APPENDICES

APPENDIX 1: DESCRIPTION OF THE INTEGRATED LAND-USE/TRANSPORT MODELS OF BRUSSELS, HELSINKI AND STUTTGART

1. Description of the model of Brussels

1.1 The TRANUS framework[1]

The Brussels case city uses the TRANUS software. TRANUS is an integrated land-use transport model developed by Modelistica based in Caracas, Venezuela. It can be applied at an urban or at a regional scale. The software has a double purpose: firstly, the simulation of the probable effects of applying particular land use and transport policies and projects, and secondly the evaluation of these effects from social, economic, financial and energy points of view. The advantages of integrating the modelling of land-use and transportation are well known and have been documented extensively in the literature. For the transport planner, land-use and transport integration provides a means of making medium and long-term demand estimates, which are impossible with transport-only models (where demand is a given input).

TRANUS has its roots in the tradition of spatial interaction theories, building on Wilson (1970) who first showed how land use and transport could be represented in a common theoretical framework. It also draws heavily on the work of Domencich and McFadden (1975) in discrete choice analysis and random utility theory. Although these authors proposed a general model, most of their work and that which followed is centred on the problem of modal choice in transport, and no specific models were proposed and developed for other elements of the urban or regional system. In TRANUS, this theoretical backbone has been extended to all decision levels, from modal split to assignment, trip generation, the location of activities, and the behaviour of property developers.

In general terms, decision theory describes social processes as sets of decisions made by individuals. The main assumption is that individuals choose rationally between the options available to them. Each individual, faced with a number of options, will rank them according to the degree of satisfaction or *utility* perceived in each case, and will choose the one that provides the greatest utility. On the other hand, utility is a subjective phenomenon - its perception will vary from one individual to another and from one choice to another.

Mathematically, utility can be represented as a *utility function* for a particular individual, which contains variables describing measured attributes of each option. Faced with a particular set of options, an individual may be assumed to evaluate each one with the same utility function, and will choose the option that yields the greatest utility. This concept provides the basis of microeconomic theory.

Aggregation introduces sources of variability, because individuals within a group are different and perceive utility in different ways. The same can be said about aggregated options and zones. Naturally, if groups are small, variability will be small also. In order to represent variability, *random utility* adds a random element to the utility function.

In the individual case, the utility function is deterministic and produces a unique result: the selection of a specific option (i.e. the one with greatest utility). In the aggregate case, since there are random elements, utility functions are probabilistic, producing a distribution of individual choices among the available groups of options. Mathematically, the probabilistic model is obtained by integrating the joint distribution. Hence, several models may be derived

---

[1] This section is widely inspired from the User Guide of TRANUS.

from the general one, according to the particular shape of the distribution. Domencich and McFadden (1975) explored several possible shapes, showing that the most appropriate was the Gumbel distribution, which after integration yields a multinomial logit model. If logit is the chosen model, then there is one and only one way of measuring the average utility of the population, the logarithmic average of the distribution, also called composite cost or *log-sum*. Furthermore, if such a model is applied in the context of two different scenarios of future conditions, the difference in utility will be equivalent to the consumers' surplus in traditional economic theory. In TRANUS, this general formulation has been improved in several ways, introducing scaled utilities and an improved formulation of the log-sum.

So far we have discussed one particular choice situation. In an urban or regional system however, long and complex *chains of decisions* may be established. An example of a typical chain would be:

place of work → residence → shopping → transport mode

Each link along the chain is conditioned by the preceding link. Thus, where to go shopping is a decision conditioned by the place of residence; the choice of place of residence is in turn conditioned by the place of work. In order to represent such a decision chain in a set of sub-models, the components must follow each other in the correct order. The problem is however complicated by the fact that each link in the chain may influence the preceding one. Thus in the example, it could well be that people decide to go shopping precisely because there is a good bus service: the choice of transport mode affects the choice of shopping place. All this means that the estimation process must work along the decision chain in both directions, backward and forward, calculating and multiplying the probabilities, until a state of equilibrium is reached. Demand elasticities also influence the process

An explicit *dynamic structure* relates the two main components of TRANUS, land use and transport. The way in which the land use relates to transport through time is shown in Figure 1, where discrete time intervals are represented as t_1, t_2, t_3, and so on. The land use and transport systems influence each other through time. Economic activities in space interact with each other, generating flows. These flows determine transport demand within the same time period, and are assigned to the supply of transport. In turn, the demand-supply equilibrium at the transport level determines accessibility, which is fed back to the land use system, influencing the location of activities and their interaction. This feedback does not, however, occur instantaneously in the same time period, but is lagged. Hence, transport accessibility in period t_1 affects the distribution of flows in the following period t_2. Since there are also elements of inertia in land use from one period to the next, the effects of transport might well take several periods to consolidate.

A change in the transport system, such as a new road, a public transport system, or changes in fares, will have an immediate effect on travel demand, but will only affect activity location, interaction, and the property market in the following time period. Changes in land use, on the other hand, such as growth in the production of particular economic sectors, a new supply of land, buildings, or investment, will result in modified interactions and change transport demand within the same time period.

