# Random graphical modelling of cross-country cultural heterogeneity

**Veronica Vinciotti**

University of Trento

Perugia, 23 February 2024

Joint work with Ernst Wit and Luca De Benedictis

# Culture and the social sciences

## Culture (now) recognised as important beyond the Humanities

- **Economics**: "cultural variables ... affect the speed of development and the wealth of nations" (Alesina and Giuliano, 2015)
- **Politics**: Political actions depend on culture (Lane and Ersson, 2016)
- **Sociology**: Culture shapes individual identity (Schwartz, 2008)
- **Management**: Organizations and local culture (Yeganeh and Su, 2006)
- **Anthropology**: Humans evolve slowly through culture (Ruck et al., 2020)
- **Psychology**: Personality traits and culture (Kashima et al., 2019)
- **International Business** is about cultural differences (Taras et al., 2009)

# Culture: Definition and Measurement

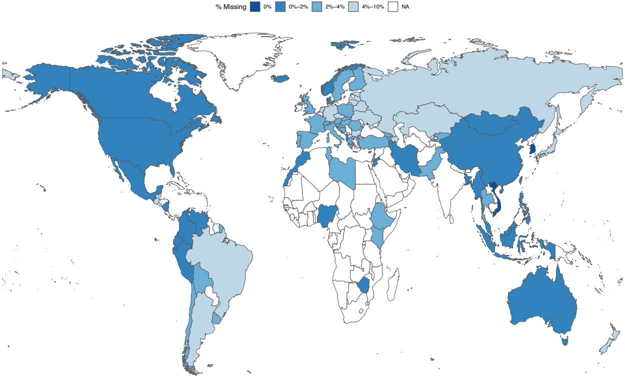
## Defining culture

- From Taylor (1871) to UNESCO (2001), defining culture is problematic
- Up to 160 possible definitions of culture (Kroeber and Kluckhohn, 1952)
- Some conclude that the very notion of cultural diversity implies that there cannot be any generally agreed definition of culture (Jahoda, 2012)

## One (operationalisable) definition

- Culture is the/a set of local norms, customs, attitudes and values (Alesina and Giuliano, 2015)
- **Implication**: Latent dimensions of culture

# Culture: Definition and Measurement

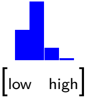



- Unit of analysis: cultural traits
- Traditionally, cultural traits are summarised by their means on survey data (Hofstede, 1980; Schwartz, 1994; Inglehart and Welzel, 2005)
- We use **European and World Values Survey data** (Wave 7, 2017-21)

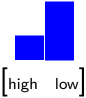
$\rightsquigarrow$ 84 countries and 10 cultural traits

# European and World Values Survey data: Cultural traits

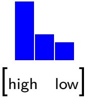
level of **happiness**
(1:4, V10 (H))



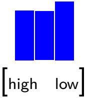

**trust** in people
(1:2, V24 (T))



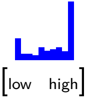

**respect for authority**
(1:3, V69 (R))



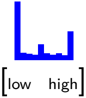

**voice** through petitions
(1:3, V85 (V))



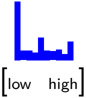

**importance of God**
(1:10, V152 (G))



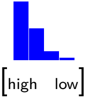

justification of **homosexuality**
(1:10, V203 (O))



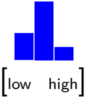

justification of **abortion**
(1:10, V204 (A))



