Urban Hierarchy Models and the Tail of the Distribution: Some Robustness Checks in Abstract and Empirical Models

Dirk Stelder
Faculty of Economics
University of Groningen
PO Box 800
9700 AV Groningen
The Netherlands
+31 503633725
t.m.stelder@rug.nl
Urban Hierarchy Models and the Tail of the Distribution: Some Robustness Checks in Abstract and Empirical Models

Abstract

Results of urban agglomeration models dramatically change with the number of cities in the model. Concentration in an \( n \)-city setting can reverse into dispersion by adding or deleting just one city. The centre-periphery model (CP) and the footloose entrepreneur model (FE) are tested for their robustness with respect to this in a set of abstract spaces and in an empirical implementation for the urban hierarchy in the Netherlands. The results show that the FE model becomes size-invariant from approximately 100 cities onwards while the CP model remains volatile, even for very large models. For the empirical model the predicted agglomeration effects of a rise of gasoline prices are not robust for models in the range of 10 to 50 cities. As larger parts of the tail of the city distribution are included, the model results become more robust. The results suggest that when building an empirical model of this kind, a sensitivity analysis on which part of the tail of the city distribution to ignore is crucial.

1. Introduction

When modelling an urban hierarchy of a country a decision has to be taken on at which point the tail of the city distribution will be ignored. The aim of this paper is to examine what the exact consequences are of that decision for the behaviour of an urban agglomeration model of the NEG type\(^1\). Section 2 of this paper will demonstrate that for a uniform city distribution in homogenous spaces like a horizontal line or a central-places type economic plane, NEG models converge to robust results when the are extended to more cities. Starting from a small number of cities (< 10) the results are very volatile every time when an extra region is added, but the outcome becomes more robust after 30 cities or more. It is argued that this result reflects the pure effect of the number of cities in the model, regardless of their size and location. This is what we would expect by intuition as well as from theory. More locations simply make more differentiated spatial configurations possible (Stelder, 2005a).

Section 3 shows that the interpretation of this analysis becomes more complex in an empirical implementation for the Netherlands using a discrete set of urban hierarchies with an increasing number of cities. First, the agglomeration parameters are different for each model because a different hierarchy has to be calibrated, although their values do not diverge very much as the number of cities in the model increases. More important, for every city that is added to the model it is always the combined effect of its size and position that changes the results. The threshold at which the model results start to converge lies around 50 cities.

To our best knowledge there is no economic literature on this subject. This is surprising because the issue is far from just academic looking at the large and increasing demand for applied spatial economic models from (international) government agencies that are needed for policy evaluation. For example, for the Directorate-General for Regional Policy of the EU, a regional model is currently under construction to evaluate the regional economic effects of EU Cohesion policy\(^2\). A prototype of this model, which will be a combination of a NEG model with CGE and an endogenous growth component, is being developed for the 53 NUTS2 regions of Germany, Poland, Czech Republic, Slovakia and Hungary. Once this prototype is successfully constructed and calibrated, the future agenda holds the following options:

1) a full extension of the model to the whole EU covering more than 300 regions.
2) an extension to the NUTS3 level
3) an extension to the city level

\(^1\) The acronym NEG (New Economic Geography) covers the original core-periphery model as well as other versions. See (Fujita, Krugman, and Venables 1999) or (Baldwin et al. 2003) for an overview and introduction to NEG models. (Ottaviano and Robert-Nicoud 2006) discuss the main features of current classes of NEG models. Annex A summarizes the main equations of the models used here.

Including the prototype this agenda could lead to four models that each might give different results for – say – the policy effects of an infrastructure investment program in the Czech Republic and Slovakia. Because of interregional interdependencies strategy 1) will most likely produce different results because not only Austrian regions nearby will enter the game, but also because the reaction of bordering German regions may become mitigated by the influence of regions further away towards Denmark, France or the Netherlands. Setting this issue aside and assuming the same country coverage, option 2) and 3) will increase the number of cities and particularly option 3) has the problem how to decide which cities should be included. Any choice on cities over 50,000 or 100,000 inhabitants is a choice on which part of the tail of the city distribution to include and is implicitly a choice on the number of cities in the model.

The key issue in this context is what we would like to call the robustness of model behavior with respect to their spatial dimension. Given a dimension of \( n \) cities, what we would like to see is that a policy evaluation does produce approximately the same results for the same model with \( n-1 \) or \( n+1 \) cities. If we can find some threshold \( n^* \) at which this robustness starts to appear we will know two things. First, this level should be the minimum number of cities in order to be reliable. Second, unless we would want further detail for other reasons, we will not need to extent the model into further detail because that would only increase data requirements without changing the outcome.

