Causes of ethnic segregation in a nineteenth century city

The case of Vyborg

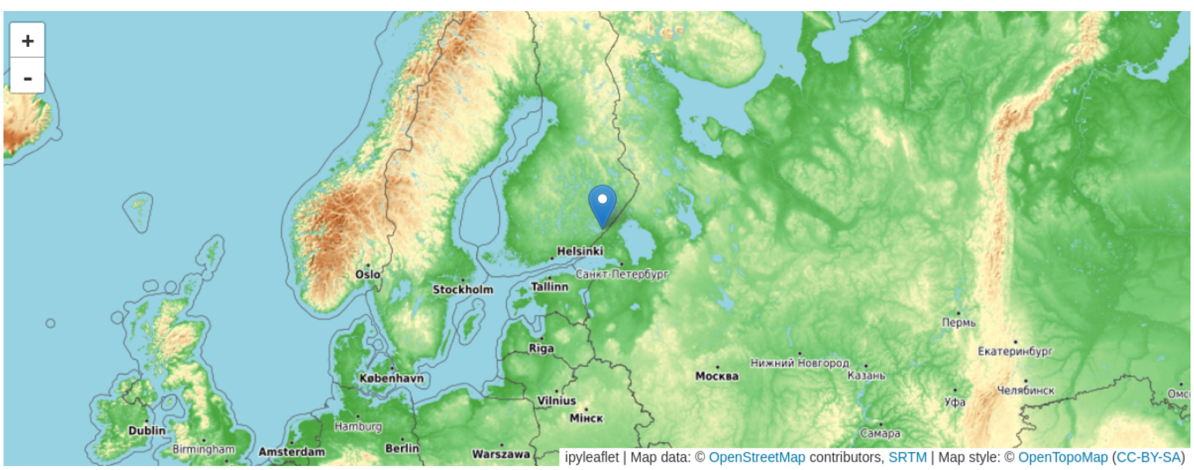
Antti Härkönen

2024-09-26

Introduction

Spatial segregation

- a classic theme of urban sociology
- implications both for individuals and society
- causes of spatial segregation studied using empirical data
- socio-economic segregation studied as a possible cause of ethnic segregation



Vyborg, a Karelian city

- castle founded in the late 13th century
- town privileges 1403

Vyborg (Viipuri, Viborg), originally a medieval trading post and stronghold founded by the Swedes, was conquered by Russians in the Great Northern War. Under Swedish rule, the city had an elite that spoke Swedish and German and commoners who spoke Finnish, just like the peasants in the surrounding area. After the Russian conquest, a large garrison was established. The military units also brought civilians with them, not only families of soldiers and other camp followers, but also higher status persons, such as retired officers or wealthier merchants and artisans. Russians remained a large and distinctive minority in Vyborg until the upheavals during and after World War One.

Causes of segregation

Several hypotheses for explaining segregation are considered, based on earlier research: policies of segregation, guild-based differentiation, discrimination from above, prejudice between groups, income-based differentiation between groups, differences in preferences, and differences of housing-market information. The main drivers of spatial segregation seem to have been the decrees enforced by both the Russian military administration and the town's civilian administration. At one time, when the administrations intended to separate Finnish and Russian lower classes on separate suburbs on the opposite sides of the city. Many of the poorer inhabitants were also de facto driven out from within the walls after fires. There are still concentrations of the Russian minority in areas which were inhabited by Russians in the eighteenth century. Segregation based on membership of guilds was not significant based on previous research and distribution of masters. Most guilds in Vyborg were tiny, only having a few masters and journeymen as members. The remaining three potential causes of segregation, namely discrimination, prejudice, and differences in housing market information cannot be studied with the data available.

Data

Sources used

The spatial data are derived from historical maps and tax records. Digitised cadastral maps provide accurate location information. The religion of the inhabitants was recorded in the poll tax registers from 1880 onward. Since every household is tied to a cadastral plot, the density of populations can be tracked in high resolution, unlike censuses. In Vyborg, the Orthodox denomination can be used as a proxy for Russian speakers. The income level can be determined based on total income tax paid. This data is provided by municipal income tax records from 1880.

