A spatial error latent class model for analysing stated preferences

A comparison

Danny Campbell Wiktor Budziński Mikołaj Czajkowski
Nick Hanley

Workshop on Discrete Choice Modelling
Warsaw, 5–6 October 2016

Presentation outline

Preference heterogeneity

For sure, current models are now better equipped to identify and address unobserved sources of preference heterogeneity.

However, aside from a few examples, the inherently spatial patterns of preferences have been rarely clarified or addressed in studies using stated choice experiments.

But, there are obvious reasons for spatial variations in preferences.

- The spatial arrangement of socio-demographic profiles of respondents is likely to impact the geographical distribution of preferences.
- The environmental conditions within a particular locality are likely to influence preferences.

Why is this relevant?

Exploratory spatial data analysis provides different insights about preferences:

- its distribution;
- regional and local outliers;
- regional trends; and,
- the level of spatial autocorrelation.

Distributions of preferences are likely to be both spatially and socially uneven, meaning that evaluating the regional nature of benefits delivers advantages from the political and policy analysis viewpoints.

- It can help policy decision makers locate areas of value and thus allows more efficient targeting of efforts.

spatial autocorrelation of preferences

Previous studies have shown that preferences are positively spatially correlated.

Therefore, it seems reasonable to expect positive spatial clustering of latent class probabilities membership (i.e., membership to latent classes is spatially related).

- Respondents who live close to one another are more likely to belong in a similar class compared to respondent who do not reside close to one another.

Knowledge gaps

The spatial dimension of preferences is now starting to gain more attention—however, there are a number of unanswered questions.

One of these questions relates to how best to accommodate the possibility that the unobserved factors that explain membership to latent classes may be spatially related.

This paper hopes to shed some light on this

Presentation outline

Latent class membership

Finite mixture models are now widely used to analyse revealed and stated preference data. Their appeal is the flexibility that they afford to the analyst.

- With the correct assumptions, they can uncover preference heterogeneity, the presence of error variance heteroscedasticity and a range of processing strategies.

At the heart of these models is the assumption that respondents belong in a given class.

Latent class membership

However, it is obviously not possible to know membership beforehand with certainty and, thus, it remains latent.

To work around this, based on observed choice behaviour, probabilistic conditions are imposed on each class.

In doing so, the presence of each class can be established up to a probability, with the full probability per respondent allocated across all classes.

Latent class membership

Typically, the class membership function, Q, associated with a given class c is given as:

$$Q_{c_n} = \gamma_c + \theta_c z_n + \varepsilon_c,$$

where γ_c is the class constant, θ_c is a vector of class-specific parameters for the z vector of individual characteristics for respondent n (e.g., socio-demographic and perhaps some attitudinal variables), and ε_c is the class-specific error term.

Assuming the error terms are *iid* type I extreme value distributed means that the unconditional class probabilities can be retrieved using a multinomial logit specification.

Latent class membership

Notwithstanding the ability to include individual characteristics in the latent class membership function, there is a possibility that the unobserved factors that explain membership to latent classes may be spatially related.

If so, the errors are spatially arranged.

- This means that the assumption that the error terms are independent of one another is violated.

Not addressing this could mean the model is mis-specified in the systematic component of the latent class membership function—in particular, the omission of variables that are spatially clustered.

Latent class membership
Spatial error latent class model

In this paper we allow spatial dependence to enter through the membership function errors.

The key assumption of this model is that spatial autocorrelation is treated as a nuisance and as an estimation problem, and as something to be estimated.

This is accommodated by decomposing the overall membership error into two components, namely a aspatial error term that is *iid* type I extreme value distributed that satisfies the standard assumption, and a spatial error term that captures the pattern of spatial dependence between errors for connected observations.

Latent class membership

Spatial error latent class model

The membership error can be rewritten as:

$$\varepsilon_c = \rho_c \zeta_c + \varphi_c,$$

where ρ_c and ζ_c are, respectively, the spatial error parameter and the term indicating the spatial component of error term for class c and φ_c is the aspatial error term.

The parameter ρ indicates the extent to which the spatial component of the errors ζ are correlated with one another for nearby observations.

- If $\rho \neq 0$, then we have a pattern of spatial dependence between the errors for connected observations, but if $\rho = 0$ we can proceed to estimate class membership by the multinomial logit specification in the conventional manner.

