9th NECTAR CONFERENCE

Porto, Portugal may 9-12, 2007

Validation of a microscopic model for firm location in LUTI models

Michiel de Bok[1]

Department of Transport and Planning, Faculty of Civil Engineering and Geosciences

Delft University of Technology, PO Box 5048, NL-2600 GA Delft, The Netherlands,

p: +31 15 2784977, f: +31 15 2783179, m.a.debok@citg.tudelft.nl

Abstract

This article presents a micro simulation model for the location of individual firms in a disaggregate urban environment. The model applies spatially detailed location factors that measure proximity to transport infrastructure, accessibility, and the composition of the firm population. Objective of the model is to quantify effects of spatial and transport planning scenarios on firm dynamics and mobility. The presented Spatial Firm-demographic Micro-simulation (SFM) model simulates transitions and events in the firm population, including firm relocations, firm growth, firm dissolution and firm formation. Accessibility is measured with proximity to infrastructure access points and logsums of business and commuting trips from a transport model. From urban economic literature, additional measures for 'Marshallian' and 'Jacobian' externalities are specified that measure specialisation and diversity in the firm population, integrated with travel time range bands.

The SFM-model is applied in a case study for the province of South Holland in The Netherlands. The area contains a firm population of about 90.000 firms that are distributed across 70.000 6-digit zip code locations. The model is estimated on an extensive longitudinal dataset with the full firm population from 1988 to 1997. This article presents the validation of the model on similar data for the successive period from 1997 to 2004. The validation shows that the SFM-model can provide reliable estimates of firm demographic development at neighbourhood level. At micro level, the results show a dynamic pattern that seems representative for coincidence in the behaviour of individual firms. The microscopic detail of the model allows the inclusion of advanced theories from urban economics and thus provides a more advanced framework for modelling firm location in integrated land use transport interaction models.

KEYWORDS: Land use transport interaction, accessibility, agglomeration, micro simulation, firm demography

[1] Present adress: Significance, Schipholweg 13-15, 2316 XB Leiden, debok@significance.nl

p +31-71-5253-881, f +31-71-5253-801

1 Introduction

The construction of significant new transport infrastructure projects, such as the A4-extension from Rijswijk-Rotterdam, has long term spatial economic effects that are difficult to predict. Integrated land use and transport models make predictions of such effects and are becoming increasingly disaggregate, both in the spatial environment and the behaviour of agents (Timmermans, 2003). Motivation for this research is to provide an improved methodology for the projection of distributive effects of infrastructure and spatial planning scenarios on firm location.

This paper presents a micro simulation approach that quantifies the effects of different spatial and transport planning scenarios on the firm population and mobility. The presented The Spatial Firm demographic Micro simulation (SFM) model is founded on firm demography (Birch, 1979; Carroll and Hannan, 2000). Similar to the approach by Van Wissen (2000), the SFM-model simulates transitions and events in the firm population. These include firm relocation, firm growth, firm dissolution and firm formation. In each time step of a model run, the firm population is processed through the separate components of the firm demographic micro simulation. The quality of locations is accounted for explicitly through the application of advanced accessibility and agglomeration measures. Accessibility is measured with proximity to infrastructure access points and logsums of business and commuting trips from a transport model. Agglomeration measures are based on theories from urban economics and integrate travel time range bands into measures for specialisation and diversification.

The chosen firm demographic approach can improve the firm behaviour in urban simulation models for three reasons. First, it represents firm specific behaviour and heterogeneity in responses. Firm behaviour is influenced both by firm internal processes and attributes of the firm's location. For example, the most important reason for firms to relocate is firm growth (Brouwer et al., 2002; Louw, 1996). Second, it allows distinctive accessibility measures as explanatory variables for each event, such as firm relocation, firm performance, formation or firm dissolution. For instance, some firms perform better in the proximity of motorway onramps (Hilbers et al., 1994), while in a relocation decision, accessibility is evaluated in a different way. A third advantage is the possibility to account for path dependency between events. For example, firm growth – e.g., triggered by a new motorway opening – might induce a firm relocation in the following years. By keeping record of the developments at the individual firm level, the life cycle of a firm and the causality between subsequent events can be modelled endogenously.

To test the assumptions underlying the presented approach, the model is applied to a case study for the province of South Holland in The Netherlands. The area contains a firm population of approximately 90.000 firms that are distributed across 70.000 6-digit zip code locations. The SFM model has been developed in three stages: model design, model estimation and validation. The model is estimated on an extensive longitudinal dataset with the full firm population from 1988 to 1997. For a more elaborate discussion on the estimation of the model that is validated in this article see De Bok (2007). The calibration of a previous model specification has been published in De Bok and Bliemer (2006). This paper discusses model validation, that has been performed on data for the successive period from 1997 to 2004.

2 The spatial firm demographic simulation model

2.1 Structure of the model

This article presents the design of a Spatial Firm demographic Micro simulation (SFM) that simulates the dynamics within a firm population in a disaggregate urban environment. Their behaviour is the outcome of their preferences and lifecycle and exogenous spatial and economic planning scenarios. Firm demographic micro simulation has a long tradition that started with Birch (1979), and many theories underlying the micro behaviour of firms have been described in the literature on demography of organisations. For a contemporary overview see Carrol and Hannan (2000). Van Wissen developed a firm demographic micro simulation for a case study in The Netherlands previously (Van Wissen, 2000). The SFM-model presented in this article is similar to a large extend but it incorporates the spatial dimension explicitly. This spatial dimension is represented with advanced accessibility and agglomeration measures.

The SFM model describes the change in the state of individual firms that are beneath aggregate economic developments. These transitions are the result of a number of firm demographic events that can occur to each individual firm, such as firm relocation, firm growth, firm dissolution, firm formation, or firm dissolution. Key principle is that these events are influenced by firm internal factors (e.g. size, age) or by location factors (available locations, accessibility, agglomeration). These events are probabilistic in nature: they can or may not occur.

The causal structure is outlined in Figure 1. From the economic, planning and mobility scenario's at the top, the dependencies between elements are drawn. Bottom up, the approach simulates the firm population with distinctive sub populations (industry sectors) and individual firms. The actual developments are simulated at the individual firm level. The firm demographic events that are simulated include firm formation, firm growth, firm migration and firm dissolution. These events are first of all influenced by the life cycle of firms. This life cycle is captured in the state of the firm and its attributes (age, size) and through other firm demographic events. For instance, firm dissolution and firm migration can both be instigated by firm growth.

Additionally, accessibility and agglomeration are included as location factors in the firm demographic events. Accessibility is derived from a transport model. Agglomeration is derived from the (sub) firm population and travel times from the transport model.

The model represents the supply of industrial real estate at each firm location in order to account for constraints in location options explicitly. Each firms requires a corresponding firm location, in other words, an office related firm can only be located at a location that has the required amount of office real estate available. And during its existence a firm might relocate to another location in case of a firm migration. Changes in the real estate supply, as a result of new industrial or office sites, are included through exogenous plans.

The regional economic development is conditional to the micro developments that are simulated in the SFM model. Thus, individual firm developments are influenced by structural developments in the industry sector or other macro influences that are outside the scope of the analysis. The implementation of exogenous scenarios for the regional economic development implies that the model cannot account for generic effects of transport infrastructure. This is assumed to be acceptable. First of all, these generative effects are much less significant compared to distributive effects (Rietveld, 1994). Moreover, the scope of this urban simulation model is primary on (re) distributive effects of infrastructure investments (allocation of employment growth and relocation of firms).

The SFM model distinguishes three geographical scale levels: the research area, zones, and locations. The research area comprises of a region, province or state. Within this research area the microscopic developments in the firm population are simulated. Interaction with regions outside this research area takes place through exogenous scenarios. Zones are applied as an intermediate level to describe the characteristics of the urban environment around a location. Moreover, zones are required in order to enable interaction with a transport model. The finest geographical scale level that is applied is at the firm location itself. Because individual firms are simulated, their exact location is known, enabling the computation of detailed spatial characteristics, such as proximity to train stations or highway onramps.

2.2 Accessibility and agglomeration in the SFM model

The spatial environment of firm locations is an explicit and important dimension in the presented simulation approach. The quality of this spatial environment is described with accessibility and agglomeration attributes.