Poll tax records

Table 1: poll tax record columns in 1894

column	description
plot_number	Plot number
taxpayer_men	Men paying poll tax
taxpayer_women	Women paying poll tax
no_tax_men	Men exempt from poll tax
no_tax_women	Women exempt from poll tax
in_russia_men	Men legally residing in Russia proper
in_russia_women	Women legally residing in Russia proper
total_men	Total men
total_women	Total women
independent	Civil servants, entrepreneurs, and financially independent
white_collar	White collar workers
worker_industry	Workers in industry
worker_other	Other workers
servants	Servants
other	Other employment status
non_resident	Resident elsewhere
orthodox	Orthodox
other_christian	Non-Lutheran and non-Orthodox Christian
other_religion	Other religions
draftable	21-year-old males eligible for draft

Estimating the size of Russian population

- over 90% of Orthodox in Vyborg Russian

Estimating the size of Lutheran population

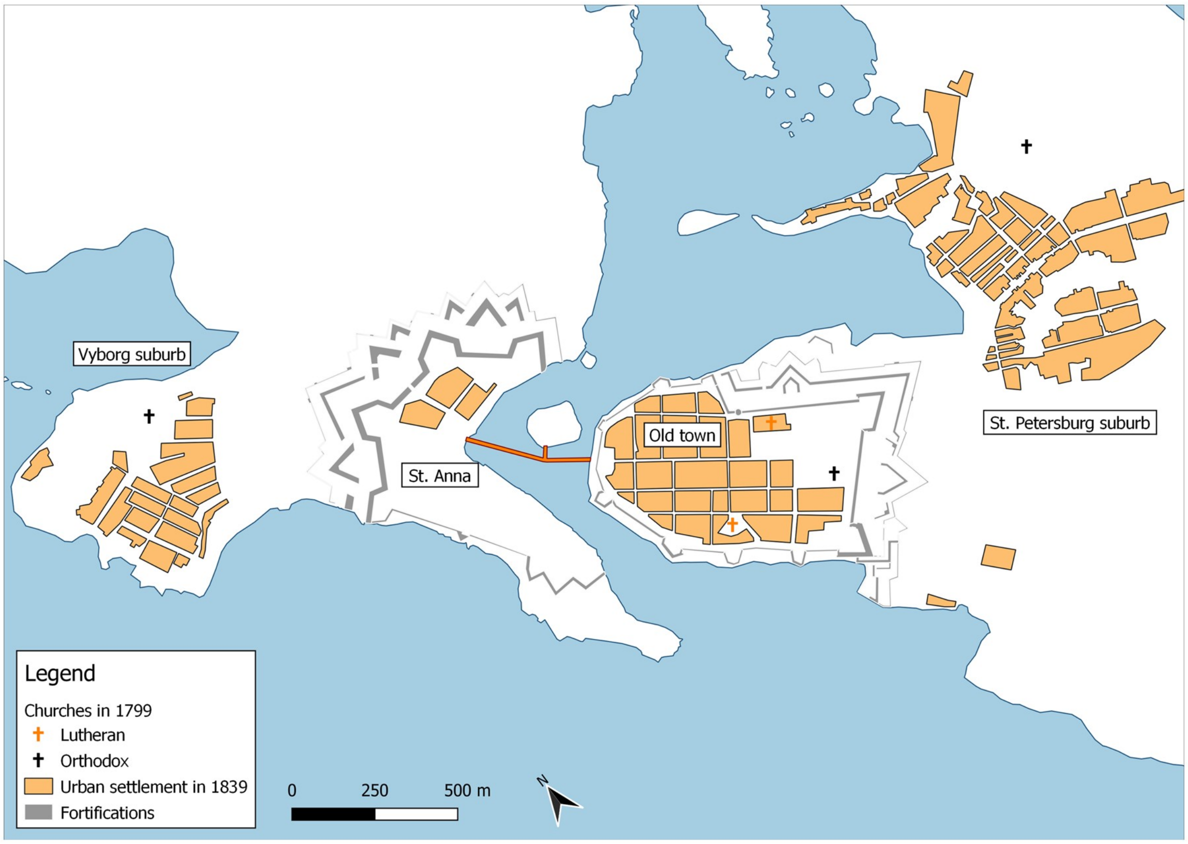
$$P_{Lutheran} = (P_{total\_men} + P_{total\_women}) - (P_{Orthodox} + P_{other\_Christian} + P_{other\_religion})$$

