

EFFECTS OF TRANSPORT IMPROVEMENTS ON COMMUTING AND RESIDENTIAL CHOICE

J. Paul Elhorst & Jan Oosterhaven¹

Prepared for the 43rd European Congress of the Regional Science Association
Jyväskylä, August 27-30, 2003

ABSTRACT

This paper develops a commuter location model able to explain and simulate residential location changes of commuters that result from transport improvements. The core model is based on the assumption of constant commuting time, while two extensions incorporate substitution possibilities having an upward effect on total commuting time. Estimation errors of the residential location of the working population with the existing transport system are limited to 7%. With the extended model, the impacts on commuting and residential choice are investigated for six higher speed rail connections between Amsterdam, located in the urban core of the Netherlands, and Groningen, located in its rural periphery. The model outcomes strongly influenced the public policy debate in the Netherlands.

KEYWORDS

Transport infrastructure, migration, commuting, the Netherlands.

¹ Department of Economics, University of Groningen, Postbus 800, 9700 AV Groningen, The Netherlands. Phone +31-50-363.3893, fax +31-50-363.7337, email: oosterhaven@eco.rug.nl.

1. INTRODUCTION

This paper focuses on the theoretical and modeling aspects of translating improvements in the transport system, expressed in terms of travel time, into effects on commuting and residential choice. To evaluate the effects of transport improvements, an increasing amount of theory and operational methods has been developed (see Vickerman, 1991; Quinet & Vickerman, 1997; Rietveld & Bruinsma, 1998; Oosterhaven & Knaap, 2002). There are three strands of literature dealing with the reactions of households.

The first strand of – gravity and entropy – models assumes that all individuals have a given residential location. This type of model is constructed to explain and predict a commuting flow matrix, provided that the marginal totals of the trip distribution matrix, i.e. the spatial distributions of employment and of population, are given. To model residential choices of commuters, however, this approach is inadequate, as these spatial distributions are not exogenous to changes in the transport system. Even in more advanced spatial interaction models, this problem still has to be solved (Naevdal et al., 1996; Thorsen et al., 1999).

A second strand of – urban economic – literature assumes that workers choose their residences to maximize utility by trading off commuting and housing costs. These models are formulated to explain the location choice of the population, population density and distance from the central business district in mono-centric cities. Recently, White (1999) has given an overview of more general, comparable models in which employment is not necessarily restricted to the central business district and in which employment locations are determined either exogenously or endogenously (see also Wilson, 2000)

The model presented here has similarities with this second strand of literature. We are also interested in the spatial pattern of the population, and we also consider multiple employment locations. The difference is that we consider a mix of urban and rural areas instead of just one urban agglomeration, and we aim at an explanation of both the intra-regional and the interregional impacts of transport improvements on commuting flows and residential choice. This difference in focus resulted from our having to answer the empirical question of the indirect impacts of railway improvements between distant mono- and polycentric urban regions, and various rural regions in the Netherlands.

A third strand of literature – elaborating on the second – raises the question whether commuting in urban areas is efficient. In a provocative article, Hamilton (1982) demonstrated that the simple mono-centric model of urban location, even when modified to allow for decentralized employment, does a poor job of predicting actual commuting in major U.S. cities. Rather, it estimates the efficient pattern of commuting, defined as the minimum average commute required for the urban area's

workers to travel from the existing housing stock to the existing jobs along the existing transportation network. To more directly compute this minimum, several authors use assignment models (White, 1988; Cropper & Gordon, 1991; Small & Song, 1992; Kim, 1995). Such models require data on (1) the actual commuting journey lengths between all pair of zones, (2) the number of residences in each zone, and (3) the number of jobs in each zone. The problem with this approach is that any commuting found beyond the minimum may easily be attributed to problems that are neglected in the specification of the assignment model used. Some of these problems are the existence of two-earner households, discrimination in the housing market and temporary jobs (see White, 1999, for more).

In this paper a kind of assignment model is developed from a different perspective. We take actual commuting behavior as given and then try to predict where people live given the location of their jobs. Besides, we use data on actual commuting times between pairs of zones, instead of data on distances. More specifically, we consider the number of jobs in each zone and the commuting time between each pair of zones by different transport modes as given, and then assign each job to a certain residential zone based on the aggregate observed commuting time distribution matrix. As opposed to the third strand of literature, our interest is not in establishing a fictitious, efficient commuting pattern, but in predicting as best as possible the actual spatial distribution of the working population. This difference in focus again resulted from our having to answer the empirical question of the indirect impacts of railway improvements in the Netherlands. Provided that the model produces acceptable historical results, it may subsequently be used to simulate the effects of future transport improvements.

The paper proceeds as follows. Section 2 introduces the model that may best be labeled as a *commuter location model*. We present its mathematical formulation and discuss the data needed to run the model. We discuss the assumption of constant commuting time, as well as two extensions that provide for variable, in fact, upward total commuting time. An application of the commuter location model is presented in section 3, where six rail projects are investigated, all improving the connection between Amsterdam Airport, located in the urban core of the Netherlands, and the City of Groningen, located in its rural periphery. Finally, some perspectives for a further development of the commuter location model are discussed in the concluding section 4.

