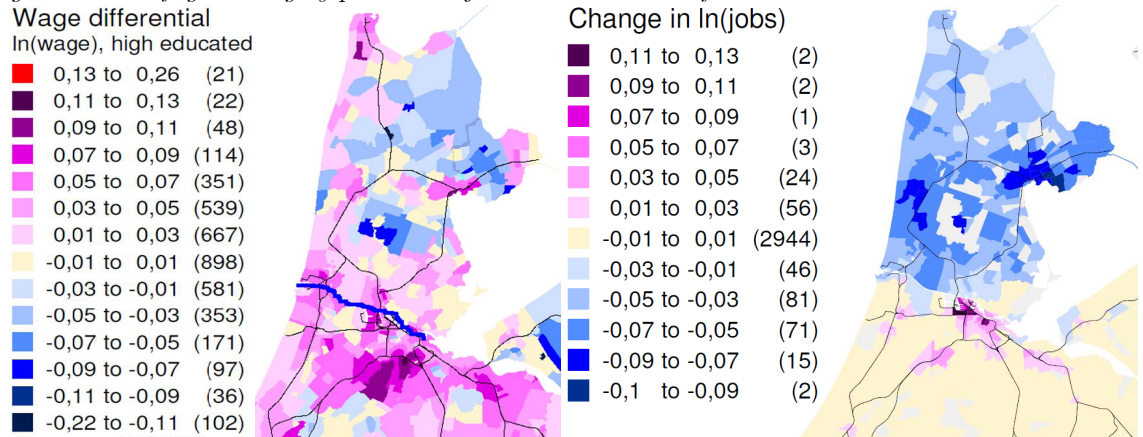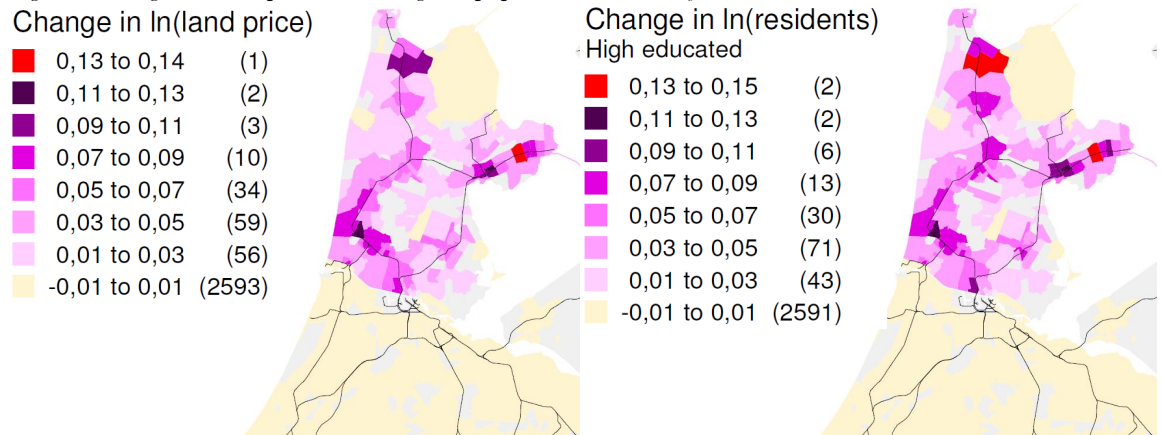

Figure 9. Higher land prices and higher population north of canal

Investments in transport infrastructure are viewed often by policy makers as a means to reduce regional and social disparities. For example, the European Commission promotes the development of a Trans-European Transport Network, while several countries considered investing in large infrastructural

projects (TGV Paris-Lyon in France, South Sea rail connection in the Netherlands, Øresund Bridge Sweden-Denmark) to stimulate the economy of peripheral regions. New economic geography models (see, for example, Baldwin et al., 2003) suggest, however, that interregional infrastructure may harm rather than help peripheral regions, leading to flight of jobs and people. Our analysis suggests that a new transportation link at commuting distance may lead to a flight of jobs from the periphery, but nevertheless it improves the region by making it a more attractive area for living, in particular for higher educated.