Digitized sources

Table 2: Sources from the National archives of Finland

Signum	Original year	Digitization process
Town plan of Vyborg. Vyborg military engineer detachment's archive of plans for fortifications and buildings, 7, 11.	1878	Georeferenced using ground control points, vectorized manually into shapefile
Vyborg province poll tax registers	1880	Digitized manually into CSV
Financial office of the city of Vyborg, Municipal tax levies and payment registers	1880	Digitized manually into CSV

Growth of Vyborg



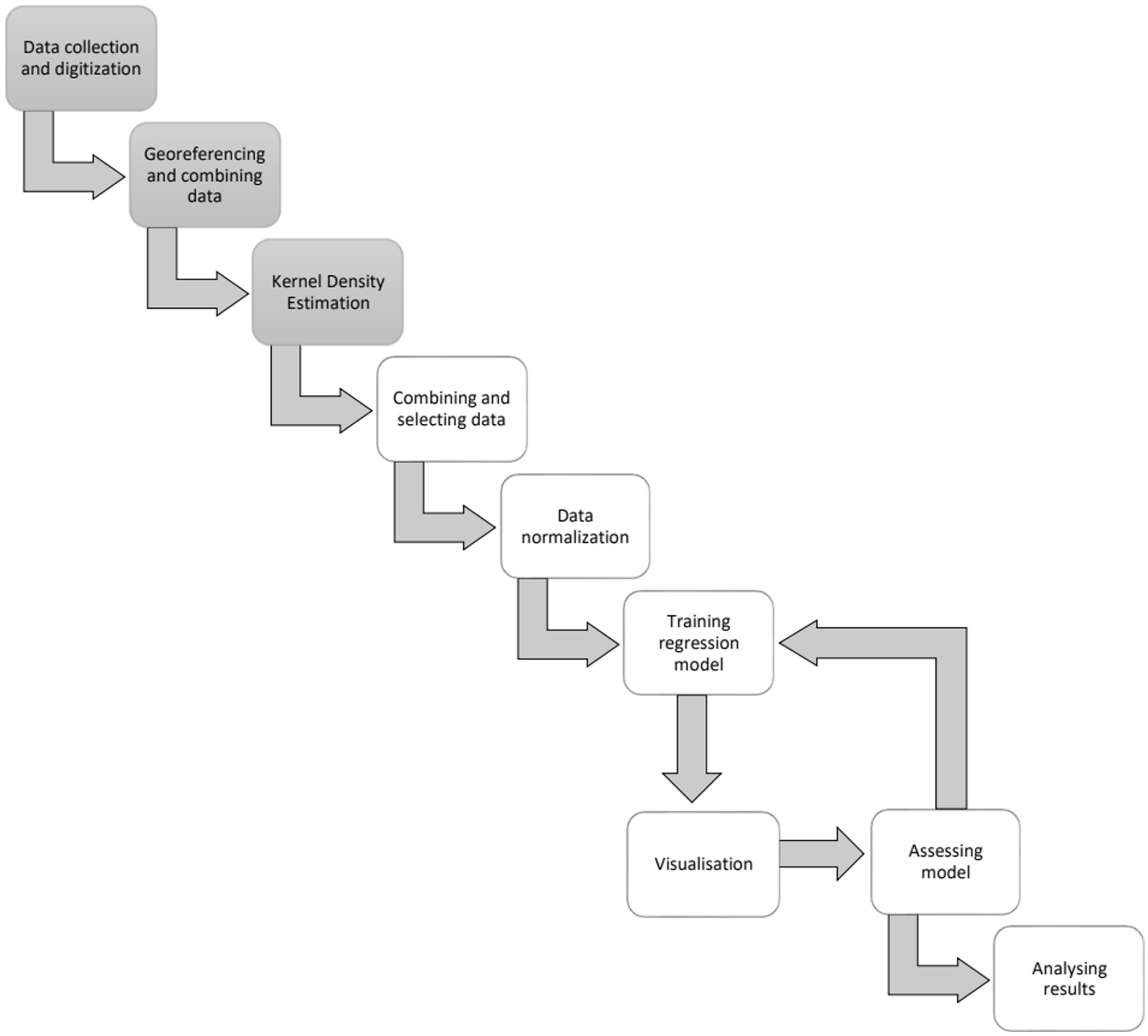
Population growth

Table 3: Population growth in key areas

District	1822	1880
Centre	1192	2506
St. Anna	244	117
Vyborg suburb	642	756
St Petersburg suburb	1512	2685

Spatial analyses

Work flow



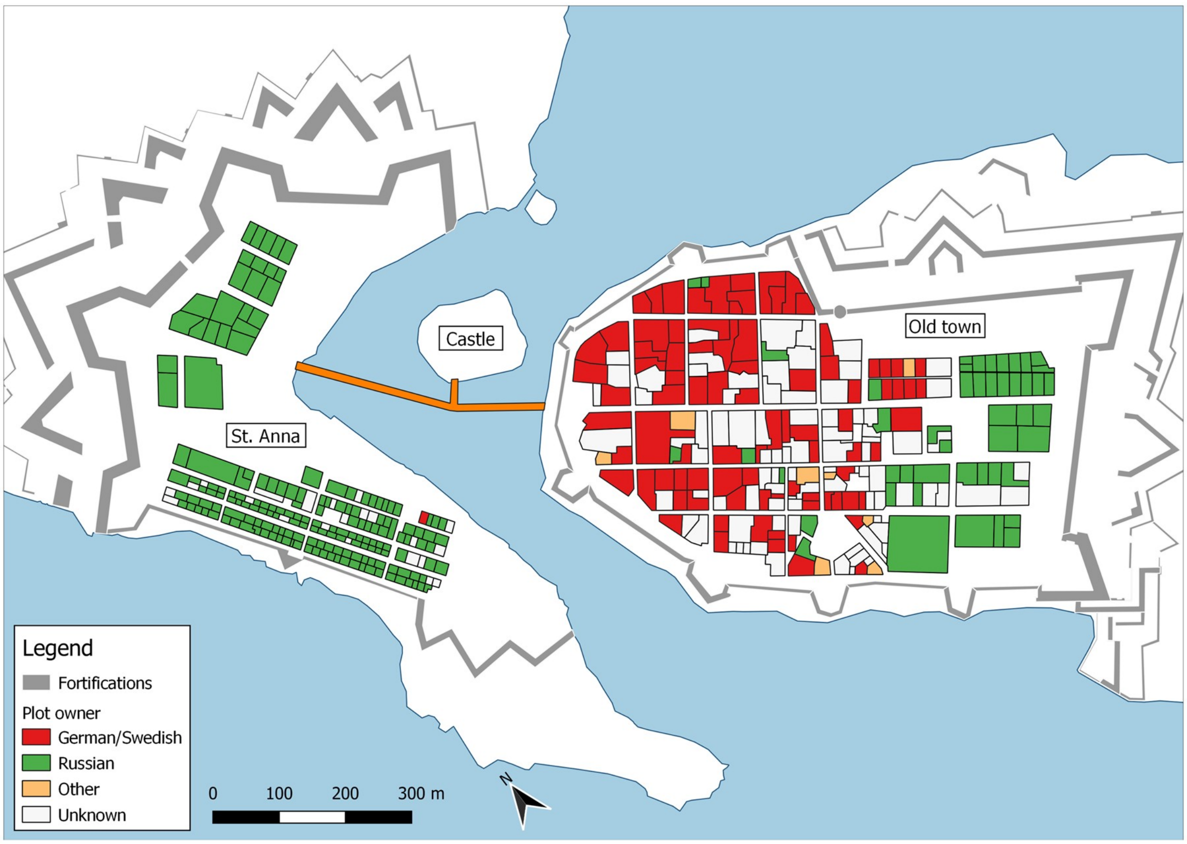
Population surface model

Population surface model

Based on Martin, Tate, and Langford (2000).