2. THE COMMUTER LOCATION MODEL

2.1 Description of the core model

The purpose of the model is to estimate the size of the working population W_i of zone i ($i=1, \dots, Z$). It is assumed that W_i depends on the employment opportunities E_j in surrounding zones j ($j=1, \dots, Z$), the commuting times T_{ijm} from i to all surrounding zones j per transport mode m ($m=1, \dots, M$), and the commuting behavior of the working population, which is measured by means of the shares S_{mk} of the working population that uses mode m with a commuting time in time class k ($k=1, \dots, K$). In sum:

$$W_i = W_i(E_j, T_{ijm}, S_{mk}, \forall j, m, k). \quad (1)$$

Actual data on the number of jobs in each zone, on travel times per mode and on commuting behavior are used to calibrate the model.²

To reach maximum accuracy, while estimating W_i , it is helpful to actually estimate C_{ijmk} , the number of people commuting from i to j using transport mode m with a commuting time that falls in time class k , because:

$$W_i = \sum_j \sum_m \sum_k C_{ijmk}. \quad (2)$$

In (2) C_{ijmk} is assumed to depend on the number of jobs in zone j , the average commuting time between i and j by mode m , the share of commuters in time and mode class mk , as well as the number and the attractiveness A_i of residential zones i located within the same modal time class around zone j , since the employment in zone j has to be shared between these residential zones. Thus:

$$C_{ijmk} = E_j S_{mk} \left[\frac{\mathbf{d}_{ijmk} A_i}{\sum_{i'} \mathbf{d}_{i'jmk} A_{i'}} \right], \quad (3a)$$

$$\text{where } \mathbf{d}_{ijmk} = 1 \text{ if } T_{ijm} \in mk \text{ and } \mathbf{d}_{ijmk} = 0 \text{ if } T_{ijm} \notin mk. \quad (3b)$$

The \mathbf{d} 's are used to specify the rings of zones around employment zone j that have the same average commuting time per travel mode. Note that the number of jobs from zone j filled by people living in an arbitrary residential zone i decreases with

² The internal travel time within a zone i may be calculated as $T_{ii} = (O_i / \mathbf{p})^{1/2} (V_i)^{-1}$, assuming that zones are circular, and that the surface O_i and the average internal speed V_i of the zone are known.

distance. This is because not only the travel time between i and j increases with distance, but also because the number of zones within a modal time ring around zone j normally increases with distance.

The commuter location model (3) can be refined by recognizing that the average commuting time between any two zones i and j hides a whole distribution of travel times between the various subzones of i and of j . On a regular grid this distribution may be approximated by a normal distribution with mean T_{ijm} and variance V_{ijm} . When data on V_{ijm} are available, (3) may be replaced with:

$$C_{ijmk} = E_j S_{mk} \left[\frac{\text{prob}(X_{ijm} \in mk) A_i}{\sum_{i'} \text{prob}(X_{i'jm} \in mk) A_{i'}} \right], \quad (4a)$$

$$\text{where: } X_{ijm} \sim N(T_{ijm}, V_{ijm}). \quad (4b)$$

Note that both commuter location models produce very detailed estimates of commuting behavior per modal time class mk , which is not normally measured at the interzonal level ij . These estimates are used here as intermediate results. The primary aim of the commuter location model is the estimation of the total working population by residential zone with (2). Appropriate aggregations of (3) and (4) produce total commuting between zones, C_{ij} , and total commuting by travel mode, C_m , which represent useful secondary results.

Several studies (Hupkes, 1977; Zahavi & McLynn, 1983; SACTRA, 1999, p.118) dealing with time spent on travelling have found that the travel time budget of the representative consumer is constant, or is characterized by a small but stable growth rate. Although the above studies point out that this only holds for all travel motives combined, some other studies have found that the separate mean commuting time has remained fairly constant too. Journey-to-work statistics from the US Census, for example, show that the mean commuting time in the largest 15 metropolitan areas increased only slightly, from 26.0 to 26.6 minutes between 1980 and 1990 (Small & Gómez-Ibánaz, 1999, p.1941).

Under the condition of constant commuting time, the effects of transport system improvements on commuting and residential choice can be simulated by adjusting the actual travel times, $T_{ijm} \Rightarrow T'_{ijm}, \forall i, j, m$, and by running the commuter location model with T'_{ijm} . The commuters' reaction to any transport improvement, holding the actual spatial distribution of jobs constant, will mainly be motivated by changes in residential opportunities, and is therefore labeled as *residential migration*:

$$W'_i - W_i = W_i(E_j, T'_{ijm}, S_{mk}) - W_i(E_j, T_{ijm}, S_{mk}). \quad (5)$$

The commuters' reaction to changes in the spatial distribution of jobs, also due to the transport system improvement, holding the adjusted travel times constant, will mainly be motivated by changes in job opportunities, $E_j \Rightarrow E'_j, \forall j$, and is therefore labeled as *labor migration*:

$$W''_i - W'_i = W_i(E'_j, T'_{ijm}, S_{mk}) - W_i(E_j, T_{ijm}, S_{mk}). \quad (6)$$

Several studies have criticized the assumption of constant travel time. The gist of the critique is that travel time budgets are not constant because travel time and money are substitutes and should therefore be considered together. Besides, changes in the composition of the population will affect the total travel time of the whole population as travel demand between different population groups (pupils, students, employed people, non-participants, aged people) are different (see Van Knippenberg, 1987; Kraan, 1996). In view of this critique, it would be wrong to consider the assumption of constant commuting time a "law".

Activity-based models nonetheless show that travel distances increase in response to higher speeds such that total travel time remains fairly constant. This is argued as follows: travel demand represents a derived demand, the utility of which depends on the utility of the basic activities that are enabled by travelling (Kraan, 1996; Fischer, 2000). Each natural day is limited to 24 hours, and, besides time to work, each employed person needs time for sleep and personal care. If commuters spend too much time or, conversely, too little time on travelling within the time being left, the number and quality of these activities diminish. The optimal amount of travel time is thus limited from above and below.

What then happens when travel times reduce due to improvements in the transport system? In that case, consumers may increase the quality of their basic activities - by increasing their commuting journey length - without changing their commuting time. This does not mean that each individual immediately moves to another municipality farther away from his or her work. Only some people are considering to move at any particular point in time, whereas the majority does not. When the transport system improves these people broaden their horizon to houses farther away from their job. On balance, after reaching a new equilibrium, the mean commuting time of all workers may remain fairly constant. The assumption of constant commuting time may thus be used as a good first approximation of the long-term effects of a change in the transport system.

