Do Dutch museums compete or cooperate?*

Thomas de Graaff [a]†, Jaap Boter [c], Jan Rouwendal [a,b],

[a] *Department of spatial economics, Faculty of Economics and Business Adminstration*
Vrije Universiteit, Amsterdam, The Netherlands

[b] *Tinbergen Institute Amsterdam*

[c] *Department of marketing, Faculty of Economics and Business Adminstration*
Vrije Universiteit, Amsterdam, The Netherlands

March 31, 2006

PRELIMINARY VERSION; DO NOT QUOTE

Abstract

Museums can serve as important magnets for attracting tourists to a city. To decide how many and which museums to fund, it is important for city planners to understand what different types of museums there are in attractive power and how this attractiveness may be interdependent on the presence of other museums. To this end, we model a generic distance decay function for all museums allowing for spatial dependence between museums to account for local competition or synergy effects. To account for heterogeneity within our sample of museums, we first adopt a spatial two error component model. Thereafter, we model the variation between museums explicitly by segmenting the museums using a finite mixture approach. We illustrate the application of this model using a unique transaction database with the visiting behavior of 80,821 museum cardholders to 108 Dutch museums, we are able to calculate market shares of each museum in all 484 Dutch municipalities. Preliminary results indicate a large variation in the effect of distance on market shares and in spatial dependence between museums.

*The authors are very grateful to the Dutch Museum Association for supplying the dataset.
†Corresponding author. Tel +31 (0)20 5986092. Email: tgraaff@feweb.vu.nl.

1 Introduction

Over the years, there has been a steady interest into the economic role of cultural amenities in urban welfare. Under changing names such as 'economic impact studies' in the 1980s (e.g., national endowment for the arts 1981, Chartrand 1984), the rise of 'city marketing' in the 1990s (e.g., Kearns and Philo 1993, Ward 1998), and the concept of 'creative cities' as the most recent guise (e.g., Landry 2000, Florida 2002), authors have investigated and/or propagated the economic contribution cultural organizations can make.

Museums are a prime example of the potential role cultural organizations may have for cities. For its inhabitants, museums serve as an important amenity for leisure pursuits (Eurobarometer 2002, national endowment for the arts 1998). Many of the larger or 'Superstar' museums also serve as magnets for attracting large crowds of tourists (Frey 1998). For instance, Bilbao, Spain, is known for its new Guggenheim Museum; The Louvre and Musée d'Orsay are important Parisian attractions; and the major tourist highlights of Amsterdam include the Van Gogh Museum, the Rijksmuseum with its large collection of Rembrandts and the Anne Frank House. Although highly appreciated by the public at large, many museums are unable to survive in an open market; admission fees and donations rarely cover operational costs. Therefore, it is up to city governments to make planning decisions on how many and what museums are required and should consequently receive public funding.

Particularly in the last few years, a number of studies have tried to determine the appropriate level of funding in these cases (see Navrud and Ready 2002, for an overview). Rationale is that public funding of any cultural organization or object is justified as long as it does not exceed its economic value. Most of these studies rely on stated preference techniques, in particular contingent valuation techniques. General approach is to put a (hypothetical) situation, such as a restoration or extra grants, to respondents and to ask how much the respondents are willing to pay in taxes or donations. Applications have covered a variety of cultural organizations, such as the National Museum of Sculpture in Valladolid, Spain (Sanz et al. 2003), the Napoli Musei Aperti in Italy (Santagata and Signorello 2000), or Lincoln Cathedral in England (Pollicino and Maddison 2001). Revealed preference techniques, on the other hand, look at behavioral data, usually the travel behavior of visitors. Commonly, travel distances of visitors are multiplied by a wage rate and summed as an expression of willingness to pay. Travel cost applications are less common, but examples include The Quebec Musée de la civilisation in Canada (Martin 1994), the historic St. Mary's City site in Maryland, United States (Poor and Smith 2004), and a comparison of the relative worth of multiple museums in The Netherlands (Boter et al. 2005).

In reviewing the literature, two observations can be made. First, none of these studies distinguish between the potentially different spatial reach of museums. The summed willingness to travel may be equal for two museums, but one based on a few tourists from far away and another based on a large group of local residents. From a planning perspective, it would be important to understand more precisely which museums have what function: which museums primarily serve a local or regional community and which museums are better at attracting visitors from elsewhere. In other words, what is the decay function in attractiveness when moving away from a museum and are there particular types of museums in terms of different types of decay functions? Second, in judging the appropriate level of funding none of the studies consider the potential interdependencies between museums. With the exception of Boter et al. (2005), all studies look into the value of single organizations. However, as museums in a city are likely to compete or strengthen each other's position, the attractiveness of a museum may be dependent on other museums. Insight into such interdependencies may help avoid suboptimalization in planning and funding decisions. The aim of this paper is to develop a model that addresses these two issues.

The remainder of the paper is as follows. First, we develop a model dealing with market shares of museums, that is able to deal with different types of decay functions and dependence between museums. We then illustrate our model by using Dutch transaction data on visiting behavior of a large number of card holders. The last section concludes.

2 Modeling spatial dependence and market segments

The problem at hand focuses on the spatial reach of market shares of museums. Basically, we are interested in the amount of visitors specific museums attract over a certain distance. We specifically focus on two main extensions of this problem definition. First, we analyse whether the distance between museums matters as well. In other words, do museums reinforce each other in terms of market shares (agglomeration or synergy effects) or is there fierce local competition between museums? Secondly, we want to identify homogeneous groups of museums, which display more or less similar behavior in terms of the relation between distance and market shares. The first subsection displays our basic model of market shares explained by a distance decay function. The second subsection continues with modeling spatial dependence and shows how to capture heterogeneity among museums. The last section deals with filling in this heterogeneity by identifying homogeneous groups of museums.