2.2.1 Transport based accessibility

The transport based accessibility measures specifically measure the attributes of the transport infrastructure and the trips that can be made with it. The quality of the transport based accessibility is measured both with distance to infrastructure access points and logsum accessibility measures. The distance to train stations or highway onramps express a specific transport infrastructure quality that is easily interpreted. Previous empirical findings suggested that proximity measures to infrastructure access points are significant location factors in the location preference of firms (De Bok and Sanders, 2005).

Logsum accessibility measures are less easily interpreted, but provide the most conclusive approach to measuring the valuation of all possibilities that can be reached from a location, taking into account individual preferences, the available modes of transport, the variation of travel times and travel costs over the day (Geurs and Van Wee, 2004; De Jong et al., 2005). These measures are well founded in micro economic theory. In this application the logsums for two trip purposes are assumed to be relevant: the logsum for (non-home based) business trips and the (reflected) logsum for commuting trips. First of all, the logsum for business trips are assumed to be a representative measure for customer and supplier accessibility. The commuting trips are used to measure labour market accessibility.

2.2.2 Agglomeration

In urban economic literature spatial externalities are identified as important factors to spatial economic development (for a contemporary overview see Rosenthal and strange, 2004). These spatial externalities are derived from input sharing, labour market pooling or knowledge transfers. The common accessibility measures cover a part of these externalities but additional information about the composition of the firm population is required to measure the full extend of spatial externalities. Therefore, the SFM model features agglomeration measures for diversification and specialisation.

Following the example of Rosenthal and Strange (2003) and Van Der Panne (2004), the agglomeration measures are computed for specific range bands. These range bands are derived from a transport model and explicitly account for the structuring impact of transport infrastructure. The level of agglomeration within each range band can be measured by analysing the level and composition of employment within the range band. A range band represents the area that can be reached within a specific range of travel times, for instance between 5 and 10 minutes. In this approach we take travel times to define our range bands instead of distance. In this way the effect of transportation developments on agglomeration is represented explicitly.

Concentration is measured as the representation of one industry within a specific travel range of a location relative to that industries share in the region. The measure is based on the commonly applied production specialisation index (PS), and is enhanced with a spatial dimension with range bands. For each location j the level of agglomeration is measured in specific range bands, R_{jb}, for a set of ranges b=0-7.5, 7.5-15. The level of agglomeration in each range band is derived from the level of employment in each industry sector in each range band. For location j the share of the employment in industry sector s in a range band R_{jb} from j is measured relative to the share of employment in that industry in the whole region. The production specialisation index for location *j* and range band R_{jb} becomes:

$$PS_{jsb} = \frac{E_{sR_{jb}} / \sum_{s} E_{sR_{jb}}}{\sum_{j} E_{sj} / \sum_{s} \sum_{j} E_{sj}}, \tag{1}$$

with $E_{sR_{jb}}$ as the employment in industry s, within range band R_{jb}. Figure 2 gives an example of a specialisation index for the business service sector in the range band 0-7.5 minutes.

Diversity externalities are measured with the similar range band concept as well. The common productivity diversity index (PD) that is computed for each range band, is based on the specification of Paci and Usai (1999). If the number of industry sectors is defined by *S*, and all industries are sorted in increasing order, the production diversity index PD_{jb} for location *j* and range band R_{jb} is defined as:

$$PD_{jb} = \frac{1}{(S-1)E_{SR_{jb}}} \sum_{s=1}^{S-1} E_{sR_{jb}}, \tag{2}$$

with $E_{SR_{jb}}$ as the employment in the largest industry within range band R_{jb}. The agglomeration attributes are visualised in Figure 2.

2.3 Firm demographic components

The firm demographic core of the SFM model consists of four components: firm migration, firm growth (either decline), firm formation and firm dissolution. Next, each component is specified in more detail.

2.3.1 Firm migration

With seven to eight percent of all firms relocating in a year (Pellenbarg, 1996), firm migration has a considerable influence on the location of the firm population. The move propensity of a firm is mainly determined by firm internal factors relating to the life-cycle of firms and to a lesser extent by site related factors (Louw, 1996; Van Dijk and Pellenbarg, 2000; Brouwer et al. 2002). The number of empirical studies in the literature is modest, so one general accepted approach to modelling the influence of transport infrastructure on firm location does not exist. Various studies report a significant influence of motorway proximity: Hilbers et al. (1994), Hunt (1997), Leitham et al. (2000), Holl (2004a, 2004b) or De Bok and Sanders (2005). In terms of urban development, this motorway orientation of economic activities has led to a suburbanisation pattern of economic activities, see Kawamura (2001), Shukla and Waddell (1991).

Similar to the simulation approach applied by Van Wissen (2000), firm migration is modelled as a stepwise procedure comprising of a joint decision to relocate, and for a new location. The relocation decision is determined by the satisfaction of a firm at its current location. Once the decision to relocate has been made, the firm will search for alternative locations. These are sampled from the available (unused) real estate supply. The probability that an alternative is being chosen depends on its attributes and the expected utility of this alternative.

For each simulation time interval t, let $P_{isj}^{M}(t)$ denote the unconditional migration probability that a firm i from industry sector s will relocate and choose location j as its new location. This probability is the product of the probability that the firm will relocate, $P_{is}^{M(1)}(t)$ and the probability that it chooses location j (out of a deterministic subset $L_{is}(t)$ with available alternatives for firm i at time t,), $P_{isj|L_{is}(t)}^{M(2)}(t)$:

$$P_{isj}^{M}(t) = P_{is}^{M(1)}(t) \cdot P_{isj|L_{is}(t)}^{M(2)}(t). \tag{3}$$

The first step in the migration procedure is to calculate the probability to relocate. This probability is described with a binary regression model. The probability of relocating for each time interval is determined by the firm characteristics and the attributes of the location at the current time instance:

$$P_{is}^{M(1)}(t) = \frac{1}{1+\exp[-(\beta_{s}^{M(1)} + \beta_{1}^{M(1)} g_{i}(t) + \beta_{2}^{M(1)} \delta_{i}(t) + \beta_{3}^{M(1)} Acc_{i}(t) + \beta_{4}^{M(1)} Agg_{i}(t))]}, \tag{4}$$

where $\beta_{s}^{M(1)}$ is an industry specific parameter, $g_{i}(t)$ is the growth rate of the firm at time t, $\delta_{i}(t)$ is a dummy variable for infrastructure proximity, and $Acc_{i}(t)$ and $Agg_{i}(t)$ are location attributes for accessibility or agglomeration. The parameters $\beta_{is}^{M(1)}$ and $\beta_{X}^{M(1)}$ are estimated. The probability to relocate is translated into a discrete relocation decision (“Yes” either “No”) through Monte Carlo simulation.

When a firm has decided to relocate, a deterministic choice set is sampled from the available and feasible real estate. Feasibility is determined by the size of the firms and the real estate object and the type of real estate. Each choice set contains maximum 20 alternatives.

The choice probability of each alternative is calculated with a spatial preference model in the form of a Competing Destination model (CD) model that accounts for the interdependencies between spatial alternatives (Fortheringham, 1983). By definition, the utility of an alternative consists of an observed component $V_{isj}(t)$ an unobserved (random) component and additionally a centrality measure $c_j(t)$, that accounts for spatial cluster membership of alternatives. The observed utility functions are industry specific and contain a number of accessibility and agglomeration attributes. The choice probabilities of each alternative in the choice context are defined as:

$$P^{M(2)}_{isj|L_{is}(t)}(t) = \frac{\exp(V_{isj}(t) + \theta^{M(2)}_s \ln c_j(t))}{\sum_{k \in L_{is}(t)} \exp(V_{isk}(t) + \theta^{M(2)}_s \ln c_k(t))}. \qquad (5)$$

The resulting choice probabilities are translated into a discrete outcome (location *j* from choice set $L_{is}(t)$) through Monte Carlo simulation.