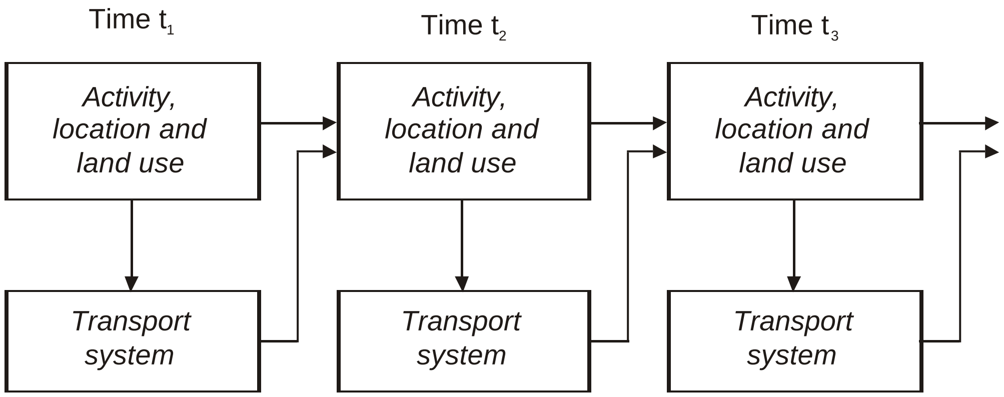

Figure 1: Dynamic relations in the land use/ transport system

1.2 The design of the Brussels land use/transport model

The initial integrated land-use/transport model for the Brussels metropolitan area has been developed in 1996 as part of the ESTEEM project, and has been used in several studies for federal, regional and local transport authorities, for the purpose of policy testing. The model has been designed to assess the major impacts of the future REN on the migrations of households and induced activities, and on the modal choice of people.

In the current version of the model, the study area covers the region that would be served by the future Regional Express Railway (about 30 km around Brussels) The area includes 19 administrative entities in Brussels-Capital Region and 116 municipalities in the suburban area.

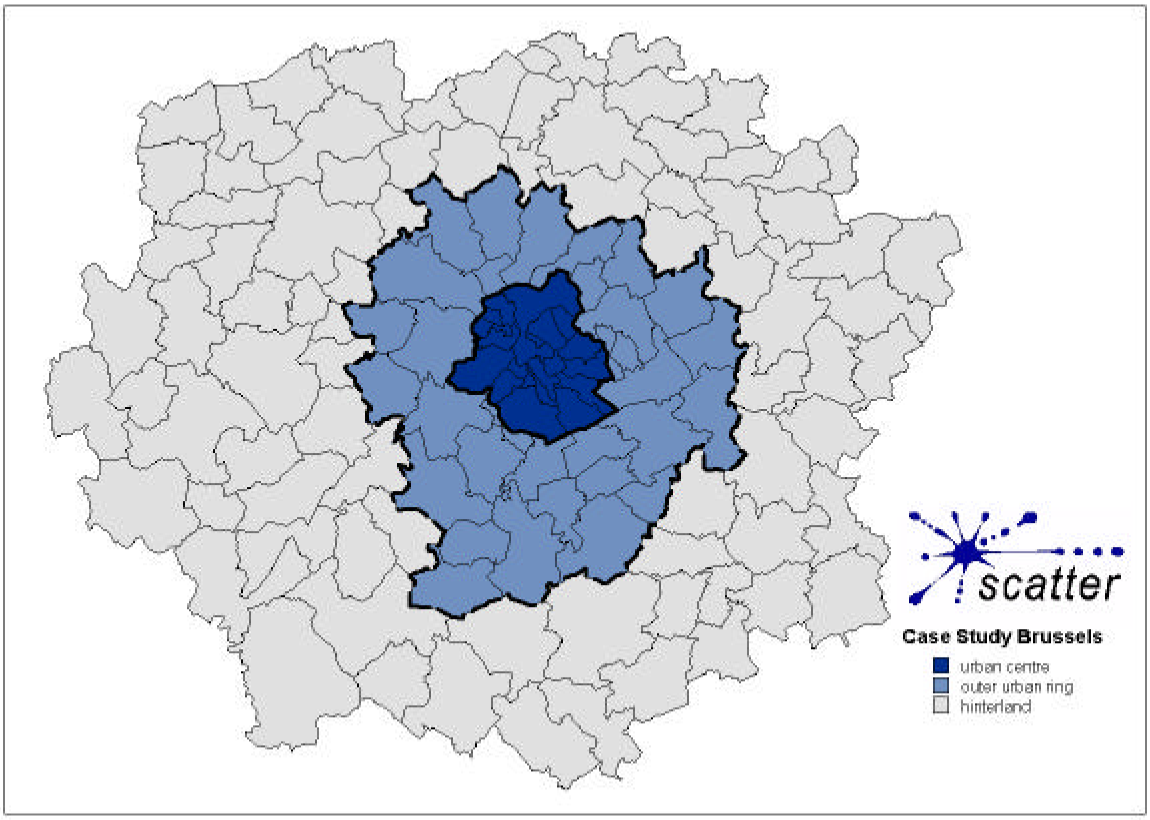

Figure 2: The Brussels model zoning system

The Brussels land-use model is based on a spatial input-output model, where economic production sectors include private local services, retail trade and business services (not allowed to locate on industrial land), as endogenous sectors, and agriculture, industry, heavy tertiary, Belgian public administration, international public administration (EC, NATO), public local services, business services (allowed to locate on industrial land), and teaching sector (primary, secondary and high education), as exogenous sectors.

Household categories (classified according to the characteristics of the household's head) consist of white collars (families or one person households), blue collars (families or one person households), non-active people, people over 65 and students living in a campus (for some of these, further distinctions are made according to the number of persons in the household). Land categories consist of 3 types: low and high density residential land, and mixed economic activities land.