2.1 A distance decay function for market shares of museums

First, assume that a country contains M museums. Typically, these museums are not uniformly distributed over a country, but display clustered patterns, especially in particular cities. Moreover, assume that this country can be subdivided in R regions or municipalities. Some of these regions contain a museum, others do not. In this setting we are able to define market shares, y_{mr} for museum m ($m \in \{1, \ldots, M\}$) in region r ($r \in \{1, \ldots, R\}$) as the percentage of all visitors of museum m that come from region r. Besides all sorts of regions and museum specific characteristics, a major determinant of the market share y_{mr} is most likely the travel time/distance between museum m and region r (denoted by d_{mr}). Usually, this relation is not linear but exponentionally decreasing instead. Thus, this distance decay function can in general form be expressed as:[1]

$$y = e^{\alpha - \beta d + \epsilon}. \tag{1}$$

where α and β are parameters, and ϵ denotes an *i.i.d.* residual term. Assuming a normal distribution for ϵ yields:

$$\ln y \sim N(\alpha - \beta d, \sigma^2). \tag{2}$$

Note that in this specification, the variable y may contain many zeros (conditional on the spatial scale of the regions). This is essentially a measurement error: because demand in region r for museum m may be very low, the market share may be measured as a zero, although it is in fact small, but positive. Apart from that, museums themselves are not identical, and usually vary widely in terms of size, exposition, focus, geographical location, etcetera. Therefore, we add initially a separate error component for each museum: $\epsilon_{mr} = \mu_m + \nu_{mr}$. Thus, μ_m describes the variation between museums in market shares and ν_{mr} describes our basic assumption; market shares are lognormally distributed.[2] Note that this means that $\sigma^2 = \sigma^2_\mu + \sigma^2_\nu$.

[1] We only use travel time here as an independent variable. Obviously, this relation can be straightforwardly extended with region and museum specific variables.

[2] Ideally, one would like to introduce another error term to capture the measurement error in the market shares. For reasons of clarity we omit this extention.

2.2 Spatial dependence between museums

If market shares of museums are spatially correlated – that is, museums are locally competitive or complementary to each other–, then it is very likely that the error term, ϵ, and the vector of market shares, $\ln y$, are correlated. Thus, ordinary regression yields biased results. To account for such spatial dependence between museums, we assume that market shares between museum A and B are related relative to the inverse distance, $1/d_{AB}$, between museum A and B. Doing this for all museums $1, \ldots, M$, we end up with an inverse distance matrix $\mathbf{W}_M$ with size $M \times M$ and with zeros on the diagonal. Using the assumed distribution of equation (2), this yields for the market shares of all museums $1, \ldots, M$ in region r the following spatial lag model:

$$\ln y_r = \lambda \mathbf{W}_M \ln y_r + \alpha - d\beta + \epsilon_r = (\mathbf{I}_N - \lambda \mathbf{W}_M)^{-1}(\alpha - d_r\beta + \epsilon_r), \tag{3}$$

where the parameter vector is thus $\phi = (\lambda, \alpha, \beta, \sigma_\mu, \sigma_\nu)'$.[3] For later purposes, we denote the density function of $\ln y$ as $f(\ln y|d, \phi)$.

Note that specification (3) resembles an error component model. To capture all market shares in all regions, we use specification (3) R times and come to the following expression:

$$\ln y = (\mathbf{I}_R \otimes \mathbf{A})^{-1}(\alpha - d\beta + \epsilon), \tag{4}$$

where $\ln y$ is now of dimension $RM \times 1$, α is $RM \times 1$, d is $RM \times 1$, β is 1×1 and ϵ is $RM \times 1$. $\mathbf{I}_R$ denotes the identity matrix with size $R \times R$, $\otimes$ is the kronecker product, and $\mathbf{A} = \mathbf{I}_N - \lambda \mathbf{W}_M$. Conform Baltagi et al. (2003), we order our observations with r being the slow running index and m the fast running index, i.e., $y' = (y_{11}, \ldots, y_{1M}, \ldots, y_{R1}, \ldots, y_{RM})$. The error term, ϵ, can now be rewritten as in Anselin (1988):

$$\epsilon = \nu + \iota_R \otimes \mu, \tag{5}$$

with ι_R a vector of ones with its index indicating its order. Note that the additional error component is observed along the slow running index, r, which – in combination with a spatial lag model – creates a full variance covariance matrix with size $RM \times RM$. Under these assumptions this variance covariance matrix takes the following form:

[3]This implies that the original model now reads as: $y_{mr} = \prod_{i \neq m}^{M} y_i r^{\lambda w_{mi}} e^{\alpha - \beta d_{mr} + \epsilon_{mr}}$, where w_{mi} denotes the inverse distance between museum m and museum i and λ a spatial interaction parameter. This specification resembles that of hedonic pricing, and indicates that when $\lambda > 0$ museums may benefit from the near presence of others and when $\lambda < 0$ museums face stiff competition from other museums near by.