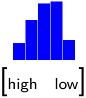

**national pride**
(1:4, V211 (P))



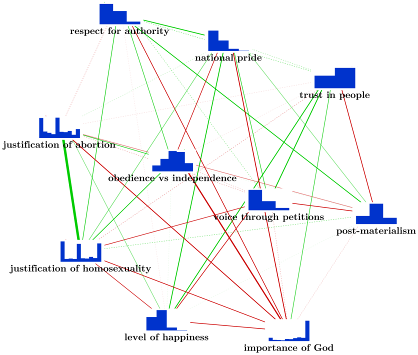

**post-materialism**
(1:3, Y002 (M))



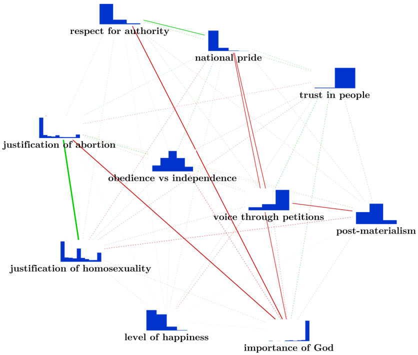

**obedience vs independence**
(1:5, Y003 (B))



# Culture as a network of cultural traits




(a) USA Cultural Network




(b) Philippines Cultural Network

$\underbrace{\texttt{Jeffreys' Divergence (Marginals)} = 5.59}_{}$, $\texttt{JD Network} = 0.87$

**H**: 0.10, **T**: 1.07, **R**: 0.24, **V**: 2.37, **G**: 0.66,
**O**: 0.14, **A**: 0.61, **P**: 0.30, **M**: 0.07, **B**: 0.05

# Cultural network heterogeneity

A cross-country comparative cultural approach to discover:

- Extent of cultural network heterogeneity
- Structural similarities between national cultures
- Drivers of cross-country cultural heterogeneity

... by modelling the process that generates the networks!

# Random graphical model

$\mathbf{Y}^{(k)} = (Y_1^{(k)}, \ldots, Y_p^{(k)})$ levels of $p$ cultural traits in country $k$

## Random graph model

Graphs $G = \left\{G^{(k)}\right\}_k$ distributed according to a joint random graph model

$$G \sim P(\mathbf{\Theta})$$

for some vector of parameters $\mathbf{\Theta}$

## Graphical Model

$\mathbf{Y}^{(k)}$ distributed according to some graphical model (GM)

$$\mathbf{Y}^{(k)} | G^{(k)} \sim GM(G^{(k)}; \mathbf{\Omega}^{(k)})$$

relative to a C.I. graph $G^{(k)}$ with associated parameters $\mathbf{\Omega}^{(k)}$

random graph model + graphical model = random graphical model

# Random graphical model for cultural networks

**random graph model**

$\boldsymbol{\Theta}$

$$\Big\downarrow\; G^{(k)} \sim \text{Bernoulli}(P(G^{(k)} = 1 \mid G^{(-k)}, \boldsymbol{\Theta}))$$

**Gaussian copula GM**

$$G^{(1)}, \ldots, G^{(K)}$$

$$\Big\downarrow\; \boldsymbol{\Omega}^{(k)} \sim W_G(3, \mathbb{I}_p)$$

$$\boldsymbol{\Omega}^{(1)}, \ldots, \boldsymbol{\Omega}^{(K)}$$

$$\Big\downarrow\; \mathbf{z_i}^{(k)} \sim N_p(\mathbf{0}, {\boldsymbol{\Omega}^{(k)}}^{-1})$$

$$\boldsymbol{z}_1^{(k)}, \ldots, \mathbf{z}_n^{(k)} \qquad\qquad F_{11}^{(k)}, \ldots, F_{np}^{(k)}$$

$$\searrow \qquad \swarrow$$

$$\mathbf{y}_1^{(k)}, \ldots, \boldsymbol{y}_n^{(k)}$$

# Random graphical model for cultural networks

**random graph model**

$\boldsymbol{\Theta}$

$$G^{(k)} \sim \text{Bernoulli}(P(G^{(k)} = 1 \mid G^{(-k)}, \boldsymbol{\Theta}))$$

**Gaussian copula GM**

$G^{(1)}, \ldots, G^{(K)}$

$$\boldsymbol{\Omega}^{(k)} \sim W_G(3, \mathbb{I}_p)$$

$\boldsymbol{\Omega}^{(1)}, \ldots, \boldsymbol{\Omega}^{(K)}$

$$\mathbf{z_i}^{(k)} \sim N_p(\mathbf{0}, \boldsymbol{\Omega}^{(k)^{-1}})$$

$\mathbf{z}_1^{(k)}, \ldots, \mathbf{z}_n^{(k)}$ $\quad$ $F_{11}^{(k)}, \ldots, F_{np}^{(k)}$