Table 12 Residents, job, and commuting North and South, in thousands

tunnels:	no				yes			
	low	middle	high	total	low	middle	high	total
residents:								
North	129	185	171	485	129	189	179	498
South	162	263	378	804	162	263	378	803
jobs:								
North	110	148	137	395	107	141	129	377
South	198	329	436	963	201	339	448	988
commuting:								
North-South	24	42	42	109	28	52	55	135
train	0	0	0	0	5	12	15	32
car	19	35	35	88	18	33	33	84
South-North	7	11	15	33	7	11	16	35
train	0	0	0	0	0.3	0.6	1	2
car	6	10	13	30	6	10	13	29
North-North	101	134	116	350	97	125	107	328
South-South	126	202	288	616	126	202	287	614

Table 13 reports the welfare gains from the improvement of the rail connection. The net welfare benefits for 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. This reduces the demand for land elsewhere in the country, thereby transferring the main part of the initial gain of the total class of landlords back to workers.

This is in line with Kuminoff and Pope (2013) and Bayer et al. (2007) who report a wedge between the capitalisation 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. In our general equilibrium framework that regards the whole country as a single land market, the negative and positive adjustments in land rents turn out to largely cancel each other.

The benefits of the tunnels are distributed unevenly among education levels: highly educated individuals benefit more, since they have the highest preference for commuting by train and the highest valuation for job availability. Their benefits are twice as high as the benefits of middle educated and eight times higher than the gains of low-educated individuals.

Table 13 Decomposition welfare effects, in mln euros

effect	education level			land owners			total
	low	middle	high	North	South	Else	
modal split	203	584	1133				1920
job relocation	66	191	461				718
wage effect	0	19	23				42
home relocation	-83	-149	-99				-331
land owners				1638	51	-1335	354
total	186	645	1518	1638	51	-1335	2703

The decomposition of welfare gains shows that the modal split plays an important role. Job relocation increases the welfare gains with about 40%, mainly due to jobs moving from the North to the Amsterdam metropolitan area.[21] The effect of changes in wages rates due to clustering of jobs in and around Amsterdam is small. This result has important implication for cost benefit analysis. In policy discussion, there is temptation to add some unspecified agglomeration benefit to the welfare gains from changes in modal split and job relocation. Our calculation shows that this is unjustified. The home relocation effect is negative. At the first sight it seems counter-intuitive that relaxing the constraint on changing the home location can decrease welfare. The reason is that part of the benefits is transferred to land owners. Low and middle educated lose rather much from home relocation, since high educated workers drive up prices at locations close to stations, crowding out low and middle educated workers.

8 Conclusion

Land rents differ widely across locations, reflecting differences in the local availability of public goods. Hence, land rents provide a useful tool for the welfare impact of changes in the supply of public goods. However, the direct effect of new public goods as estimated from a hedonic land rent regression potentially ignores agglomeration benefits. Moreover, since the valuation of public goods differs widely across education levels, their benefits are distributed unequally between levels of education. We have developed a spatial general equilibrium model for the valuation of the indirect effects of local public goods on home and job location choice, agglomeration benefits, and land use. Our application of the model to a particular policy experiment, closing down the railway tunnels connecting Amsterdam to the region north of the city, shows that the direct effect does indeed ignore substantial indirect effects, up to 30% of the total effect. Moreover, the benefits of the railway tunnels are distributed highly unequally across education levels indeed, the gains for higher educated being ten times larger than for lower educated. This unequal distribution of benefits poses a political economy challenge.

---

[21] We most likely overestimate the sensitivity of the location of jobs to changes in infrastructure. Firms' location in our model is positively influenced by a larger labour supply at a location; we do not model the countervailing impact of higher land prices however.