The interrelations between the different factors are characterised by the coefficients of an input - output matrix that were derived from the national census and national surveys on labour force and on household expenses. These coefficients are elastic for land consumed by economic sectors and households .

Transport supply is represented by a single integrated multimodal network, whose details are adjusted to the scale of the zones under consideration. The modelled transport network consists of the primary road network (ca. 1300 links), the railway network serving the study area (ca. 500 links), the metro and pre-metro (i.e. tramway in tunnels) networks (ca. 100 links), representing 85 different link types. Buses are not modelled explicitly: they are represented by links gathering zone centroïds to railways or metro stations. However, separate bus links have been designed for the express buses running mainly in segregated lanes, modelled in several scenarios.

The Brussels model considers passenger transport only. The multi-path search is based on a multimodal shortest path search procedure (i.e. a path between a given O-D pair may include several modes), and the assignment of demand on the paths is based on a conventional multinomial logit procedure based on the path *generalised cost*, considering travel time and cost. Available modes in the reference scenario are car (with a distinction between single-occupancy car and high-occupancy car), metro and pre-metro (i.e. tramways in tunnels), train, REN and express buses.

The demand for travel is represented by a set of O-D matrices of flow volumes in the morning peak hour (7h00 - 9h00): high/medium income home-to-work trips, low income home-to-work trips, non-regular trips (i.e. other than home-to-work or home-to-school), as endogenous matrices, and home-to-school trips, commuting from outside the study area and transit trips, as exogenous matrices.

1.3 Calibration of the Brussels model

In the Brussels case, the model has been calibrated on situation 2001 and the reference scenario at horizon 2021 (without the REN) was built exogeneously (outside the model). The 2001 situation has been calibrated with observed data whose sources were:

- for transport data: the Regional Mobility Plan of Brussels (IRIS1 1990-1996 – STRATEC), road countings carried out in 1997 by STRATEC (on a cordon around the city) and the National Survey on the mobility of the households (1998-1999 – Facultés universitaires Notre-Dame de la Paix de Namur);
- for land-use data: the national census of the population and the residences (1991, INS, National Institute of Statistics), the register of the population (2001), a specific analysis of the ONSS (national office of social security) statistics, the Survey on labour force (INS, 1993 and 1999), statistics of the Ministry of National Education and of University foundation, the survey on the household's budget (INS 1995), the register statistics on buildings giving data on the land surfaces (1991 and 2000), and finally the housing prices of the STADIM data (1991 and 2000).

The 2021 reference scenario was also built up exogenously[2], on the basis of various sources and data mentioned previously. These included the recent socio-economic tendencies observed in the study area between 1991 and 2001, demographic forecasts set up by the National Institute of Statistics, macro-economic forecasts set up by the National Planning Office, and the strategic planning objectives expressed in the Master Plans of the 3 Regions (Brussels-Capital Region, Flemish Region, Walloon Region), especially with regard to the spatial structure.

---

[2] The model doesn't provide the situation at horizon 2021 starting from the base year 2001. This results partly from the fact that the model is rather complex (7 household segments, 13 activity sectors) and that the part of endogenous actors is high (in the Brussels model, 72 % of the total number of households and 45 % of the total employment are endogenous, i.e. their location is determined by the model).

The Brussels model reference scenario 2021 (002B) includes the project of Regional Express Railway Network (REN), which will provide high quality, rapid and frequent train services between the periphery and the central area. Other transport investments, such as 19 new lines of express buses on radial highways giving access to the central area, are not included in the reference scenario but are tested in some of the accompanying measures, such as the so-called “local investment plans” (711B) policy.

Most of the policies were tested on this 2021 reference scenario (002B, without population/employment growth). Some policies were tested on the third reference scenario (003B) constituting the local investment plan (711B). Results are provided for the 2021 horizon, in comparison with the adequate reference.

2. Description of the model of Helsinki

2.1 The MEPLAN framework

The Helsinki region case study was carried out using the existing MEPLAN model application of the region (see Echenique et al 1995 and Moilanen, 2000). MEPLAN is a comprehensive land-use and transport interaction modelling package that can represent strategic multi-modal networks/services and estimate transport demand based on the spatial economic interactions between the households, employment and land use (see *e.g*. Echenique, 1994; Williams, 1994; Harris, 1996). The modelling process follows closely the TRANUS process (see the section on TRANUS).

The urban model applications of the MEPLAN framework follow a traditional four-step transport model supplemented with a land-use location model. The various overall phases (from demand to supply) the model predicts for a given period are as follows:

(i) The location of the households and firms (employment);

(ii) The generation of trips from the interactions between households and employment;

(iii) The distribution of the trips between zones in the area;

(iv) The mode split of the trips into car, public transport and slow modes trips;

(v) The assignment of the vehicles on the transport networks.

The process modelling the economic interactions and socio-economic characteristics of the region of step (i) is based on a spatial input-output framework for endogenous employment and population that also have elastic (Stone-Geary) consumption functions. The chain of production and consumption is started based on the exports and other exogenous employment and inactive (non-working) households. Various constraints (e.g. rents based on available floorspace) increasing the costs of location affect this process in addition to the accessibilities due to transport system demand and supply characteristics (steps ii-v above).