$\mathbf{y}_1^{(k)}, \ldots, \mathbf{y}_n^{(k)}$

# Graph generative model: latent network (probit) model

Joint modelling $\{G^{(k)}\}_k$ to discover the extent of structural similarities between the different countries:

$$P({G_{j_1 j_2}}^{(k)} = 1 | G^{(-k)}) = \Phi\Big(\alpha_k + \underbrace{\beta^t \mathbf{w}}_{\text{covariates}} + \underbrace{\mathbf{c}_k^t \sum_{k' \neq k} \mathbf{c}_{k'} (1_{\{{G_{j_1, j_2}}^{(k')} = 1\}} - 1_{\{{G_{j_1, j_2}}^{(k')} = 0\}})}_{\text{latent space environments}}\Big)$$

where

- ${G_{ij}}^{(k)} = 1$: edge between node $Y_i$ and node $Y_j$ in condition $k$
- $\mathbf{w} \in \mathbb{R}^d$: edge-specific and/or country-specific covariates
- $\mathbf{c}_1, \ldots, \mathbf{c}_K \in \mathbb{R}^2$: latent space variables for each environment
- $\alpha_k$: sparsity level of graph $G^{(k)}$

Given $\boldsymbol{\Theta} = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{c})$ and $G^{(-k)}$, edges in $G^{(k)}$ become independent

# Random graphical model for cultural networks

$\Theta$

**latent network probit model**

$G^{(k)} \sim \text{Bernoulli}(P(G^{(k)} = 1 \mid G^{(-k)}, \Theta))$

**Gaussian copula GM**

$G^{(1)}, \ldots, G^{(K)}$

$\boldsymbol{\Omega}^{(k)} \sim W_G(3, \mathbb{I}_p)$

$\boldsymbol{\Omega}^{(1)}, \ldots, \boldsymbol{\Omega}^{(K)}$

$\mathbf{z_i}^{(k)} \sim N_p(\mathbf{0}, \boldsymbol{\Omega}^{(k)^{-1}})$

$\boldsymbol{z}_1^{(k)}, \ldots, \mathbf{z}_n^{(k)}$ $\qquad$ $F_{11}^{(k)}, \ldots, F_{np}^{(k)}$

$\mathbf{y}_1^{(k)}, \ldots, \boldsymbol{y}_n^{(k)}$

# Gaussian copula graphical model

$\mathbf{Y}^{(k)} = (Y_1^{(k)}, \ldots, Y_p^{(k)})^\top$ in condition $k$ have a joint distribution:

$$F(Y_1^{(k)} \leq y_1, \ldots, Y_p^{(k)} \leq y_p) = C(F_1^{(k)}(y_1), \ldots, F_p^{(k)}(y_p)),$$

with marginal distributions $F_j^{(k)}$ and copula $C$

## Gaussian copula

$$C(F_1(y_1), \ldots, F_p(y_p) \mid \mathbf{K}) = \Phi_p(\Phi^{-1}(F_1(y_1)), \ldots, \Phi^{-1}(F_p(y_p)))$$

with $\Phi_p$ the CDF of $\mathcal{N}_p(0, \mathbf{K})$ with correlation $\mathbf{K}$ and $\Phi$ the $N(0,1)$ CDF

Conditional independence graph $G^{(k)}$ associated to $\boldsymbol{\Omega}^{(k)} = (\mathbf{K}^{(k)})^{-1}$

# Marginals: just a nuisance?

- Marginals are typically estimated non-parametrically
- But covariates are often available (e.g., age and gender of the respondent) and may have an effect on the response to a cultural trait
- The dependence structure ($\boldsymbol{\Omega}$) may not depend on $\boldsymbol{x}$, but ignoring $\boldsymbol{x}$ may distort the estimation of $\boldsymbol{\Omega}$
- Sample configurations may be different between countries, so ignoring $\boldsymbol{x}$ may distort cross-country comparisons