In addition, we nonetheless decided to extend the commuter location model with two types of substitution to meet those studies that have found that the total amount of commuting time shows a positive time trend (Van Ommeren, 1996; CBS, 1999; Small & Gómez-Ibánaz, 1999). The first extension relates to introducing different commuting time distribution matrices for different types of municipalities

and the second regards introducing substitution between modes of transport. Details of these extensions are spelled out in section 2.3.

2.2 Implementing the model for the Netherlands

Discussing the data needed to run the commuter location model, running ahead of its implementation for the Netherlands, further elucidates its working.

Employment (E)

The numerical starting point of the model is the number of jobs per employment zone, which equals the number of incoming commuters. The latter is then ‘translated’ into an equal number of outgoing commuters over all residential zones. This presupposes that people only hold one job and do not commute in from outside the study area (i.c. from foreign countries).

The commuting time distribution matrix (S) and the number of modes (M)

Table 1 shows an example of the type of commuting time distribution matrix needed. The percentages are adjusted on a continuous sliding scale³ and are based on more than 70,000 observations. Each transport mode consists of a number of submodes:

- Car: car drivers, car passengers, taxi, motor and scooter;
- Public transport: train, tram, bus and metro;
- Slow transport: walking, bike and moped.

Analysis of the data showed that both the use of and the travel times between these three main modes of transport differ widely across different commuting distances. The differences between the submodes within each main mode are substantially smaller and are therefore neglected.

Note that the matrix S not only reflects the heterogeneous preferences of the working population with respect to commuting time and transport mode, but indirectly also aspects related to them, such as the availability of adequate housing, preferences for urban versus rural housing environments etc.

Travel times (T)

The construction of the travel time matrices for each mode before and after changes in a transport system is a complex task. In the Netherlands the following data are available:

- AND Data Solutions supplies travel times between zip code areas by car during normal hours. The strength of these data is that they are available on a very low

³ This adjustment has been made because part of the respondents report their travel time only in multiples of 10 minutes or quarters of an hour instead of in multiples of 5 minutes as was required.

level (zip codes). The shortcoming is that they only relate to the current transportation network and that they do not correct for peak hours.

- OVR (public transport information center) supplies travel times by public transport. The strength of these data is that they are available between any two places in the Netherlands. The shortcoming is that they only relate to the current transportation network and assume that the whole trip is made by public transport.
- CBS (Central Bureau of Statistics) supplies travel times by all modes of transport, amongst others for commuters. The strength of these data is that they are actually observed during peak hours as well as off-peak hours, and relate to total travel times, i.e. they sum waiting and travel times over the whole transport chain. The shortcoming is that the number of observations is far too low to construct full matrices at lower spatial levels of analysis (e.g. municipalities).
- HCG (Hague Consultancy Group) supplies travel times between 345 zones by car and by train, inter alia for commuters during peak and off-peak hours. The strength of these data is that they are available both before and after the proposed changes in the transportation network. The shortcomings are that data on slow transport are lacking and that the data on public transport only cover journeys by train, and not by bus, tram and metro.

Starting with the HCG data and utilizing the strength of the other three sources, we have been able to construct travel time matrices for the three main transport modes for commuters during peak hours both before and after the proposed changes in the transport network (see Elhorst et al. 2000, for details).

Zoning (Z) and the variance in travel times (V)

To assign numbers of jobs in different work areas to people in residential areas, a zoning of the study area is necessary. In the Netherlands, possible zonings include the 12 provinces (nuts-2), 40 so-called corop regions (nuts-3), 548 municipalities and about 4000 zip code areas. To have access to sufficient data on employment and residential attractiveness, using administratively defined zonings is the only option. Besides, administratively defined areas are also preferred from a policy point of view. Both arguments cancel a zip code zoning scheme. Another requirement is that the zoning should be sufficiently refined. This cancels the zoning into nuts-2 regions, since these regions are simply too large to simulate commuting flows adequately.

To moderate the possible mis-estimation related to choosing a specific zoning scheme, a selective use of data from subzones may be helpful. The following illustrates why and how. From the travel times between municipalities it is possible to determine the travel time between nuts-3 regions as the mean travel time between each subzone in the origin nuts-3 region and each subzone in the destination nuts-3 region. Similarly, it is possible to determine the internal travel time of a nuts-3 region as the mean travel time between and within all municipalities within that region. The

advantage of such a set-up is that not only the mean travel time but also the variance in travel times between and within regions can be computed. Thus, (4a) may be used instead of the less adequate (3a).

In view of the above considerations, a choice was made between a zoning of the Netherlands into the 40 corop regions with a subzoning into 548 municipalities, and a zoning into 548 municipalities with a subzoning into 4000 zip code areas. In testing the fit of the model, it appeared that the results of the latter zoning were significantly better, especially with regard to commuting over shorter distance. Hence, we worked with the mean travel time between all pairs of Dutch municipalities, and with the variance of the mean travel time between these municipalities based on a subzoning of the Netherlands into zip code areas.

Residential attractiveness (A)

When different residential zones i are located within the same modal time class mk around an employment zone j , the employment that is assigned to class mk has to be shared between these residential zones. These shares are best specified by $A_i / \sum_i A_i$, where A_i denotes the attractiveness of zone i as a place to live. Two variables are the most likely candidates for A : the number of houses and the available land. The first variable best approximates the existing physical opportunities to live within a certain zone. For this reason this variable is best suitable to test the fit of the commuter location model. Research, however, has shown that the majority of the working population has a preference for larger lots in areas with more green (Elhorst et al. 1999; VROM, 2000; Rouwendal and Meijer, 2001). For this reason the second variable is more suitable to simulate long-run residential changes, provided that the housing market is not supply-constrained due to spatial policy interventions and has got time to adjust itself to changes in the transport system and the consequent changes in residential preferences.