$$\begin{aligned}\mathbf{\Omega} &= \sigma_\nu^2 \mathbf{\Psi} &(6)\\ \mathbf{\Psi} &= \iota_R \iota_R' \otimes \frac{\sigma_\mu^2}{\sigma_\nu^2}\mathbf{I}_M + \mathbf{I}_R \otimes \mathbf{I}_M, &(7)\end{aligned}$$

So, that we may come to the following loglikelihood function, apart from a constant:[4]

$$\ln L = -\frac{RM}{2}\ln(\sigma_\nu^2) + R\ln|\mathbf{A}| - \frac{1}{2}\ln|\mathbf{\Psi}| - \frac{1}{2\sigma_\nu^2}e'\mathbf{\Psi}^{-1}e, \tag{8}$$

with

$$e = (\mathbf{I}_R \otimes \mathbf{A})\ln y - (\alpha - d\beta) \tag{9}$$

Although $\mathbf{\Psi}$ has a rather simple structure, its sheer size $(RM \times RM)$ makes it oftentimes uncomputational. Therefore, we adopt a classical trick from Wansbeek and Kapteyn (1982). First, let $\mathbf{J}$ be a square of ones, its index indicating its order and let ω be $\sigma_\mu^2/\sigma_\nu^2$. Subsequently, define $\overline{\mathbf{J}}_\mathbf{R} \equiv \mathbf{J}_R/R$ and $\overline{\mathbf{J}}_\mathbf{M} \equiv \mathbf{J}_M/M$. Note that these matrices are idempotent. Finally, define the following centering operator $\mathbf{C} \equiv \mathbf{I} - \overline{\mathbf{J}}$, where the corresponding index indicates it order. Now, we are able to rewrite the covariance matrix as:

$$\begin{aligned}\mathbf{\Psi} &= \mathbf{I}_R \otimes \mathbf{I}_M + \overline{\mathbf{J}}_R \otimes \omega \mathbf{I}_M\\ &= \left(\mathbf{C}_R + \overline{\mathbf{J}}_R\right) \otimes \left(\mathbf{C}_M + \overline{\mathbf{J}}_M\right) + R\omega\left(\overline{\mathbf{J}}_R \otimes \left(\mathbf{C}_M + \overline{\mathbf{J}}_M\right)\right)\\ &= \mathbf{C}_R \otimes \mathbf{C}_M + \mathbf{C}_R \otimes \overline{\mathbf{J}}_M + (1+R\omega)\left(\overline{\mathbf{J}}_R \otimes \mathbf{C}_M\right) + (1+R\omega)\left(\overline{\mathbf{J}}_R \otimes \overline{\mathbf{J}}_M\right)\\ &= \sum_{i=1}^{4}\xi_i \mathbf{M}_i, &(10)\end{aligned}$$

where ξ_i and $\mathbf{M}_i$ are implicitly defined. Basically, equation (10) is a spectral decomposition, with ξ_i as the eigenvectors, as M_i are mutually orthogonal, symmetric idempotent and sum up to the unit matrix. This leads us to the very useful result that $\mathbf{\Psi}^a = \sum_{i=1}^4 \xi_i^a \mathbf{M}_i$, or:

$$\mathbf{\Psi}^{-1} = \sum_{i=1}^{4}\xi_i^{-1}\mathbf{M}_i, \tag{11}$$

and for the determinant finally yields:[5]

---

[4]We used here a well-known determinant property (see, e.g, Magnus and Neudecker 1988). Namely, it is easy to see that the Jacobian is equal to $|\Psi^{-1/2}(\mathbf{I}_R \otimes \mathbf{A}))|$, which can be decomposed in $|\Psi|^{-1/2}|\mathbf{A}|^R$.

[5]This property immediately follows from the properties of the $\mathbf{M}_i$ matrices. So, the determinant can be

$$|\mathbf{\Psi}| = (1 + R\omega)^M, \tag{12}$$

which leaves us with the complete (and computationable) likelihood.

2.3 Creating market segments by finite mixture modeling

Estimation of specification (3) and interpretation of the parameters is only meaningful when the sample of museums is homogeneous. Usually, however, this is not the case given the wide range in types of museums around, such as art, handicraft, or (natural) history museums. Therefore, we divide our population of museums in a, a priori unknown, number of different subpopulations of museums. Thus, assume that observations on $\ln y$ arise from a population that is a mixture of S segments in proportions $\pi_1, \ldots, \pi_S$, where we do not know in advance from which segment observations on $\ln y$ arise. Then, the unconditional density function as expressed above can now be decomposed in its various segments as follows:

$$f(\ln y|d, \theta) = \sum_{s=1}^{S} \pi_s f_s(\ln y|d, \phi_s), \tag{13}$$

where the vector of parameters is now denoted for each segment s as $\theta_s = (\lambda_s, \alpha_s, \beta_s, \pi_s, \sigma_\phi)'$. Thus, the distance-decay and spatial dependence parameters are assumed to be segment specific, where the variance term is common to all segments.[6]

To estimate the loglikelihood of (13), we apply the EM-algorithm, made popular by the seminal contribution of Dempster et al. (1977). The EM-algorithm was oginally constructed to deal with missing observations and proceeds as follows (see Wedel and Kamakura 2000, for a more detailed description). First, introduce unobserved data, z_{ms}, indicating whether observation vector $y_m = (y_{ms})$ on museum m belongs to segment s. That is, $z_{ms} = 1$ if m comes from segment s and $z_{ms} = 0$ otherwise. The z_{ms} are assumed to be i.i.d. multinomial:

$$f(z_m|\pi) = \prod_{m=1}^{M} \pi_m^{z_{ms}}, \tag{14}$$

where the vector $z_m = (z_{m1}, \ldots, z_{mS})'$. Now denote the matrix $(z_1, \ldots, z_M)$ by $\mathbf{Z}$ and the vector f_m as the conditional distribution of $\ln y_m$ given z_m. With z_s considered as missing data, the complete loglikelihood function can be formed as follows:

calculated as: $|\mathbf{\Psi}| = \prod_{i=1}^{4} \xi_i^{\text{rank}(\mathbf{M}_i)}$.