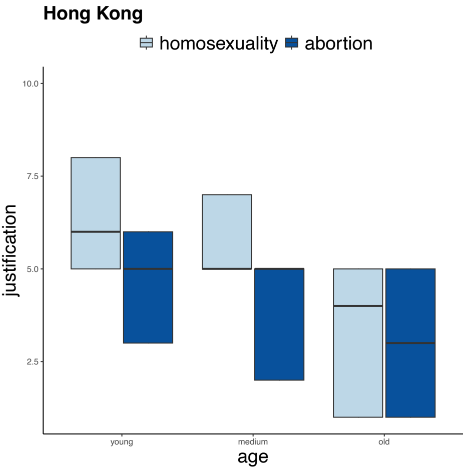
Covariates $\rightsquigarrow$ parametric models at the level of marginals

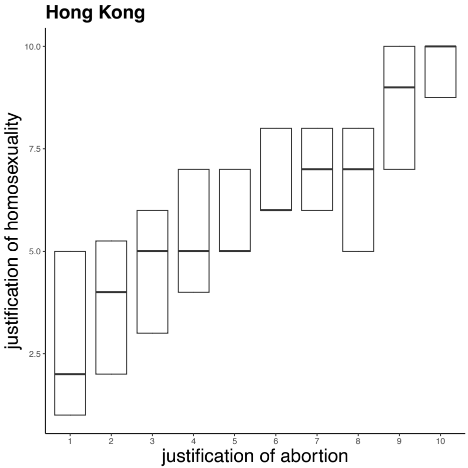
Ordinal data $\rightsquigarrow$ ordinal regression

# Marginal covariate adjustment matters!

$Y_j^{(k)}$ for cultural trait $j$ in country $k$ is modelled by

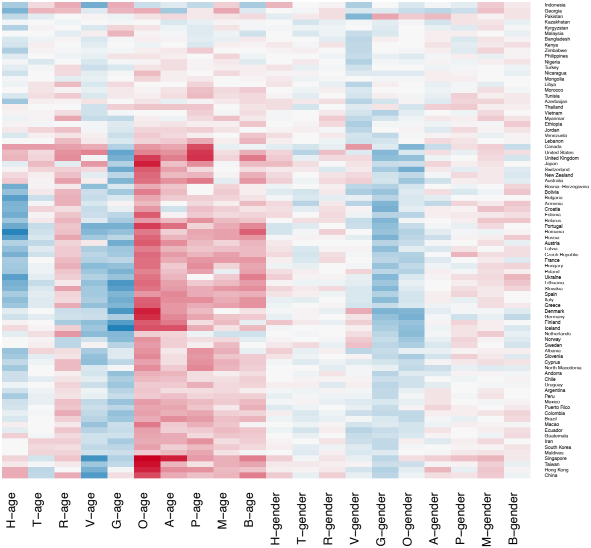
$$F_j^{(k)}(c|\boldsymbol{X}=\boldsymbol{x}) = \eta_{jc}^{(k)} + \gamma_j{}^{(k)}\mathbf{x}, \quad \boldsymbol{x} = (\text{age}, \text{gender})$$