The paper is structured as follows. First, section 2 gives the model structure and explains the setup of the simulations. Section 3 presents the results for abstract spaces in one- and two dimensional space. Section 4 turns to the empirical implementation for a range of samples from the 400 largest cities in the Netherlands. Section 5 contains a summary and conclusions. Annex A presents a special urban Herfindahl index developed for this paper that is invariant to the number of locations.

2. Model structure and setup of the simulations

We use two different model types which are known in the literature as the Core-Periphery (CP) model and the Footloose Entrepreneur (FE) model (Baldwin et. al.; 2003). The description can be brief here because we use the standard models without any alterations from our part. Every region is divided into a uniform fraction \( \delta \) of manufacturing employment in firms in a monopolistic competition framework producing under increasing returns and a fraction \((1-\delta)\) perfectly competitive agricultural sector producing under constant returns to scale. The same fraction \( \delta \) of income is spent on manufacturing goods and \((1-\delta)\) on agricultural goods. Every manufacturing firm produces one variety. Consumers like varieties which makes their utility of consuming manufactured goods increasing with the number of varieties from which to choose.

The derivation of the equilibrium equations can be found in Fujita, Krugman & Venables (1999) or Brakman, Garretsen & van Marrewijck (2008). Following the notation of the latter the core equations of the CP model are:

\[
\begin{align*}
(1) \quad y_r &= \delta \lambda_r w_r + (1-\delta)\phi_r \\
(2) \quad I_r &= \sum_{s=1}^{R} \lambda_s T_{rs}^{1-\epsilon} w_s^{1-\epsilon} \\
(3) \quad T_{rs} &= T^{\delta(r,s)} \\
(4) \quad w_r &= \sum_{s=1}^{R} y_s T_{rs}^{1-\epsilon} I_r^{1-\epsilon}
\end{align*}
\]

In (1) for every region \( r \), \( y_r \) is total income, \( \delta \) is the fraction of income spent on manufacturing goods, \( w_r \) is manufacturing wages, \( \lambda_r \) is the regional fraction of national manufacturing employment and \( \phi_r \) is the regional fraction of national agricultural employment. Next, in (2) \( I_r \) is
the regional price index and $T_{rs}$ is the iceberg transport costs indicating the number of goods needed to be shipped from region $r$ in order to have one unit of goods arriving in region $s$. Equation (3) shows that this is equal to the transport cost parameter $T$ in the case of a two-region model with $D(1,2) = 1$ and raised to the power $D(r,s)$ in an $n$-region model. Finally, in (2) and (4) $\varepsilon$ is the elasticity of substitution derived as $\varepsilon = 1/(1 - \rho)$ with $\rho$ being the substitution parameter in the aggregate utility function for consumers.

System (1)-(4) determines the short term equilibrium value of $w_r$ given the parameter values for $\delta$, $\varepsilon$ and $T$ space (this is the economic parameter space $E_0$ mentioned in the paper) the and the initial values of $\lambda$ ($\Lambda_0$).

One of the assumptions behind the core equations of the CP model is the production function for manufacturing (M) firms in which manufacturing labor is the only product factor with increasing returns to scale. Following (Robert-Nicoud 2005) this defines total costs $C$ for the typical M-firm producing $x(i)$ quantities of manufacturing good $i$ as

$$C(x(i)) = w_m [F + \beta x(i)]$$

with $F$ and $\beta$ as the fixed and variable costs parameters$^3$ and $w_m$ indicating the wage in the manufacturing sector. It is in this equation where the CP model differs from the footloose entrepreneur model (FE) which uses

$$C(x(i)) = w_m F + w_a \beta x(i)$$

where $w_a$ is the wage in the agricultural sector. The production function (6) has fixed costs of skilled (manufacturing) labor and variable costs of unskilled (agricultural) labor (Forslid & Ottaviano, 2003). This makes the equations (2) and (4) no longer recursive because in the FE model the price index equation (2) simplifies to:

$$I_r = \left[ \sum_{s=1}^{R} \lambda_s T_{rs}^{1-\varepsilon} \right]^{1/(1-\varepsilon)}$$

The FE model equations (1), (3), (4) and (7) enable an analytic solution in stead of the numerical solution that is needed in the CP model$^4$.

Finally, the determining of the long term equilibrium is the same for both models. Given the short term equilibrium for $w_r$ the model assumes that manufacturing workers migrate to regions with the highest real wage $\omega_r$ according to

$$\omega_r = W_r I_r^{-\delta}$$

and a migration reaction relative to the regional deviation from the average national real wage $\bar{\omega}$:

$$\lambda_{r,t+1} = \eta \lambda_{r,t} / \omega_r / \bar{\omega}$$

Here $\eta$ is a migration sensitivity parameter and index $t$ indicates the iteration number. The model simulation stops when real wages are equalized across the regions and convergence is reached to the long term equilibrium $\Lambda_t$ when for $\forall r$ $\lambda_{r,t+1} / \lambda_{r,t} < (1 + \kappa)^{-\delta}$.