[6]Because we want to explain the variation between museums we drop from now on the specific museum error term, σ_μ.

$$\ln L(\theta|\ln y, d, Z) = \sum_{m=1}^{M}\sum_{s=1}^{S} z_{ms} f_{m|s}(\ln y|d,\theta_s) + \sum_{m=1}^{M}\sum_{s=1}^{S} z_{ms}\ln \pi_s. \tag{15}$$

In the E-step, we first estimate the class probabilities for each museum (*cf.* Leisch 2004). Thus the probability that museum m belongs to segment s is:

$$\hat{z}_{ms} = \frac{\pi_s \prod_{r=1}^{R} f(\ln y_{mr}|d_{mr},\theta_s)}{\sum_{k=1}^{S}\pi_k \prod_{r=1}^{R} f(\ln y_{mr}|d_{mr},\theta_k)}. \tag{16}$$

The probabilities for each segment are now derived as:

$$\hat{\pi}_s = \frac{1}{M}\sum_{m=1}^{M}\hat{z}_{ms}. \tag{17}$$

Inserting $\hat{z}_{ms}$ and $\hat{\pi}_s$ now enables us to estimate the complete loglikelihood of (15) in the M-step. The E- and M-steps are repeated until the loglikelihood of (15) stops improving.

The actual number of segments is a priori unknown and must be inferred from the data. To this end we use information criteria, which balance the increase in fit against the larger number of segments – and thus more parameters – used. Basically, these criteria impose a penalty on the likelihood, which is related to the number of parameters estimated: $C = -2\ln L + P\rho$. Here, P is the number of parameters estimated and ρ is some constant, reflecting the penalty imposed on the likelihood. We use two widely used information criteria. The first one is the classical Akaike information criterion (AIC), where $\rho = 2$. The second one is the bayesian information criterion (BIC), where $\rho = \ln(N)$ (N = number of observations). Note that the last criterion usually penalize the likelihood more heavily.

The next section first decribes the data and then gives the results for both the (unsegmented) spatial error component model and the fully segmented spatial model.

3 Application

3.1 Data

The basis of our dataset is the transaction data of Dutch National Museum Card holders used by Boter et al. (2005). This Museum Card is an important tool in promoting museum attendance in The Netherlands. In return for an annual fee of €25 for adults or €12.50 for anyone younger than 26 years, card holders get free access to 442 museums in this country; the only remaining cost per visit being the cost of traveling. At the 150 largest participating museums, card holder visits are logged electronically. These data are collected and stored on a central server to aid reimbursement to the museums. The dataset provided by the organization was limited to customer number, type of card (youth or adult), the museum, the date and time of the visit, and the zip codes of both museum and visitor. Using a commercial GIS database that contains travel distance and travel time by road for every zip code combination in The Netherlands, travel distance and travel time were added to the dataset for each recorded visit. Similar to Boter et al. (2005), we only use the visits of one full year (2002) to exclude seasonal effects on demand. Also, museums with missing data or that faced incidental closure were excluded. The remaining 108 museums are a representative variety in size, type of collection and location.

To capture market areas, we calculate the market shares of the museums in each of the 484 municipalities in the Netherlands. Thus, our full dataset comprises 52,272 observations of market shares. However, as Table 1 shows, most of these market shares are zero (about 60%). This is a direct consequences of the small size of some of the Dutch municipalities, so that most of the (smaller) museums are never visited by any inhabitant of these small-sized municipalities. On the other hand, the size of this dataset has the distinct advantage that it captures a wide range of different museums, regions, competitive situations and travel distances. As Table 1 shows, on average, 717 citizens from each municipality are recorded to visit a museum. A preliminary analysis of the dataset reveals that within the common willingness to travel of 44.19 minutes, the average card holder has 29.5 out of the 108 museums to choose from. The museums visited are therefore likely to reflect a real utility to the card holder.

Table 1: Descriptive statistics

Number of museums participating (M)	108
Number of regions (municipalities) (R)	484
Number of cardholders in the dataset	80,821
Number of visits recorded in the dataset	346,978
Average number of visits per region	716.9
Percenage of non-zero observations	39.89
Average travel time from a region to a museum (in minutes)	97.12
Average observed travel time from cardholders to a museum (in minutes)	44.19

However, when applying the travel cost method to these data, some complications have to be taken into account. The model assumes that observed museum visits are the result of trips that have this visit as their single (or at least most important) purpose so that travel costs can indeed be regarded completely as (part of) the price of this visit. Moreover, it assumes that preferences for a particular museum are independent of the household's location.

Table 2 displays the top 10 museums in our database with the highest amount of visitors and with the highest average travel time, respectively.