Empty time classes

For the commuter location model to be fully operational one technical problem has to be taken into account. Jobs can only be assigned to a particular time class if there is at least one residential zone in that time class from which people can commute to these jobs. In practice, not each time class always has such a zone, especially not when (3) is used. Empty time classes will occur when an employment zone is spatially isolated from other zones, as is the case with the Dutch Wadden Islands. This problem may also occur when a certain employment zone and the residential zones in one of its surrounding modal time rings both happen to consist of single zip code areas. In that case, the variance of the travel time between these zones is zero by definition, which increases the likelihood of empty time classes. When the emptiness problem arises for a particular time class k ($k > 1$), the best solution is to assign the jobs of time class k to

the inner time class $k-1$. In the isolated island case, this implies that almost all jobs are assigned to that island and only a few jobs to residential areas on the mainland.

2.3 Modifying the assumption of constant commuting time

The core commuter location model may relatively easily be modified to meet those studies that have found that the total amount of commuting time shows a small positive time trend (Van Ommeren, 1996; CBS, 1999; Small & Gómez-Ibánaz, 1999). Here we discuss and apply the two most obvious extensions.

Substitution due to municipal differences in transport systems

In section 2.1 one commuting time distribution matrix is used for all municipalities, whereas it is likely that this matrix differs from one municipality to another because of differences in the supply of transport modes. Analysis of the primary data showed that the model's estimation error might well reduce if the following distinction is made between the employment zones:

1. The four biggest cities, i.e. Amsterdam, Rotterdam, The Hague and Utrecht;
2. Urbanized municipalities with a train station; and
3. Urbanized municipalities without a train station and rural municipalities.

The four biggest cities have the best local public transport since they are the only Dutch cities that supply tram and/or metro services. In these cities, 22% of the people commute by public transport, 50% by car and 28% by slow transport. In the second type of municipalities 7% commute by public transport, 59% by car and 34% by slow transport, whereas in the third type only 3% commute by public transport, 64% by car and 33% by slow transport. Besides, the average commuting time amounts to 34 minutes in the first, 24½ minutes in the second and 21 minutes in the third type of municipality.

Using different commuting time distribution matrices for different types of municipalities implies replacing S_{mk} with S_{jmk} in (4a). This leads to an implicit relaxation of the assumption of constant commuting time for two reasons. First, a municipality will shift from the third to the second type if it gets connected to the railway network. As a result, total commuting time will rise. Second, jobs may shift to a different type of municipality. Total commuting time may either rise or fall depending on the type of shift. In the case study shifts towards the four biggest cities and shifts from municipalities without a train station to municipalities with a train station dominate, which also increases total commuting time.

Substitution due to shifts between modes of transport

When the travel time between two zones for one particular mode of transport reduces, substitution towards this mode may occur at the expense of others. These modal shifts

have been modeled with the almost ideal demand system (AIDS, see Deaton and Muellbauer, 1980). Its general form can be specified as:

$$w_i = \mathbf{a}_i + \sum_j \mathbf{g}_{ij} \ln P_j + \mathbf{b}_i \ln (T / P), \quad (7a)$$

where: w_i is the share of total expenditure allocated to the i -th good ($w_i = P_i Q_i / T$), P_i is the price of the i -th commodity, T is total expenditure on the group of commodities being analyzed, and P is the price index for the group, which in turn is defined as:

$$\ln P = \mathbf{a}_0 + \sum_j \mathbf{a}_j \ln P_j + 0.5 \sum_j \sum_i \mathbf{g}_{ij} \ln P_i \ln P_j, \quad (7b)$$

In the case study, the variables w , P and T are taken from the commuting time distribution matrix S : w_{mk} measures the share of each transport mode within each time class (S_{mk} / S_k), P_{mk} measures the time needed to travel one kilometer by each transport mode in each time class ($T_{mk} / Distance_{mk}$), and T_k measures total expenditure (read: time spent) on commuting.⁴

The parameters of this model are estimated by Maximum Likelihood taking into account the usual adding-up, homogeneity and symmetry conditions to be consistent with consumer theory (see Deaton and Muellbauer, 1980). Note that the parameters are not estimated by micro-economic data but by aggregated data taken from the commuting time distribution matrix in Table 1. Our approach, however, is also based on the assumption of a representative consumer. An estimation of the AIDS system by micro-economic data, in our case, would complicate the analysis for two reasons. The first reason is that individuals only choose one transport mode, which would require a limited-dependent variable model. The second reason is that at the individual level only the price of the chosen transport mode is observed and not those of the alternative transport modes.⁵

Table 2 reports the parameter estimates and the price elasticities that can be derived from these parameter estimates (see Green and Alston, 1990, for the exact formulas). It appears that if the rail connection between two zones becomes faster, both the numbers of people travelling by car and by slow transport decreases. The

⁴ For each mode of transport the time needed to travel 1 kilometer is found to decrease with the commuting trip time. For the slow transport mode it starts with about 8 minutes for a short trip and decreases to 5 minutes for a long trip. For public transport it starts with about 4 minutes and decreases to 1,5 minutes and for car transport it starts with about 3 minutes and also decreases to 1,5 minutes.

⁵ In this respect it should be noted that a typical commuting model of modal choice contains personal characteristics and attributes of the residential environment, but not prices of the different transport modes (see Dieleman et al., 2002, for a recent example based on the same data set being used in this study).

modal shift in the case study has been approached with the help of this model by replacing S_{mk} in (4a) with:

$$S_{mk} = S_{mk}(P_{1k}, \dots, P_{Mk}, T_k), \quad (8)$$

using (7). The strength and novelty of this approach is that it takes account of substitution between transport modes at each single pair of zones (548×548 in total). When the transport system improves, as in the case study, this again has an upward effect on total commuting time.

2.4 The commuter location model's fit

Some simulations to judge the model fit have been discussed above. In this section the observed and predicted working population estimated at the level of 548 municipalities is compared for nuts-2 regions, nuts-3 regions and 400 municipalities in 1998. A comparison at the level of all 548 municipalities is impossible, as data on the working population are available for only 400 municipalities.