The econometric models utilised are of the stochastic discrete choice type with a nested logit formulation. The hierarchical structure is adopted extensively in the transport model from trip distribution to modal split and generation of trip matrices (see McFadden 1978). This gives a strong theoretical foundation of utility maximising and also leads to a consistent evaluation based on consumer surplus calculation considerations inherent in economic welfare theory (Williams 1977).

2.2 The design of the Helsinki land use/transport model

The land use/transport model for the Helsinki region has been developed in several phases. The current updated model is designed for carrying out practical tests of the transport and land-use policy proposals in the area. The structure follows a traditional (extended) four-step transport modelling approach described above. The transport sub-model simulates both peak and inter-peak travel conditions using all modes (car, public transport and low modes) for 15 separate trip types (purpose/period/SEG) that had different characteristics in the calibration of the model.

The trip generation is modelled in the demand model with a land-use generation and location process of the economic interactions in the region, which is affected by the characteristics of the transport conditions presenting a given modelled 5-year period. The demand model encompasses various types of households and employment sectors that are in economic interaction with each other (through an input-output framework), and the floorspace that they use for locating in a zone. The lack of floorspace will turn up as higher rents affecting in turn the location of the households and employment in addition to transport accessibilities and other consumption costs in a zone.

The study area consists of 81 zones, and covers an area of 14,400 square kilometres that includes not only the Helsinki Metropolitan Area but also the surrounding cities. The model area includes the cities of Helsinki, Espoo, Vantaa and Kauniainen, which form the Helsinki metropolitan area. In addition to the metropolitan area, also the surrounding region is included in the model (the provinces of Uusimaa and Itä-Uusimaa as well as major parts of the provinces of Kanta-Häme, Päijät-Häme and Pirkanmaa). The model has therefore some characteristics of an inter-urban model.

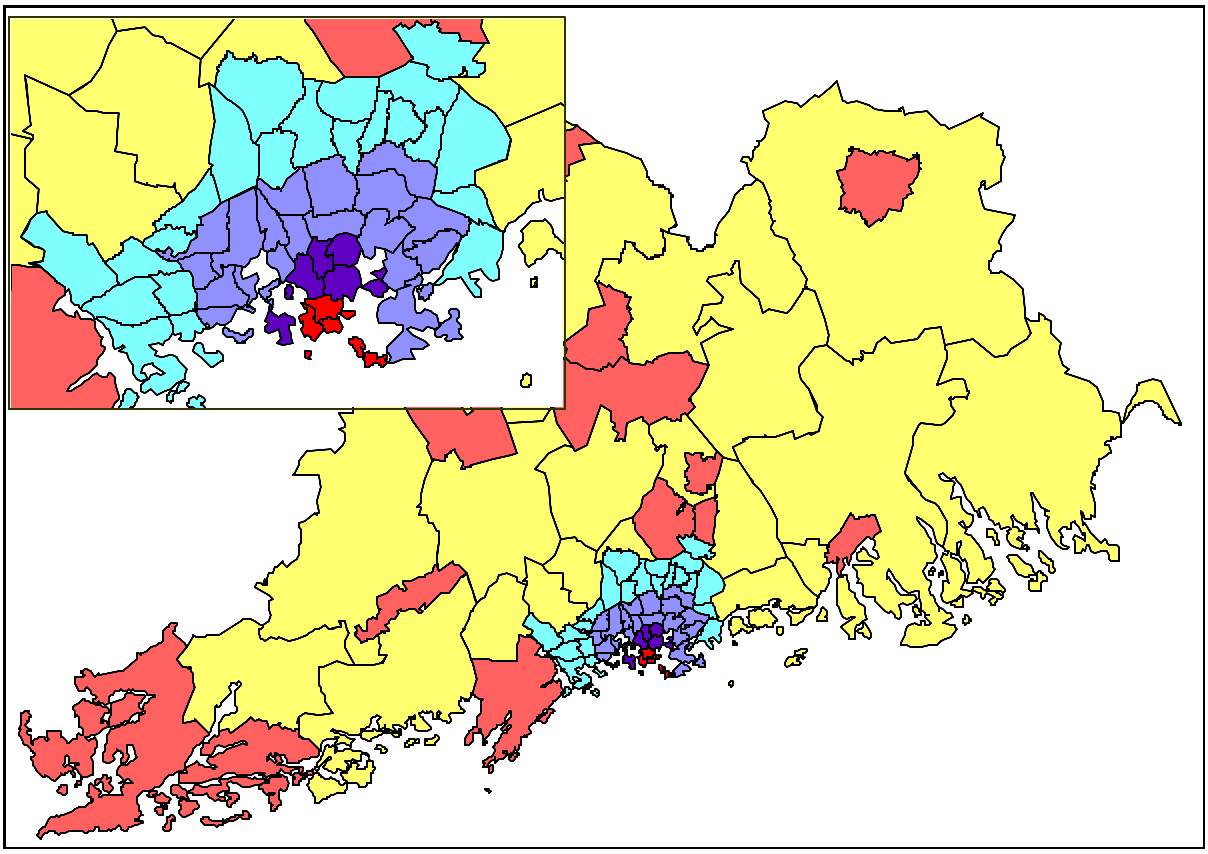

Figure 3: The study area and the super-zone definitions used in the analysis of sprawl : Helsinki centre (red), Inner Helsinki Metropolitan Area (HMA) (dark blue), outer HMA (mid-blue), HMA suburbs (light-blue), other urban conurbations outside HMA (orange) and rural municipalities (yellow).

