A spatial factor model for summarizing area-level Townsend-like Index

Demetris Lamnisos Nicos Middleton

Department of Nursing, School of Health Sciences
Cyprus University of Technology

April 24, 2013

Introduction

- The complex concept of socio-economic deprivation is often desribed by area-level census-based composite measures
- These composite measures are used to investigate socio-economic inequalities in health

Introduction

- A “traditional” measure of socio-economic deprivation in UK is the Townsend Index
- This is a simple index calculated by summing the normalised values of four cencus variables
- Townsend Index is correlated with mortality/morbidity and its correlations are similar to a more complex alternatives such as the Index of Multiple Deprivation (includes 33 variables classified in 6 domains) discussed in Jordan et al. (2004)
- Townsend Index is also often used to assess the convergent validity of recently developed multi-dimensional indices through factor analysis in several European countries

Objectives

- Explore for first time the geographical variability of the components of a Townsend-like Index across Cypriot communities
- Investigate the construct validity of a Townsend-like Index in Cyprus through a spatial factor model that enables us to assess the extend to which components share a common latent factor
- In contrast to factor analysis, the spatial factor model takes into account the spatial auto-correlation of census socio-economic characteristics

Methods

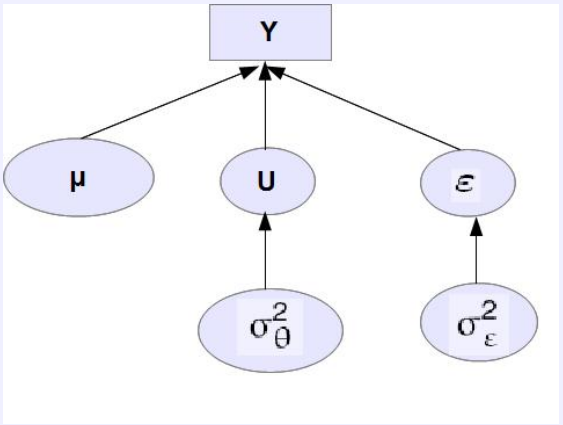
Three components of the Townsend Index were available at a community level ($n = 370$) from the 2001 census

- Unemployed economically active population (%)
- Not owner occupied households (%)
- Households with $>$ 1 person/room (%)
- No access to a car (very uncommon in Cyprus and not recorded in the census) was replaced with No access to a personal computer (%)

Univariate spatial model

- The geographical patterning and the amount of spatial variability in each indicator were investigated through a Bayesian Hierarchical model

$$
\begin{aligned}
y_i &= \mu + U_i + \varepsilon_i \\
\boldsymbol{U} &\sim \text{CAR}(\boldsymbol{W}, \sigma^2_U) \\
\varepsilon_i &\sim \text{N}(0, \tfrac{\sigma^2_\varepsilon}{m}) \\
\sigma^2_U, \sigma^2_\varepsilon &\sim \text{IGamma}(0.5, 0.005)
\end{aligned}
$$

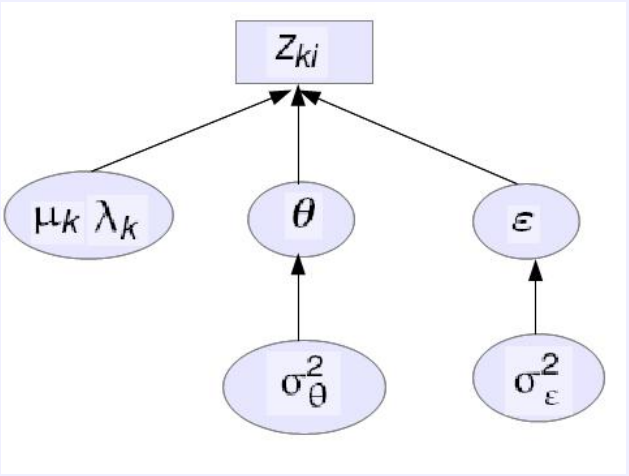


- The random effect $\boldsymbol{U}$ modeled the spatially structured variability and ε the unstructured variability
- $\boldsymbol{W}$ is the adjacent matrix and m is the appropriate total count of either households or persons

Spatial latent factor model

- The construct validity of a Townsend-like Index was investigated through a spatial latent factor model

$$
\begin{aligned}
z_{ki} &= \mu_k + \lambda_k \theta_i + \varepsilon_{ki} \\
\boldsymbol{\theta} &\sim \text{CAR}(\boldsymbol{W}, \sigma^2_\theta) \\
\varepsilon_{ki} &\sim \text{N}(0, \tfrac{\sigma^2_{\varepsilon_k}}{m_k}) \\
\lambda_k &\sim \text{TN}(0, 100) \\
\sigma^2_{\varepsilon_k} &\sim \text{IGamma}(1, 0.001) \\
\sigma^2_\theta &= 1
\end{aligned}
$$



- This model allows us to assess the extent to which components share a common latent factor representing the socio- economic deprivation