The observed utility of each alternative, V_{isj}, is specified with an industry specific linear additive utility function. It includes the accessibility and agglomeration attributes, as well as the Euclidean distance between the original location and the location alternative, d_{ij}:

$$V_{isj}(t) = \beta^{M(2)}_{1s} d_{ij} + \beta^{M(2)}_{2s} \delta_i(t) + \beta^{M(2)}_{3s} Acc_i(t) + \beta^{M(2)}_{4s} Agg_i(t). \qquad (6)$$

The competing destinations model includes a centrality measure $c_j(t)$ that accounts explicitly for the interdependency between spatial alternatives. This measure recognises the more proximal alternatives are in space, the more likely they are to be substitutes for one another. Similar to Fotheringham (1983), a sum of weighed distances is calculated from one alternative to all others, where the weight is the size of each alternative:

$$c_j(t) = (1/(K(t)-1)\sum_{k \neq j} w_k(t)/d_{kj}(t)) \qquad (7)$$

with $K(t)$ as the number of available firm locations at time *t*, $d_{kj}(t)$ as the distance between alternative *k* and *j* and $w_k(t)$ as the size of alternative *k*. The size of an alternative is specified as the available floorspace or industrial area at a firm location at time *t*.

2.3.2 Firm growth

The scope of the presented model is primary on distributive effects of transport infrastructure; regional economic development is exogenous input. Rietveld (1994) has shown that for the case of a new motorway opening, the distributive impacts are far more substantial compared to the generative effects. It is stressed that our presented approach neglects the generative effects of transport infrastructure, but accounts for structural developments within the industry sectors through economic scenarios. A two step approach is applied in computing the expected firm size. First, a tentative firm size is estimated based on the firm- and location attributes. Next, this tentative firm size is corrected to fit the total employment in the firm sector to the exogenous regional totals.

First of all, firm performance is assumed to be influenced by the proximity of transport infrastructure. Hilbers et al. (1994) provide evidence for better performances of specific industry sectors in the proximity of motorway onramps. However, firm size is mostly explained by firm attributes, relating to the size and lifecycle of a firm. Following recent studies, such as Audretsch et al. (2002) and Van Wissen and Huisman (2003), an autoregressive model is applied that estimates the size of a firm, relative to the average firm size in the industry sector. The z-value of firm size, $z_{is}(t)$, is defined as the deviation of the log of the firm size from the average log of firm size:

$$z_{is}(t) = \log s_{is}^{*}(t) - \frac{\sum_{k=1}^{N_s(t)} \log s_{ks}(t)}{N_s(t)}, \tag{8}$$

where $s_{is}^{*}(t)$ is the tentative size of firm *i*, that belongs to sector *s* at time *t* and $N_s(t)$ is the number of firms in sector *s*. Following Audretsch et al. (2002) and Van Wissen and Huisman (2003), $z_{is}(t)$ is derived from the firm specific z-values for the two previous years and a location attribute. This first order autoregressive model includes a slope coefficient β_0^S and a first-order autocorrelation coefficient ρ_0^S:

$$\begin{aligned} z_{is}(t) = (\beta_{0s}^{S} + \rho_{0s}^{S}) z_{is}(t-1) + (-\beta_{0s}^{S}\rho_{0s}^{S}) z_{is}(t-2) + \\ + \beta_{1s}^{S}\delta_i(t) + \beta_{2s}^{S} Acc_i(t) + \beta_{3s}^{S} Agg_i(t) + \varepsilon_{is}(t), \end{aligned} \tag{9}$$

where $\delta_i(t)$ is a dummy variable for infrastructure proximity and $\varepsilon_{is}(t)$ is a stochastic disturbance term. The parameters β_{0s}^{S}, ρ_{0s}^{S}, β_{xs}^{S} and $\varepsilon_{is}(t)$ are industry specific and are estimated. The tentative firm size $s_{is}^{*}(t)$, can be derived from rewriting the previous two equations. Next, this tentative firm size is corrected with a sector specific regional balancing factor in order to fit the total employment to the regional economic scenario:

$$s_{is}(t) = s_{is}^{*}(t) \frac{E_s(t)}{E_s^{*}(t)}, \tag{10}$$

where $s_{is}(t)$ is the firm size at time *t*, $E_s(t)$ is the regional employment in industry sector *s*, and $E_s^*(t)$ is the initially estimated regional employment.

2.3.3 *Firm dissolution*

In firm demographic literature it can be found that firm dissolution is mainly determined by firm characteristics: size, sector, age and firm growth (Carroll and Hannan, 2000; Van Wissen, 2000; Ekamper, 1996). Clearly, larger firms have a higher probability of surviving a time interval. With respect to age, the 'survival of the fittest' is apparent among young firms: the dissolution probability is high among young firms. Even though no empirical evidence has been found for a relation between firm dissolution and accessibility, accessibility attributes will be included in the estimation of the dissolution probability.

The procedure includes a binary regression model to calculate the probability of dissolution, $P_{is}^D(t)$. This probability of dissolution is determined by firm- and location attributes:

$$P_{is}^D(t) = \frac{1}{1 + \exp[-(\beta_s^D + \beta_1^D g_i(t) + \beta_2^D \ln s_i(t) + \beta_3^D (1/a_i(t)) + \beta_4^D \delta_i(t) + \beta_5^D Acc_i(t) + \beta_6^D Agg_i(t))]}, \tag{11}$$

where β_s^D is an industry specific parameter, $g_i(t)$ is the growth rate of the firm, $s_i(t)$ the size of the firm, $a_i(t)$ the age of the firm and $\delta_i(t)$ a dummy variables for infrastructure proximity. β_{is}^D and all β_X^D parameters are estimated. During micro simulation the dissolution probability is determined for each firm. This dissolution probability is translated into a discrete dissolution event ("Yes" either "No") through Monte Carlo simulation. In case of a firm dissolution event, the status of the corresponding firm is set to 'Dissolved'.

2.3.4 *Firm formation*

Firm formation concerns a complex process of starting up a new firm. An important engine behind firm formation is the firm population itself (Van Wissen, 2000). An existing firm can induce a firm start up, for instance by splitting up or by starting a new branch. Another instigator for firm birth is the labour population: firms can also be formed by firm employees, school-leavers or an unemployed. The urban environment is also regarded to have an effect on firm formation. Van Oort et al. (1999) found evidence for a distinctive relationship between firm formation rates and the urban environment for different industry sectors. Firm formation in the non-basic sector seemed more likely to occur in the urban area, which is explained by the incubation and seedbed theory. Furthermore, it was found that sectoral diversification has a positive effect on the firm formation rate. Not many empirical examples have been found on the influence of infrastructure on the firm formation event. Positive exceptions are Holl (2004a) and Holl (2004b), with firm level data on the location of new manufacturing establishments in Spanish municipalities. Results show that new motorways affect the spatial distribution of formations of manufacturing establishments.

Firm formation is an exception to the micro simulation approach for it simulates a firm start up as sequential macro to micro steps. First the number of firm formations is determined by each industry sector. The number of firm births is derived from the firm population in the respective industry sector for the whole study area. Next, an initial firm size is randomly drawn from an observed distribution of firm sizes at start up. Finally the firm is allocated to a random location that is drawn from all available locations that are feasible as well.

3 Data and estimation of the SFM model

The research approach is applied to a case study for the South Holland region. This case study provides an empirical test bed to analyse the assumptions underlying the model and to gain more knowledge into the influence of accessibility and agglomeration on spatial firm dynamics.

3.1 Data

The model has been estimated on a longitudinal dataset, covering all developments in the firm population. This dataset has been constructed by linking the annual LISA datasets (National Information System of Employment) from 1988 to 2004. The model is estimated on period 1988 to 1997. This paper discusses model validation on the successive period: from 1997 to 2004. The observations include the following attributes for each firm: industry sector, size (in full time employment units), the change in size compared to previous year, the location (6 digit zip code) and dummy's for firm demographic events. The spatial detail of firm locations allows a detailed analysis of spatial attributes of each location. The firm population is segmented to 12 industry sectors.

The accessibility and agglomeration attributes are linked to the locations. The distance measures to highway onramps and trainstation are derived from the location of each firm (6 digit zip code) and a GIS analysis. The logsum accessibility attributes, and the travel time between zones in the study area, are provided by backcasting data from the National Modelling System (NMS), the national transport model for the Netherlands (Hague Consulting Group, 2000). The attributes for diversification or specialisation in the direct surroundings of a location or in specific range band from that location, are computed from the travel times from the NMS and the location of all firms in the LISA-dataset.

