Spatial heterogeneity of willingness to pay for forest management

Mikołaj Czajkowski,

Wiktor Budziński, Danny Campbell, Urška Demšar, Marek Giergiczny, and Nick Hanley

czaj.org

References

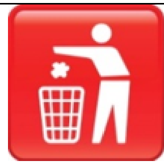
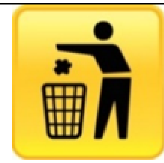
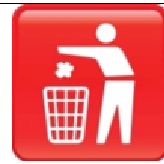
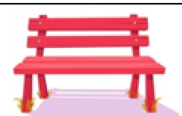
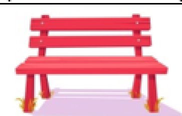
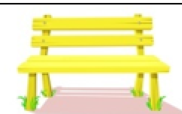
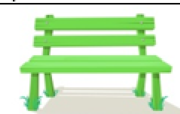
1. Czajkowski, M., Budziński, W., Campbell, D., Giergiczny, M., and Hanley, N., forthcoming. Spatial heterogeneity of willingness to pay for forest management. Environmental and Resource Economics.
2. Budziński, W., Campbell, D., Czajkowski, M., Demšar, U., and Hanley, N., Using geographically weighted choice models to account for spatial heterogeneity of preferences.

Spatial patterns of preferences for environmental goods

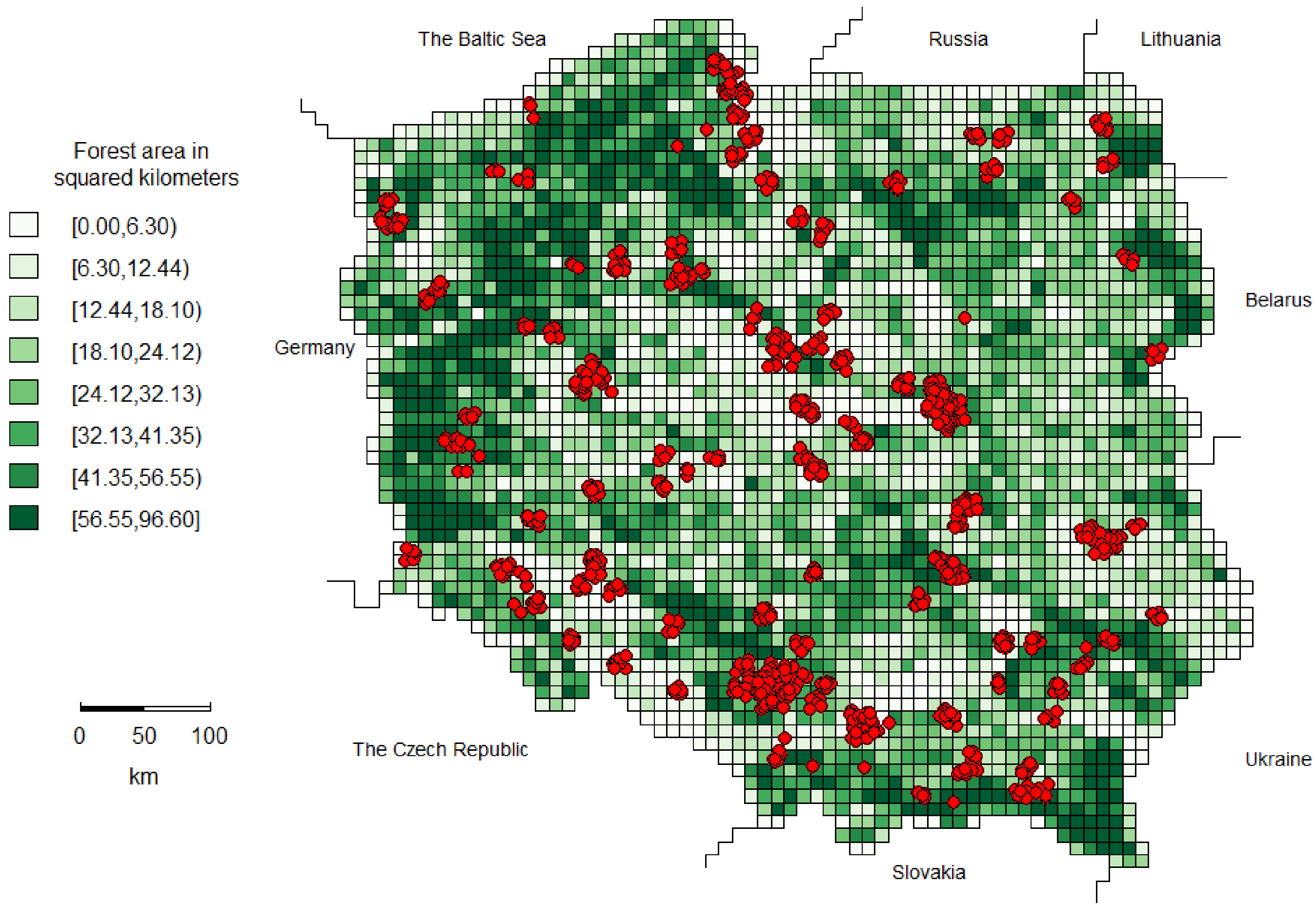
- Preferences for environmental goods likely to display spatial patterns
- Why?
  - Differences in the spatial configuration of goods
  - Availability of substitutes
  - Peoples’ preferences adapt to their local environments
  - Residential sorting
- So what?
  - Information on spatial distribution of preferences / welfare measures provides important information for improving the economic efficiency of land management
  - A source of observed preference heterogeneity that can be accounted for
- We want to be able to investigate spatial patterns in stated and revealed preferences for environmental goods
- ‘Traditional’ vs. ‘new’ method (geographically weighted choice models)