The choice for the CP model and the FE model in this paper is motivated by two reasons. The original CP model is well known original and the most basic NEG model with relatively strong

---

$^3$ These parameters do not show up in (1)-(4) due to parameter normalizations. See Brakman et. al. (2008) chapter 4 for details.

$^4$ As any programmer will quickly find out, the reduced form equation for the wages in the two-region case as presented in Brakman et. al. (2008) or (Forslid and Ottaviano 2003) becomes too long and cumbersome for 3 or more regions so numerical techniques are needed after all to find the analytical solution.

$^5$ In all simulations used in this paper $\eta$ is set to 1 and the break-off condition $\kappa$ is set to 0.0001.
agglomeration tendencies but nevertheless has been proven to be able to produce very realistic differentiated urban hierarchies when applied to large empirical geographical models (Stelder, 2005a). The FE model is its main competitor as it is also simple but has an analytical solution and less strong agglomeration tendencies.

The long term equilibrium of both models depends on what can be called the three basic components economics, geography and history. The economic parameter space \( E \) is defined by the three parameters \( \delta, \varepsilon, \) and \( T \), or \( E = \{ \delta, \varepsilon, T \} \). Geography is given by the number of locations \( n \) and their distance matrix \( D \). History is the initial distribution \( \Lambda_0 = \{ \lambda_1, \lambda_2, \ldots, \lambda_n \} \) which co-determines the outcome due to path dependencies\(^6\).

The use of these three components is different in section 3 and 4. First, in section 3 the CP model and the FE model are applied to abstract spaces of different types and increasing size using two different economic parameter spaces \( E_1 \) and \( E_2 \) defined as \( E_1 = \{ \delta=0.3, \varepsilon=3, T=1.35 \} \) and \( E_2 = \{ \delta=0.4, \varepsilon=4, T=1.45 \} \). These chosen values are typical for a medium-high and a medium-low agglomeration outcome and stem from simulations with many economic parameter configurations on one and the same geography as done by Stelder (2005a) and Brakman et.al. (2009). This gives four cross-combinations of \( \{ E_1, E_2 \} \) and \( \{ \text{CP}, \text{FE} \} \) which are labeled as model CP1, CP2, FE1 and FE2. As will be shown, the four models not only produce different size distributions but also their spatial outcome is different. Next, for the history component \( \Lambda_0 \) only the flat initial distribution is used in which \( \lambda_1 = \lambda_2 = \ldots = \lambda_n \). This “no history assumption” means that all regions are assumed to be equal in size from the start and simplifies the analysis enabling us to examine the general agglomeration tendencies of specific geographical spaces\(^7\). Finally, as the geography component, the distance matrix \( D \) is simply the shortest path distance assuming direct connections between horizontal and vertical neighbors on the networks used (see section 3).

The empirical implementation in section 4 is different. Here we use as history \( \Lambda_0 \) the urban distribution of employment over the 443 largest cities of the Netherlands in 2007. Instead of a fixed economic parameter space \( E \), each model is calibrated separately to find the \( E^* \) that reproduces this distribution most closely as a long term equilibrium. Each model of different size \( n \) increasing from the top 10 largest cities to the whole distribution thus has its own calibrated economic parameter space \( E_n^* \). The distance matrix here \( D \) contains the shortest path distance in kilometers over the road network of 2009. After calibration we take a general rise of gasoline prices of 10% as a simple policy evaluation example which is implemented as a proportional rise of all entries in the distance matrix \( D \) with 10%. The spatial agglomeration effects of this policy turn out to be substantially different for each model. Section 5 concludes and discusses options for future research.

3. The robustness of abstract agglomeration models

3.1. Types of abstract spaces

For our purpose we need to make a choice about which type of space to use. Figure 1 shows the main types that can be distinguished. The relevant types of geographies are here classified according to their euclidean dimension and the properties neutrality and symmetry which are particularly relevant for a NEG model which is basically a network with \( n \) nodes and \( m \) connections.

\(^6\) In the words of Krugman (1991): “history matters”.