Latent class membership

Spatial error latent class model

The spatial component of error term is defined as:

$$\zeta_c \sim \text{MVN}(0, \Sigma_c),$$

where Σ_c is the symmetrical $N \times N$ covariance matrix.

The (Euclidean) distances between sampled respondents are represented by a symmetrical $N \times N$ distance matrix D.

The value of covariance between two locations can be obtained by scaling the distance matrix by a negative exponential function, as is common practice in spatial models:

$$\Sigma_{c_{ij}} = \exp\left(-\lambda_c D_{ij}\right).$$

Latent class membership

Spatial error latent class model

The unconditional class probability for a given class and respondent, π_{c_n}, is the weighted average of the class probability evaluated at different values of ζ_c, with the weights given by $f(\zeta_c)$:

$$\pi_{c_n} = \int \frac{\exp\left(\gamma_c + \theta_c z_n + \rho\zeta_{c_n}\right)}{\sum_{\forall c} \exp\left(\gamma_c + \theta_c z_n + \rho\zeta_{c_n}\right)} f(\zeta_c)\, d(\zeta_c).$$

The unconditional class probabilities are approximated through simulating the probabilities over a large number of draws, R, and averaging the results.

This average is the simulated unconditional class probability:

$$\pi_{c_n} = \frac{1}{R}\sum_{r=1}^{R} \frac{\exp\left(\gamma_c + \theta_c z_n + \rho\zeta_{c_{r_n}}\right)}{\sum_{\forall c} \exp\left(\gamma_c + \theta_c z_n + \rho\zeta_{c_{r_n}}\right)}.$$

Data generating process

This paper seeks to assess the bias caused by estimating aspatial latent class models.

A natural way to achieve this is to compare the performance of the modelling approaches on simulated data with a known data generating process.

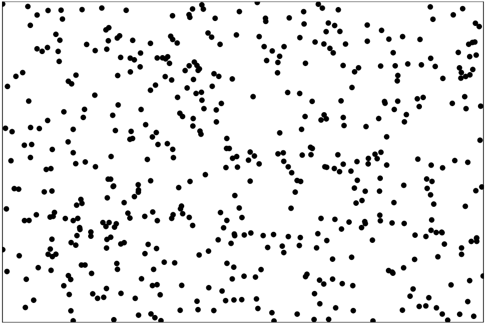
Sample spatial point pattern

The data generating process begins by generating a spatial point pattern to depict the irregularly distribution of sampled residential locations within a study region S.

The study region is 100 km by 150 km and include a sample of 500 respondents.

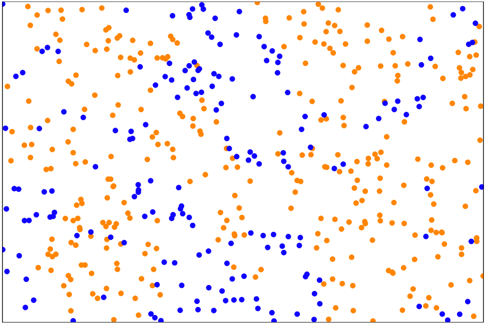
The sample is assumed to be distributed according to a homogeneous Poisson process.

Sample spatial point pattern



The point pattern has stationary and isotropic characteristics and resembles ***complete spatial randomness***, meaning that intensity of the simulated residential locations do not vary spatially.

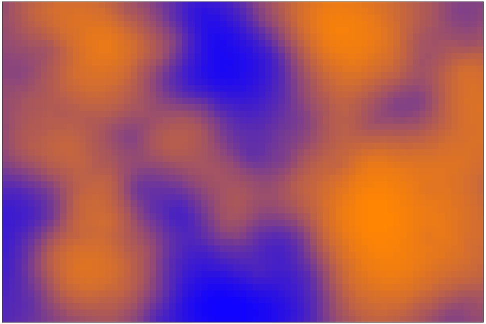
Sample spatial point pattern



Respondents belong in one of two classes.

- Approximately 70 percent belong in class 1.
- A degree of spatial clustering is assumed.

Sample spatial point pattern



Kriged surface of interpolated class membership.

Generation of choice data

The simulated choice experiment consists of three attributes: two environmental attributes and a cost attribute, denoted respectively using *A*, *B* and *C*.

- The two environmental attributes are qualitative dummy variables.
- The cost levels are €4, €8, €12 and €16.

The choice tasks are defined as having three alternatives.

An orthogonal main-effects experimental design was generated to produce 32 such choice tasks.