Model factors are divided into three broad categories: employment factors (agriculture, industry, construction, wholesale, retail sale, private and public services), household categories (according to the social economic characteristics of the household head) and building stock (housing and employment floorspace) that is regulated in the land use model affecting the location of the new stock. Households in each income group are further divided into two types: active (i.e. working) and inactive (non-working). For inactive households, accessibility to work does not affect their location (like rents and other consumption costs).

2.3 Calibration of the Helsinki model

The calibration of the model is based on a wide set of census and other zone-based socio-economic and transport data supplied by the Finnish authorities. The household consumption (which determines for example the home based shopping trips) was modelled using a standard household expenditure survey by Statistics Finland as a source. The connection from employment to households was in turn developed based on a sample of anonymised records from the Finnish census. The design of the input-output framework in the model of Helsinki Region did benefit from a regional input-output table that was a constructed in a separate thesis study. It enabled the model to represent the whole economic structure of the modelled region with intermediate demand and good estimates of the exports and final consumption.

The network model of the multi-modal transport system is based on the Helsinki Metropolitan Area Council data. The road network includes the major arteries and the minor streets. The road network outside the Metropolitan area is a more coarse description of the connections between municipalities based on Finnish Road Administration data. The flow-delay functions have been calibrated to match the observed speed during periods of congestion. For public transport a 595-zone network of all bus and rail services within the Metropolitan area is used. The description of services includes the lines, stops, speeds/times, headways and the type of service (rail/metro/tram/bus). A simple distance-based (mostly intrazonal) network is used for slow modes.

For the region outside the Metropolitan area a coarser model is used based on the services between the municipalities and the Helsinki Metropolitan Area. The route assignment takes into account the in-vehicle, waiting, access and interchange times that are weighted according to the perceived time/inconvenience by the traveller. The disutility of travelling is also dependent on the type of service. The model has been calibrated to roughly match the observed ridership on the lines.

The assigned travel times and costs between origin and destination zones are used in the travel demand model that estimates the modal split of the trips generated by the land-use model according to the theory described above. The level of overall demand between zones is based on the input-output (e.g. working) relationships in the land use model.

The land-use model has been calibrated to model the location of the 1.6 million inhabitants (2000) in the area (out of approximately 5 million overall in Finland). The first stage has the objective to calibrate the model parameters in such a way as to reproduce the situation in the base year, whereas the following stages for each 5-year period following it aim at forecasting the interaction of transport demand and supply at the Master Plan until the final horizon year of 2020.

3. Description of the model of Stuttgart

3.1 The STASA modelling framework

The commuter flows are modelled via the master equation framework. In order to analyse both inter- and intra-regional flows the STASA-transport/land-use model (Weidlich/Haag 1988; BMBVW 1999; Weidlich/Haag 2000; Haag 2001, Binder, Haag, Rabino 2003) had to be modified. In the following, a rather short description of the general modelling framework is presented. Especially the differences with other "integrated approaches" should become obvious.

Investments into the transport sector and communication networks improve the accessibility and attractiveness of suburban areas. This may lead to a redistribution of migration flows and traffic flows and is discussed as one possible reason for urban sprawl. The quantitative treatment of those nested processes of the different subsystems (transport-, population-, communication-subsystem) and its interactions require an integrated modelling.

On the one hand the dynamics on the macrolevel - i.e. the development of the traffic subsystem and of the urban/regional subsystem - is determined by the behaviour of the individuals on the micro-level. On the other hand "attractivity" differences between the spatial units (traffic cells), which depend on the macro-variables, influence the decisions of the individuals as well. Apart from rational motives of the actors several elements of uncertainty, e.g. irrational behaviour as a result of insufficient information, have to be taken into account. Hence, the description of decision processes is based on a stochastical and dynamical decision model within the master equation approach.

The traffic subsystem as well as the urban/regional subsystem form a complex intertwined system. Its dynamics take place on different time scales but are modelled making use of the same principles:

a) The daily flows of traffic in the region of Stuttgart are the result of very quick decision processes of the actors to realize a trip between two traffic cells (origin - destination) with a special purpose. Decision processes for a certain destination, the moment of the setting out, the mode of transportation, the choice of the route etc. take place on a very short time scale.

b) The development of the urban/regional subsystem (e.g. spatial population distribution) is a process on a long-term time scale. The population distribution changes because of migration acts of individuals. The equations of motion which describe the migratory behaviour contain transition rates, i.e. migration flows between the cells. These flows depend on accessibility measures (coupling to the transport subsystem) and "attractivity" differences as a result of different regional advantages.

The total number of in- and out-commuters of communities has increased steadily in Germany for the last twenty years. In particular, this trend appears for far distance commuting but nevertheless the willingness to commute diminishes sharply at a time distances larger than approximately 45 minutes (Johansson et al 2001). Commuting offers the possibility to choose a new home by retaining the workplace and vice versa. This is often combined with the acquisition of real estate or the improvement of working conditions by retention of the place of work without simultaneously losing the advantages of an existing residence (Steierwald/Kühne 1993).

The organisation of the traffic network is of crucial importance for commuting, and thus for urban sprawl, because the decision to commute depends among others on accessibility measures. Commuter flows represent an essential part of traffic flows in the rush hours. Considering this, it is obvious that changes in the socio-economic environment and investments in the traffic infrastructure has an impact on the number of commuter and distance of commuter flows, with feedback effects to the housing market.

The imbalance between labour demand and labour supply on the level of communities is partially compensated by the commuter dynamics. The development of regional employment, the regional gross wage payment and parts of the community revenues are strongly dependent on the spatial distribution of commuter flows (Binder/Haag/Koller 2001).