\(^7\) Note that this assumption can only be used in non-neutral space because in neutral space a flat initial distribution is an immediate long term equilibrium. It's important advantage is that we do not need many model runs with random initial distributions for our analysis as is common practice for neutral geographies (Fujita, Krugman & Venables, 1999). In non-neutral space the resulting equilibria from many random runs will evidently form a normal distribution around locations that have a structural spatial agglomeration advantage. When the number of simulations goes to infinity these normal distributions will obviously converge to “spikes” on exactly the same places as the simulation with \( \Lambda_0 \) predicts directly (Stelder, 2005a).
First, we will define spatial \textit{neutrality} as the case in which every region has the same potential \( P \): \( P_i = P_j \) \( \forall i,j \) with \( P_i = 1/\sum_j d_{ij} \). Figure 1a shows a neutral space of 6 locations on an equidistant hexagonal network in \( \mathbb{R}^2 \). Note that in order for the neutrality condition to hold not all connections need to be of equal distance because the example in Figure 1b is also spatially neutral.

Next, the simplest introduction of non-neutral space is to delete one connection in Figure 1a which cuts the hexagon into a line segment congruent to a straight line in \( \mathbb{R} \) (Figure 1c). The assumed equal distances between all direct neighboring locations leads to a special property which we will here define as \textit{symmetry}. The symmetry condition holds when \( \forall i \) with \( P_i < \max (P_i) \) \( \exists j \neq i \) for which \( P_i = P_j \). Our results will show that the (non)existence of symmetry is important because it has a dominant influence on the model results when the number of regions is relatively small.

In the same way, the continuous economic plane in \( \mathbb{R}^2 \) can be approximated by a symmetric grid network as depicted in Figure 1d, or as a hexagonal central places structure in Figure 1e. In real geographies symmetry no longer exists and the network can be of any form. Figure 1f is a very simple example where only one location of Figure 1d is deleted which gives location \( B \) a unique value \( P_B \) lower than the highest potential value \( P_A \).

The next three sections present our results for the horizontal line (Figure 1c), the square regular grid (Figure 1d) and the simplest asymmetric space (Figure 1f) as three cases of increasing geographical complexity while keeping all cities of the same size. Theoretically, the analysis is also possible for neutral spaces like figure 1a, but then the assumption of identical cities has to be dropped which would blur the picture.

3.1. Results for one-dimensional space: the horizontal line

The four models CP1, CP2, FE2 and FE2 were tested on the horizontal line model in Figure 1c for the range \( n=3 \) to 100 using the initial “no history” flat distribution \( \Lambda_0 \) with \( \lambda_1 = \lambda_2 = \ldots = \lambda_n \). Due to the non-neutral geography some regions have a relative better access to surrounding markets

\[ \text{Note that the space types 1c and 1d can have more than one vertices } i \text{ for which } P_i = \max (P_i) \text{ depending on whether } n \text{ is an even or uneven number} \]

\[ \text{This is because a neutral space has to be feeded with many random initial distributions in order to analyze its behavior. See footnote 7} \]
than others. This provides them with a lower price index and higher wages following (1)-(4). Typically this leads to a long term equilibrium with $m < n$ agglomerations where all manufacturing labor has become concentrated ("cities") and $(n-m)$ remaining locations from which all manufacturing labor has moved away and have only agricultural labor left ("villages")\textsuperscript{10}.

For a correct measurement of the models behavior as $n$ increases we have constructed a special modified H"{e}fndahl index that is invariant to $n$ for otherwise identical urban hierarchies. This \textit{Urban Herfindahl Index (UHI)} is introduced in Appendix A and is calculated for the long term equilibrium values of the spatial distribution of manufacturing employment in $\Lambda_t$. Figure 2 gives the $UHI$ of CP1, FE1, CP2 and FE2 for $n$ increasing from 3 to 100. There is a strong effect of going from even to uneven numbers for small values of $n$ with the $UHI$ jumping from high values for $n=3,5,7,9$ to low values for $n=4,6,8,10$. This effect becomes less important for larger values of $n$, although CP1, FE1 and CP2 remain to show substantial fluctuations going from $n$ to $n+1$. The most stable model is FE2 that has an almost constant $UHI$ for $n>30$. The $UHI$ for CP2 and FE1 also seems to converge to a constant value, but not until $n$ has become 60 or larger. The average level of the UHI is different for each model. As expected, CP1 and FE1 have a significantly higher centripetal tendency compared with CP2 and FE2 because higher values for $\sigma$ and $\tau$ lead to more spreading\textsuperscript{11}.

\textbf{Figure 2. Urban Herfindahl Index for a horizontal line with increasing size}

As an example of the full range of all 97x4 simulations presented in Figure 2, Figure 3-4 show the location and size of the cities in the long term equilibrium for $n=22$ and $n=23$\textsuperscript{12}. Clearly the models are no soccer game: even going from 22 to 23 players still has a very sensitive effect on the long term equilibrium. In this simple example the main cause is the effect of changing from even to uneven numbers. The CP1 model predicts two small second order cities in the centre when $n=22$ but one large central agglomeration when $n=23$. The central market captured by location 12 in the latter case has to be shared between location 11 and 12 in the first case who have a mirrored but equally competitive spatial market power. The same type of shift happens for model FE2.