- This was blocked in four, such that each simulated respondent completed 8 tasks.

Generation of choice data

The underlying data is generated based on:

	Class 1	Class 2
β_A	1.6	-0.6
β_B	1.0	0.9
β_C	-0.2	-0.1

Presentation outline

Estimation results

	Actual	Aspatial LC	Spatial LC
Class 1			
β_A	1.600	1.680 (0.126)	1.605 (0.126)
β_B	1.000	0.929 (0.073)	0.942 (0.071)
β_C	-0.200	-0.193 (0.011)	-0.190 (0.010)
Class 2			
β_A	-0.600	-0.596 (0.123)	-0.690 (0.131)
β_B	0.900	0.958 (0.103)	0.916 (0.108)
β_C	-0.100	-0.105 (0.010)	-0.103 (0.011)
$\gamma_{c=1}$		0.496 (0.182)	1.010 (0.324)
$\rho_{c=1}$			2.143 (0.232)
$\lambda_{c=1}$			3.476 (0.193)
Log-likelihood		-3,566.343	-3,565.918

Both models produce equivalent model fits.

- But this is to be expected: the unconditional class probabilities are averaged over the draws.

Estimation results

	Actual	Aspatial LC	Spatial LC
Class 1			
β_A	1.600	1.680 (0.126)	1.605 (0.126)
β_B	1.000	0.929 (0.073)	0.942 (0.071)
β_C	-0.200	-0.193 (0.011)	-0.190 (0.010)
Class 2			
β_A	-0.600	-0.596 (0.123)	-0.690 (0.131)
β_B	0.900	0.958 (0.103)	0.916 (0.108)
β_C	-0.100	-0.105 (0.010)	-0.103 (0.011)
$\gamma_{c=1}$		0.496 (0.182)	1.010 (0.324)
$\rho_{c=1}$			2.143 (0.232)
$\lambda_{c=1}$			3.476 (0.193)
Log-likelihood		-3,566.343	-3,565.918

Importantly though, ρ and *lambda* are both significant.

This confirm the incidence of positive spatial clustering of latent class probabilities.

Estimation results

	Actual	Aspatial LC	Spatial LC
Class 1			
β_A	1.600	1.680 (0.126)	1.605 (0.126)
β_B	1.000	0.929 (0.073)	0.942 (0.071)
β_C	-0.200	-0.193 (0.011)	-0.190 (0.010)
Class 2			
β_A	-0.600	-0.596 (0.123)	-0.690 (0.131)
β_B	0.900	0.958 (0.103)	0.916 (0.108)
β_C	-0.100	-0.105 (0.010)	-0.103 (0.011)
$\gamma_{c=1}$		0.496 (0.182)	1.010 (0.324)
$\rho_{c=1}$			2.143 (0.232)
$\lambda_{c=1}$			3.476 (0.193)
Log-likelihood		-3,566.343	-3,565.918

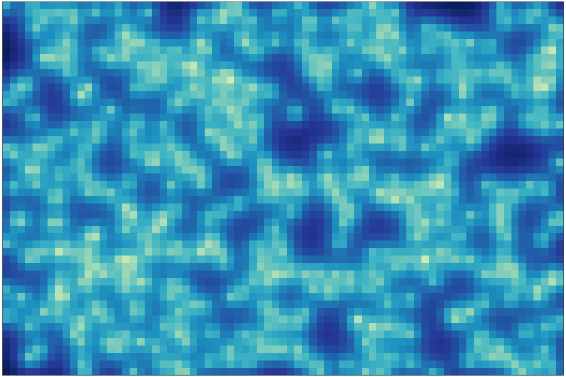
Class membership errors for an observation, therefore, tend to vary systematically in size with the errors for other nearby observations.

Estimation results

	Actual	Aspatial LC	Spatial LC
Class 1			
β_A	1.600	1.680 (0.126)	1.605 (0.126)
β_B	1.000	0.929 (0.073)	0.942 (0.071)
β_C	-0.200	-0.193 (0.011)	-0.190 (0.010)
Class 2			
β_A	-0.600	-0.596 (0.123)	-0.690 (0.131)
β_B	0.900	0.958 (0.103)	0.916 (0.108)
β_C	-0.100	-0.105 (0.010)	-0.103 (0.011)
$\gamma_{c=1}$		0.496 (0.182)	1.010 (0.324)
$\rho_{c=1}$			2.143 (0.232)
$\lambda_{c=1}$			3.476 (0.193)
Log-likelihood		-3,566.343	-3,565.918