$$P_i = \sum_{j=1}^{N} P_j w_{ij}$$

Biweight kernel



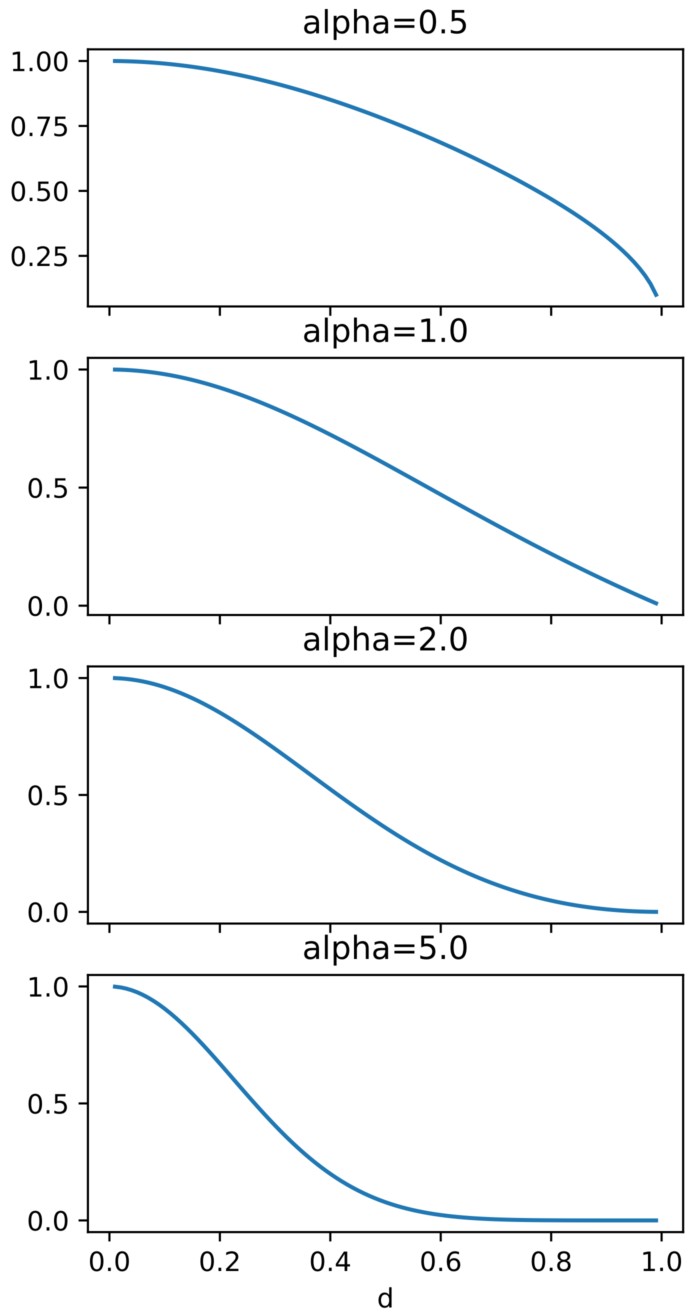


Figure 1: Kernel function

Vyborg suburb
St. Anna
Old town
Drivers' sloboda
Soldiers' sloboda
St. Petersburg suburb
Orthodox churches (1799)
Highways
Fortifications (1839)
Orthodox, % (1880)
> 35
30 - 35
25 - 30
20 - 25
15 - 20
10 - 15
5 - 10
< 5
0
250
500 m
N

Explaining segregation

To test the impact of income on the location of the Russians, a spatial regression analysis is performed. The predicted variable is the proportion of the Russians in a location (N=540), and the predictors are the natural logarithm of the average local income and distance to the nearest Orthodox church. The form of the model is a Bayesian multilevel linear regression model with spatial correlation between observations. The coefficients of the linear regression are different for each of the three areas of Vyborg. These are the western suburb, the centre within the walls, and the eastern suburbs. This means that the effects of predictors on Russian population density can vary. There is also hyperparameter that acts as a restraint on the regression coefficients of the areas. In other words, the observations are partially pooled, which combines the flexibility of treating areas as separate (unpooled observations) with the robustness of using all observations (pooled observations).