Case study – public preferences over management options for national forests in Poland

- Attributes used to describe future management options
  1. Passive protection of the most ecologically valuable forests
  2. Reducing the amount of litter (garbage, rubbish) in forests through tougher law enforcement and by increasing forest cleaning services
  3. Increasing the level of recreational infrastructure, such as improved signposting of forest trails
  4. Cost
  5. No change (status quo)
- Representative sample of 1001 Poles
  - 253 distinct locations
- 4 alternatives per choice task
- 26 choice tasks per respondent

	Alternative 1	Alternative 2	Alternative 3	Alternative 4
Protection of ecologically valuable forests	 **Status quo** Passive protection of **50%** of the most ecologically valuable forests (1.5% of all forests)	 **Status quo** Passive protection of **50%** of the most ecologically valuable forests (1.5% of all forests)	 **Status quo** Passive protection of **50%** of the most ecologically valuable forests (1.5% of all forests)	 **Substantial improvement** Passive protection of 100% of the most ecologically valuable forests (3% of all forests, 100% increase)
Litter in forests	 **Status quo** No change in the amount of litter in the forests	 **Partial improvement** Decrease the amount of litter in the forests by half (50% reduction)	 **Status quo** No change in the amount of litter in the forests	 **Partial improvement** Decrease the amount of litter in the forests by half (50% reduction)
Infrastructure	 **Status quo** No change in tourist infrastructure	 **Status quo** No change in tourist infrastructure	 **Partial improvement** Appropriate tourist infrastructure in an additional 50% of the forests (50% increase)	 **Substantial improvement** Appropriate tourist infrastructure available in twice as many forests (100% increase)
Cost	0 PLN	10 PLN	25 PLN	100 PLN
Your choice	☐	☐	☐	☐

Respondents and forest area spatial distribution



The baseline for the comparison – traditional 2-step approach

1. Use mixed logit choice models (e.g., latent class, random parameters) to retrieve individual-specific conditional distributions
   - Every individual has a separate, independent set of parameters
   - Individual parameters are not directly observed but we know:
     - (1) population-level estimates of parameter distributions
     - (2) each individual’s choices
   - It is possible to estimate their individual-specific values using the Bayes theorem
2. Use the predicted (expected) individual-specific parameters as dependent variables in spatial lag / spatial error/ spatial Durbin model etc.
   - Simple (linear) regression models in which dependent variable / error term in one location depends on (distance weighted) dependent variables / error terms in other locations or fixed effects for geographically-defined clusters are included
   - Allows for correlations between nearby locations
   - GIS and socio-demographic explanatory variables are used

Location- and individual-specific MXL model results (EUR/year)