\textsuperscript{10} A detailed discussion on the behavior of non-neutral NEG models is given by Stelder (2005b)

\textsuperscript{11} The higher value of $\delta$ in CP2 and FE2 works in the opposite direction, but this was done because keeping $\delta$ at 0.3 made the centripetal forces too low compared with CP1 and FE1.

\textsuperscript{12} The full set of all simulations is downloadable from: (link)
3.1. Results in two-dimensional space

The analysis has been repeated for two-dimensional symmetric geographies of type 1d in Figure 1 using an $m \times m$ square grid for $m = \{3, 4, \ldots, 30\}$ which makes the increase of the model size quadratic following $n = \{9, 16, 32, \ldots, 900\}$. The results are given in Figure 5 and lead to two conclusions. First, from $n = 9$ to 225 we can see the returning pattern of alternating high and low agglomeration when changing from even to uneven numbers. This corresponds with the fluctuating values of the UHI for $n = 1$ to 25 in Figure 2. Apparently this “geographical” effect is more dominant in the two-dimensional case because CP1 and FE1 behave exactly the same over the range $n = 9$ to 144 ($m=1$ to 12), while in Figure 2 this is only the case for $n(m) = 1$ to 8.
Second, Figure 5 shows more or less the same stability properties of the four models as Figure 2 but convergence to stability starts to appear for much larger models than in the one-dimensional case. FE2 becomes stable for \( n > 100 \). For the other three convergence (if any at all) should not be expected for values of \( n < 1000 \). The intuitive explanation of this different behavior is that agglomeration forces have more degrees of freedom in two-dimensional space than on a horizontal line. Because centripetal and centrifugal forces have their influence on manufacturing labor in other locations in all directions, more spatial hierarchy configurations become possible. The stable value of the UHI to which the FE2 model in Figure 5 converges is 19, which is ten times higher than the stable value of 1.9 found in Figure 2.

Finally, how do the results look like when asymmetry is entered in this model? As discussed in section 3.1, a model of type 1f in Figure 1 is the most simple form of an asymmetric two-dimensional space. Figure 6 shows how increasing size is interpreted in this case. For \( m = 3 \) the upper right location is deleted which makes \( n = 8 \). Next, for \( m = 5 \) the three upper right locations are removed which makes \( n = 21 \) etc. As in the previous case a series of 30 models was created for \( m = \{3, 5, 7, \ldots, 57\} \) which now in total size go up to the largest model of 2465 locations. The results given in Figure 7 show approximately the same fluctuations over the range \( n = 400 \) to \( 900 \) as in Figure 5. Note that the steep jumps in the range of small models due to even and uneven numbers have disappeared because \( m \) only increases over uneven numbers. FE2 is again by far the most stable model but on average all UHI values are 40% or more higher than in
the symmetric case (see table 1). The size range where the models start to become stable is also higher. FE2 reaches stability around \( n=200 \) and for FE1 and CP2 stability, or at least fluctuations around some stable level, starts to emerge beyond \( n=1500 \). This did not show up in Figure 5 were the 30 simulations did not go beyond \( n=900 \). CP1 shows no stability at all.

### Table 1. Average value of the Urban Herfindahl Index

*Average over the 15 largest models of Figure 5 and Figure 7*

<table>
<thead>
<tr>
<th>Model</th>
<th>CP1</th>
<th>CP2</th>
<th>FE1</th>
<th>FE2</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) symmetric type 1d (fig 5)</td>
<td>132</td>
<td>42</td>
<td>69</td>
<td>19</td>
</tr>
<tr>
<td>b) asymmetric type 1f (fig 7)</td>
<td>183</td>
<td>74</td>
<td>96</td>
<td>26</td>
</tr>
<tr>
<td>ratio ( b/a )</td>
<td>1.4</td>
<td>1.8</td>
<td>1.4</td>
<td>1.4</td>
</tr>
</tbody>
</table>

### 4. Increasing the number of regions: an empirical example for the Netherlands

As mentioned in section 2, the analysis has been repeated with a real model of the agglomeration structure of the Netherlands. We have used employment data of 2007 for the 443 Dutch municipalities as a proxy of the initial agglomeration structure \( \lambda_0 \). The distance matrix \( D \) between the municipalities has been calculated in kilometers over the road network of 2009 using Google maps. Apart from reasons of data availability, the Netherlands was chosen because of its relatively homogeneous city distribution without one or two dominant agglomerations (See figure 8). Only going beyond the four largest cities Amsterdam, Rotterdam, The Hague and Utrecht creates some discontinuity in the rank-size distribution.