Changes in the location of population and workplaces can be classified as a slow adjustment process. This means that commuting patterns should be expected to adjust itself to new conditions on a much faster time scale than household location pattern. Moreover econometric studies indicate that the location of firms adjusts faster than the location of households, and that workplaces tend to follow households (Mills/Carlino 1989, Holmberg et. al.2001, Oppenheim 1995, Schnabel/Lohse 1997, Johansson 2001).

The "empirical data base" consits of traffic flows (all modes) between the traffic cells, as well as population numbers and migration flows on a yearly base between all cells, and data on the traffic network.

These data base is used to estimate the system parameters of the transport and urban/regional subsystems, e.g. "transport attractivities" and "distance" (or resistance) parameters for the transport subsystem and "regional attractivities" for the urban/regional subsystem. In a further step these "attractivities" have been connected with appropriate macro-variables (key-variables) making use of a multiple regression.

The calibrated integrated model was applied to the region of Stuttgart (Haag/Binder 2001).

3.2 The design of the Stuttgart land use/transport model

The population distribution is denoted by $\vec{n} = \{n_1, ..., n_i, ..., n_L\}$, where n_i is the number of individuals (households, persons) living in community *i*. n_i will be modified by the decisions of the people to commute between community *i* and any one of the other communities. Therefore, the population distribution $\vec{n}$ is connected via commuter related activities or migration events with individual decision processes.

Let $P(\vec{n},t)$ be the configuration probability to find a certain population distribution $\vec{n}$ at time *t*, taking into account the complicated interactions of those agents. Of course this probability $P(\vec{n},t)$ must satisfy the normalisation condition

$$\sum_{\vec{n}} P(\vec{n},t) = 1 \qquad (1)$$

where the sum extends over all possible population configurations $\vec{n}$.

The temporal evolution of the probability distribution $P(\vec{n},t)$ can be described by the master equation (Weidlich/Haag 1983; Haag 1989)

$$\frac{d}{dt}P(\vec{n},t) = \sum_{\vec{k}} F_t(\vec{n},\vec{n}+\vec{k})P(\vec{n}+\vec{k},t) - \sum_{\vec{k}} F_t(\vec{n}+\vec{k},\vec{n})P(\vec{n},t) \qquad (2)$$

where the sum on the right hand side extends over all $\vec{k}$ with non vanishing configurational transition rates $F_t(\vec{n}+\vec{k},\vec{n})$ and $F_t(\vec{n},\vec{n}+\vec{k})$. Hereby the transition rate $F_t(\vec{n}+\vec{k},\vec{n})$ (transition probability per unit of time) specifies the transition from any population distribution $\vec{n}$ to a neighbouring distribution $\vec{n}+\vec{k}$.

The master equation (2) has a very direct and intuitively appealing interpretation. The change in time of the configuration probability $\frac{dP(\vec{n},t)}{dt}$ is due to two effects of opposite direction: first to the probability flux from all neighbouring configurations $\vec{n}+\vec{k}$ into the considered configuration $\vec{n}$ namely $\sum_{\vec{k}} F_t(\vec{n},\vec{n}+\vec{k})\,P(\vec{n}+\vec{k},t)$ and second to the probability flux out of the configuration $\vec{n}$ into all neighbouring configurations $\vec{n}+\vec{k}$, namely $\sum_{\vec{k}} F_t(\vec{n}+\vec{k},\vec{n})\,P(\vec{n},t)$. Consequently, the master equation represents a balance equation for probability fluxes. The transition rates in the master equation are directly associated with the evolution of the conditional probability.

The transition rate $F_t(\vec{n}+\vec{k},\vec{n})$ from the population distribution $\vec{n}$ to the neighbouring distribution $\vec{n}+\vec{k}$ is the sum of all the contributions $F_{ij}^{a}(\vec{n}+\vec{k},\vec{n})$:

$$F_t(\vec{n}+\vec{k},\vec{n}) = \sum_{a=1}^{A}\sum_{i,j=1}^{L} F_{ij}^{a}(\vec{n}+\vec{k},\vec{n}) \tag{3}$$

where $F_{ij}^{a}(\vec{n}+\vec{k},\vec{n})$ indicates the number of commuter trips[3] (transitions) between the communities $i \to j$ for a member of subpopulation α. The explicit dependence of the individual terms on $\vec{n}$ indicate that all contributions related to a change of the population distribution $\vec{n} \to \vec{n}+\vec{k}$ have been summed up. In this way a summation of all such terms yields the total transition rate.

In the next step the transition rates have to be specified for the decision process to commute (Fischer et. al. 1988, Weidlich/Haag 1988). If $n_i(t)$ persons are at time *t* in community *i*, the "probability to commute" to another community will be proportional $n_i(t)$. In this way the number of trips between *i* and *j* is given by

$$F_{ij}^{a}(\vec{n}+\vec{k},\vec{n}) = n_i(t)\cdot p_{ij}^{a}(\vec{x};t), \tag{4}$$

where $p_{ij}^{a}(\vec{n},\vec{x})$ is the individual transition rate from *i* to *j* for a member of the subpopulation a, $\vec{k}=\{0,\dots,\ 1_j,\dots,\ 0,\dots,\ (-1)_i,\dots,\ 0,\dots\}$ and $F_{ij}^{a}(\vec{n}+\vec{k},\vec{n})=0$ for all other $\vec{k}$. Of course, this transition rate depends among others on the explicit spatial distribution of the population $\vec{n}$ and specific characteristics $\vec{x}$ of the communities, e.g. labour demand, labour supply, housing market, accessibility measures, specific location factors, services available for companies and households as well as leisure facilities (Domencich/McFadden 1975, Pumain/Saint-Julien 1989).

