

The overview of methods to account for spatially explicit preference heterogeneity

Wiktor Budziński

Danny Campbell, Mikołaj Czajkowski, Urška Demšar,
Nick Hanley



1. Two-step method

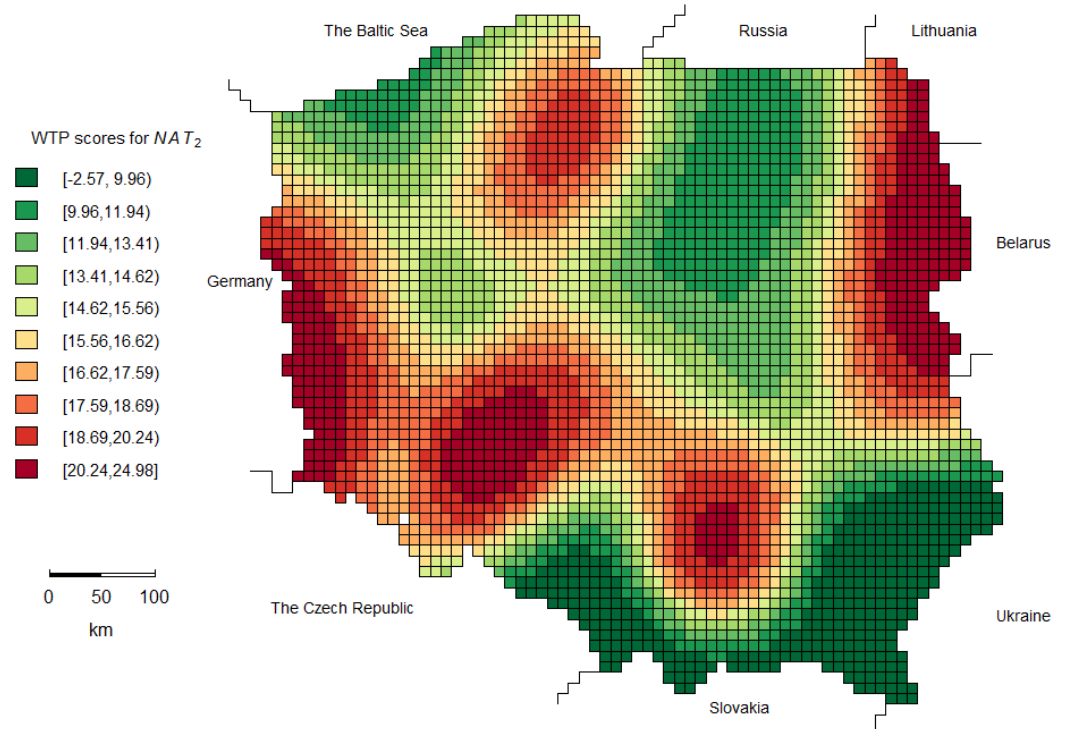
- Estimating Mixed logit model as a first step
 - Predicting individual-specific WTP, by using posterior means of random parameters given by the Bayes formula:

$$E(\alpha_n | y_n, X_n, \theta) = \int \alpha_n \frac{p(y_n | X_n, \theta, \alpha_n, \beta_n^{\text{cost}}) f(\alpha_n, \beta_n^{\text{cost}} | \theta)}{p(y_n | X_n, \theta)} d(\alpha_n, \beta_n^{\text{cost}})$$

- Estimating (panel) regression on these estimates
 - Abildtrup, J., Garcia, S., Olsen, S. B., & Stenger, A. (2013). Spatial preference heterogeneity in forest recreation. *Ecological Economics*, 92, 67-77.
- Or estimating spatial (panel) regression models
 - Spatial lag or spatial error
 - 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*.

1. Two-step method

- Or using kriging to obtain WTP map
 - Campbell, D. (2007). Willingness to Pay for Rural Landscape Improvements: Combining Mixed Logit and Random-Effects Models. *Journal of Agricultural Economics*, 58(3), 467-483.



1. Two-step method

- Advantages:
 - No additional programming needed, only standard models used
 - Takes different sources of preference heterogeneity into account (not only spatial)
- Disadvantages:
 - Conditional on parametric distributions
 - First step ignores spatial dependencies
 - Posterior means may not describe individual-specific parameters well

2. One-step method

- It is possible to estimate Mixed logit model which directly accounts for spatial autocorrelation of preferences
- Likelihood function is complicated, therefore it is useful to apply Bayesian estimator
- In simulation setting it recovers parameters well
- Metropolis Hastings algorithm (using Hamiltonian (Hybrid) Monte Carlo) is inefficient
 - Solutions?

2. One-step method

Means

True	coef.	st.dev.	P>0	nse	IEF	M*	CD
-5	-4.9461	0.2207	0	0.0371	169.2323	35.4542	-0.0979
3	3.0961	0.1482	1	0.0219	131.6301	45.5823	0.3517
-2	-1.8303	0.2182	0	0.0724	660.3757	9.0857	-1.3936

Variances

True	coef.	st.dev.	P>0	nse	IEF	M*	CD
2	1.6113	0.2983	1	0.0428	123.3458	48.6437	-0.1791
2	2.0849	0.2334	1	0.0253	70.4524	85.1639	0.1332
2	1.8768	0.2132	1	0.0205	55.599	107.9156	0.3606

Spatial autocorrelations

True	coef.	st.dev.	P>0	nse	IEF	M*	CD
0.7	0.749	0.0794	1	0.0101	97.8658	61.3085	1.2352
0.7	0.5529	0.0875	1	0.0085	56.3056	106.5614	-0.2861
0.7	0.7147	0.0674	1	0.0085	94.9473	63.1929	-0.5094

2. One-step method

- Advantages
 - Directly accounts for spatial dependencies
 - Allows for other sources of heterogeneity
- Disadvantages
 - Computationally intensive
 - Current algorithm is inefficient
 - Conditional on parametric distributions

3. Geographically weighted choice models

- Growing interest in so called local-models
 - Non-linear effects of attributes on choices
 - Preference dynamics
 - Spatial dependencies
- A separate model is estimated for every location using weighted Maximum Likelihood

$$\beta_i = \arg \max \left(\sum_j w_{ij} LL_j \right)$$

- We use Geographically Weighted Multinomial logit
 - Budziński, W., Campbell, D., Czajkowski, M., Demšar, U., and Hanley, N., Using geographically weighted choice models to account for spatial heterogeneity of preferences.

3. Geographically weighted choice models

- Different weighting schemes

- Gaussian weighting:

$$w_{ij} = \exp\left(-0.5 \frac{(Lat_i - Lat_j)^2 + (Long_i - Long_j)^2}{b^2}\right)$$

- Spatially varying kernel: $w_{ij} = \exp\left(-\frac{R_{i,j}}{b}\right)$

- Depends on so called bandwidth parameter b which cannot be estimated

- Needs to be determined based on some penalized fit function
 - Or by eyeballing

3. Geographically weighted choice models

- Advantages
 - Non parametric method
 - Directly accounts for spatial dependencies
- Disadvantages
 - Does not account for other sources of heterogeneity
 - Possible solution – local latent class/mixed logit models (work in progress)
 - Choice of the bandwidth is quite arbitrary
 - It is not easy to include socio-demographic variables
 - Bias-variance trade-off