Typically, at this low spatial level every agglomeration over 5000 inhabitants - here referred to as cities - is represented by one of these municipalities. In the range of the 200 largest cities (> 25,000) the size of the model was increased as \( n, n+k, n+2k \) etc. with \( n=10 \) and \( k=10 \) resulting in 20 simulations. The lower part of the tail has been included by four extra runs with \( n=250, n=300, n=350 \), and finally the whole distribution with \( n=443 \). In stead of fixed parameters in CP1, CP2, FE1 and FE2, all 24 simulations were calibrated separately for the CP model and the FE model\(^{13}\).

---

\(^{13}\) The calibration was done on a set of \( 20^3 \) cross combinations of values for \( \delta, \varepsilon \) and \( T \) varying around the default values \( \delta=0.4, \varepsilon=4 \) and \( T=1.25 \).
Figure 8. The urban structure of the Netherlands 2007

Figure 9. Calibration ratios for the ten largest cities in simulations of increasing size
Calibrated city size/real city size
Figure 9 shows the ratio between the calibrated city size and the real city size for the ten largest cities plotted against the size of the model, ranging from the top 10 to the top 200 cities. In the range of the small models, the CP model has calibrated values for the four largest cities that are 5% to 10% below their real size while the 6 smaller cities are slightly larger than in reality. For the larger models all top 10 cities are calibrated below their real size, of which the three largest cities Amsterdam, Rotterdam and the Hague are 15%-20% too small. The FE model shows a similar trend but much more moderate. This is a typical result for the Netherlands because of its satellite structure of four relatively large cities close to each other. In order to match the real distribution as much as possible the calibrated parameter space must have a relatively weak agglomeration tendency because otherwise the attraction forces between these four cities become too strong which leads to the collapse of one of them into the other14.

With all 48 models ready on the test bank (24 CP models and 24 FE models) we can now turn to the key question of this paper: to what extent do models with few locations behave differently from models with many locations when we have to evaluate policy? For this purpose we assume a very simple policy that touches upon all regions: an (additional) fuel tax which leads to a rise of gasoline prices of 10%. This policy is implemented by raising the distance matrix $D$ to $D^*=1.1^*D$ and run each calibrated model one more time again with $D^*$ in stead of $D$ as the counterfactual. The resulting city distribution is then compared with the original one which gives the change in the agglomeration structure due to this policy.

As is well known from the NEG literature, a rise or fall of transport costs does not have unambiguous effects on the agglomeration equilibrium (Baldwin et. al., 2003). Whether it will lead to more concentration or spreading depends on the interplay of $D$ (and implicitly $T_{nn}$) with the other calibrated values for $\delta$ and $\varepsilon$ and with the initial distribution $\Lambda$. Also a specific geography may lead to spreading in one part of the country and concentration in other parts. Figure 8 shows the isolated position of the 7th city Groningen in the North which - contrary to Amsterdam, Rotterdam, The Hague and Utrecht - has no competitors of comparable size in it direct surroundings.

Figure 10 shows the results of the 24 CP models for the ten largest cities which are part of all 24 models. The same setup is given in Figure 11 for the 24 FE models. Because of the large scale difference between the top10 to top50 simulations and others in Figure 10, their rescaled results are given in an enlarged inlay. The two smallest models (top10 and top20) predict a concentration effect of the simulated policy because the three largest cities Amsterdam, Rotterdam and the Hague gain relative to the other 7 cities, but this picture is reverse in the top30, top40 and top50 model. In these model runs all three cities lose including the 4th city Utrecht, even the 5th city Eindhoven in the top50 model. Next, in the range top60-top90 the picture again turns around with a positive policy effect on most of the largest cities. In the middle range model from top90 to top140 the results remain relatively stable and are about the same is for the top30 to top50 models: a small decline of the ten largest cities with the two largest cities Amsterdam and Rotterdam losing more than the others. If we increase the model further the results become very volatile with major jumps when going from the top140 to the top150 and top160 model. Including the last four parts of the tail again changes the picture a lot, with only the top350 model showing a result comparable to the stable results in the middle range top90-top140.

A researcher who is familiar with the Dutch geography and urban distribution would be tempted to find geographical ad hoc explanations for the volatility in Figure 10. For example, most of the larger cities are located in the center and western part of the country. Going from the top20 to the top30 model then basically means more nearby competition for the top 10 cities, which may explain the switch from positive to negative effects for Amsterdam, Rotterdam and The Hague at that point. The opposite effect could appear for the large models top160-top190 where all new players are peripheral cities which might lead to an increasing counter balancing effect of the North, East and South relative to the West. To put it more general: the jumps in Figure 10 could be due to discontinuous expansions of the economic plane because every next ten cities may accidently be located in specific spatial clusters.