When the number of houses is used to model actual housing opportunities, the estimation error of (4), inclusive of the first extension,⁶ appears to be relatively small. After aggregating the predicted working population to the nuts-3 level, the mean absolute error is almost 12,000 people. After aggregating to the nuts-2 level, the mean error is almost 30,000 people. In percentages, the mean error amounts to just 7% at both spatial levels of analysis, which implies that the commuter location model maintains its accuracy when scaling down from the 12 nuts-2 regions to the 40 nuts-3 regions. With 400 municipalities, the mean absolute error is 2,350 people or 17% of the working population per municipality. This last partial result shows that the accuracy of the commuter location model declines at the level of municipalities. For this reason inferences are only drawn at the nuts-3 level (cf. Figure 1).

When the available land is used to model housing preferences, the estimation error of (4) increases to 14%. This was expected, as the Dutch housing market is supply-constrained due to spatial policy interventions that restrict the fulfillment of people's preferences for space. Overall, it may be concluded that the commuter location model produces quite acceptable results.

⁶ The second extension is only relevant when simulating the effects of transport improvements.

3. AN APPLICATION TO SIX PROPOSED NEW RAIL LINKS

The Dutch government will soon decide on the construction of a new or better rail connection between Schiphol (i.e. Amsterdam Airport) and the City of Groningen in the North of the country. Amsterdam is part of the Randstad region, the economic core of the Netherlands. This region covers 42% of Dutch employment and more than 48% of Dutch GDP, while its surface only covers 15% of national total.⁷ By contrast, the North is considered an economically lagging region, covering 8% of Dutch employment and 11% of Dutch GDP, while its surface covers 27% of national total.⁸ Attempts to stimulate the northern economy with infrastructure, investment premiums and relocation of central government offices have not solved the problem yet. The unemployment rate is still relatively high and part of the once-relocated offices, after privatization, returned to the Randstad (see Oosterhaven, 1996, for an overview).

The construction of a new rail link between the Randstad and the North, especially a high-speed rail link, is thought to remedy at least part of this problem. Arguments in favor of this link focus on two indirect effects. The first is related to the commuting behavior of workers in the Randstad region. With a fast rail link, they could relocate to the North while keeping their current jobs. This increases demand for locally produced goods, which initiates a multiplier process leading to a higher level of regional production and employment.

The second effect is related to the changes in the size and location of production that can be brought about by the link. Many companies that start in the North eventually move to the Randstad, quoting their desire to be close to other companies as a reason for leaving. This desire is strong enough to overcome the higher prices of land, the tighter labor market and greater congestion in the Randstad. With a rail link in place, the prices of the services to and from firms located in the North would be lower, possibly shifting the balance in favor of a location outside the Randstad. This induced activity is seen as the key to further economic development.

We have evaluated six rail variants in relation to a baseline scenario. The latter projects the number of jobs in each municipality for 2020 (see further TNO et al., 2000). The six rail variants are briefly described in Table 3 and shown in Figure 1. Each rail variant leads to exogenous changes in the travel time matrices for the 548 municipalities. These are fed into the commuter location model two times. First, they are combined with the exogenous baseline employment of 2020 for each municipality, to estimate the number of *residential migrants* as in (5). Second, the new travel times

⁷ This region consists of the Amsterdam/Noordzeekanaal region, the province of Utrecht and the province of South Holland with Rotterdam and The Hague. GDP statistics are taken from RUG/CBS (1999, p.16), the other statistics from the 1998 LISA database.

⁸ The North consists of the three northern provinces of Groningen, Friesland and Drenthe. Its GDP figure is misleadingly high since the huge natural gas exploitation in the North is part of it, whereas the North hardly benefits from it.

are combined with the endogenous transport cost induced employment for 2020 in each municipality, as determined by a separate spatial CGE model (see Knaap and Oosterhaven, 2000),⁹ to estimate the number of *labor migrants* as in (6).

Figure 1 shows the spatial distribution of the total number of migrants at the nuts-3 level, while Table 4 reports the separate numbers of migrants for four aggregate regions. In interpreting the numbers three remarks should be made. First, when some rail lines are made faster, competing lines may have to become less frequent and stop more often due to the consequent reduction in their number of passengers. In the case study, this indeed happens on some competing old rail lines. Second, if rail links improve, competing car links may also improve due to less congestion. Third, not all migrants do actually move. Part of the predicted in-migration into the North relates to prevented out-migration, foreseen in the baseline scenario. Usually, a large part of the university and college graduates leave the North to take up jobs in the Randstad. With the extra employment in the North some of them will decide to stay, while others may take the opportunity to commute to the Randstad instead of migrating to it.

One of the issues in the public debate about the high-speed rail links is whether the North would thus be able to catch up with the rest of the country. Table 4 shows that in case of a Maglev (the MZM and MZB variants), the working population increases by 4,000-4,740 people,¹⁰ which is about 0.7% in relation to the total working population. Approximately 75% of this increase is due to labor migration. In case of a TGV link (the ZHS variant), the total number of migrants amounts to 2,250 people, 83% of which is due to labor migration. In sum, the North indeed catches up (in terms of both population and jobs), but whether the effect found is large or small is difficult to assess, since this involves a comparison with subjective expectations that are highly personal.

Another issue in the public debate is whether a high-speed rail link relieves the market tensions in the Randstad region. Transport networks, industrial and office sites, housing and labor markets are all under pressure in the heavily urbanized Randstad, resulting in congestion and higher land, labor and housing prices compared to the more peripheral North. As these costs are partly external to private decision-makers, they lead to an over-concentration of people and economic activities, since such costs are not taken into account in private location decisions (cf. Elhorst et al., 1999). Table 4 shows that in case of a Maglev, the working population in the Randstad decreases by 3,580-7,630 people, which is approximately 0.1 to 0.2% in

⁹ This spatial GCE model fits in the new economic geography literature (cf. Fujita, Krugman and Venables, 1999) and directly extends the model introduced by Venables (1996). Its basic structure resembles a similar model developed for the European Union by Bröcker (1999), but it is more detailed in that fourteen different sectors and their input-output linkages have been specified, while people and freight transport cost are separated.