It has been tested, that mainly three sets of indicators are of importance for the transition rate to commute $p_{ij}^{a}(\vec{x};t)$:

- attractiveness indicators $u_i^{a}(\vec{x};t)$ of the particular community *i* for the subpopulation a, which depend across-the-board on the distribution of labour demand and supply of the communes. It is commonly known that individuals (commuters) compare the

---

[3] If panel data are available on the commuter-decision behaviour of the different agents of the system (micro-level), the configurational transition rates can directly be calculated via $F_{ij}(\vec{n}+\vec{k},\vec{n})=\sum_{l\in\Gamma_i} p_{ij}^{(l)}(\vec{n},\vec{k}_l)$, where one has to sum up over all individual trips of all commuter from community *i* to community *j*. This procedure is however very extensive, because of the required immense data base (Courgeau 1985).

attractiveness of the communities with respect to certain characteristics such as working and housing conditions. The probability not to work at the place of home *i* instead to work in community *j* (place of work) increases with increasing differences $(u_j^a(\vec{x};t) - u_i^a(\vec{x};t)) > 0$ of attractiveness (Fechner 1877; Weber 1909; Fazio/Zanna 1981). Without any loss of generality the attractiveness can be scaled

$$\sum_{i=1}^{L} u_i(\vec{x};t) = 0 \cdot$$

- resistance function $g_{ij}^a(t_{ij}, v_i^a; t)$, representing the spatial interrelation (accessibility) of communities, depending on travel time t_{ij}, as well as regional shadow-costs $v_i^a(\vec{x};t)$. The resistance function is modelled via

$$g^a(t_{ij}, v_i^a; t) = \exp\left(\frac{-\boldsymbol{b}^a(t)\, t_{ij}}{1 + g^a(t)\, t_{ij}} - v_i^a(\vec{x};t) \right) \tag{5}$$

with deterrence parameters b^a, g^a and shadow-costs $v_i^a(\vec{x};t)$. The shadow-costs $v_i^a(\vec{x};t)$ take into account the heterogeneity of the communities. Shadow costs act as barriers and reduce the attractiveness of a region for commuters. By definition, the shadow-costs have to fulfil the constraint

$$\sum_{i=1}^{L} v_i(\vec{x};t) = 0. \tag{6}$$

- a time-dependent scaling parameter $e^a(t)$ which correlates with the global mobility to commute.

This leads to the following commuter trip model (trip distribution):

$$\begin{aligned} F_{ij}^a(\vec{n},\vec{x};t) &= n_i(t)\, p_{ij}^a(\vec{x},t) \\ &= n_i(t) \cdot e^a(t) \cdot g_{ij}^a(t_{ij}, v_i^a; t) \cdot \exp(u_j^a(\vec{x};t) - u_i^a(\vec{x};t)) \end{aligned} \tag{7}$$

where $F_{ij}^a(\vec{n},\vec{x};t)$ indicates the number of commuter trips (transitions) between the communities $i \rightarrow j$ for a member of subpopulation α.

The probability distribution $P(\vec{n},t)$ contains a huge amount of information compared with the empirical information (data base). Therefore, a less comprehensive description in terms of mean values is adequate. The mean population number in community *i* at time *t* defined as

$$\overline{n}_i(t) = \sum_{\vec{n}} n_i\, P(\vec{n},t)\,. \tag{8}$$

It is possible to derive equations of motion for the mean values drectly from the master equation. For this purpose the master equation is multiplied by $\vec{n}$ from the left and summed up via all states $\vec{n}$. However, the resulting equations are not yet self-contained, since the determination of the right hand side (rhs) requires the knowledge of the probability distribution $P(\vec{n},t)$. However, if one assumes that the probability distribution is a well behaved, sharply peaked uni-modal distribution quasi-closed approximate mean value equations can be derived:

$$\begin{aligned}\frac{d\bar{n}_i(t)}{dt} &= \sum_{a=1}^{A}\sum_{j=1}^{L} F_{ji}^{a}(\vec{n},\vec{x};t) - \sum_{a=1}^{A}\sum_{j=1}^{L} F_{ij}^{a}(\vec{n},\vec{x};t) \\ &= \sum_{a=1}^{A}\sum_{j=1}^{L} \bar{n}_j(t)\, p_{ji}^{a}(\vec{n},\vec{x};t) - \sum_{a=1}^{A}\sum_{j=1}^{L} \bar{n}_i(t)\, p_{ij}^{a}(\vec{n},\vec{x};t) \\ &= E_i^W(t) - E_i^H(t) = NC_i(t)\end{aligned} \tag{9}$$

Therefore the master equation provides the link between decisions to commute on the micro-level and the commuter flows on the macro-level. The dynamics of the mean population number $\bar{n}_i(t)$ of community *i* can be calculated on the basis of the commuter flows $F_{ij}^a(\vec{n},\vec{x};t)$ and $F_{ji}^a(\vec{n},\vec{x};t)$ between the different communities *i* and *j*. The first sum on the rhs of (13) equals the employees at the place of home $E_i^H(t)$, the second sum equals the employees at the place of work $E_i^W(t)$. Therefore the dynamics of the population redistribution on the macro-level depends on the development of the net commuters $NC_i(t)$.