Regression model (1)

$$O_i \sim MvNormal(\mu, \mathbf{K})$$

$$\mu_i = \beta_{0,k[i]} + \beta_{1,k[i]} ln(W) + \beta_{2,k[i]} C_i$$

$$k \in 1,2,3,4 \qquad i,j \in 1,2,3,...539$$

$$\beta_k \sim MvNormal\left(\theta, \begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0.1 \end{bmatrix}\right)$$

$$\theta \sim MvNormal\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0.1 \end{bmatrix}\right)$$

Regression model (2)

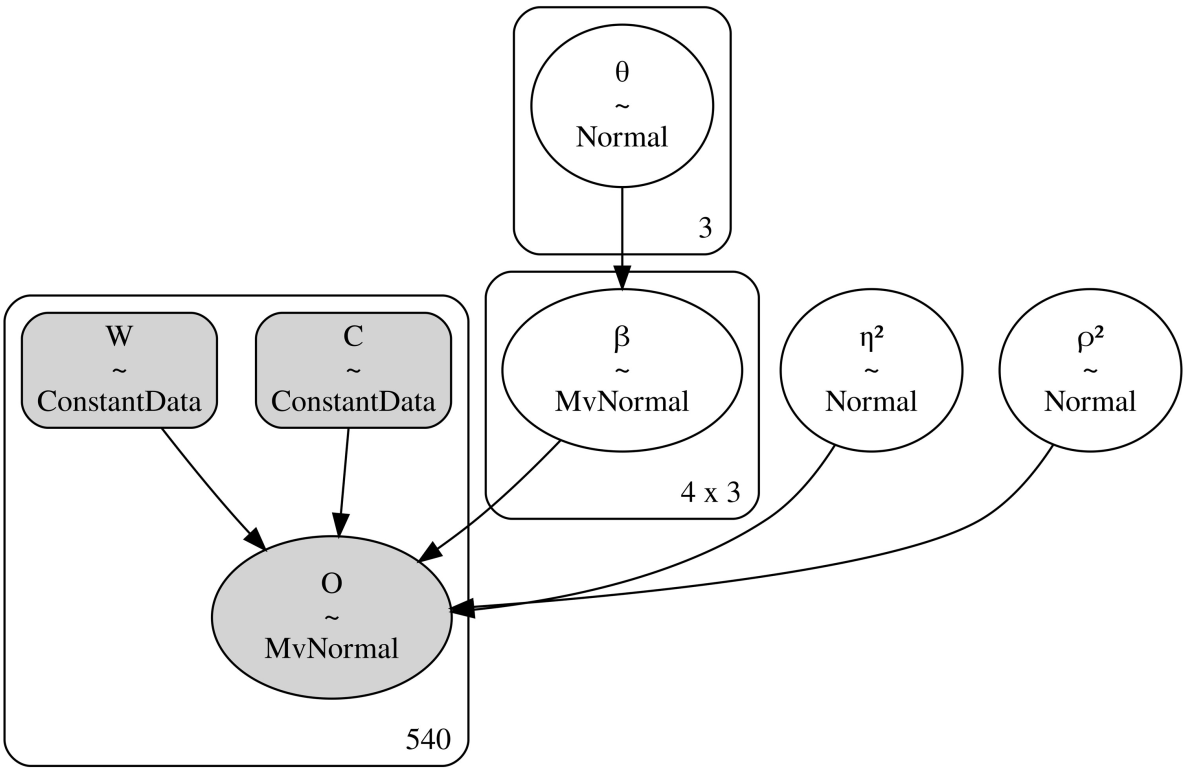
$$\mathbf{K}_{ij} = \eta^2 exp(-75\rho^2 d_{ij}^2) + 0.01 \times I_{540}$$

$$\eta^2 \sim Normal(1, 0.2)$$
$$\rho^2 \sim Normal(1, 0.2)$$

Multilevel Bayesian regression

Variable	Shape	Description
O	540	Normalized proportion of Russian Orthodox of the local population
W	540	Smoothed total income in a location in öre
C	540	Distance to nearest Orthodox church in 1799 in kilometres
d	540 x 540	Distance matrix holding pairwise distances between plots