The input to the validation runs consists of the firm population in the baseyear, the economic development and the real estate development scenario. In base year 1996 the study area hosted 95 thousand firms. The validation runs are based on the observed regional economic development in the study area between 1996 and 2004. The macro economic development is input to the firm growth model and in such forms constraints to the simulation and influences the outcomes at micro level. The supply of industrial real estate determines the constraints to what locations are available to a firm. Changes in this industrial real estate supply are implemented through synthesized real estate data for the corresponding period.

3.2 The estimated SFM model

Table 1 presents the estimated parameters for each module in the SFM-model. For a more elaborate discussion of the estimation procedure see De Bok (2007). In this article, the most important findings from the estimations are discussed.

The relocation probability can mostly be explained from the firm attributes, which is in line with firm demographic literature: bigger firms are less likely to relocate (Carroll and Hannan, 2000; Brouwer et al., 2002) and firms with relative large growth rates are more likely to relocate (Carroll and Hannan, 2000; Pellenbarg 1996; Louw, 1996). The relocation probability varies across industry sectors, but accessibility appears to have a limited influence. Agglomeration however, does have an effect: firms at diverse locations are more likely to relocate. This is interpreted as a pattern of successful firms that leave their breeding areas.

When firms relocate and search for a new location, they have a significant preference for locations in the proximity of their original location. This is interpreted as evidence for keep-factors: a relocating firm strives to maintain their existing spatial network. Moreover, the spatial clustering of location alternatives proves to have a significant influence on the choice behaviour of firms: alternatives that are clustered in space, individually have a smaller choice probability. This is in line with findings by Fortheringham and Pelligrini (2002). Furthermore, outspoken differences in location preference between industry sectors are measured for highway and/or train station proximity. Moreover, some industry sectors have a preference for locations with a good labour market accessibility. The estimations provide strong evidence that firms prefer locations that have a relatively high representation of firms from their own industry sector. This is interpreted as evidence for the existence of Marshall externalities and consistent with the findings of Duranton and Puga (2000) and Holl (2004b).

The firm growth model revealed a dominant influence of the growth pattern in the previous years on the expected firm size, consistent with firm demographic literature (Dunne and Hughes, 1994; McCloughan, 1995; Audretsch et al., 2002; Van Wissen and Huisman, 2003). The estimated autoregressive coefficients were smaller than unity for all industry sectors, implying that large firms are expected to grow more slowly than small firms. All estimated ρ_0^s parameters have a negative sign, which implies a negative correlation between firm growth in subsequent years. In other words: a firm with a relative substantial growth in one year, is expected to grow less quickly the next year. Infrastructure proximity has a significant effect on firm performance for various industry sectors. Most estimated coefficients for diversity are negative, indicating relatively less growth at diversified locations. The estimation results for the specialisation coefficient reveal a positive influence of specialisation on the expected growth.

Firm dissolution can be explained with the firm attributes for the largest part: age and size. Firm dissolution is high among young firms. This confirms the liability of newness hypothesis (Stinchcombe, 1965). Larger firms have smaller dissolution probabilities, as well as firms that are growing in size. Both findings are in line with empirical literature (Hannan and Freeman, 1977; Ekamper, 1996; Van Wissen, 2000; Brouwer, 2004). Next, structural developments within industry sectors are significant as well. Firm dissolution appears to be higher among firms at locations with higher diversity indices. Firm dissolution appears to be a little bit higher in the proximity of highway onramps (γ-locations), probably indicating to a higher dynamics at industrial sites at such locations.

In conclusion the estimations confirm a complex interdependency between firm demographic events and confirm a significant influence of accessibility and agglomeration throughout the various events. In general the estimations provide a consistent pattern that can be understood from the life cycle of firms, as outlined in Duranton and Puga (2000). It is argued that the presented approach provides an advanced methodology to better understand the effect of infrastructure and spatial planning scenarios on firm dynamics, by taking into account the theories on urban economics and industrial organisations.

4 Validation

The SFM-model features individual behaviour, spatial externalities and geographic detail, with the objective to provide accurate and valid projections of firm dynamics under different spatial planning scenarios. The assumptions were first tested with the estimation of the firm-demographic models. This section will analyse if the estimated SFM-model actually leads to valid simulation results. For this purpose, a number of validation runs are performed with the estimated coefficients. These runs are performed for the period 1996-2004. The simulated results from the model are compared with the observed developments, across different levels of aggregation.

4.1 Population aggregates

First, the aggregate results from the validation runs are compared to the aggregate development in the observed LISA data. Figure 3 shows the total number of simulated firms from 1996 to 2003 (indicated with black symbols) and the observed size of the firm population (indicated with grey squares). In general the average trend of the size of the simulated firm population is similar compared to the observed firm population. However, the size of the simulated firm population shows a more steady growth trend compared to the observed firm population. This can be explained from the estimation of the firm formation and firm dissolution models: these are estimated on a period of a few years. As a result, the estimated coefficients yield average formation and dissolution models. As a consequence, the dynamic trend is averaged out but the average trend in the population size is reproduced.

4.2 Micro results

The SFM model simulates spatial economic developments at the most disaggregated level possible: individual firms. This detailed level of analysis might suggest that the model aims at predicting individual behaviour. It is stressed that in this case micro-simulation is used to predict the behaviour of the whole spatial economic system. The micro-simulation accounts for the disaggregate dynamics underneath spatial economic development. The objective is to find answers to research questions that apply to intermediate levels of aggregation.

Figure 4 shows the microscopic results for six random firms. Their simulated size is compared to their observed size in the simulation period. The stochastics in the various demographic events prove to lead to very distinct developments at the micro level. For example: in case of firm 200000 (General services) in run 4 it grows from a firm with 10 employees over 8 years to a firm with 15 employees, while in run 5 the same firm is dissolved in the simulation period. It is argued that the distinctive developments at micro level, are representative for the element of coincidence in firm-demographic events. From these microscopic results however, it can not be concluded if the behaviour of the spatial economic system as a whole is represented correctly.

4.3 Neighbourhood results

The scope of the presented micro-simulation model is to make accurate predictions of the behaviour of the whole system of the spatial economy. The associated research questions apply to intermediate levels of aggregation: what is the transport demand from a particular neighbourhood? Or what is the demand for real estate supply in particular neighbourhoods? Therefore, it is most relevant to analyse the validity of the simulation results at neighbourhood level. This validity is analysed with correlation coefficients and scatter plots of the observed and simulated development at neighbourhood level.

Table 2 presents the Pearson correlation of the simulation output for 2004 and the observed neighbourhood population in 2004. The correlations are computed for the firm population and employment totals across all neighbourhoods, and for each industry sector after simulating firm-demographic development for 8 time intervals (8 years). These correlation coefficients are used as an indication for the goodness of fit between the simulation output and the observed firm population.

The correlation figures indicate a respectable match between the simulation results and the observed population or employment totals. The correlation coefficients between the predicted and observed number of firms in a neighbourhood are relatively close to unity. Particular for the total number of firms: around 0.96. The correlation coefficients for sector specific outputs are less high, meaning that the exact type of firms at a location is less easy to predict. The correlation coefficients between the predicted and observed employment in a neighbourhood reveal a satisfactory match between simulated and observed employment totals. The correlation between the total employment predicted and observed varies between 0.88-0.90 across the validation runs. Compared to the correlation coefficients for the number of firms, the correlation coefficients for employment are less high. This can be explained from a stronger dynamics behind employment change.

The correlation coefficients are more or less an abstract measure for similarity, therefore Figure 5 presents scatter plots that give a more intuitive presentation of the similarity between observed and simulated developments. Each dot represents a neighbourhood and its position indicates the match between observed and simulated employment in this neighbourhood: if the fit is perfect, the dot is on the diagonal. The left hand figures compare the levels of employment for total employment and the employment for one of the industry sectors. To improve the visibility of the fit in smaller neighbourhoods, the x- and y-axis are on a logarithmic scale. The total employment predicted for each zone seems to correspond quite well to the observed employment. The dots appear to be located across the diagonal. However, the predictions for the smaller neighbourhoods are less accurate (the dots are further from the diagonal) and seem to be higher then the observations (the dots are above the diagonal). This pattern becomes even more apparent if the sectoral outputs are visualised.