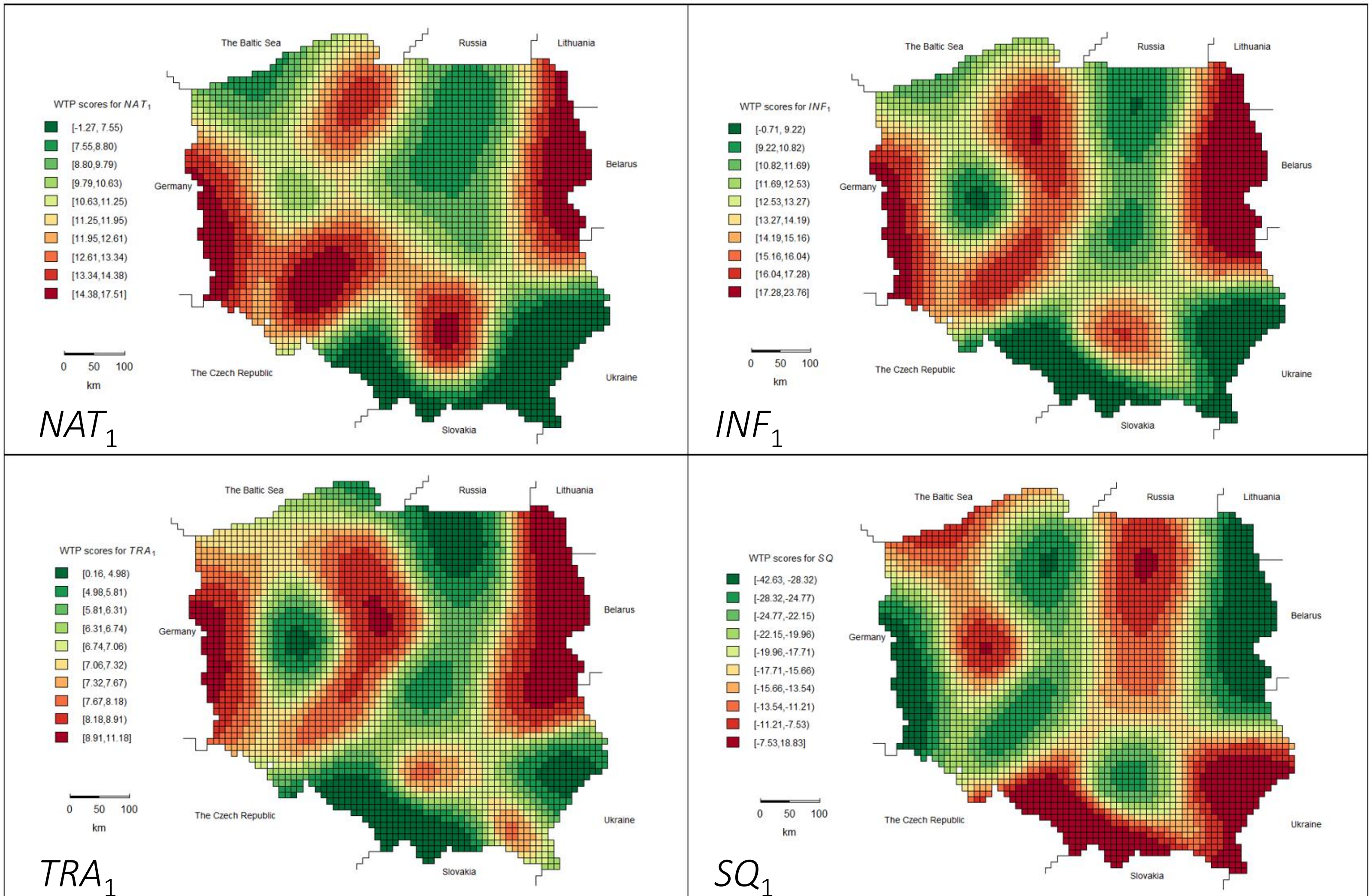
Variable	MNL model	Location specific MXL model		Individual specific MXL model	
	coef.	Mean	Std. Dev.	Mean	Std. Dev.
NAT_1	14.83***	11.79***	7.61***	9.89***	11.86***
NAT_2	21.82***	16.86***	12.17***	13.54***	17.35***
TRA_1	26.66***	17.44***	8.33***	11.55***	12.88***
TRA_2	35.67***	25.23***	14.13***	17.68***	21.48***
INF_1	12.14***	8.26***	4.58***	6.23***	6.14***
INF_2	19.55***	12.11***	6.51***	8.63***	8.61***
SQ	37.24***	-3.24***	43.27***	-13.74***	30.90***
COST	0.05***	-2.24***	0.70***	-1.57***	1.09***
Model characteristics					
LL	-29,708.27	-22,632.30		-17,169.76	
AIC/n	2.2836	1.7426		1.3228	
k	8	44		44	

Are individual WTP-scores spatially autocorrelated?

	NAT_1	NAT_2	TRA_1	TRA_2	INF_1	INF_2
	(passive protection of most valuable forests – partial improvement)	(passive protection of most valuable forests – substantial improvement)	(the amount of litter in forests – partial improvement)	(the amount of litter in forests – substantial improvement)	(tourist infrastructure – partial improvement)	(tourist infrastructure – substantial improvement)
Moran's I statistic	0.1519	0.1563	0.25601	0.25517	0.246	0.2347
p-value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

Let’s have a look:

WTP scores extrapolated to Poland using regression Kriging method



2’nd step results:
decompose the estimated WTP using GIS characteristics

- 7 regressions (1 for each attribute) in which WTP explained by the same GIS variables
- Spatial lag models for conditional expected values of random parameters from location-specific MXL and individual-specific MXL

GIS data

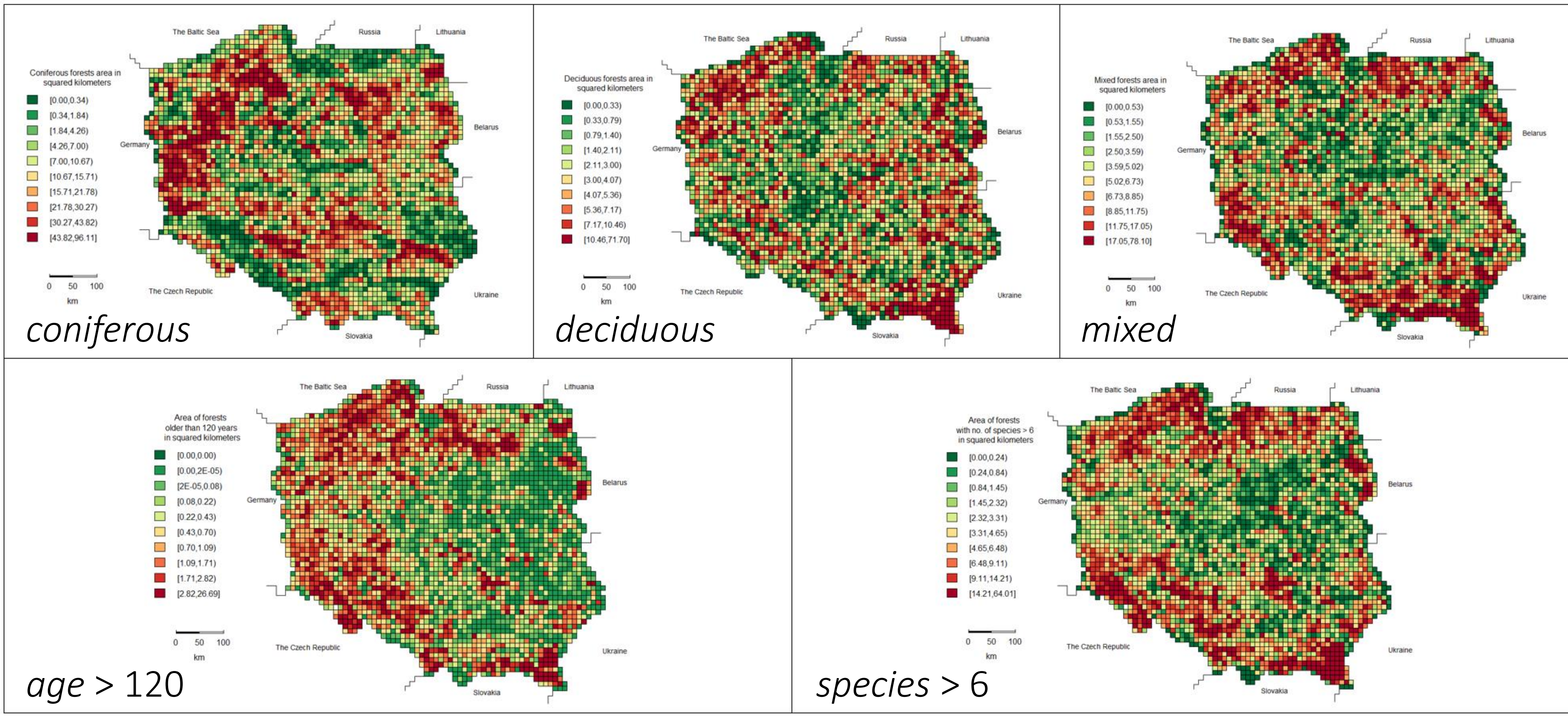
- Information on local forest characteristics:
  - CORINE Land Cover dataset
  - Polish Information System of State Forests (very precise data about the characteristics of forests in Poland)