Table 2: Top 10 museums by total number of cardholders and by average travel time of visiting cardholders

Museum	Visitors	Museum	Average travel time in min.
1) Rijksmuseum Amsterdam	34,236	A) Natuurcentrum Ameland	233.1
2) Stedelijk Museum Amsterdam	23,067	B) Industrion	130.3
3) Haags Gemeentemuseum	22,250	C) Bonnefantenmuseum	119.6
4) Groninger Museum	18,527	D) Zeeuws Biologisch Museum	117.8
5) Van Gogh Museum	17,301	E) Groninger Museum	101.7
6) Cobra Museum Amstelveen	12,540	F) Natura Docet Natuurmuseum	95.9
7) Singer Museum	11,343	G) Marine Museum	86.1
8) Mauritshuis	10,173	H) Fries Museum	80.4
9) Amsterdams Historisch Museum	9,580	I) Limburgs Museum	78.6
10) Joods Historisch Museum	8,695	J) Hannema-De Stuers Fundatie	78.0

As Figure 1 clearly shows, 9 of the 10 museums ranking highest in the number of visits are all located in or near the 'Randstad', the western, most densely populated area of the Netherlands, formed by the four largest cities of the Netherlands. In fact, 6 of the top 10 museums are all located in or very near the capital city of Amsterdam. As can also be observed from Figure 1, many of the cardholders live in the 'Randstad' area as well, so that many museums find it convenient to be based close to their largest market areas. However, the museums with highest travel time are all located in the periphery of the country. This might reflect the large willingness to travel of a particular group of cardholders to these museums. E.g., many cardholders find it worthwhile to travel to the city of Groningen to visit the 'Groninger' Museum (in both lists).

It might also reflect a different function of these museums. E.g., cardholders might only find it interesting to visit the 'Natuurcentrum Ameland', when they are already on the island for their holidays. Thus, accounting for heterogeneity, in the willingness to travel to a museum, or in the functionality of a museums is rather important when estimating market areas.

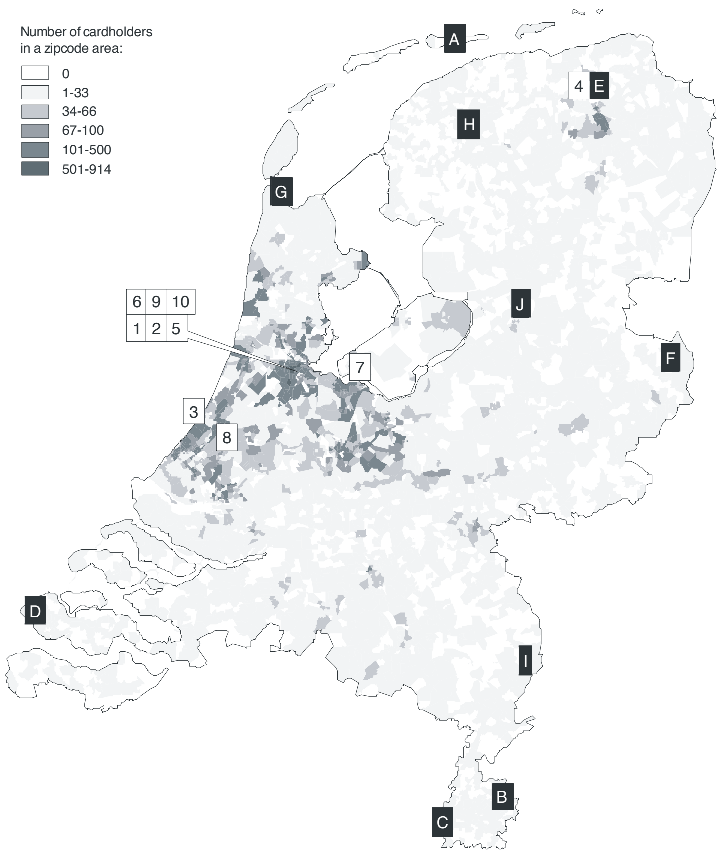

Figure 1: Number of Museum Cardholders by four-digit zip code area and locations of the museums in top 10 of both visitors and average travel time

3.2 Results

To analyse the spatial reach of Dutch museums, we first estimate the 'ordinary' distance-decay function (1). To capture spatial dependence we yhen estimate the loglikelihood of (8), with specific focus on the specific museum error term (σ_μ) as well. In the last part of this subsection, we then repeatedly estimate the loglikelihood of (15) to explain (part of) the specific variation in market shares between museums until the information criteria show that convergence has been reached.

3.2.1 Distance decay and spatial dependence

Table 3 present the results for the non-segmented case, both for the non-spatial (the 'ordinary' distance decay function) and the spatial case (with and without the museum-specific error term, σ_μ).

Table 3: Distance decay function of museum shares (N = 52,272)

	OLS		Spatial SUR		Spatial ECM	
Variable	Coeff.	St. err	Coeff.	St. err	Coeff.	St. err
α	-8.353	0.050	-2.880	0.076	-2.594	0.265
β	-0.032	0.000	-0.019	0.000	-0.018	0.000
λ			0.608	0.007	0.644	0.006
σ_ϕ	5.464	0.017	5.060	0.016	4.277	0.013
σ_μ					2.687	0.180
Mean Logl.		-0.407		-0.343		-0.182

Clearly, introducing spatial dependence has a significant impact on the distance decay function. Figure 2 shows the distance decay function ($e^{\beta d}$) for each of the three models shown in Table 3.