¹⁰ Multiplying by about 2.2 gives the comparable numbers for the total population.

relation to the total working population. By contrast, in case of a TGV system the working population increases by 720 people. This is because the number of labor immigrants exceeds the number of residential emigrants. In other words, the pull of a TGV train, contrary to a Maglev, appears to work in the wrong direction.

When looking at the other regions, it appears that the spatial effects are the largest for the province of Flevoland. The reason is that this region is located along the line in-between the North and the Randstad. In case of a Maglev, its working population increases by 7,590-12,750 people, which is as much as 5 to 10% in relation to its working population. As opposed to the net migration into the North, the greatest part of the increase in Flevoland consists of residential migrants (87%). This increase comes on top of an already large autonomous growth foreseen in the baseline projection. In fact, its most populous municipality (Almere) is already the fastest growing Dutch city with an increase from 100,000 to 150,000 people over the last six years. For the second most populous city of Flevoland (Lelystad) this development will probably be very welcome, as this city has to cope with vacant houses and high unemployment.

The results for the rest of country, finally, illustrate that the relocation of the working population is a zero-sum game.

Three other results figuring in the public debate are of interest. The first regards the intra-regional migration flows. Migrants, especially residential migrants, are interested in locations close to train stations. However, when the time distance to a train station would become too large, the attractiveness of that location gets lost. The effect is that especially those municipalities that are called at by the high-speed rail link will benefit. One of the disappointments in this respect is that the spatial effects for Eastern Groningen - a backward region with its center approximately 20 kilometers eastward of the City of Groningen (cf. Figure 1) - are completely neutral, whereas one hoped for positive effects.

Another interesting result regards the impacts of differences in schedules, such as frequency and waiting time. The most striking differences in this respect are found for the city of Groningen. In case of the Maglev metro variant (MZM, 6 trains per hour with 6 stops between Groningen and Amsterdam), 2,000 residential migrants leave the city. As the number of labor immigrants amounts to 2,260, the working population will nonetheless increase slightly. In case of the Maglev high speed variant (MZB, 2 trains with 6 stops and 2 trains with only 3 stops per hour), only 40 residential migrants leave the city, whereas the number of labor immigrants increases to 2,600. In other words, the spatial effects appear to be strongly influenced by differences in service levels.

Finally, the modal shifts are of interest. In case of the Maglev metro variant, 6.4% of the commuters travelling by car between the eight cities with stations will switch to public transport. For the Netherlands as a whole, it concerns 7,800

commuters or 0.01% of the total working population (see Table 5). A small number of commuters switch from public transport to slow transport. The reason is that the public transport in some parts of the country deteriorates as a consequence of the new rail link.

4. CONCLUSIONS

In this paper we have developed, tested and applied a new model to simultaneously predict the intra-regional and the interregional distribution of the working population, dependent on the spatial distribution of employment and the commuting times in a transport network. The model is able to simulate any transport improvement in the private as well as the public transport network. The first tests of the commuter location model are encouraging. The estimation error for the working population of the 12 Dutch nuts-2 regions and the 40 Dutch nuts-3 regions is only 7%, while the results of the case study appeared to be plausible and very informative from a policy point of view. The estimation error, however, did increase to 17% at the level of 400 Dutch municipalities.

Still a great deal of research may be done. A major challenge is to fully integrate the extensions with the core model into one single framework; a model in which individuals maximize the utility of their total activity pattern, including travelling, under a time and a money budget constraint. Such a framework should include a model explaining the modal split in response to travel costs. The commuting time distribution matrix, which is used to run the present model, is based on current prices. But if prices change, the shares in this matrix may also change. To cope with this problem the Almost Ideal Demand System, adapted here for the first time to estimate modal shifts at pairs of zones, besides travel time, also needs to include user cost and other quality aspects that affect modal choice.

Such a single framework, however, requires a database that will not always be available. Using the present commuter location model then provides a good, less data-intensive alternative.

ACKNOWLEDGEMENTS

The authors thank Ward E. Romp and Dirk Stelder for helping with the idea of the core model, and for processing data and running the extended version of the commuter location model.

REFERENCES

- Bröcker J. (1999) *Trans-European Effects of "Trans-European Networks": A Spatial CGE Analysis*. Mimeo, University of Technology, Dresden.
- CBS (1999) *De Mobiliteit van de Nederlandse Bevolking in 1998*. Central Bureau of Statistics, Voorburg/Heerlen.
- Cropper M.L. & Gordon P.L. (1991) Wasteful Commuting: A Re-examination. *Journal of Urban Economics* 29, pp.2-13.
- Deaton A. & Muellbauer J. (1980) *Economics and Consumer Behaviour*. Cambridge University Press, Cambridge.
- Dieleman F.M., Dijst M. & Burghouwt G. (2002) Urban Form and Travel Behaviour: Micro-level Household Attributes and Residential Context. *Urban Studies* 39, pp. 502-527.
- Elhorst J.P., Oosterhaven J., Sijtsma F. & Stelder D. (1999) Welfare Effects of Spatial Deconcentration: A Scenario for the Netherlands. *Tijdschrift voor Economische en Sociale Geografie* 90, pp. 17-31.
- Elhorst J.P., Knaap T., Oosterhaven J., Romp W., Stelder T.M., & Gerritsen E. (2000) *Ruimtelijk Economische Effecten van Zes Zuiderzeelijn Varianten*. REG-publication 22, University of Groningen.
- Fischer M.M. (2000) Travel demand. In: Polak J.B. & Heertje A. (eds.) *Analytical Transport Economics*, pp. 51-78. Edward Elgar, UK.
- Fujita M., Krugman P.R. & Venables A.J. (1999) *The Spatial Economy, Cities, Regions and International Trade*. MIT Press, Cambridge, MA.
- Green R. & Alston J.M. (1990) Elasticities in AIDS Models. *American Journal of Agricultural Economics* 72, pp. 442-445.
- Hamilton B.W. (1982) Wasteful Commuting. *Journal of Political Economy* 90, pp. 1035-1053.
- Hupkes G. (1977) *Gasgeven of Afremmen, Volume 1 and 2*. Kluwer, Deventer.
- Kim S. (1995) Excess Commuting for Two-Worker Households in the Los Angeles Metropolitan Area. *Journal of Urban Economics* 38, pp. 166-182.
- Knaap T. & Oosterhaven J. (2002) *The welfare effects of new infrastructure: An economic geography approach to evaluating a new Dutch railway link*. Paper presented to the North American RSAI meetings, Chicago, November 9-12.
- Knippenberg C. van (1987) *Time in Travel*. University of Groningen (PhD).
- Kraan M. (1996) *Time to Travel?* University of Twente (PhD).