Variable name	Description
Area of coniferous forests	Sum of areas of all coniferous forests [km^2]
Area of deciduous forests	Sum of areas of all deciduous forests [km^2]
Area of mixed forests	Sum of areas of all mixed forests [km^2]
Average Euclidean distance to forest	It is average distance from any point in 10x10 km square to the nearest forest
Area of forests with age > 120	Sum of areas of all forests older than 120 years [km^2]
Area of forests with no. of species > 6	Sum of areas of all forests with no. of tree species greater than 6 [km^2]
Built-up area	Built-up area [km^2]

GIS and SD data used as explanatory variables

GIS variables			Socio-demographic variables		
Name	Mean	St. Dev.	Name	Mean	St. Dev.
Area of coniferous forests	11.3202	13.3060	Age	44.2957	16.0257
Area of deciduous forests	4.2290	3.9805	Higher Education	0.2288	0.4203
Area of mixed forests	6.5767	6.1084	Income	3.2777	0.9984
Area of forests with age >120	0.9586	1.3336	No. of forests visited in last 12 months	2.4076	4.5873
Average euclidean distance to the forest	1.3075	0.8921	Number of trips to the forests in last 12 months	49.4276	68.5458
Bulit-up area	19.5532	19.3520	Sex	0.4216	0.4941
Area of forests with no. of species > 6	5.9285	7.1911	Household size	2.9501	1.3811

Spatial distribution of forest characteristics



Bayesian posterior mean WTP from the location-specific MXL model regressed on GIS and SD variables

	NAT_1	NAT_2	TRA_1	TRA_2	INF_1	INF_2
Constant	12.69***	17.81***	5.02***	7.29***	10.86***	17.25***
Area of coniferous forests	-0.08***	-0.12***	-	-	-0.08**	-0.14***
Area of deciduous forests	-0.45***	-0.66***	-0.12***	-0.19***	-0.41***	-0.71***
Area of mixed forests	-0.25***	-0.37***	-0.08***	-0.13***	-0.24***	-0.40***
Area of forests with age >120	1.27***	1.86***	0.30**	0.48**	1.11***	1.93***
Average Euclidean distance	-1.77***	-2.63***	-0.50***	-0.76***	-1.72***	-2.87***
Age	-0.10***	-0.14***	-0.02***	-0.04***	-0.07***	-0.11***
Higher education	-	-	-0.75**	-1.11**	-1.46*	-2.33*
Income	0.91***	1.35***	0.48***	0.71***	1.17***	1.91***
No. of forests visited (log)	1.96***	2.88***	0.52***	0.87***	1.39***	2.32***
No. of trips to forests (log)	0.59***	0.85***	-	-	0.58**	1.06***
ρ	0.21***	0.21***	0.37***	0.37***	0.35***	0.33***
Model characteristics						
R^2	0.0203	0.0187	0.0256	0.0239	0.0234	0.0198

Did this help?
Are individual WTP-scores spatially autocorrelated?