Variable	Shape	Description
	3	Hyperparameter for
	4 x 3	Linear regression coefficients for each district
2	1	Parameter for the covariance function
2	1	Parameter for the covariance function



Results

Variable	Mean	SD	HDI, 95%	
0	−0.027	0.096	−0.227	0.15
1	0.027	0.085	−0.142	0.193
2	−0.135	0.096	−0.309	0.067

Variable	Mean	SD	HDI, 95%	
0,0	−0.609	0.299	−1.162	−0.013
0,1	0.104	0.056	−0.009	0.209
0,2	−1.076	0.314	−1.702	−0.487
1,0	0.097	0.3	−0.46	0.743
1,1	0.142	0.14	−0.117	0.433
1,2	−0.037	0.316	−0.625	0.626
2,0	0.118	0.299	−0.509	0.677
2,1	0.119	0.074	−0.024	0.261
2,2	−0.287	0.312	−0.905	0.306
3,0	0.016	0.272	−0.54	0.515
3,1	0	0.069	−0.141	0.135
3,2	−0.496	0.248	−0.991	−0.024
scaled ^{2}	0.93	0.04	0.852	1.006
^{2}	1.0	0.099	0.812	1.194

Change of segregation

Spline model (1)

$$S_i \sim Normal(\mu_i, \sigma)$$

$$\mu_i = \alpha + \sum_{k=1}^{K} w_k B_{k,i}$$

$$\alpha \sim Normal(0.45, 0.01)$$

$$\sigma \sim HalfNormal(0.05)$$

Spline model (2)

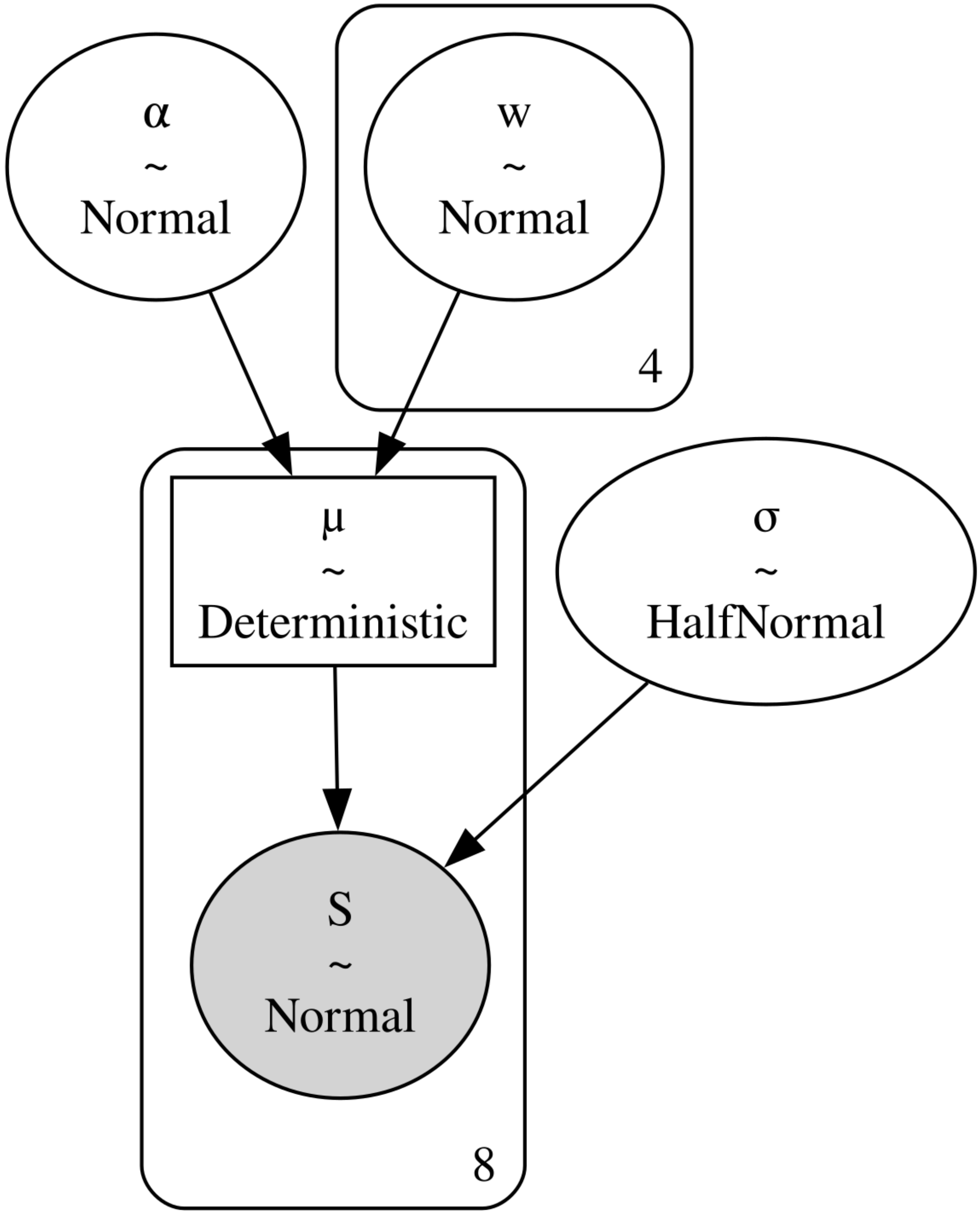
$$B = \begin{bmatrix} 1 & 0.687 & 0.295 & 0.02 & 0 & 0 & 0 & 0 \\ 0 & 0.299 & 0.601 & 0.612 & 0.367 & 0.276 & 0.007 & 0 \\ 0 & 0.015 & 0.104 & 0.367 & 0.612 & 0.658 & 0.209 & 0 \\ 0 & 0 & 0 & 0 & 0.02 & 0.066 & 0.784 & 1 \end{bmatrix}$$