Like any model, our analysis has its limitations. We mention three. First, 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 leads to an underestimation of the benefits. Luckily, travel by car did not massively increase in our policy experiment, so this does not affect our conclusions much. Second, we allow for changes in land intensity within the area available for residential use, but we do not allow for transferring land from agricultural to residential use. Similarly, we do not allow for the local supply of amenities to be adjusted to changes in the local population density. Again, ignoring this margin of adjustment leads to an underestimation of the benefits of public goods. Our research does not provide information about the magnitude of this effect. Finally, we have not modelled commercial land use. Hence, the elasticity of labour demand at each location is overestimated, making it easier to benefit from new infrastructure. However, the agglomeration elasticities applied in our study are on the low side of the range found in the literature. Both effects might cancel. Getting a better feel for the magnitude of these additional effects is a challenge for future research.

Appendices
Appendix A Derivation of the logit models

Workers choose the transport mode maximizing v_{ihjm}. Hence, by equation (1)

$$
\begin{aligned}
\Pr\left[m|ihj\right] &= \Pr\left[v_{ihjm} > v_{ihjk}|\forall k \in M\right] \\
&= \Pr[c_{shjm} + \widetilde{e}_{ihjm} > c_{shjk} + \widetilde{e}_{ihjk}|\forall k \in M].
\end{aligned}
$$

This is a standard logit model, see equation (7). Applying the formula for a maximum of a number of type I extreme value stochasts with equal variance yields

$$
\max_{m\in M}\left[c_{shjm} + \widetilde{e}_{ihjm}\right] = c_{shj} + \widetilde{e}^*_{ihj}, \tag{17}
$$

$$
c_{shj} \equiv \ln\left(\Sigma_{m\in M}c_{shjm}\right).
$$

$\widetilde{e}^*_{ihj}$ follows the same distribution as $\widetilde{e}_{ihjm}$. Define $v_{ihj} \equiv \max_{m\in M}\left(v_{ihjm}\right)$. Substitution of equation (17) in (1) yields

$$
\begin{aligned}
v_{ihj} &= \mu_{Hs}^{-1}\left(z_h + \alpha'_s a_h + \ln K_h\right) + \mu_{Js}^{-1}g_{shj} - \omega_s\rho_s\ln R_h^* \\
&\quad + \widetilde{e}_{ih} + \widetilde{e}_{ihj} + \widetilde{e}^*_{ihj}, \tag{18}
\end{aligned}
$$

$$
g_{shj} \equiv \mu_{Js}\left(c_{shj} + y^*_{sj}\right).
$$

Workers choose the job location maximizing v_{ihj}. Hence, by equation (18)

$$
\begin{aligned}
\Pr\left[j|ih\right] &= \Pr\left[v_{ihj} > v_{ihk}|\forall k \in H\right] \\
&= \Pr[\mu_{Js}^{-1}g_{shj} + \widetilde{e}_{ihj} + \widetilde{e}^*_{ihj} > \mu_{Js}^{-1}g_{shk} + \widetilde{e}_{ihk} + \widetilde{e}^*_{ihk}|\forall k \in H].
\end{aligned}
$$

$\widetilde{e}_{ihj} + \widetilde{e}^*_{ihj}$ follows the same distribution as $\widetilde{e}_{ihj} + \widetilde{e}_{ihjm}$. Hence, $\Pr\left[j|ih\right]$ satisfies equation (8). Similar to equation (17)

$$
\max_{j\in H}\left[\mu_{Js}^{-1}g_{shj} + \widetilde{e}_{ihj} + \widetilde{e}^*_{ihj}\right] = \mu_{Js}^{-1}g_{sh} + \widetilde{e}^*_{ih}, \tag{19}
$$

$$
g_{sh} \equiv \ln\left(\Sigma_{j\in H}g_{shj}\right).
$$

$\widetilde{e}^*_{ih}$ follows the same distribution as $\widetilde{e}_{ihj} + \widetilde{e}^*_{ihj}$.