Proportion of variability explained

- A Gibbs algorithm was implemented in WinBUGS to generate a sample from the posterior distribution of the parameters. **Univariate spatial model**
- The empirical variances of the spatially structured and unstructured random effects

$$s^2_U = \frac{1}{n-1} \sum_{i=1}^{n} (U_i - \bar{U})^2$$

$$s^2_\varepsilon = \frac{1}{n-1} \sum_{i=1}^{n} (\varepsilon_i - \bar{\varepsilon})^2$$

were calculated in each iteration of the Gibbs algorithm
- The proportion of variability explained by the spatially structured random effect is given by the ratio

$$\frac{s^2_U}{s^2_U + s^2_\varepsilon}$$

Proportion of variability explained

Spatial latent factor model

- The empirical variances of the spatially structured common latent factor and unstructured random effect for each indicator k

$$s_\theta^2 = \lambda_k^2 \sum_{i=1}^{n} (\theta_i - \bar{\theta})^2$$

$$s_\varepsilon^2 = \frac{1}{n-1} \sum_{i=1}^{n} (\varepsilon_{ki} - \bar{\varepsilon}_k)^2$$

  were calculated in each iteration of the algorithm
- The proportion of variability explained by the spatially structured common latent factor for each indicator k is given by the ratio

$$\frac{s_\theta^2}{s_\theta^2 + s_\varepsilon^2}$$

Results of univariate spatial analysis

Census variable	Unem	Crowd	NOO	No PC
Unemployment	1.00	0.37	0.24	-0.26
Crowding		1.00	0.23	-0.06
Not owner occuppied			1.00	-0.21
No PC				1.00

Table : Bivariate correlations between variables

- Pairwise correlations were generally low
- Internal consistency between the variables was insufficient (Cronbach's $\alpha = 0.55$ even when % No PC was excluded)

Results of univariate spatial analysis

Census variable	Mean	2.5%	97.5%
Unemployment	25.50	18.34	33.82
Crowding	26.63	19.26	34.87
Not owner occuppied	44.37	34.67	53.77
No PC	97.15	92.40	99.93

Table : Posterior mean and 95% credible intervals for the proportion of spatially structured variability

- PC ownership displayed a striking spatial structure; however, more resembling of an urban-rural dichotomy
- It was not considered further since its correlations with the rest were in the opposite direction

Results of univariate spatial analysis

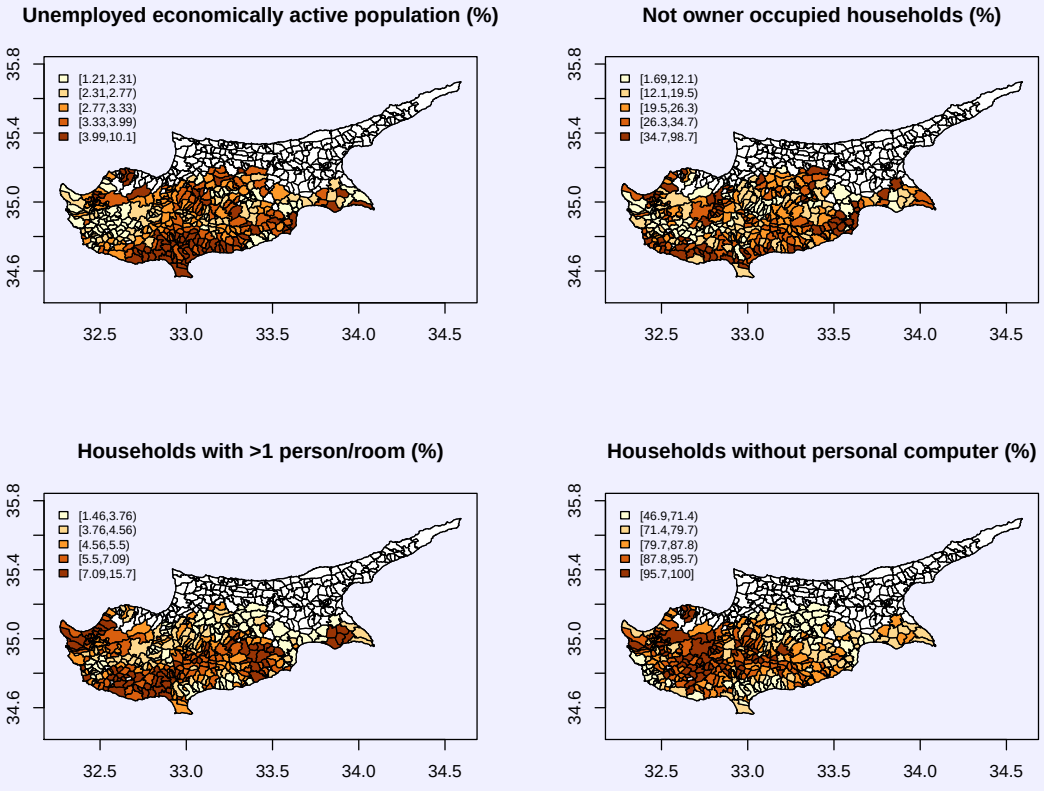


Figure : Spatially smoothed choropleth maps of indicator variables across Cypriot communities in quintile class intervals

