Mixed Logit, Geographically Weighted Choice Models, or One-Step Bayesian Estimation?

Wiktor Budziński

Mikołaj Czajkowski

Fundacja na rzecz Nauki Polskiej

Two-step method (Mixed Logit)

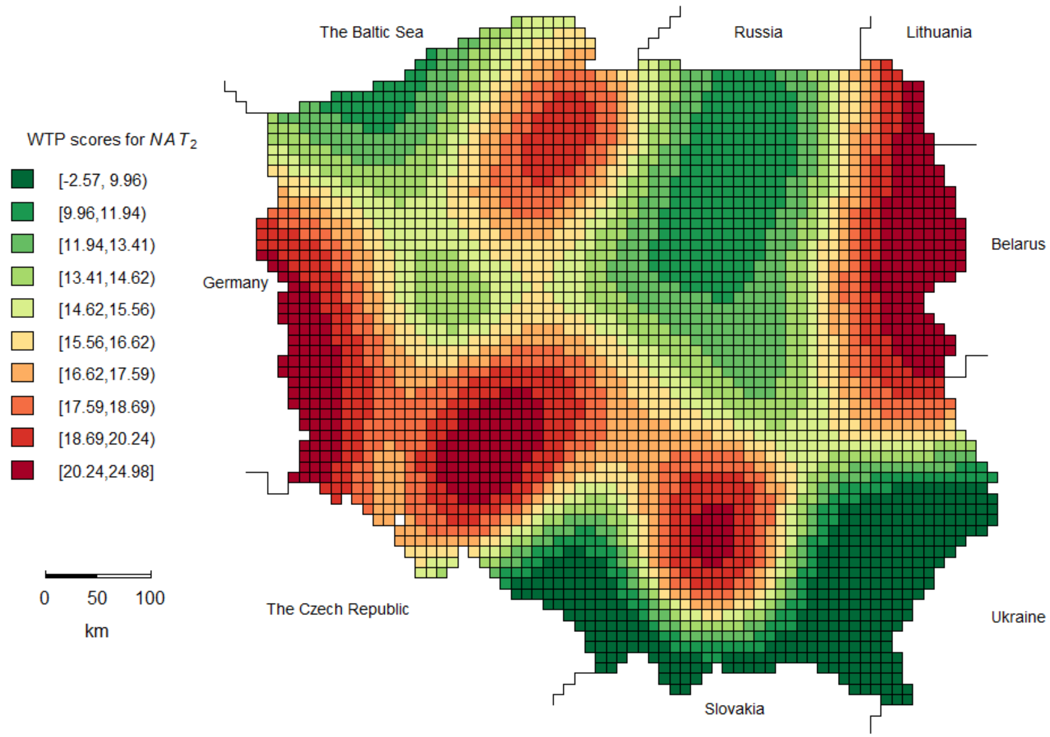
- Estimating Mixed logit model as a first step
  - Predicting individual-specific WTP, by using posterior means of random parameters given by the Bayes formula:

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

- 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.

Two-step method (Mixed Logit)

- 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.



Two-step method (Mixed Logit)

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

Geographically weighted choice models

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

$$\boldsymbol{\beta}_i = \arg\max\left(\sum_j w_{ij} LL_j\right)$$

- 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.

Geographically weighted choice models

- Different weighting schemes
  - Gaussian weighting:

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

  - Spatially varying kernel: $\mathrm{w}_{ij} = \exp\left(-\frac{R_{\mathrm{i,j}}}{b} \right)$
- Depends on so called bandwidth parameter b which cannot be estimated
  - Needs to be determined based on some penalized fit function
  - Or by 'eyeballing'

Geographically weighted choice models

- Advantages
  - Non-parametric method
  - Directly accounts for spatial dependencies
- Disadvantages
  - GW-MNL does not account for other sources of heterogeneity
    - Possible solution – local latent class/mixed logit models
  - Choice of the bandwidth is quite arbitrary
  - It is not easy to include socio-demographic variables

One-step method (Bayesian)

- It is possible to estimate Mixed logit model which directly accounts for spatial autocorrelation of preferences

$$U_{ij} = \beta_i X_{ij} + \varepsilon_{ij}$$

$$\beta_i = \beta + \theta_i$$

$$\theta_i = u_i + \rho \sum w_{ik} \theta_k \quad \textit{where} \;\; u_i \sim N(0, \sigma)$$

- Likelihood function is complicated, therefore it is convenient to apply Bayesian methods