14 Other satellite structures like the Ruhr area in Germany are also difficult to approach model with NEG models, See Stelder (2005b).
Figure 10. Policy effects of a 10% increase of gasoline prices in CP models of increasing size
% change of city size for the ten largest cities

Figure 11. Policy effects of a 10% increase of gasoline prices in FE models of increasing size
% change of city size for the ten largest cities
There are two reasons why we believe this is not true. First, a sensitivity analysis did not reveal any spatial specific clustering of additional cities. More important, each model has been calibrated separately before simulating the policy effects. It is therefore not strictly correct to interpret every next model in Figure 10 as being the same as the former one with only more cities. It is the sensitivity of every separately calibrated equilibrium to the policy that changes.

Figure 11 gives another argument why the discontinuity in Figure 10 may not have much to do with geographical discontinuities: all switches from concentration to spreading that the CP model produces do not show up with the FE model. Instead, the behavior of the FE model changes very gradually, with the top50 as the only major exception. In all models Amsterdam, Rotterdam and Utrecht lose some share in the distribution. In the small models up to the top50 model other top10 cities gain but less in larger models. From the top50 onwards, the adding of smaller cities leads to the relatively stable result that larger cities loose so one could safely conclude that the general effect of the policy is spreading. It is not until we add the three last tails of the distribution that the results starts to be reverse: Only the top300, the top350 and the full top433 models predict concentration. This is probably due to increased competition between many small cities in the tail of the distribution which in its turn gives the larger agglomeration an advantage. We have seen a comparable situation in the results of model CP1 and FE2 showed earlier in Figure 3: two competing centre locations keep each other small so other locations can grow larger.

Finally, another difference between the CP model and the FE model to be noted here is that the order of magnitude of the predicted effects in Figure 11 are about a 100 times smaller then in Figure 10. This is not surprising because our simulations in section 3 also showed that the CP model leads to higher UHI values than the FE model. Indeed, a relative decline of 2% for the city of Amsterdam predicted by the CP top160 model due to a 10% rise of gasoline prices seems not very realistic. The FE top160 model predicts a decline of 0.015%.

5. Concluding remarks

Despite different degrees of homogeneity of the (additional) cities, the robustness analysis in this paper shows a remarkable similarity between abstract and real agglomeration models: their is a minimum number of cities at which stability starts to appear. Cutting off the tail of the distribution before that level leads to volatility and the possibility that an n-city model produces conclusions that contradict conclusions of an n+1 city model. For the FE model this threshold value is remarkably the same in both the abstract models as in the real urban model for the Netherlands: around a 100 regions in the two-dimensional abstract models and around 80 regions in the real model.

Because of their geographical and player homogeneity, the abstract models can be extended to thousands of cities without changing the outcome. This is not true for the real model because of the heterogeneity of the city distribution. Adding the very lower ends of the tail reveals that there is also a maximum number of cities worth modelling. For the Netherlands this maximum falls between 250 and 300 cities, or 85% to 90% of the distribution. In other countries this value will probably be different depending on their agglomeration level. As a comparison: in the USA the 85%-90% corresponds with the 163rd and 193rd city respectively.

The analysis in this paper also shows that it matters which economic model is used. In the abstract models a CP model with a modest concentration parameter configuration reaches less stability than the FE model and at higher numbers of cities. In the real implementation its behavior remains unpredictable over the whole range of cities.

To our best knowledge, most spatial agglomeration models are implemented at spatial levels way below a 100 regions. We do not wish to claim here that a 100 locations should be the minimum for every model. What we would suggest is that when building an empirical model of this kind, a sensitivity analysis on its optimal spatial resolution is crucial.

15 With GIS for every additional cluster of ten cities the gravity point was calculated which was always located within 50 miles of the centre city of Utrecht. The lower right panel of Figure 8 also shows that there is only a major break in the size distribution after Utrecht (nr4) and one other smaller one after Den Bosch (nr 18 ).

16 Calculated from http://www.citypopulation.de/USA-Cities.html
Annex A. An Urban Herfindahl Index for comparing urban hierarchies of different size

Let an urban structure $U = \{u_1, \ldots, u_N\}$ be defined by $u_i = \frac{p_i}{\sum_i p_i}$ with $p$ being population or any other absolute size indicator. Then the Herfindahl Index $HI = \Sigma_i (u_i)^2$ can be interpreted as a measure of agglomeration ranging from 1 (full agglomeration into 1 city) to $1/N$ (no agglomeration: all cities are of equal size). Because the minimum of $HI$ depends on $N$, two distributions $U_1$ and $U_2$ can not be compared unless $N_1 = N_2$. The Normalized Herfindahl Index $NHI$ solves this problem by subtracting the minimum $1/N$ and rescaling the interval back to $(0,1)$ with $NHI = \frac{1}{1-1/N} (HI - 1/N)$.