- Naevdal G., Thorsen I. & Uboe J. (1996) Modeling Spatial Structures through Equilibrium States for Transition Matrices. *Journal of Regional Science* 36, pp. 171-196.
- Ommeren J.N. van (1996) *Commuting and Relocation of Jobs and Residences: A Search Perspective*. Free University of Amsterdam (PhD).
- Oosterhaven J. (1996) Dutch Regional Policy Gets Spatial. *Regional Studies* 30, pp. 527-32.
- Oosterhaven J. & Knaap T. (2000) Spatial Economic Impacts of Transport Infrastructure Investments. Forthcoming in: A. Pearman, P. Mackie. J. Nellthorp & L. Giorgi (eds.), *Transport Projects, Programmes and Policies: Evaluation Needs and Capabilities*, European Union, Brussels.
- Quinet E. & Vickerman R. (eds.) (1997) *The econometrics of major transport infrastructures*, Applied Econometrics Association Series. MacMillan, UK.
- Rietveld P., Bruinsma F. (1998) *Is Transport Infrastructure Effective? Transport Infrastructure and Accessibility: Impacts on the Space Economy*. Springer, Berlin.
- Rietveld P. & Nijkamp P. (2000) Transport Infrastructure and Regional Development. In: Polak J.B. & Heertje A. (eds.) *Analytical Transport Economics*, pp. 208-232. Edward Elgar, UK.
- Rouwendal J. & Meijer E. (2001) Preferences for Housing, Jobs, and Commuting: A Mixed Logit Analysis. *Journal of Regional Science* 41, pp. 475-505.
- RUG, CBS (1999) *Regionale Samenhang in Nederland, Bi-regionale Input-output Tabellen en Aanbod- en Gebruiktabellen voor de 12 Provincies en de Twee Mainport Regio's*. REG-publication 20, University of Groningen/Central Bureau of Statistics, Groningen.
- SACTRA (1999) *Transport and the Economy*. Standing Advisory Committee on Trunk Road Assessment, TSO, London.
- Small K.A. & Song S. (1992) "Wasteful" Commuting: A Resolution. *Journal of Political Economy* 100, pp. 888-898.
- Small K.A., Gómez-Ibánaz J.A. (1999) Urban Transportation. In: Mills E.S., Cheshire P. (eds.) *Handbook of Regional and Urban Economics, Volume 3*, pp. 1937-1999. Elsevier, Amsterdam.
- Thorsen I., Uboe J. & Naevdal G. (1999) A Network Approach to Commuting. *Journal of Regional Science* 39, pp. 73-101.
- TNO, RUG, VU & TUD (2000) *Indirecte Effecten Zuiderzeelijn, Hoofdrapport*. TNO Inro, University of Groningen, Free University, Technical University Dresden, Delft.

- Venables A.J. (1996) Equilibrium Locations of Vertically Linked Industries. *International Economic Review* 37, pp. 341-359.
- Vickerman R.W. (ed.) (1991) *Infrastructure and Regional Development*. Pion, London.
- VROM (2000) *Nota Wonen; Mensen, Wensen, Wonen* (Summary). Ministry of Spatial Affairs, The Hague.
- White M.J. (1988) Urban Commuting is not 'Wasteful'. *Journal of Political Economy* 96, pp. 1097-1110.
- White M.J. (1999) Urban Areas with Decentralized Employment: Theory and Empirical Work. In: Mills E.S., Cheshire P. (eds.) *Handbook of Regional and Urban Economics, Volume 3*, pp. 1375-1412. Elsevier, Amsterdam.
- Wilson, A.G. (2000) *Complex Spatial Systems: The Modelling Foundations of Urban and Regional Analysis*. Pearson Education, Harlow.
- Zahavi Y. & McLynn J.M. (1983) Travel Choices under Changing Constraints as Predicted by the UMOT Model. *Colloquium Vervoersplanologisch Speurwerk*, Delft, pp. 137-154.

Table 1. The Dutch commuting time distribution matrix for 1998 (%)