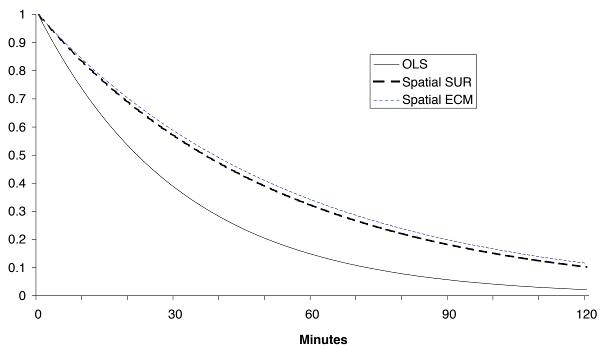

Figure 2: Distance decay functions

Regarding the highly positive value of λ ($\approx$ 0.6), it seems that there is a large correlation between spatial clustering and large market shares. The large amount of much visited museums in Amsterdam is an example of this phenomenon.[7] After the correction for spatial dependence,

[7]This does not directly mean that clustering of museums causes high market shares. But when one does not account for spatial dependence when analyzing these market shares, regression coefficients may be severely biased

the distance decay function remarkably flattens out, indicating that museums have a far larger spatial reach then assumed previously (Boter et al. 2005). The interpretation is clear. Taking into account spatial dependence reduces travel costs as perceived by the visiting cardholder. Thus, positive spatial dependence can also be seen as a positive spatial externality. More museums in a limited spatial area increases options for the visiting cardholder and the possibility to visit more museums per city visit.

The third model (Spatial ECM) introduces an additional error component. Because introducing the additional museum error component only affects efficiency, the coefficient values from the Spatial ECM model do not differ much from the spatial SUR model. However, the mean loglikelihood improves significantly, indicating a large amount of heterogeneity present, which can be fully allocated to the incorporation of the various museums in our dataset (Table 2 and Figure 1 already indicated a large amount of heterogeneity between museums). Thus, to obtain further insight in the spatial reach of more or less homogeneous groups of museums, we estimate the market shares for specific subgroups of museums by segmenting our data.

3.2.2 Finite mixture modeling with spatial dependence

The approach we adopt here is to expand the number of segment until the likelihood and the information criteria stop improving. Because we want to explain the variation between museums, we leave the specific museum error component (μ) out of the analysis.[8]

Table 4: Distance decay function of segmented museum shares ($N = 52{,}272$ and standard errors between parentheses)

	Number of segments				
Variable	2	3	4	5	6
Segment 1					
α_1	-4.873 (0.086)	-5.407 (0.091)	-2.741 (0.000)	-3.581 (0.083)	-2.775 (0.117)
β_1	-0.016 (0.000)	-0.014 (0.000)	-0.019 (0.000)	-0.018 (0.001)	-0.018 (0.001)
λ_1	0.594 (0.008	0.585 (0.008)	0.699 (0.005)	0.666 (0.008)	0.705 (0.010)
π_1	0.685	0.587	0.369	0.396	0.319
Segment 2					

continued on next page

as Table 3 clearly shows.

[8]Moreover, estimating a spatial error component model with large variance-covariance matrices using decomposition methods as in equation (10) is extremely cumbersome. Estimation time for the unsegmented case is about nine hours and increases exponentionally when the number of segments increases.

continued from previous page

	Number of segments				
Variable	2	3	4	5	6
α_2	1.418 (0.115)	0.771 (0.122)	-7.646 (0.109)	-8.414 (0.112)	-6.733 (0.134)
β_2	-0.018 (0.001)	-0.024 (0.001)	-0.010 (0.001)	-0.008 (0.001)	-0.012 (0.001)
λ_2	0.701 (0.011)	0.768 (0.011)	0.488 (0.010)	0.453 (0.011)	0.527 (0.012)
π_2	0.315	0.271	0.298	0.215	0.263
Segment 3					
α_3		1.430 (0.168)	1.247 (0.067)	0.629 (0.065)	0.815 (0.143)
β_3		-0.014 (0.001)	-0.022 (0.000)	-0.002 (0.000)	-0.021 (0.001)
λ_3		0.578 (0.016)	0.777 (0.006)	0.789 (0.005)	0.793 (0.012)
π_3		0.141	0.202	0.200	0.180
Segment 4					
α_4			1.365 (0.138)	2.271 (0.080)	2.230 (0.160)
β_4			-0.013 (0.000)	-0.022 (0.000)	-0.021 (0.001)
λ_4			0.569 (0.012)	0.684 (0.008)	0.680 (0.015)
π_4			0.132	0.161	0.154
Segment 5					
α_5				0.139 (0.080	-10.648 (0.298)
β_5				-0.012 (0.000)	-0.007 (0.002)
λ_5				0.262 (0.008)	0.328 (0.028)
π_5				0.028	0.056
Segment 6					
α_6					0.137 (0.358)
β_6					-0.013 (0.002)
λ_6					0.258 (0.038)
π_6					0.028
σ_ϕ	4.498 (0.014)	4.381 (0.014)	4.327 (0.013)	4.295 (0.013)	4.285 (0.013)
AIC	23,863	21,250	20,094	19,370	19,194
BIC	23,907	21,312	20,173	19,476	19,309
Mean Logl.	-0.228	-0.203	-0.192	-0.185	-0.183

Obviously, there is large variation present between the segments in the distance decay parameters and the spatial dependence parameters. Note that the mean loglikelihood converges to that of the spatial error component model, indicating that almost all variation that was fully allocated to the various museums is taken into account by creating six separate segments of museums. The distance decay functions for each segment are depicted in Figure 3.