The correct interpretation of the NHI, however, is that it really measures market concentration for different numbers of firms. If one firm achieves 75% of a market with just one competitor taking the remaining 25%, market concentration is considered to be higher than in a 4-firm situation where two pairs of firms take each 37.5% and 12.5%. Table 1 shows that the NHI declines from 0.250 to 0.083, and further down to 0.036 in the comparable 8-firm situation.

If a distribution is interpreted as an urban hierarchy, however, the NHI does not give identical values for identical agglomeration structures. Figure A1 shows the three distributions of table 1 on a horizontal line market with locations at a constant distance from each other. The distinctive property of this urban hierarchy is an alternating pattern of a first order city and a second order city at distance 1. A correct urban agglomeration index would have to be invariant to whether we define the urban hierarchy over locations 1-2, 1-4 or 1-8. This is achieved by deleting the rescaling of the maximum value to 1 in NHI which gives us an Urban Herfindahl Index $UHI = \frac{N}{1-1/N} (HI - 1/N)$. In this example, the UHI has a value of 0.250 invariant to $N=2$, $N=4$ or $N=8$.

As the NHI, the UHI has a minimum of zero when all cities are of the same size but its maximum is $N-1$. This follows our intuitive interpretation: suppose we start to travel from a large city at location 1 to the right along the horizontal line. Then, the more “empty” locations we pass, the more impressive the agglomeration at location 1 becomes. If all locations 2-8 are empty, the UHI over the whole space covering 1-8 is 7. The interpretation of this maximum is straightforward: one player has become a monopolist winning a game with 7 other players. Contrary to the NHI which is always 1 for monopoly, the UHI indicates that winning a game with more players makes the winner more competitive.

A normal application of the UHI will be to compare two urban size distributions $U_1$ and $U_2$ with $N_1 = N_2$ regardless of their spatial pattern. Table A1 shows that alternatives like the Zipf coefficient $\beta$ or Gini coefficient $\gamma$ have the same problem as the NHI. An identical UHI will only imply an identical $\beta$ or $\gamma$ when $N_1=N_2$.

---

17 An “empty” city with $u_i = 0$ is comparable with a firm with a zero market share. Such a firm exists selling other products or selling the product on other markets. Empty cities play a role in New Economic Geography models as locations with agricultural employment, but with a zero share in national manufacturing employment.

18 The standard Zipf regression equation is $\log(R_i) = \alpha - \beta \log(P_i)$ with ranking number $R_i$ and size $P_i$ for $i=1..N$. The coefficient $\beta$ changes with $N$ unless the urban hierarchy follows the pure Zipf distribution with $\beta=1$. As $N$ increases, the cumulative distribution of $U$ converges to a continuous curve which makes the Gini coefficient $\gamma$ also $N$-dependent.
**Table A1. Measures of concentration in markets of different size**

<table>
<thead>
<tr>
<th>% market share</th>
<th>2</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>city/firm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>75.0</td>
<td>37.5</td>
<td>18.8</td>
</tr>
<tr>
<td>2</td>
<td>25.0</td>
<td>12.5</td>
<td>6.3</td>
</tr>
<tr>
<td>3</td>
<td>37.5</td>
<td></td>
<td>18.8</td>
</tr>
<tr>
<td>4</td>
<td>12.5</td>
<td></td>
<td>6.3</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td>18.8</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td>6.3</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td>18.8</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td>6.3</td>
</tr>
<tr>
<td>HI</td>
<td>0.625</td>
<td>0.313</td>
<td>0.156</td>
</tr>
<tr>
<td>NHI</td>
<td>0.250</td>
<td>0.083</td>
<td>0.036</td>
</tr>
<tr>
<td>UHI</td>
<td>0.250</td>
<td>0.250</td>
<td>0.250</td>
</tr>
<tr>
<td>Zipf $\beta$</td>
<td>1.585</td>
<td>0.908</td>
<td>0.838</td>
</tr>
<tr>
<td>Gini $\gamma$</td>
<td>0.500</td>
<td>0.333</td>
<td>0.286</td>
</tr>
</tbody>
</table>

**Figure A1. Identical agglomeration in systems of different size**

Absolute population size on a horizontal line market

---

**N = 2**  
**N = 4**  
**N = 8**
References