	NAT_1	NAT_2	TRA_1	TRA_2	INF_1	INF_2
	(passive protection of most valuable forests – partial improvement)	(passive protection of most valuable forests – substantial improvement)	(the amount of litter in forests – partial improvement)	(the amount of litter in forests – substantial improvement)	(tourist infrastructure – partial improvement)	(tourist infrastructure – substantial improvement)
LM statistic	30.940	33.087	112.757	110.810	94.180	83.064
p-value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

Geographically weighted regression

- Geographically weighted regression belongs to the general class of "locally estimated" models
  - Rationale: while estimating a model for one location, take into account (distance-weighted) data from other locations; separate models for different locations but each time all data is used
  - It recognizes nonlinear relationships with respect to spatial dimensions
    - Relationship between analyzed variables may be highly nonlinear and, therefore, is difficult to determine it parametrically
- Early applications based solely on linear local models
  - They were used for analysis of morbidity, house price data, economic growth, school performance and urban temperatures
  - In the context of non-market valuation – hedonic price models of house prices

Geographically weighted choice models

- In the choice models context – use weighted maximum likelihood estimators for inference (local likelihood models)
  - Rationale is the same: while estimating a model for one location, take into account (distance-weighted) data from other locations; separate models for different locations but each time all data is used
  - Estimation differs (weighted maximum likelihood)
- Earlier applications of locally estimated models
  - They were used to recover WTP distribution non-parametrically, to analyze behavioral tendencies such as the implications of prospect theory, and analyze preference dynamics
- We use local discrete choice models to analyze spatial heterogeneity
  - We aim at exploring the advantages and limitations of this approach in the context of understanding the spatial heterogeneity of environmental values

Geographically weighted multinomial logit model

- Standard MNL model
- Estimated in WTP-space
- The GWMNL model is conducted by estimating *L* 'local' models
  - *L* is a number of distinct locations
  - In our case, 253 distinct locations of respondents (unique postal codes)
  - Estimated via the weighted maximum likelihood method

$$WL^{l} = \sum_{n=1}^{N} \sum_{j=1}^{J} \lambda\left(Lat_{n}, Long_{n}, b, l\right) \log\left(L^{l}_{j,n}\right)$$

- $\lambda\left(Lat_{n}, Long_{n}, b, l\right)$ – geographical weight (kernel), which depends on latitude and longitude of individual n's location, *b* which is called the 'bandwidth parameter' and the location *l* for which the local model is estimated

The choice of kernel (does not matter much)

- A few functional forms of $\lambda(\cdot)$ proposed in the literature
  - We use the Gaussian kernel defined as:

$$\lambda\left(Lat_n, Long_n, b, l\right) = \exp\left(-0.5\,\frac{\left(Lat_n - Lat_l\right)^2 + \left(Long_n - Long_l\right)^2}{b^2}\right)$$

  - Simply an exponential function of minus half of squared Euclidean distance of individual *n*'s location from location *l* divided by the square of the bandwidth parameter
- We also tried different weighting functions, such as the spatially varying kernel:

$$\lambda\left(Lat_n, Long_n, b, l\right) = \exp\left(-\frac{R_{n,l}}{b}\right)$$

  - Where $R_{n,l}$ is the rank of the *n-th* location from *l-th* location, in terms of the distance *n* is from *l*
  - The results were not much different from the Gaussian kernel

The choice of bandwidth (does matter a lot)

- The choice of bandwidth may have a greater impact on the results than the choice of a specific weighting scheme (Fosgerau, 2007)
- Several methods for choosing the bandwidth parameter available in the literature, with no apparent dominant approach
- We tried:
    1. Corrected Akaike Information Criterion (Dekker, Koster and Brouwer, 2014)
    2. The lowest bandwidth for which all local models converge
    3. Leave-one-individual-out cross-validation criterion (Fotheringham, Brunsdon and Charlton, 2003)
- To evaluate them, we used simulated data which utilized the designs utilized in our study
- Conclusion: the available methods are unsatisfactory and lead to either under or over-smoothing (Koster and Koster, 2015)
- So we used the 'eye-balling' approach (Koster and Koster, 2015):
    - Choose the lowest bandwidth for which the model estimates satisfy a set of *a priori* specified conditions (e.g., achieving identification of all the models or avoiding extreme estimates)

Comparison of the approaches

Traditional 2-step approach

- Spatial correlation accommodated indirectly
- Individual-specific results can include different sources of unobserved preference heterogeneity
- Need to assume parametric distribution of population-level parameters

Geographically weighted multinomial logit model

- Spatial correlation accommodated directly
- Individual-specific results account for spatial heterogeneity only
- No need to specify a distribution from which the parameters are drawn (non-parametric approach)

Summary statistics of the estimated parameters for the GWMNL models (EUR/year)

	Mean	Std. Dev.
NAT_1	15.71***	6.87***
(passive protection of most valuable forests – substantial improvement)	[0.12]	[0.17]
NAT_2	23.07***	10.01***
(passive protection of most valuable forests – partial improvement)	[0.18]	[0.25]
TRA_1	28.30***	11.02***
(the amount of litter in forests – partial improvement)	[0.20]	[0.23]
TRA_2	37.85***	14.75***
(the amount of litter in forests – substantial improvement)	[0.27]	[0.36]
INF_1	12.71***	5.12***
(tourist infrastructure – partial improvement)	[0.09]	[0.10]
INF_2	20.60***	8.29***
(tourist infrastructure – substantial improvement)	[0.13]	[0.15]
SQ	39.37***	26.29***
(alternative specific constant for the no-choice alternative)	[0.38]	[0.45]
***COST* (preference-space equivalent)**	0.05***	0.02***
(annual cost – tax increase)	[0.00]	[0.00]

Location- and individual-specific MXL model results (EUR/year) – again (for comparison)