To analyse the over- or under estimation of the level of employment in bigger or smaller neighbourhoods, the right hand figures indicate the deviation between the simulated and observed neighbourhood employment. A dot above the x-axis represents a neighbourhood with too many simulated jobs; a dot below the axis represents a neighbourhood whose employment is under predicted. The large zones in terms of total employment (e.g. employment around 10.000), are underestimated by thousands of jobs. There is not an unambiguous explanation for the underestimation of the larger zones at this point, but a possible explanation might be found in the agglomeration forces in the model. Specialisation is a driving factor for many sectors, what leads to leading to the concentration of particular sectors and the driving out of others.

5 Discussion

It is argued that the presented approach provides an advanced methodology to better understand the effect of infrastructure and spatial planning scenarios on firm dynamics. First of all the approach is based on industrial organisation theory. Hence, individual firm characteristics are applied in the models, allowing heterogeneity in responses. A second distinctive feature is the inclusion of urban economic theories and advanced agglomeration measures that until now have only been applied marginally in projection models. A number of disaggregate agglomeration attributes are included that measure the level of specialisation or diversification in the composition of the local firm population. The transport dimension is included in these measures through travel times from a transport model. The estimation of the firm demographic sub models proved that these accessibility and agglomeration measures are important factors in the various firm demographic events.

The validation of the SFM-model across different levels of aggregation yielded interesting results. The simulated firm population size proves to follow the trend of the observed population size, however the observed growth pattern is more dynamic compared to the average pattern of the simulations. At neighbourhood level the results indicate a good match between observed and simulated development, however larger neighbourhoods seem to be under predicted by the simulation model. A sample of microscopic results showed the effect of the stochastic elements in the firm-demographic models. These stochastics prove to lead to very distinct developments at micro level, but these are argued to be representative for the element of coincidence in firm-demographic events. It is concluded that the presented approach provides reliable estimates of future firm location but further enhancement of the models are desirable. The disaggregate nature of the approach yields detailed results that allow detailed analysis. This provides many ways in extending the approach in future research efforts.

Acknowledgements

The author would like to acknowledge their appreciation for the following persons or institutes that made a valuable contribution to the research: Prof. Frank van Oort from Utrecht University and the Netherlands Institute for Spatial Research, Michiel Bliemer and Piet Bovy from Delft University of Technology, Prof. Leo Van Wissen from the University of Groningen and the Netherlands Interdisciplinary Demographic Institute and the Geo-Database Management Center at Delft University of Technology.

References

Audretsch, D.B., L. Klomp and A.R. Thurik (2002) *Gibrat's law: are the services diffirent?* ERIM: ERIM Report Series No. ERS-2002-04-STR, Rotterdam.

Birch, D. (1979) *The job generating process*. Cambridge University Press, Cambridge (Mass.).

Brouwer, A. (2004) The inert firm; why old firms show a stickiness to their location. *Proceedings of the 44th ERSA conference*, Porto.

Brouwer, A.E., I. Mariotti and J.N. Van Ommeren (2002) The firm relocation decision: a logit model. *Proceedings of the 42nd ERSA conference*, Dortmund.

Carroll, G.R. and M.T. Hannan (2000) *The demography of corporations and industries*. Princeton University Press, Princeton.

De Bok (2007) *Infrastructure and firm dynamics*. PhD-Thesis. Delft University of Technology.

De Bok, M. and F. Sanders (2005) Firm relocation and the accessibility of locations: empirical results from The Netherlands. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1902, pp. 35-43.

De Bok, M. and M.C.J. Bliemer (2006) Land Use and Transport Interaction: Calibration of a Micro Simulation Model for Firms in the Netherlands. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1977, pp. 132-144.

De Jong, G., A. Daly , M. Pieters and T. Van Der Hoorn (2005) The Logsum as an evaluation measure: review of the literature and new results. Proceedings of the 45th ERSA conference, Amsterdam.

Dunne, P. and A. Hughes (1994) Age, size, growth and survival. *The Journal of Industrial Economics*, Vol. 42, pp. 115-140.

Duranton, G. and D. Puga (2000) Diversity and specialisation in cities: why, where and when does it matter? *Urban Studies*. Vol. 37(3), pp. 533-555.

Duranton, G. and D. Puga (2004) Micro-foundations of urban agglomeration economies. In: J. V. Henderson and J. F. Thisse, *Handbook of regional and urban economics*. Vol 4, pp. 2064-2086.

Ekamper, P. (1996) Opheffing van bedrijfsvestigingen: een sterftetafelbenadering. *Planning*. Vol. 48, pp. 12-21.

Fortheringham, A.S. (1983) A new set of spatial-interaction models: the theory of competing destinations. *Environment and Planning A*, Vol. 15, pp. 15-36.

Geurs, K.T. and B. van Wee (2004) Accessibility evaluation of land-use and transport strategies: review and research directions. *Journal of Transport Geography*, Vol. 12(2), pp. 127-140.

Hague Consulting Group (2000) Het Landelijk Model Systeem versie 7.0. The Hague (in Dutch).

Hannan, M.T. and J.Freeman (1977) The population ecology of organizations. *American Journal of Sociology*, Vol. 82, pp. 929-964.

Hilbers, H., P. Jorritsma and P. Louter (1994) The relationship between accessibility and regional development explored for the region of Amersfoort. *Proceedings of the 21st Colloquium Vervoerplanologisch Speurwerk*, Zeist.

Holl, A. (2004a) Manufacturing location and impacts of road transport infrastructure: empirical evidence from Spain. *Regional Science and Urban Economics*, Vol. 34, pp. 341-363.

Holl, A. (2004b) Start ups and relocations: manufacturing plant location in Portugal. *Papers in Regional Science*, Vol. 83, pp. 649-668.

Hunt, J.D. (1997) Stated-preference examination of location choice behavior of small retail firms. *Transportation Research Record: Journal of the Transportation Research Board*. No. 1606, pp. 25-32.

Kawamura, K. (2001) Empirical examination of the relationship between firm location and transportation facilities. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1747, pp. 97-103.

Leitham, S., R.W. McQuaid, and J.D. Nelson. (2000) The influence of transport on industrial location choice: a stated preference experiment. *Transportation Research Part A: Policy and Practice*. Vol. 34, pp. 515-535.

Louw, E. (1996) *Kantoorgebouw en vestigingsplaats*. Ph.D. Thesis, OTB Research Institute, Delft.

McCloughan, P. (1995) Simulation of Concentration Development from Modified Gibrat Growth-Entry-Exit Processes. *The Journal of Industrial Economics*. Vol 43(4), pp. 405-433.

Paci, R. and S. Usai (1999) Externalities, knowledge spillovers and the spatial distribution of innovation. *GeoJournal* Vol 49, pp. 381-390.

Pellenbarg, P.H. (1996) Structuur en ontwikkeling van bedrijfsmigratie in Nederland. *Planning*, Vol. 48, pp. 22-32.

Pelligrini, P.A. and A.S. Fortheringham (2002) Modelling spatial choice: a review and synthesis in a migration context. *Progress in Human Geography*, Vol. 26(4), pp. 487-510.

Rietveld, P. (1994) Spatial economic impacts of transport infrastructure supply. *Transporation Research Part A: Policy and Practice*, Vol. 28, pp. 329-341.

Rosenthal, S.S. and W.C. Strange (2003) Geography, industrial organisation, and agglomeration. *The Review of Economics and Statistics*, Vol. 85(2), pp. 377-393.

Rosenthal, S.S. and W.C. Strange (2004) Evidence on the nature and sources of agglomeration economies. In: J. V. Henderson and J. F. Thisse, *Handbook of regional and urban economics*. Vol 4, pp. 2119-2167.

Shukla, V. and P. Waddell (1991) Firm location and land use in discrete urban space: A study of the spatial structure of Dallas-Fort worth. *Regional Science and Urban Economics*, Vol. 21, pp. 225-253.

Stinchcombe, A.L. (1965) Social structure and organisations. In: J. March, *Handbook of organisations*. Chicago: Rand McNally. Pp. 142-193.

Van Der Panne, G. (2004) *Entrepreneurship and localised knowledge spillovers*. Ph.D. Thesis, Delft University of Technology.