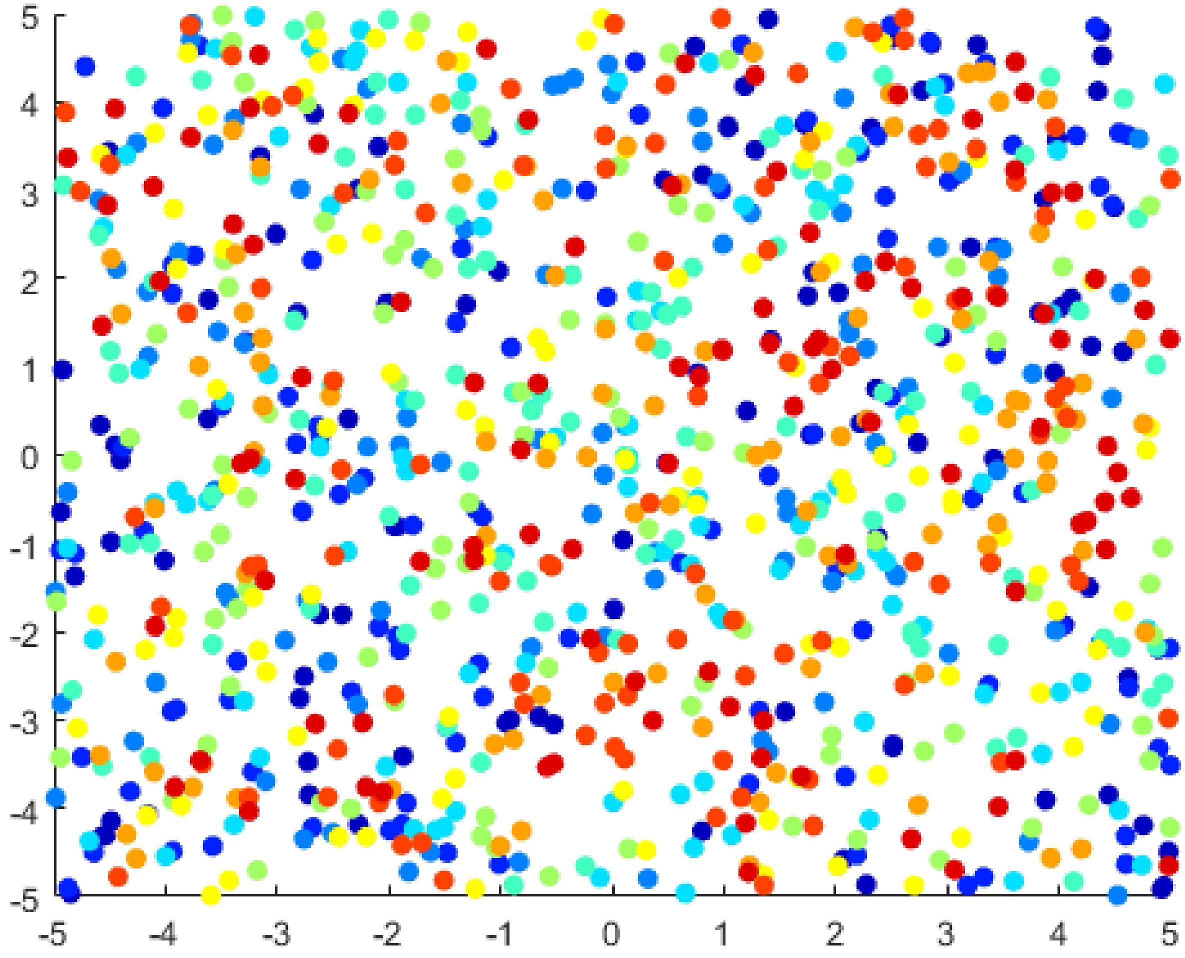
One-step method (Bayesian)

- Advantages
  - Directly accounts for spatial dependencies
  - Allows for other sources of heterogeneity
- Disadvantages
  - Computationally intensive
  - Current algorithm is inefficient
  - Still relies on parametric distributions
  - Also relies on posterior means