$$w_k \sim Normal(0, 0.1)$$

Spline model code

```
import pymc as pm

with pm.Model() as model:
    a = pm.Normal(" ",  _a,  _a)
    w = pm.Normal("w", mu= _w, sigma= _w, shape=B.shape[1])
      = pm.Deterministic(
      " ", a + pm.math.dot(np.asarray(B, order="F"), w.T
    ))
      = pm.HalfNormal(' ',  _ )
    S = pm.Normal("S",  ,  , observed=regression_data['200'])
    idata = pm.sample(1000, tune=1000, chains=2)
```


α
~
Normal
w
~
Normal
4
μ
~
Deterministic
σ
~
HalfNormal
S
~
Normal
8

Results

The results indicate that neither the different preferences of Russians and others nor the income differences between areas explain the distribution of Russians. The posterior distributions of regression coefficients are relatively wide, but they tend to be around zero. In other words, predictors have little effect on the proportion of Russian population. According to the classic models of segregation, segregation gradually diminishes due to social diffusion. However, segregation-driving policy decisions of the eighteenth and early nineteenth centuries were still visible in the data from 1880. Interestingly, while the segregation of Russians decreases during fin de siècle, it begins increasing around 1900. One explanation for this may be the political battle between Finnish nationalists and the Imperial regime, which intensified after 1899. The disappearance of old segregation patterns may be related to the changes in the build environment, since the new concentrations of Russians were different than those in 19th century.

Conclusions

Segregation

To conclude, segregation in Vyborg cannot be explained by any single cause. The explanations behind segregation are most likely a complex system of causal links that are hard to untangle with empirical research. However, the use of high-quality spatial data allows the rejection of overly simplistic explanations.

Bayesian multilevel models

References

Martin, David, Nicholas J. Tate, and Mitchel Langford. 2000. "Refining Population Surface Models: Experiments with Northern Ireland Census Data." *Transactions in GIS* 4 (4): 343–60. https://doi.org/https://doi.org/10.1111/1467-9671.00060.