Results of multivariate spatial factor analysis

Census variable	Mean	2.5%	97.5%
Unemployment	25.03	18.06	33.06
Crowding	0.28	0.00	1.38
Not owner occuppied	9.23	9.23	14.11

Table : Posterior mean and 95% credible intervals for the proportion of variability explained in each indicator by the common latent factor

- The remaining three indicators exhibited a different geography since the shared component only accounted for a small proportion of total variability in each indicator
- The shared component was mainly driven by Unemployment

Results of multivariate spatial factor analysis

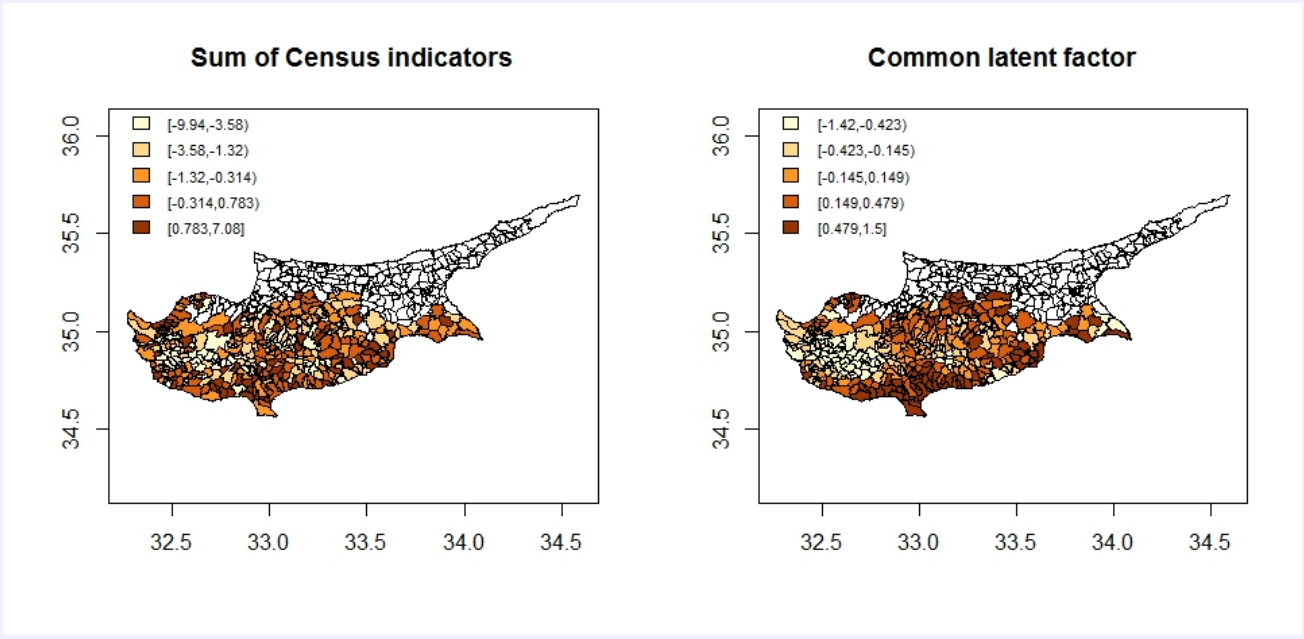


Figure : Choropleth maps of the Sum of Census Indicators(left) and the Common Latent Factor (right) across Cypriot communities in quintile class intervals

Assocation of the composite measures with mortality

- The Pearson correlation of the Common Latent Factor with Standartised Mortality Ratios (SMR) was
- The Pearson correlation of the Sum of Census Indicators with SMR was

Conclusions

- A spatial factor model has been employed to investigate the construct validity of a Townsend-like Index in Cyprus
- A Townsend-like Index does not appear to be an adequate measure of socio-economic deprivation in Cyprus
- Efforts are concentrated in developing a home-grown index from a wider set of possible indicators and exploring its predictive ability based on its association with health outcomes

References


Abellan J.J. et al.
Bayesian Analysis of the Multivariate Geographical Distribution of the Socio-Economic Environment in England.
Environmetrics, 18: 745–758, 2007.


Hogan, J.W., and R. Tchernis.
Bayesian Factor Analysis for Spatially Correlated Data, With Application to Summarizing Area-Level Material Deprivation from Census Data.
Journal of the American Statistical Association, 99: 314–324, 2004.


Jordan H., P. Roderick and D. Martin.
The Index of Multiple Deprivation 2000 and Accesibility Effects on Health.
J Epidemiological Community Health, 58: 250–257, 2004.


Havard S et al.
A Small-Area Index of Socio-Economic Deprivation to Capture Health Inequalities in France .
Social Science & Medicine, 67: 2007–2016, 2008.