Variable	MNL model	Location specific MXL model		Individual specific MXL model	
	coef.	Mean	Std. Dev.	Mean	Std. Dev.
NAT_1	14.83***	11.79***	7.61***	9.89***	11.86***
NAT_2	21.82***	16.86***	12.17***	13.54***	17.35***
TRA_1	26.66***	17.44***	8.33***	11.55***	12.88***
TRA_2	35.67***	25.23***	14.13***	17.68***	21.48***
INF_1	12.14***	8.26***	4.58***	6.23***	6.14***
INF_2	19.55***	12.11***	6.51***	8.63***	8.61***
SQ	37.24***	-3.24***	43.27***	-13.74***	30.90***
COST	0.05***	-2.24***	0.70***	-1.57***	1.09***
Model characteristics					
LL	-29,708.27	-22,632.30		-17,169.76	
AIC/n	2.2836	1.7426		1.3228	
k	8	44		44	

Comparison of the WTP results

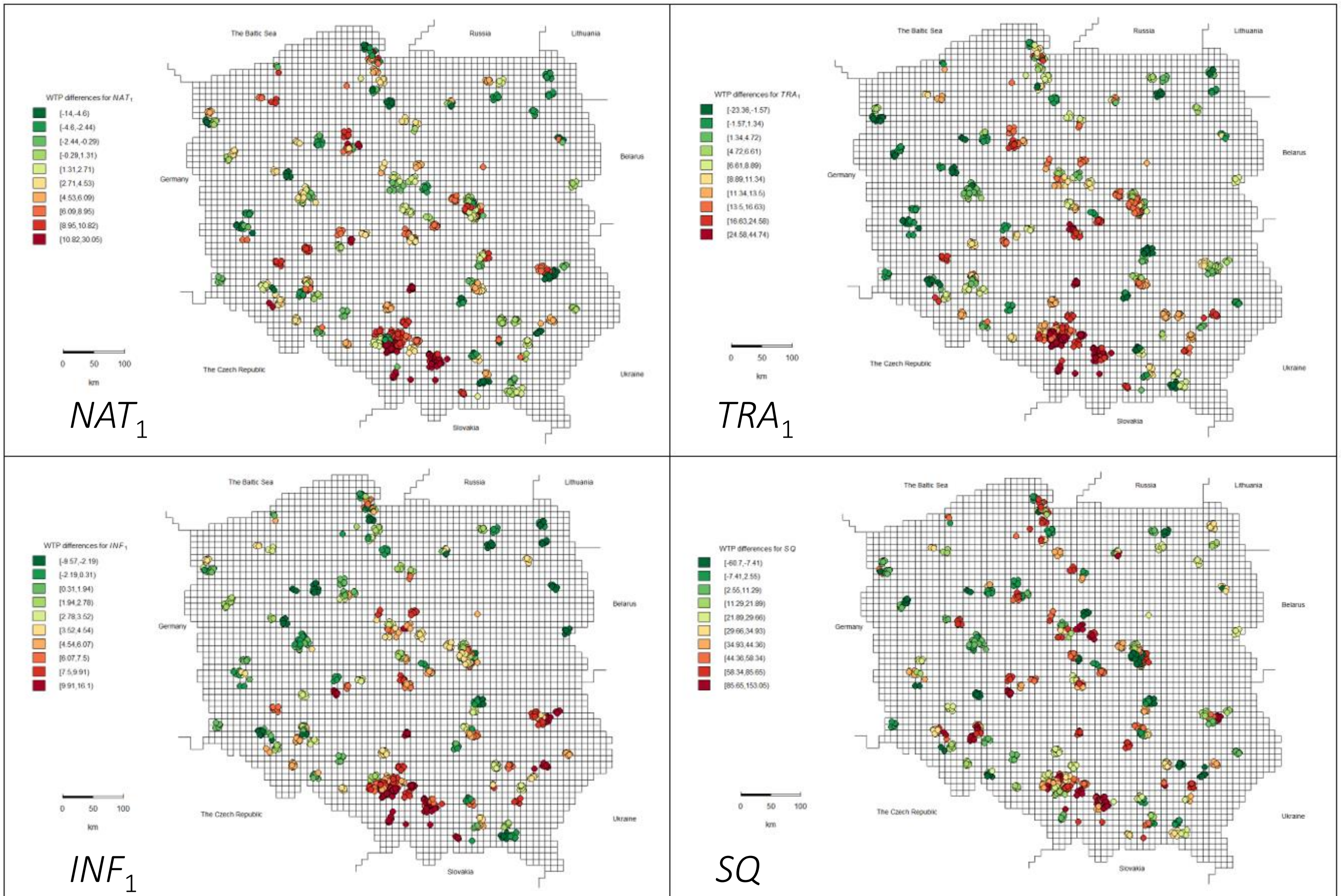
- Different approaches can lead to significant changes:
  - Location- and individual-specific MXL – lower mean WTP values than the GWMNL
  - SQ parameter appears to have a reversed sign
- Why?
  1. Not allowing for spatial correlation in the specification of the MXL model may lead to biased estimates (GWMNL superior)
  2. The assumption of the MNL model form of local models in GWMNL may not be justified (e.g., we assume independence of error terms – likely to not be true for non-SQ alternatives; MXL superior)
  3. Distributional assumptions of the MXL model (cost*scale parameter log-normally distributed, marginal WTP distributions all normally distributed; the GWMNL model is a non-parametric approach and thus makes no such assumptions – superior)