Substitution of equation (19) in (18) yields

$$
v_{ih} = \mu_{Hs}^{-1}\left(z_h + q_{sh} + \ln K_h\right) + \widetilde{e}_{ih} + \widetilde{e}^*_{ih}, \tag{20}
$$

$$
q_{sh} \equiv \alpha'_s a_h + \mu_{Hs}\mu_{Js}^{-1}g_{sh} - \mu_{Hs}\omega_s\rho_s\ln R_h^*.
$$

Hence

$$
v_{ih} = \mu_{Hs}^{-1}\left(z_h + q_{sh} + \ln K_h\right) + \widetilde{e}_{ih} + \widetilde{e}^*_{ih} \tag{21}
$$

Workers choose the home location maximizing v_{ih}. Hence, by equation (21)

$$
\begin{aligned}
\Pr\left[h|i\right] &= \Pr\left[v_{ih} > v_{ik}|\forall k \in H\right] \\
&= \Pr[\mu_{Hs}^{-1}\left(z_h + q_{sh} + \ln K_h\right) + \widetilde{e}_{ih} + \widetilde{e}^*_{ih} > \mu_{Hs}^{-1}\left(z_k + q_{sk} + \ln K_k\right) + \widetilde{e}_{ik} + \widetilde{e}^*_{ik}|\forall k \in H].
\end{aligned}
$$

Hence, $\Pr[h|i]$ satisfies equation (9).

Appendix B Data on transport cost and wages

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 kilometre traveled plus toll costs.[22] 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 cost for bus/tram/metro is calculated from the number of urban transit zones traveled.[23] Biking and walking travel times are calculated by using the travel distances calculated for car trips, assuming an average speed of 16 km/hr. The cost of these trips is set equal to zero. We deleted implausible observations, e.g. for which the actually chosen travel mode 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; home-work distances smaller than the home-work straight line).

In the dataset for wages, the ZIP code of the employer is unknown for a substantial part of the individuals. There are two types of employees:

(1) workers employed by a firm located in a single ZIP code (57% of the workers);

(2) workers employed by a firm with establishments at multiple ZIP codes (35% of the workers) or a firm for which ZIP codes are missing.

For this second group, ZIP codes are attributed by an optimization routine using the commuting distance or according to the share of each ZIP code in total employment at the firm level. This procedure yields accurate job locations for 86% of the workers in firms that employ less than 20 workers, but for only 40% of the workers in firms that employ between 1000 and 5000 workers. The level of education is derived from the Education Registry. This is highly reliable for workers younger than 40 years, but less so for older workers. Both the level of education and ZIP code for the job location are available for 2.75 million workers (one third of the workforce). We include only workers in our analysis aged 18-65 that have been in their current job for at least one month, working at least 12 hours per week, and being paid the adult minimum wage or above. Hourly wages are calculated as the monthly wages minus incidental compensation and dividing that by the sum of hours worked.

Table 14 presents the descriptive statistics of the mode choice by education level. Table 15 presents the number of ZIP codes with 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.

---

[22] This includes fuel, amortization, insurance, maintenance, and taxes for a car in a medium-price range, using http://www.autoweek.nl/kostenberekening.php?id=35685&jaar=2005) 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.

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

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.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. 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 C Estimation results first-stage transpose logit

Table 17 Estimation results home location logit

education level	low			middle			high		
variable	coef	t-val	t-val (clust)	coef	t-val	t-val (clust)	coef	t-val	t-val (clust)
job availability g^*_{rh}	0.637	(129.5)	(3.6)	0.449	(86.2)	(1.6)	0.434	(86.6)	(2.4)
$\ln R^*_h$	-0.177	(108.6)	(2.9)				0.138	(102.8)	(2.3)
# mon.1km/1000	-0.131	(12.5)	(0.4)				-0.062	(9.7)	(0.2)
# mon.1-5km/1000	-0.017	(7.3)	(0.2)				0.020	(10.4)	(0.3)
share nature 5km	0.089	(9.5)	(0.3)				0.409	(49.6)	(1.3)
dummy.uni.10km	-0.001	(0.3)	(0.0)				0.138	(59.0)	(1.5)
# rest. 1km/100	-0.331	(30.1)	(1.0)				0.097	(12.5)	(0.3)
# rest.1 to 5km/100	0.062	(43.5)	(1.1)				-0.020	(16.8)	(0.4)
intercept	-0.587	(94.2)	(2.5)				0.276	(55.1)	(1.3)
# observations			2.75 mln						