| Time class | Car | Public transport | Slow transport | Total |
|-----------------|-------|------------------|----------------|--------|
| 0- 5 minutes | 6.50 | 0.02 | 9.85 | 16.38 |
| 5- 10 minutes | 9.31 | 0.08 | 8.40 | 17.78 |
| 10- 15 minutes | 10.06 | 0.17 | 6.29 | 16.52 |
| 15- 20 minutes | 8.39 | 0.22 | 2.73 | 11.34 |
| 20- 25 minutes | 6.55 | 0.29 | 2.06 | 8.90 |
| 25- 30 minutes | 4.91 | 0.79 | 1.39 | 7.08 |
| 30- 35 minutes | 3.61 | 0.73 | 0.88 | 5.22 |
| 35- 40 minutes | 2.65 | 0.67 | 0.54 | 3.86 |
| 40- 45 minutes | 1.95 | 0.60 | 0.34 | 2.90 |
| 45- 50 minutes | 1.46 | 0.53 | 0.22 | 2.20 |
| 50- 55 minutes | 1.10 | 0.45 | 0.14 | 1.70 |
| 55- 60 minutes | 0.84 | 0.39 | 0.10 | 1.32 |
| 60- 65 minutes | 0.64 | 0.33 | 0.00 | 0.97 |
| 65- 70 minutes | 0.50 | 0.27 | 0.00 | 0.77 |
| 70- 75 minutes | 0.38 | 0.23 | 0.00 | 0.61 |
| 75- 80 minutes | 0.30 | 0.19 | 0.00 | 0.49 |
| 80- 85 minutes | 0.23 | 0.16 | 0.00 | 0.39 |
| 85- 90 minutes | 0.18 | 0.13 | 0.00 | 0.31 |
| 90- 95 minutes | 0.13 | 0.11 | 0.00 | 0.24 |
| 95-100 minutes | 0.10 | 0.09 | 0.00 | 0.19 |
| 100-105 minutes | 0.07 | 0.07 | 0.00 | 0.14 |
| 105-110 minutes | 0.05 | 0.06 | 0.00 | 0.11 |
| 110-115 minutes | 0.03 | 0.05 | 0.00 | 0.07 |
| 115-120 minutes | 0.01 | 0.04 | 0.00 | 0.05 |
| > 120 minutes | 0.26 | 0.21 | 0.00 | 0.47 |
| Total | 60.20 | 6.86 | 32.94 | 100.00 |

Source: CBS (1999).

Table 2. Substitution between transport modes estimated with AIDS.

| Parameter estimates | | | |
|---|--------------------------|--------------------------------|----------------|
| $\alpha_0 = 10.682$ (1.46) | $\beta_1 = 0.010$ (0.31) | $\gamma_{11} = -0.188$ (-1.97) | |
| $\alpha_1 = 0.590$ (2.00) | $\beta_2 = 0.141$ (3.86) | $\gamma_{12} = 0.176$ (2.25) | |
| $\alpha_2 = 1.398$ (1.43) | | $\gamma_{22} = -0.033$ (-0.31) | |
| | | | $R^2 = 0.928$ |
| Own and cross price elasticities per transport mode | | | |
| Mode by price: | Car | Public transport | Slow transport |
| Car | -1.33 | 0.27 | 0.04 |
| Public transport | 0.30 | -1.73 | -0.09 |
| Slow transport | 0.86 | 0.26 | -0.97 |

T-values in parentheses, McElroy R^2 for the AIDS system.

Table 3. Description of the proposed rail variants,
with travel times Schiphol-Groningen (in minutes)*

| | | |
|------|---|-----|
| BASE | Baseline scenario, including the so-called Hanzeline from Lelystad to Zwolle, which is not yet constructed | 118 |
| HIC | Hanzeline intercity, the only difference with BASE is its higher speed | 102 |
| HHS | Hanzeline high speed, including rail shortcuts and a TGV only calling at major stations | 71 |
| ZIC | Zuiderzeeline intercity, a new rail track from Lelystad to Groningen with conventional trains | 89 |
| ZHS | Zuiderzeeline high speed, as ZIC but including a TGV only calling at three major stations | 65 |
| MZM | Transrapid metro, a new Magnetic levitation rail system from Schiphol to Groningen calling at all 6 intermediate stations 6 times per hour | 59 |
| MZB | Transrapid high speed, as MZM but calling at the 6 stations 2 times, at the 3 major stations 2 times, and running a shuttle 2 times from Schiphol to Almere | 45 |

* See figure 1 for the trajectories of the two H-variants and the four Z-variants.

Table 4. Residential and labor migration in deviation of the baseline in 2020.

| Region | | HIC | HHS | ZIC | ZHS | MZM | MZB |
|---------------------|----------------------|-------------|--------------|--------------|--------------|--------------|--------------|
| The North | Residential migrants | -310 | -430 | -110 | 380 | 1030 | 1370 |
| | Labor migrants | 650 | 1810 | 830 | 1870 | 2970 | 3370 |
| | <i>Total</i> | <i>340</i> | <i>1380</i> | <i>720</i> | <i>2250</i> | <i>4000</i> | <i>4740</i> |
| The Randstad | Residential migrants | -520 | -810 | -520 | -920 | -9420 | -5290 |
| | Labor migrants | 60 | 190 | 1060 | 1640 | 1790 | 1710 |
| | <i>Total</i> | <i>-460</i> | <i>-620</i> | <i>540</i> | <i>720</i> | <i>-7630</i> | <i>-3580</i> |
| Flevoland | Residential migrants | 630 | 390 | 1570 | 1580 | 10850 | 6000 |
| | Labor migrants | 300 | 480 | 300 | 660 | 1900 | 1590 |
| | <i>Total</i> | <i>930</i> | <i>870</i> | <i>1870</i> | <i>2240</i> | <i>12750</i> | <i>7590</i> |
| Rest of the country | Residential migrants | 210 | 840 | -940 | -1050 | -2470 | -2080 |
| | Labor migrants | -1030 | -2530 | -2230 | -4230 | -6700 | -6750 |
| | <i>Total</i> | <i>-820</i> | <i>-1690</i> | <i>-3160</i> | <i>-5280</i> | <i>-9170</i> | <i>-8820</i> |

Table 5. Modal shifts with six rail variants in deviation of the baseline in 2020.

| | HIC | HHS | ZIC | ZHS | MZM | MZB |
|---------------------------------------|-----|-----|------|------|------|------|
| Substitutions car to train | 600 | 600 | 2200 | 2800 | 7800 | 5600 |
| Substitutions slow transport to train | -40 | -70 | -170 | -170 | +30 | -50 |

Figure 1. Total migration effects of six rail variants in deviation of the baseline scenario in 2020.