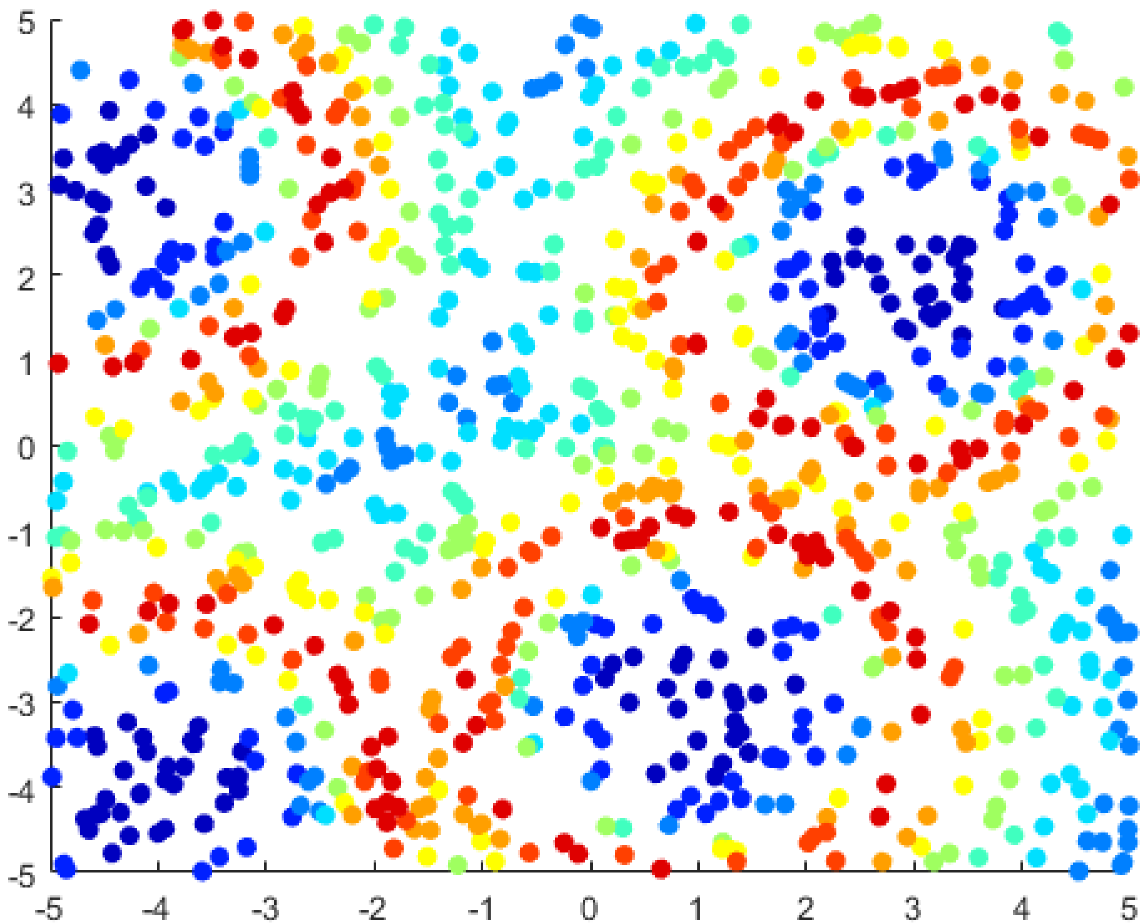
Simulation

- Comparison of three approaches
- 30 repetitions (for now)
- Datasets with
  - 1000 respondents
  - 6 choice tasks
  - 3 alternatives
  - 2 attributes (quality and cost)
- Two preference heterogeneity types
  - Spatially autocorrelated
  - Distance decay type (deterministic)

Spatially autocorrelated



Distance decay type (deterministic)



Spatially autocorrelated

	WTP distribution percentiles				
	5	25	50	75	95
True	-0.2063	0.5050	0.9206	1.5628	4.429
GWR	0.2836	0.6699	0.9687	1.3877	2.8988
One step (Bayesian)	-0.7152	0.6282	1.026	1.6182	4.2101
Two Step (MXL)	-1.5685	0.5449	0.9757	1.6319	4.7847

	Median absolute Errors		
	Min	Mean	Max
GWR	0.3904	0.4555	0.5140
One step (Bayesian)	0.4358	0.5067	0.6059
Two Step (MXL)	0.4522	0.5694	0.7176

Spatially autocorrelated

- Are MXL-based models really worse?
  - Posterior mean WTP may not exist…
  - They recover parameters quite well:

	True	**One step (Bayesian)**	**Two Step (MXL)**
Means	2	1.9654	1.9656
	-2	-2.0005	-2.0073
Variance	1	0.9900	1.1850
	1	1.0335	1.2622
Spatial autocorrelation	0.6	0.5351	0.0000
	0.6	0.5159	0.0000

	Median absolute Errors		
	Min	**Mean**	**Max**
One step (Bayesian)	0.5222	0.5606	0.5919
	0.5039	0.5681	0.6033
Two Step (MXL)	0.5215	0.5837	0.6446
	0.5272	0.5964	0.6716

Distance decay type (deterministic)

	WTP distribution percentiles				
	5	25	50	75	95
True	-2.5945	0.7282	1.567	2.3236	2.7286
GWR	-2.2624	0.5207	1.3597	1.8408	2.9322
One step (Bayesian)	-5.2994	-0.1851	1.3466	2.2252	6.6373
Two Step (MXL)	-3.4842	0.0651	1.6719	3.1186	5.5447

	Median absolute Errors		
	Min	Mean	Max
GWR	0.3320	0.4006	0.4678
One step (Bayesian)	0.7042	0.9106	1.0871
Two Step (MXL)	0.9894	1.2303	1.6007

Conclusions

- It seems that Geographically weighted choice model performs best in both scenarios
  - Possibly influenced by the choice of bandwidth
- Errors significantly higher when using mean absolute errors
- Other types of spatial heterogeneity?