Appendix D Solution method and welfare analysis

The solution method proceeds as follows. First, we take W_{sj} to be constant and eliminate L_{shj} using equation (3). The final two equations of **(15)** can be written as a system of supply and demand equations for houses at each location h:

$$K^d_h = \sum_{s \in S} \Pr[h|s, R, W_s, K^s] N_s,$$

$$K^s_h = A_h \left[\rho_s \sum_{s \in S} \sum_{j \in H} \Pr[s|h, R, W_s] \Pr[j|sh, W_s] \left(1 - x^c_{shj}\right) \frac{\ln R^+_h}{R_h} W_{sj} \right]^{-1}.$$

where W_s, R, and K^s are vectors of W_{sj}, R_h, and K^s_h respectively. The dependence of the probabilities on these vectors is made explicit in the notation. K^s_h and K^d_h measure supply and demand for houses at location h, keeping R_h and W_{sj} and hence land use fixed. The task is to adjust R_h to set equal demand and supply at each location. The land rent elasticity of demand is approximately equal to

$$d \ln K^d_h \cong -\rho^*_s d \ln R^*_h \cong -\rho^*_s \ln R^+_h d \ln R_h,$$

see equation (9). This elasticity is used to calculate an update for land rents $\Delta \ln R_h$:

$$\Delta \ln R_h = \xi (\ln K^d_h - \ln K^s_n),$$

where ξ is a smoothing parameter.[24] This new value of land rents is used to

---

[24] ξ is set equal to 0.2. This procedure for updating R ignores the price elasticity of supply, but that does not hamper our ability to find a solution. In the empirical application in Section 7, the system converges to an equilibrium in about 110 iterations.

repeat this calculation. Next, the value of W_s is updated for each s, using equation (5) and (15) (for N_{sj}). The new values of W_{sj} are used for a new round of iterations on R.

Next, the derivation of the partial derivatives required for the welfare analysis. By equation (18) and (20), we have

$$
\begin{aligned}
g_{shj}\left(W_{sj}\right) &= \mu_{Js}\left(c_{shj} + y_{sj} + \omega_s \ln W_{sj}\right), \\
q_{sh}\left(R_h, W_{sj}\right) &= \alpha'_s a_h + \mu_{Hs}\mu_{Js}^{-1}g_{sh}\left(W_{sj}\right) - \rho_s^* \ln R_h^*\left(R_h\right),
\end{aligned}
$$

where we explicitly account for the dependence of $\ln R_h^*$ on R_h. Hence, by equation (8) and (9)

$$
\begin{aligned}
\frac{\partial g_{shj}}{\partial W_{sj}} &= \mu_{Js}\omega_s W_{sj}^{-1}, \\
\frac{\partial g_{sh}}{\partial W_{sj}} &= \exp(g_{shj} - g_{sh})\frac{\partial g_{shj}}{\partial W_{sj}} = \mu_{Js}\omega_s \Pr[j|sh, W_s]W_{sj}^{-1}, \\
\frac{\partial q_{sh}}{\partial W_{sj}} &= \mu_{Hs}\mu_{Js}^{-1}\frac{\partial g_{sh}}{\partial W_{sj}} = \mu_{Hs}\omega_s \Pr[j|sh, W_s]W_{sj}^{-1}, \\
\frac{\partial q_s}{\partial W_{sj}} &= \sum_{h \in H}\exp(z_h + q_{sh} + \ln K_h - q_s)\frac{\partial q_{sh}}{\partial W_{sj}} = \sum_{h \in H}\Pr\left[h|s, R, W_s\right]\frac{\partial q_{sh}}{\partial W_{sj}}.
\end{aligned}
$$

References

- Ahlfeldt, G.M., Redding, S.J., Sturm, D.M. and N. Wolf (2012). The Economics of Density: Evidence from the Berlin Wall, CEP Discussion Papers dp1154, Centre for Economic Performance, LSE.