Van Dijk, J. and P. Pellenbarg (2000) Firm relocation decisions in The Netherlands: An ordered logit approach. *Papers in Regional Science*, Vol. 79, pp. 191-219.

Van Oort, F., B. van der Knaap and W. Sleegers (1999) New firm formation, employment growth and the local environment: emperical observations in South-Holland. In: J. Van Dijk and P. Pellenbarg (Ed.) *Demography of firms; Spatial Dynamics of Firm Behavior*. Utrecht/Groningen: Netherlands Geographical Studies 262, pp. 173-204.

Van Wissen, L. (2000) A micro-simulation model of firms: Applications of concepts of the demography of the firm. *Papers in Regional Science*, Vol. 79, pp. 111-134.

Van Wissen, L. and C. Huisman (2003) Firm growth across regions: a regional test of Gibrat's law: are the services different? 43rd European Regional Science Association Congress, Jyvaskyla.

Waddell, P. and G.F. Ulfarsson (2003) Accessibility and Agglomeration: Discrete-Choice Models of Employment Location by Industry Sector. Proceedings of the 82nd Annual Meeting of the Transportation Research Board, Washington, D.C..

Table 1: Parameters in the estimated SFM-model.

	Finance	Business services	Government	Education	Health services	General Services	Manufactering	Construction	Transport, Warehousing & Comm	Agriculture	Trade & retail	Restaurants & Food services
	β	β	β	β	β	β	β	β	β	β	β	β
RELOCATION PROBABILITY						(unsegmented)						
Constant						-3.68						
Individual firm attributes:												
Log of size						-0.04						
Growth rate						0.40						
Industry sector	0.41	0.74	0.53	0.08	0.18	0.00	0.22	0.53	0.71	-0.48	-0.08	-1.27
Accessibility attributes:												
α-location; near trainstation						0.027						
β-location; near trainstation & highway onramp						0.046						
γ-location; near highway onramp						-0.055						
ρ-location; neither						-0.018						
Agglomeration attributes:												
Diversity Rb < 7,5 min.						0.238						
Specialisation Rb < 7,5 min.						-0.008						
LOCATION CHOICE MODEL												
Migration attribute:												
Distance to original loc.[$km^{1/2}$]	-2.37	-1.89	-1.83	-2.70	-2.59	-2.06	-1.73	-1.91	-1.62	-2.65	-2.07	-2.79
Accessibility attributes:												
α-location; near trainstation	-0.10	-0.07	0.35	0.47	0.24	-0.08	-0.37	-0.15	0.02	-0.28	-0.21	0.15
β-location; near trainstation & highway onramp	0.50	0.33	0.26	0.48	0.23	0.30	0.01	0.32	0.11	0.41	-0.12	0.90
γ-location; near highway onramp	-0.01	0.00	-0.10	-0.40	-0.15	0.15	0.31	0.07	0.17	-0.06	0.49	-0.20
ρ-location; neither	-0.39	-0.26	-0.50	-0.55	-0.32	-0.38	0.05	-0.25	-0.30	-0.07	-0.16	-0.85
Reflected logsum commuting trips	0.67	0.41	0.17	-0.03	0.10	0.15	0.39	0.12	0.32	-0.45	0.36	0.24
Logsum business trips	-2.30	-0.58	-1.30	0.57	0.62	0.42	-1.01	-0.90	-1.02	0.71	-1.06	-0.79
Agglomeration attributes:												
Diversity Rb < 7,5 min.	0.19	0.01	0.48	-0.42	0.72	-0.31	0.36	0.37	0.42	0.67	-0.07	-2.42
Specialisation Rb < 7,5 min.	0.02	0.64	-0.37	-0.04	0.00	-0.11	0.49	0.18	0.31	0.04	1.05	-0.07
Centrality parameter:												
Teta	-0.53	-1.38	0.44	-1.68	-1.42	-0.97	-0.59	-0.68	-1.25	-3.35	-1.96	0.11
FIRM GROWTH MODEL												
Firm attributes:												
Size previous years: slope coefficient BETA	0.982	0.982	0.979	0.994	0.988	0.983	0.983	0.980	0.978	0.952	0.978	0.954
Size previous years: autocorrelation coefficient RHO	-0.163	-0.117	-0.062	-0.153	-0.188	-0.156	-0.120	-0.123	-0.133	-0.153	-0.157	-0.161
Accessibility attributes:												
α-location; nearby trainstation [-]	0.017	-0.001	0.011	-0.001	0.000	0.004	0.004	0.003	0.002	-0.007	0.001	0.017
β-location; nearby trainstation & highway onramp [-]	-0.005	0.003	0.003	0.003	0.004	-0.001	-0.006	0.001	0.003	0.019	0.000	-0.001
γ-location; nearby highway onramp [-]	-0.004	0.002	-0.012	0.005	-0.001	-0.001	0.004	0.001	0.004	-0.006	0.001	-0.008
ρ-location; neither [-]	-0.008	-0.004	-0.002	-0.007	-0.003	-0.002	-0.001	-0.005	-0.009	-0.006	-0.003	-0.009
Agglomeration attributes												
Diversity Rb < 7,5 min. [-]	-0.036	0.001	-0.060	0.002	-0.006	-0.015	-0.031	-0.010	-0.031	-0.032	-0.042	-0.056
Specialisation Rb < 7,5 min. [-]	0.011	0.011	0.007	0.001	0.003	0.002	0.005	0.005	0.008	0.000	0.006	0.003
FIRM DISSOLUTION MODEL						(unsegmented)						
Constant						-2.83						
Individual firm attributes:												
Log of size						-0.35						
Growth rate						-0.27						
1 / age						0.74						
Industry sector	0.53	0.47	1.09	0.43	-0.13	0.00	0.46	0.28	0.48	-0.13	0.35	-0.18
Accessibility attributes:												
α-location; near trainstation						0.012						
β-location; near trainstation & highway onramp						-0.030						
γ-location; near highway onramp						0.031						
ρ-location; neither						-0.014						
Agglomeration attributes:												
Diversity Rb < 7,5 min.						0.190						
Specialisation Rb < 7,5 min.						-0.012						
FIRM FORMATION MODEL												
Yearly birth rate (% of population)	6.58	15.70	7.70	5.24	6.49	7.74	6.55	6.96	8.39	2.44	7.11	4.92
Start up size distribution (% of start ups):												
0-4 employees	86.3	88.6	30.4	75.6	73.8	92.4	75.7	82.8	80.5	78.7	89.1	85.0
5-9 employees	6.6	6.0	15.0	8.3	12.2	4.5	11.5	8.0	8.8	14.9	7.0	9.1
10-24 employees	4.3	3.5	20.5	10.4	9.8	2.1	7.3	5.8	6.1	5.2	2.6	4.1
25-49 employees	1.9	1.1	13.9	3.6	2.2	0.7	3.2	2.2	2.2	0.6	0.8	0.8
50-99 employees	0.6	0.4	8.7	1.9	1.1	0.4	0.9	0.8	1.1	0.4	0.4	0.6
>100 employees	0.3	0.3	11.5	0.3	1.0	0.0	1.4	0.4	1.2	0.1	0.1	0.3

Table 2: Pearson correlation

between observed and simulated zone employment for 2004.