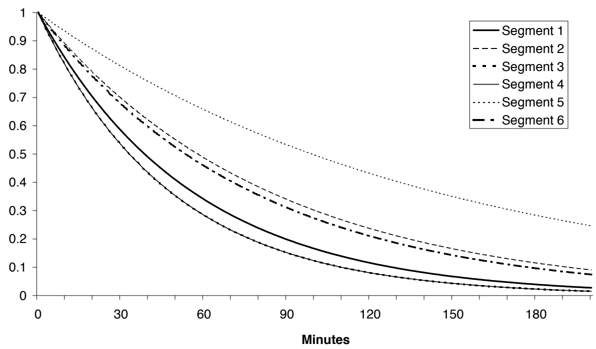

Figure 3: Distance decay functions for the various segments

Table A in appendix A offers the top 10 museums by estimated probability parameters, $\hat{z}_{ms}$, for each segment. On the basis of this classification, and the results in Table 4 and Figure 3 we can finally give the following interpretation to the six segments:

Segment 1 This is a large homogeneous group of smaller specialized museums, mostly centrally located in the 'Randstad' area and benefitting most of large museums close by (hence the large spatial interaction parameter of ≈ 0.7).

Segment 2 This segment closely remembles Segment 1, but the museums generally attract less visitors, mainly because they are located more in the periphery of the Netherlands. However, the distance-decay parameter is smaller, indicating that these small and specialized museums attract visitors from a larger distance.

Segment 3 This segment contains somewhat larger museums, which are mainly based in the 'Randstad', and which benefit highly from the larger museums in their vicinity (the spatial interaction parameter here is about ≈ 0.8).

Segment 4 This group closely resembles the museums in Segment 3. However these museum are usually large to very large, and include very influential museums as the Bonnefantenmuseum, Boijmans van Beuningen, Cobra museum, Stedelijk musuem and the van Gogh museum (which is ranked 11 with probability 0.99). Note that these museums still benefit significantly from each others presence.

Segment 5 This segment contains the smallest and very specialized museums. Usually, these museums can be found outside the 'Randstad'. Note the very small distance-decay parameter (≈ -0.007), indicating that these museums attracts visitors from large distances, which raises the suspicion that these museums are usually visited as part of a multi-purpose visit. The other purposes might include holidays, visiting friends and relatives, etcetera.

Segment 6 This segment only contains a few museums, which are ranked amongst the largest and most important museums in the Netherlands. Its distance-decay parameter is rather low, just as the spatial dependence parameter. The last coefficient indicates that these museums do not benefit anymore from other (smaller) museums closeby and that these museums are growing into what Frey (1998) describes as 'Superstar' museums.

4 Conclusion & further research

The aim of this study was twofold. First, we wanted to investigate the different spatial reach of museums. Thus, which museums primarily serve a local community and which museums have a much wider (national) reach to attract visitors? Secondly, we wanted to take spatial interdepencies between museums into account to analyse interdepencies between museums, which might result in clustering phenomena, such as can be witnessed in Amsterdam, Paris and London. The results of the analysis can be straightforwardly interpreted: namely, (i) not taking spatial dependence into account creates a downward bias in the distance decay function (because positive spatial dependence lowers perceived travel costs), and (ii) not correcting for heterogeneity in the characteristics of museums ignores specific behavior of groups of museums (such as the museums that display 'Superstar' museum characteristics).

However, as pointed out earlier, the dataset has – although large in size – some disadvantages; the most important one being the lack of information on the nature of the trip. Especially the assumption of single-purpose is a strong one. It might well be that cardholders combine museum visits in Amsterdam with a shopping-trip or use the opportunity to visit friends or relatives. Visits to museums in the periphery of the Netherlands might very well be related with (short) holidays in that region. So, the observed travel cost to museums does not necessarily reflect the real travel cost. Although the finite mixture approach partly corrects for these disadvantages, some bias in the travel cost measurement remain. Possible extensions to solve this problem is to take the date on which the cardholder visits a museum into account and to correct for possible multiple museum visits per day.

It might seem a bit of a conundrum why we immediately tackled heterogeneity by segmenting our data and not by first adding additional region – and especially – museum specific characteristics. Of course, a first extension of the paper involves adding additional data (with the usual suspects, such as size of the region/municipality, total amount of visitors per year for each museum, or dummies indicating the nature of the museum). However, this paper shows as well that without additional control variables, much of the variation can be tackled by the original data itself. A feature which is rather attractive when additional control variables are hard to find.

A Appendix

Table A: Top 10 museums by estimated parameter for each segment

Museum	$\hat{z}_{ms}$	Museum	$\hat{z}_{ms}$
Segment 1		*Segment 2*	
Verweyhal/De Hallen	0.994	Museum Gevangenpoort	0.969
Aboriginal Art Museum	0.993	Natuurmuseum Groningen	0.930
Nationaal Glasmuseum	0.992	Theater Instituut Nederland	0.922
Kasteel Groeneveld	0.990	Museum van het Nederlandse Uurwerk	0.921
Museum van het Boek	0.986	Nationaal Bevrijdingsmuseum '44–'45	0.919
Stedelijk Museum Zwolle	0.984	Fries Natuurmuseum	0.910
Allard Pierson Museum	0.981	Museum Kempenland	0.904
Muiderslot	0.978	Verzetsmuseum Amsterdam	0.901
Museum Mesdag	0.973	Natuurmuseum Rotterdam	0.900
Goud-, Zilver- en Klokkenmuseum	0.967	Het Nederlands Vestingmuseum	0.897
Segment 3		*Segment 4*	
Tropenmuseum	1.000	Bonnefantenmuseum	1.000
Stedelijk Museum De Lakenhal	1.000	Singer Museum	1.000
Nederlands Textielmuseum	0.999	Museum Catharijneconvent	1.000
Teylers Museum	0.999	Museum Boijmans Van Beuningen	1.000
Museon	0.999	Cobra Museum Amstelveen	1.000