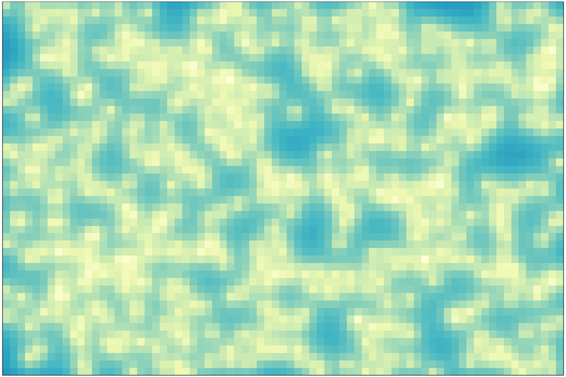
There is also some evidence that estimated parameters are closer to the true values.

Implications for spatial interpolation



The Kriged surface of interpolated WTP have considerably more variance under the aspatial LC model.

Implications for spatial interpolation



The Kriged surface of interpolated WTP have considerably more variance under the aspatial LC model.

Conclusion

This paper explores accommodating spatial clusters when analysing discrete choice experiments.

Given that this clustering of residuals violates the assumption that the error terms are independent of one another, it raises concerns on the appropriateness of the widespread use of aspatial latent class models.

Conclusion

This paper explores ...

... concerns

... use of aspatial latent

cl...

However ... more work is needed!

The paper is very much a work in progress.

The results are not as convincing as hoped.

Much more work is needed!

Spatial sampling strategies

When planning a sample survey of spatial units, it is important to appropriately determine the sample size.

- If it is too large, a huge amount of resources are required; if it is too small, the results may become inefficient and as a consequence not useful.

Key to this is the concept of spatial autocorrelation: values at positions near to one another are more likely to be similar (and thus have less variance) than values at distances further apart from one another.

Spatial sampling strategies

Creating an optimal sampling design requires balancing accuracy of prediction (requiring more samples) with minimizing the cost of sampling (limiting the number of samples and attendant cost of gathering them).

If the variable of interest is spatially correlated (i.e., values nearby are more similar than values farther apart), then taking samples close to one another may not increase prediction accuracy and will increase costs.

- A well spread out sample is sometimes called spatially balanced.

With knowledge of spatial autocorrelation it is possible to improve the sampling efficiency in terms of the estimator error variance in relation to sample design and sample size.

Results so far

	samp50	samp100	samp150	samp200	samp250	samp500
rand.smp	7.520534	7.404686	7.318954	7.319755	7.299141	7.237014
strat.50	7.361887	7.423310	7.257290	7.292310	7.260606	7.229819
strat.25	7.573648	7.258952	7.381485	7.289712	7.249509	7.186121
buff0.05	7.486357	7.373092	7.283127	7.269754	7.222936	7.193185
buff1.25	7.455260	7.381288	7.298650	7.235954	7.231267	7.234802
buff2.00	7.482806	7.343476	7.280118	7.218792	7.184569	7.094057
spatbald	7.585846	7.370522	7.278839	7.197978	7.190398	7.155839

Mean of WTP predictions

Results so far

	samp50	samp100	samp150	samp200	samp250	samp500
rand.smp	0.7226391	0.3374207	0.2852943	0.1326973	0.1340184	0.04482396
strat.50	0.7253969	0.3239714	0.2414581	0.1278017	0.1429852	0.04413522
strat.25	0.9412702	0.3623901	0.1866670	0.1606991	0.1443116	0.03979999
buff0.05	0.5992019	0.3380530	0.3548016	0.2410333	0.1755851	0.04729270
buff1.25	0.6145421	0.3392602	0.3416403	0.1933354	0.1376302	0.03736202
buff2.00	0.5889807	0.4018739	0.3555331	0.2066023	0.1305430	0.01991544
spatbald	0.5378573	0.4260316	0.2438908	0.1309698	0.1007679	0.03060867

Variance of WTP predictions

Conclusion

The main disadvantage of a classical random sampling approach is that it ignores any spatial dependence.

If spatial dependence exists, random sampling may lead to data redundancy.

UNIVERSITY *of* STIRLING

Management School

Danny Campbell

Economics Division
Stirling Management School
University of Stirling
Scotland

Email: danny.campbell@stir.ac.uk

www.dannycampbell.me

Twitter: @DCampbell_econ