	Firm population *)				Employment *)			
	Run 1	Run 2	Run 3	Run 4	Run 1	Run 2	Run 3	Run 4
Total population	0.96	0.96	0.96	0.96	0.89	0.90	0.88	0.90
Agriculture	0.83	0.81	0.84	0.83	0.65	0.59	0.60	0.62
Manufacturing	0.90	0.90	0.90	0.89	0.83	0.82	0.80	0.80
Construction	0.84	0.85	0.86	0.85	0.66	0.72	0.69	0.69
Trade and Retail	0.95	0.95	0.95	0.95	0.90	0.91	0.90	0.91
Restaurants and Food service	0.97	0.97	0.98	0.98	0.95	0.95	0.95	0.95
Transp. Wareh. and Comm.	0.85	0.84	0.81	0.83	0.75	0.79	0.71	0.79
Finance	0.86	0.85	0.87	0.87	0.83	0.80	0.81	0.76
Business services	0.88	0.87	0.86	0.88	0.75	0.81	0.75	0.80
Government	0.83	0.86	0.87	0.87	0.78	0.79	0.79	0.82
Education	0.81	0.84	0.82	0.85	0.82	0.81	0.80	0.75
Health service	0.89	0.90	0.88	0.88	0.56	0.57	0.56	0.56
General service	0.92	0.93	0.93	0.93	0.84	0.85	0.78	0.88
N (number of zones)	433	433	433	433	433	433	433	433

*) Presented are the Pearson Correlations for the neighbourhood aggregates

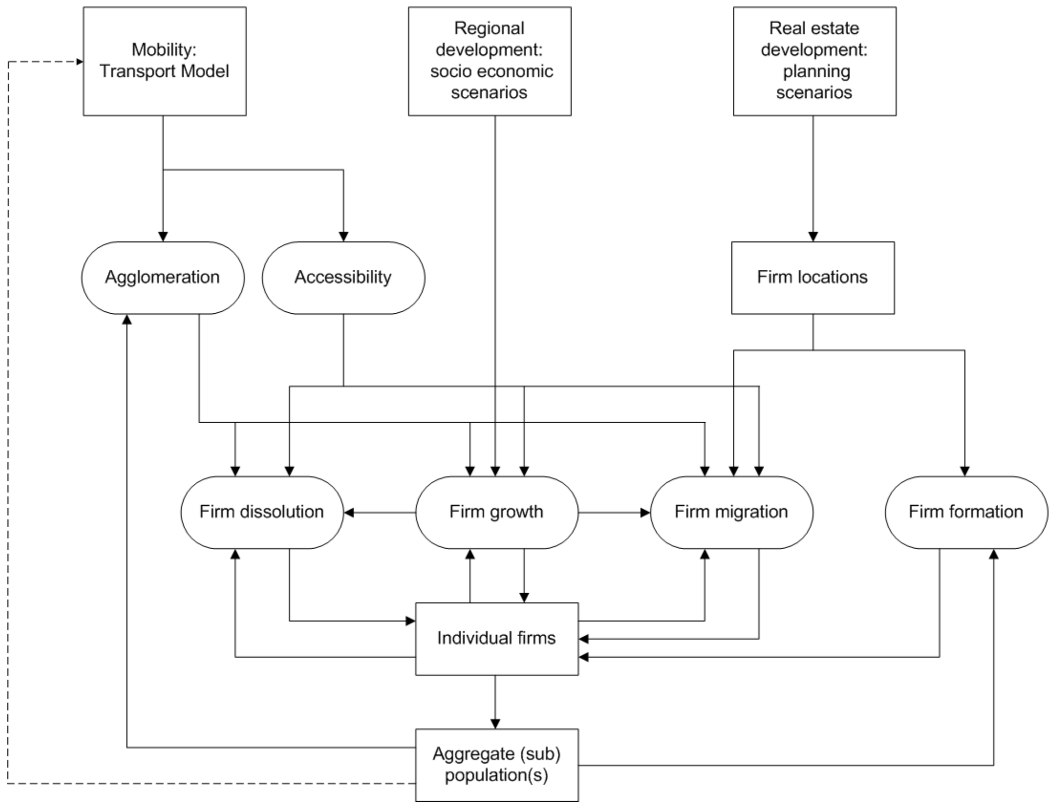

Figure 1: Causal structure of the SFM framework

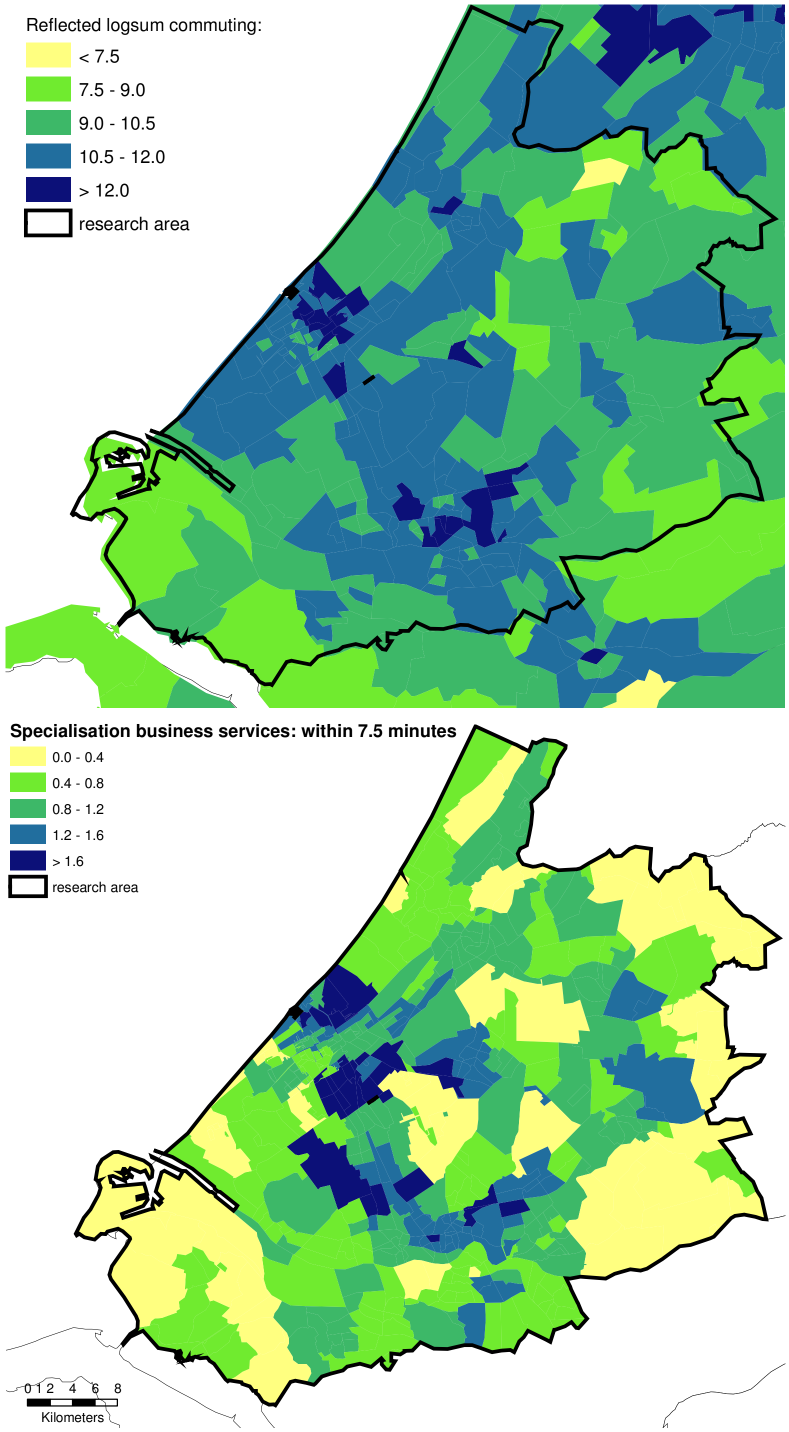

Figure 2: Examples of logsum accessibility (top) and specialisation index (bottom)