Correlation between the WTP estimates the GWMNL model vs. the posterior means from the individual-specific and location-specific MXL models

	MXL location-specific			**MXL individual-specific**		
	Pearson product-moment correlation	Spearman's rank correlation	Mann-Whitney U test	Pearson product-moment correlation	Spearman's rank correlation	Mann-Whitney U test
NAT_1	0.34***	0.31***	850,540***	0.14***	0.14***	849,407***
NAT_2	0.37***	0.34***	829,556***	0.17***	0.17***	835,692***
TRA_1	0.27***	0.31***	702,578***	0.03	0.05	643,904***
TRA_2	0.31***	0.34***	754,979***	0.09***	0.10***	726,605***
INF_1	0.35***	0.35***	761,014***	0.04	0.03	694,666***
INF_2	0.32***	0.37***	718,047***	0.02	0.04	638,309***
SQ	0.28***	0.31***	728,741***	0.16***	0.12***	557,149***

– Correlation coefficients are positive, although they are lower than one could have hoped

Spatial distribution of differences between WTP estimates from GWMNL and conditional expected values from location-specific MXL



The comparison of the relative fit to the data

- Relative fit to the data compared using the Ben-Akiva-Lerman's pseudo-R^2, adapted to the panel character of our data
  - A measure of predicted probabilities of choosing the alternatives which were actually chosen by respondents

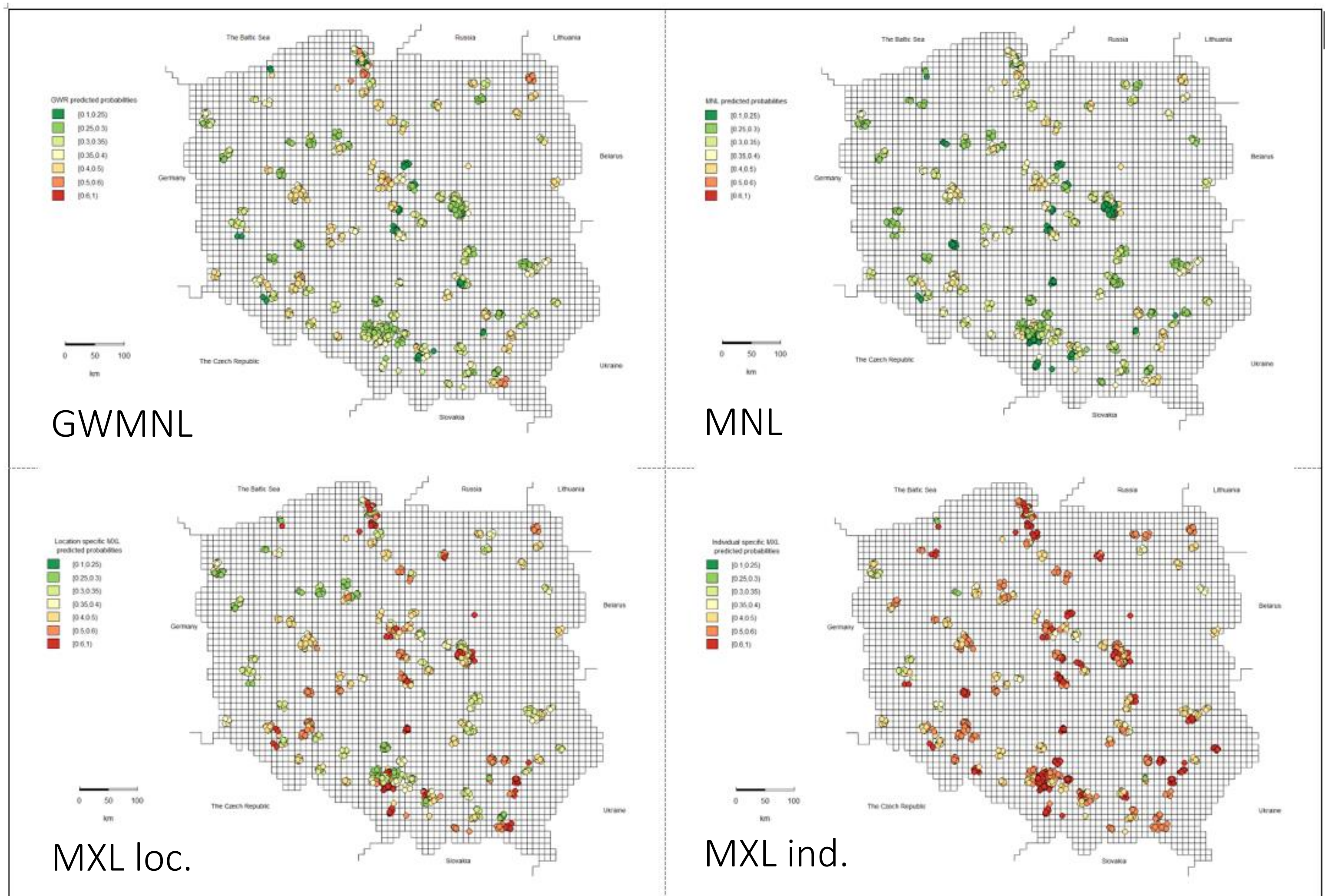
	GWMNL	MNL	Location specific MXL model	Individual specific MXL model
Mean	0.3550	0.3282	0.4626	0.5266
2.5'th percentile	0.1874	0.2066	0.2682	0.2727
97.5'th percentile	0.5294	0.4396	0.9097	0.9370