Long-term effects, e.g. structural development effects have to be considered as well. It is reasonable to assume that the attractiveness $u_i^a(\vec{x},t)$ and shadow-costs $v_i^a(\vec{x};t)$ of a community depend on a set of socio-economic variables $\vec{x}$ (among other things also on the population distribution $n_i(t)$). Therefore, the impact of commuting on the population redistribution is of importance (Figure 4). Depending on the initial conditions, such as the distribution of population at a given time and the further system parameters, the non-linear dynamics lead to self-organised commuter flow pattern (Nijkamp/Reggiani 1992; Goodwin 1994).

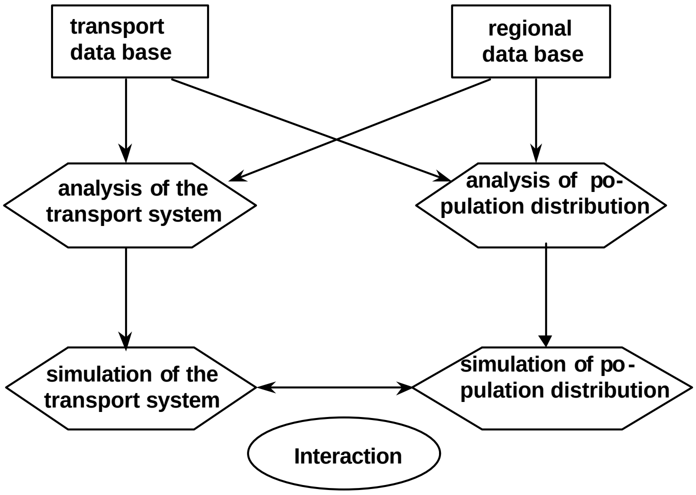

Figure 4: Principle structure of the nested transport and urban model

3.3 Calibration of the Stuttgart model

The system parameters (11), such as mobility $e^a(t)$, attractiveness $u_i^a(\vec{x},t)$, shadow-costs $v_i^a(\vec{x};t)$ and the resistance function parameters $b^a(t)$ and $g^a(t)$ can directly be linked to the empirical (statistical registered) commuter flow matrices $F_{ij}^{a\ emp}(t)$ (index *emp*), and the population numbers $n_i^{emp}(t)$, respectively. The minimisation of the functional (entropy-estimation)[4] (Wilson 1970, 1981)

$$G(u_i, v_i, e, b^a, g^a; t) = MIN\left(\sum_{a=1}^{A} \sum_{i,j=1}^{L} F_{ij}^{a\ emp}(t) \cdot \ln\left(\frac{F_{ij}^{a\ emp}(t)}{F_{ij}^{a}(t)}\right)\right)$$
$$\approx MIN\left(\sum_{a=1}^{A} \sum_{i,j=1}^{L} \left(\frac{F_{ij}^{a\ emp}(t) - F_{ij}^{a}(t)}{F_{ij}^{a}(t)}\right)^2\right) \qquad (10)$$

with the constraint

$$\sum_{a=1}^{A} \sum_{i,j=1}^{L} F_{ij}^{a\ emp}(t) = \sum_{a=1}^{A} \sum_{i,j=1}^{L} F_{ij}^{a}(t) \qquad (11)$$

enables one to calculate an optimal set of system parameters ($u_i^a(\vec{x}, t)$, $v_i^a(\vec{x}; t)$, $e^a(t)$, $b^a(t)$, $g^a(t)$).

In a second step, the estimated attractiveness and shadow-costs are linked to particular location factors (key-factors) $x_i^n, n = 1, \ldots$ gained from two different fields: from a class of the so-called synergy variables, describing general group effects (positive and negative network externalities), and from a sequence of potential explanatory indicators, e.g. number of available jobs, the number of vacant dwellings, regional income per capita, shop distribution and other local infrastructure depending factors. The set of explanatory variables and its corresponding elasticity's are determined via a multiple regression analysis:

$$u_i^a(\vec{n}, \vec{x}) = \sum_n a_n^a x_i^n \qquad v_i^a(\vec{n}, \vec{x}) = \sum_n b_n^a x_i^n \qquad (12)$$

The elasticity's a_n^a, b_n^a assigned to the socio-economic variables x_i^n are dimensionless numbers and indicate the influence of the independent variables on the dependent variable. The selection of relevant indicators is performed using appropriate statistical characteristics (T-values, other significance tests).

The results of the regression for the attractiveness (shadow costs) explains about 83% (75%) of the data variation (Haag, Binder 2001b). Because commuting describes trips from home to work and vice versa, the explaining variables are related to the following :

- Labour market (distribution of work places, wages,…)
- Housing market (distribution of apartments, dwellings, houses, price of land, rent level,…)
- Accessibility (travel time, travel costs, parking possibilities,…)

---

[4] Using test series it was determined that the least-square-estimation represent the single flows $F_{ij}(t)$ much better then the entropy procedure estimation. On the other hand, the specific entropy-estimation takes into account that the origin and destination flows of each community are equal to the employees at the place of home (empirical origin flows) and employee at the place of work (empirical destination flows). This property of the entropy procedure can also be derived analytically by geometric programming (Kádas, Klafszky 1967).

- Other specific indicators (availability of different services, neighbourhood, environment, recreation possibilities,…)