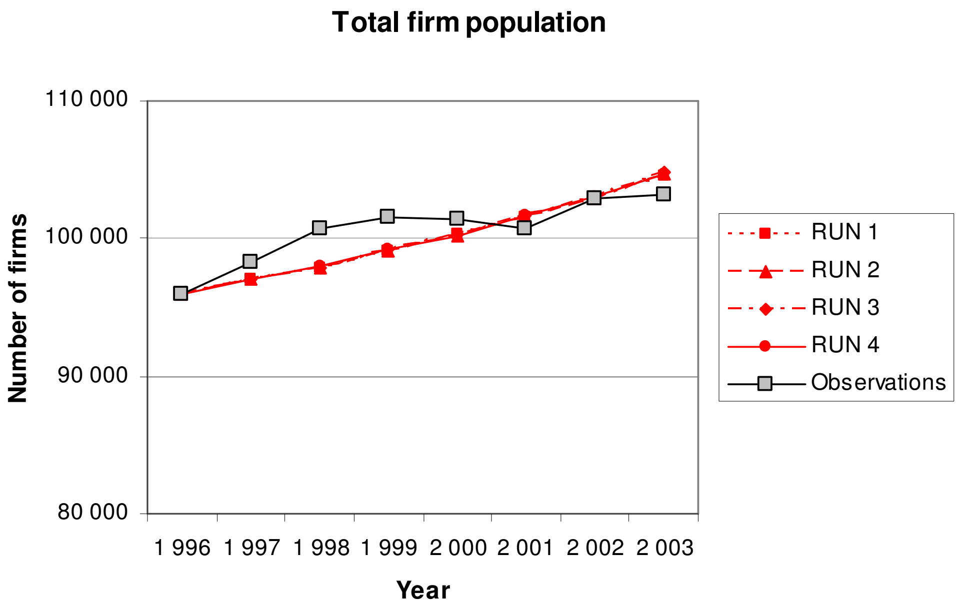

Figure 3: Observed and simulated size of firm population

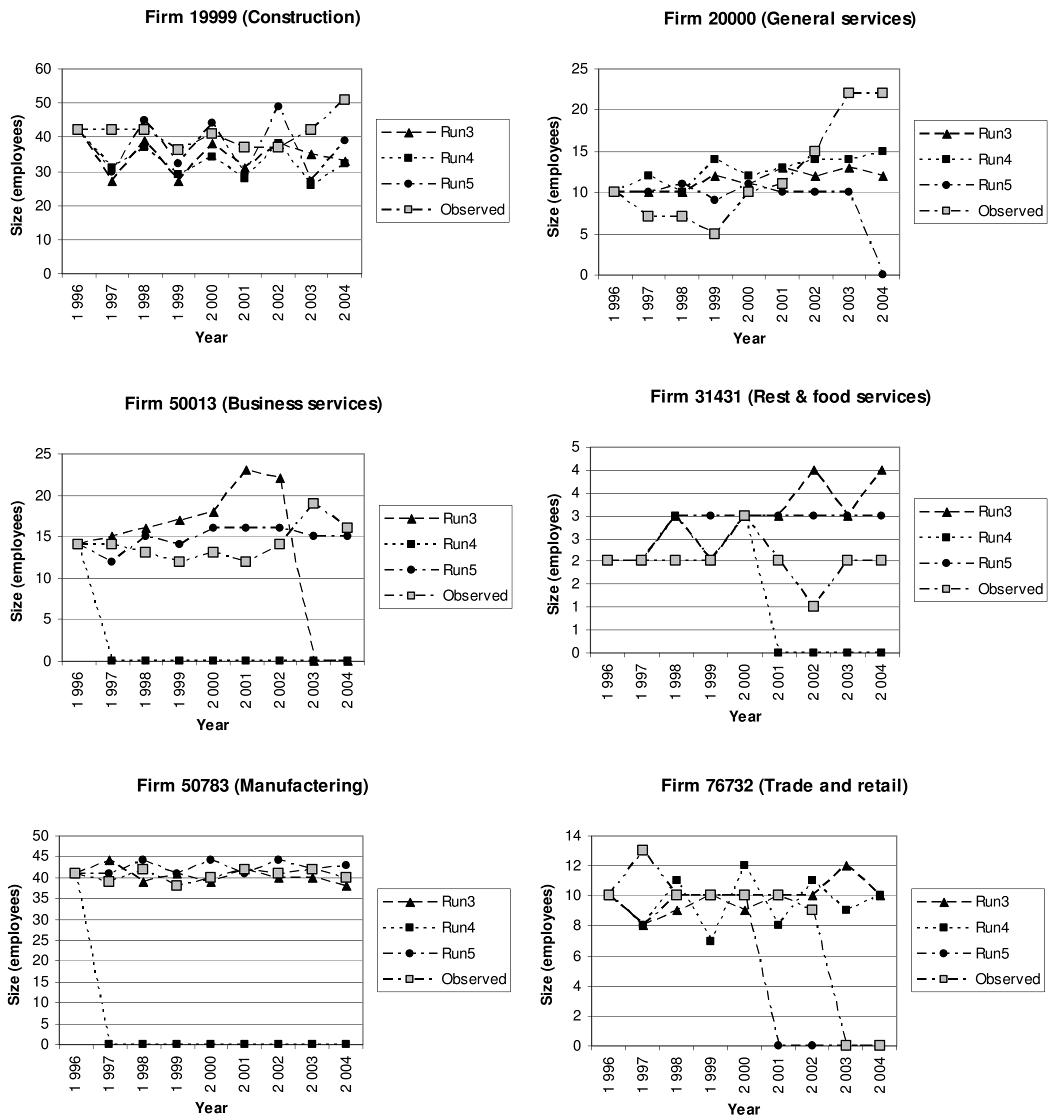

Figure 4: Sample of micro results, observed vs. simulated size development

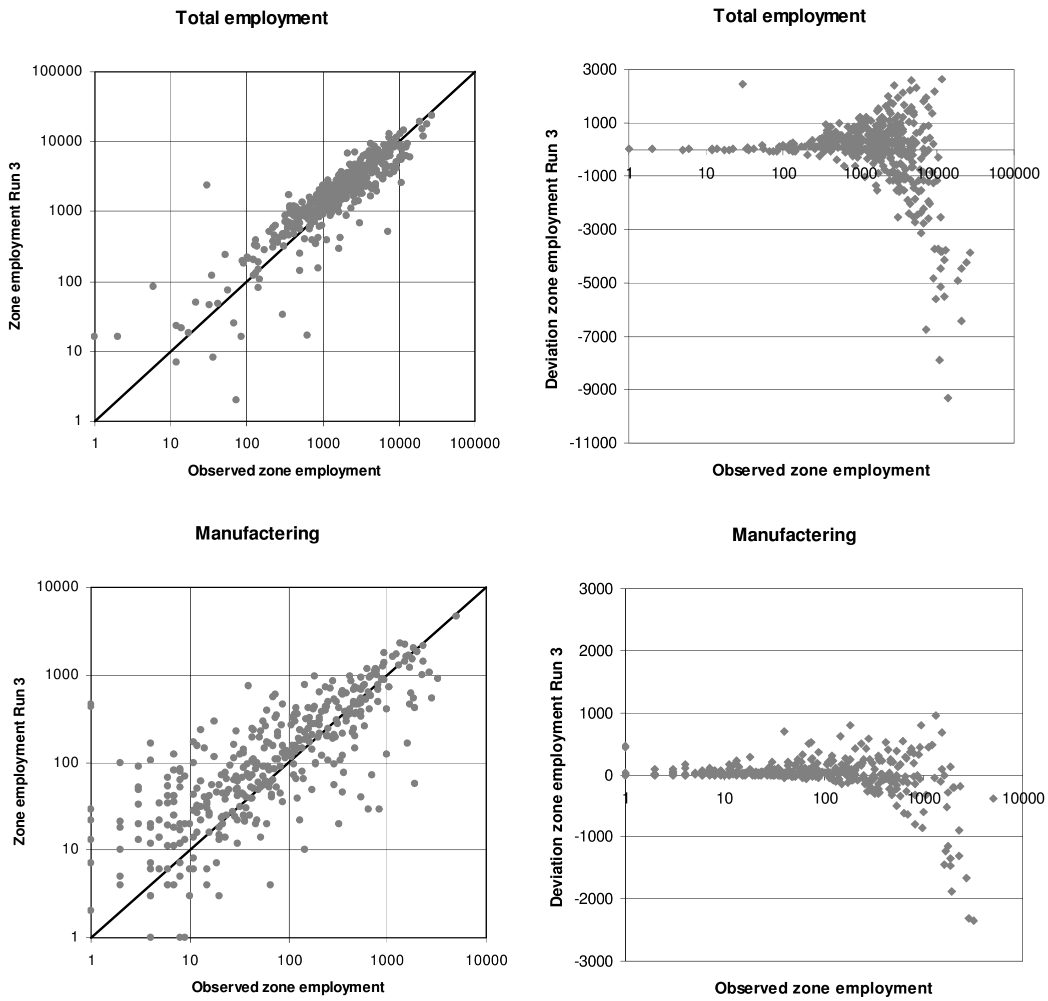

Figure 5: Scatter plot observed vs. simulated zone employment