- Albouy, D. and G. Ehrlich (2012). Metropolitan Land Values and Housing Productivity. NBER Working Papers 18110, National Bureau of Economic Research, Cambridge, MA.

- Alonso, W. (1964). Location and land use: toward a general theory of land rent. Cambridge, MA: Harvard University Press.

- Anas, A. and Y. Liu (2007). A Regional Economy, Land Use, and Transportation Model (Relu-Tran©): Formulation, Algorithm Design, and Testing. Journal of Regional Science, 47(3): 415-455.

- Arnott, R. and J. MacKinnon (1977). The effects of urban transportation changes: a general equilibrium simulation. Journal of Public Economics 8: 19-36.

- Arzaghi, M. and J.V. Henderson (2008). Networking off Madison Avenue. Review of Economic Studies 75: 1011-1038.

- Autor, D., L. Katz and M. Kearney (2006). The polarization of the U.S. labor market. American Economic Review 96(2): 189-194.

- Baldwin, R.E., R. Forslid, P. Martin, G.I.P Ottaviano and F. Robert Nicoud (2003). Economic Geography and Public Policy. Princeton: Princeton University Press.

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

- 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. (2007). Did highways cause suburbanization? Quarterly Journal of Economics 122(2): 775-805.

- Baum-Snow, N. and B.F. Lutz (2011). School Desegregation, School Choice, and Changes in Residential Location Patterns by Race. American Economic Review 101: 3019-3046.

- Bayer, P., McMillan, R. and K. Rueben (2004). An equilibrium model of sorting in an urban housing market. NBER Working Paper 10865, National Bureau of Economic Research, Cambridge, MA.

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

- Ben-Akiva, M. and S.R. Lerman (1985). Discrete choice analysis, The MIT Press, Cambridge Massachusetts.

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

- Berry, S. (1994). Estimating discrete choice models of product differentiation. Rand Journal of Economics 25: 242-262.

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

- Ciccone, A. and R.E. Hall (1996). Productivity and the Density of Economic Activity. American Economic Review, 86(1): 54-70.

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

- Desmet, K. and E. Rossi-Hansberg (2013). Urban Accounting and Welfare. American Economic Review 103(6): 2296-2327.

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

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

- Gamper-Rabindran, S. and Ch. Timmins (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.

- Gennaioli, N., LaPorta, R., Lopez-de-Silanes, F. and A. Shleifer (2013). Human capital and regional development. Quarterly Journal of Economics 128(1): 105-164.

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

- Glaeser, E.L. and J.D. Gottlieb (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. and M.E. Kahn (2004). Sprawl and urban growth. In: J.V. Henderson and J.F. Thisse (eds), Handbook of Regional and Urban Economics, volume 1, pp. 2481-2527, Amsterdam: Elsevier.

- 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.

- Inoa, I., Picard, N. and A. de Palma (2014). Effect of an Accessibility Measure in a Model for Choice of Residential Location, Workplace, and Type of Employment. Mimeo.

- Klaiber H.A. and V.K. Smith (2010). Valuing incremental highway capacity in a network. NBER Working Paper 15989, National Bureau of Economic Research, Cambridge, MA.

- Kuminoff, N.V. and J.C.Pope (2013). Do "Capitalization Effects" for Public Goods Reveal the Public's Willingness to Pay? Forthcoming in International Economic Review.

- Lucas, R.E. and E. Rossi-Hansberg (2002). On the internal structure of cities. Econometrica 70(4): 1445-1476.

- Roback, J. (1982). Wages, rents and the quality of life. Journal of Political Economy 90: 1257-1278.

- Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. Journal of Political Economy, 82 (1): 34-55.

- 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.

- Schelling T.C. (1969) Models of segregation, American Economic Review, 1969, 59(2): 488–493.

- Tiebout, Ch. (1956). A pure theory of local expenditure. Journal of Political Economy 64: 416-424.