continued on next page

continued from previous page

Museum	$\hat{z}_{ms}$	Museum	$\hat{z}_{ms}$
Museum Het Rembrandthuis	0.998	Paleis Het Loo Nationaal Museum	1.000
Nederlands Spoorwegmuseum	0.998	Joods Historisch Museum	1.000
Frisia Museum, Magisch Realisme	0.997	Amsterdams Historisch Museum	1.000
Bijbels Museum	0.997	Zuiderzeemuseum	0.999
Nederlands Architectuur Instituut	0.995	Stedelijk Museum Amsterdam	0.999
Segment 5		*Segment 6*	
Molenmuseum	0.991	Groninger Museum	1.000
Streekmuseum Crimpenerhof	0.959	Haags Gemeentemuseum	1.000
Stedelijk Molenmuseum De Valk	0.843	Rijksmuseum Amsterdam	0.934
Museum Beeckestijn	0.777	Mauritshuis	0.034
Mariniersmuseum der Koninklijke Marine	0.478	Van Gogh Museum	0.009
Zeemuseum	0.365		
Natura Docet Natuurmuseum	0.323		
Nationaal Schoolmuseum	0.304		
Historisch Museum Apeldoorn	0.198		
Techniek Museum Delft	0.122		

References

Anselin, L.: 1988, *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers, Dordrecht.

Baltagi, B. H., Song, S. H. and Koh, W.: 2003, Testing panel data regression models with spatial error correlation, *Journal of Econometrics* **117**, 123–150.

Boter, J., Rouwendal, J. and Wedel, M.: 2005, Employing travel time to compare the value of competing cultural organizations, *Journal of Cultural Economics* **29**, 19–33.

Chartrand, H.: 1984, An economic impact assessment of the canadian fine arts. Canada Council, Research and Evaluation Section, Ottawa, Ontario.

Dempster, A. P., Laird, N. M. and Rubin, D. B.: 1977, Maximum likelihood from incomplete data via the em algorithm, *Journal of the Royal Statistical Society B* **39**, 1–38.

Eurobarometer: 2002, *Europeans' participation in cultural activities*, Eurostat, Brussels/Strasbourg.

Florida, R.: 2002, *The rise of the creative class: and how it's transforming work, leisure, community and everyday life*, Basic Books, New York (NY).

Frey, B. S.: 1998, Superstar museums: An economic analysis, *Journal of Cultural Economics* **22**, 113–125.

Kearns, G. and Philo, C.: 1993, *Selling Places: The city as cultural capital, past, and present*, Pergamon Press, Oxford.

Landry, C.: 2000, *The creative city: a toolkit for urban innovators*, Comedia, Near Stroud.

Leisch, F.: 2004, Flexmix: A general framework for finite mixture models and latent class regression in r, *Journal of Statistical Software* **11**, 1–18.

Magnus, J. R. and Neudecker, H.: 1988, *Matrix Differential Calculus with Appliciations in Statistics and Econometrics*, John Wiley & Sons Ltd., Chippenham.

Martin, F.: 1994, Determining the size of museum subsidies, *Journal of Cultural Economics* **18**, 255–270.

national endowment for the arts: 1981, Economic impacts of arts and cultural institutions: Case studies in columbus, minneapolis/st. paul, st. louis, salt lake city, san antonio, springfield. National endowment for the arts, Washington D.C.

national endowment for the arts: 1998, 1997 survey of public participation in the arts. Rep. No. Research Division Report 39, National Endowment for the arts, Washington D.C.

Navrud, S. and Ready, R. C.: 2002, *Valuing Cultural Heritage; Applying Environmental Techniques to Historic Buildings, Monuments and Artifacts*, Edward Elgar, Cheltenham.

Pollicino, M. and Maddison, D.: 2001, Valuing the benefits of cleaning lincoln cathedral, *Journal of Cultural Economics* **25**, 131–148.

Poor, P. J. and Smith, J. M.: 2004, Travel cost analysis of a cultural heritage site: the case of historic st. mary's city of maryland, *Journal of Cultural Economics* **28**, 217–229.

Santagata, W. and Signorello, G.: 2000, Contingent valuation of a cultural public good and policy design: The case of "napoli musei aperti", *Journal of Cultural Economics* **24**, 181–204.

Sanz, J. A., Herrero, L. C. and Bedate, A. M.: 2003, Contingent valuation and semiparametric methods: A case study of the national museum of sculpture in valladolid, spain, *Journal of Cultural Economics* **27**, 241–257.

Wansbeek, T. and Kapteyn, A.: 1982, A class of decompositions of the variance-covariance matrix of a generalized error components model, *Econometrica* **50**, 713–724.

Ward, S. V.: 1998, *Selling Places: The marketing and promotion of towns and cities*, Routledge, New York (NY).

Wedel, M. and Kamakura, W.: 2000, *Market Segmentation: Conceptual and Methodological Foundations*, Kluwer Academic Publishers, Boston.