  - Although the GWMNL approach provides a better fit than the MNL model, it is worse than the location- and individual-specific MXL models
  - The ability to generically account for the unobserved preference heterogeneity offers more of an improvement in fit than explicitly accounting for spatial correlations in the MNL model

Predicted probabilities are highly correlated

	GWMNL	MNL	Location specific MXL model	Individual specific MXL model
GWMNL	1.0000	0.7782	0.4110	0.2380
MNL	0.7782	1.0000	0.2313	0.0922
Location specific MXL model	0.4110	0.2313	1.0000	0.7780
Individual specific MXL model	0.2380	0.0922	0.7780	1.0000

The regions in which respondents' choices are relatively better or worse predicted are unchanged across the four models



WTP estimates from the GWMNL model regressed on GIS variables

	SQ	NAT_1	NAT_2	TRA_1	TRA_2	INF_1	INF_2
Constant	34.76***	17.61***	27.01***	30.88***	44.85***	11.86***	22.09***
Area of coniferous forests	-0.15	-0.07*	-0.11*	-0.07	-0.15*	0.02	-0.04
Area of deciduous forests	1.07**	-0.12	-0.34*	-0.04	-0.34	0.23***	0.21
Area of mixed forests	-0.16	-0.24***	-0.30**	-0.29**	-0.43**	-0.03	-0.22**
Area of forests with age >120	-7.04***	-0.49	-0.53	-1.37**	-1.09	-1.40***	-2.09***
Average Euclidean distance to a forest	-1.78	-1.16*	-1.71*	-2.31**	-3.81***	0.03	-1.10
Built-up area	0.29***	0.08***	0.11***	0.14***	0.17***	0.01	0.04
Area of forests with no. of species > 6	1.04***	0.28***	0.31**	0.35**	0.41**	0.11*	0.38***
Model characteristics							
R^2	21.95%	12.15%	9.80%	12.93%	9.25%	15.56%	16.48%
n (observations)	253	253	253	253	253	253	253
k (parameters)	8	8	8	8	8	8	8

Bayesian posterior mean WTP from the location-specific MXL model regressed on GIS variables

	SQ	NAT_1	NAT_2	TRA_1	TRA_2	INF_1	INF_2
Constant	-35.08***	15.21***	23.24***	16.28***	26.62***	5.94***	8.97***
Area of coniferous forests	0.48**	-0.08**	-0.14**	-0.06*	-0.15**	0.01	-0.01
Area of deciduous forests	2.54***	-0.37***	-0.64***	-0.19	-0.53***	0.04	0.00
Area of mixed forests	0.76*	-0.18**	-0.31**	-0.20**	-0.34**	-0.06	-0.10*
Area of forests with age >120	-0.63	0.60	0.99	0.46	0.94	-0.07	0.09
Average Euclidean distance to a forest	10.84***	-1.68***	-2.89***	-1.60**	-3.26***	-0.29	-0.68
ρ	0.25***	0.18**	0.18**	0.33***	0.33***	0.37***	0.40***
Model characteristics							
R^2	21.95%	12.15%	9.80%	12.93%	9.25%	15.56%	16.48%
n (observations)	253	253	253	253	253	253	253
k (parameters)	8	8	8	8	8	8	8

The 2’nd step results – comparison

- Most of GIS variables are highly significant
- R^2 in all models is very low – suggest that the GIS variables we used explain only a small fraction of the observed variance
  - Most of the preference heterogeneity is caused by some other factors, which were not accounted for
  - Large differences between forests which lie next to each other – significant variance in their values may occur even on a local level
  - Could GWMNL perform better in the case of preferences for environmental goods which change more gradually?
- Substantial discrepancies with regard to spatial patterns recovered with the two methods
  - Differences in the signs and significance of coefficients
  - WTP distributions differ in structure between the two approaches
- As the GWMNL model explicitly deals with spatial heterogeneity (rather than trying to recover it indirectly post estimation) it could be considered more reliable

Summary

- We investigate spatial patterns in stated preferences for forest management
- We try a different approach to addressing spatial patterns in WTP – the geographically weighted choice (multinomial logit) model
  - Allow for new insights regarding the spatial distribution of preferences
  - Difficult to conclude whether this approach is superior to using the Bayesian posterior means from an MXL model – both have several shortcomings
    - GWMNL addresses spatial correlations directly (observed location-specific WTPs are not conditional on the MXL assumptions and the distributions)
    - But does not allow for other sources of (unobserved) heterogeneity
- Compare to a "two-step" approach (the MXL model + posterior Bayesian means of random parameters)
  - There are similarities but …
  - Significant differences in the estimates of WTP (particularly wrt the SQ)
  - Structural differences in the observed effect of GIS variables

Future directions

- Try more complicated weighting functions, accounting for the respondents' socio-demographics
  - E.g., using more than one bandwidth parameter
- Allow for unobserved heterogeneity in the local models
  - E.g., latent class or random parameters instead of MNL
- Investigate the reliability of Bayesian posterior means and their vulnerability to MXL assumptions
- Underdeveloped methods of choosing an appropriate bandwidth parameter