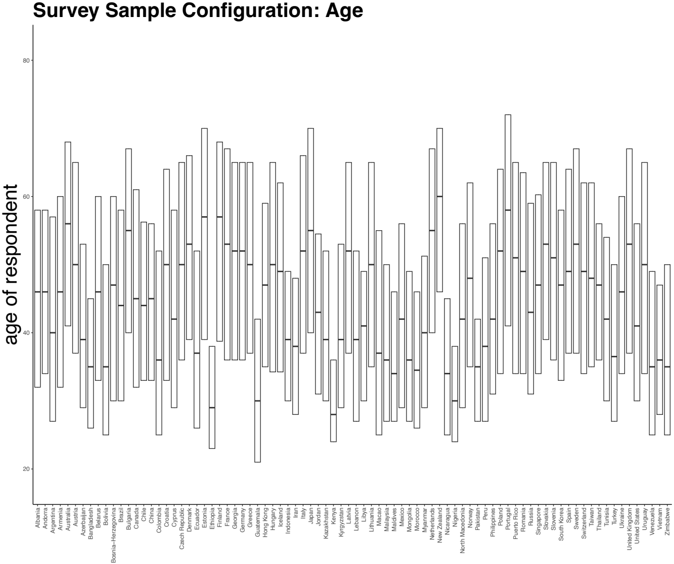



# Age and gender significant across most countries/traits

standardized $\gamma$ coefficients: **red** $< 0$, **blue** $> 0$

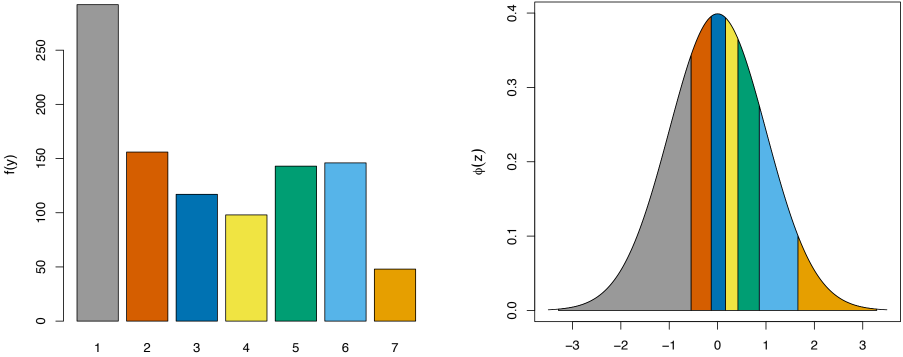


# Sample configuration varies across countries



# From discrete data to latent Gaussian intervals

$Y_j$ discrete $\rightsquigarrow F_j$ not injective $\rightsquigarrow$ projection to the latent Gaussian space $z_j = \Phi^{-1}(F_j(y_j))$ of the copula not unique




Ordinal $Y$      Gaussian $Z = \phi^{-1}(F(Y))$

Each observation $(y_j, \boldsymbol{x})$ associated to an interval in the latent space

$$\mathcal{I}_{F_j}(y_j|\boldsymbol{x}) = \left(\Phi^{-1}(F_j(y_j - 1|\boldsymbol{x})), \Phi^{-1}(F_j(y_j|\boldsymbol{x}))\right]$$

# Bayesian inference: $G^{(k)}$, $\mathbf{\Omega}^{(k)}$, $\mathbf{\Theta}$

Following Hoff(2007), the likelihood function (for condition $k$) is:

$$P(\boldsymbol{Z} \in \mathcal{I}_F(\boldsymbol{y}) \mid \mathbf{\Omega}, G, \mathbf{\Theta}) = \int_{\mathcal{I}_F(\boldsymbol{y})} P(\boldsymbol{Z}|\mathbf{\Omega}, G) \; dZ$$

with $P(\boldsymbol{Z}|\mathbf{\Omega}, G, \mathbf{\Theta})$ the profile likelihood in the Gaussian latent space:

$$P(\boldsymbol{Z}|\mathbf{\Omega}, G, \mathbf{\Theta}) \propto |\mathbf{\Omega}|^{n/2} \exp\left\{-\frac{1}{2}\mathrm{Trace}(\mathbf{\Omega}\mathbf{U})\right\}$$

with $\boldsymbol{U} = \boldsymbol{Z}^t \boldsymbol{Z}$ the sample moment

Likelihood is combined to priors to lead to the posterior:

$$P(\mathbf{\Omega}, G, \mathbf{\Theta} \mid \boldsymbol{Z} \in \mathcal{I}_F(\boldsymbol{y})) \propto \underbrace{P(\boldsymbol{Z} \in \mathcal{I}_F(\boldsymbol{y}) \mid \mathbf{\Omega})}_{\substack{\text{Truncated} \\ \text{Normal}}} \underbrace{P(\mathbf{\Omega}|G)}_{\text{G-Wishart}} \underbrace{P(G|\mathbf{\Theta})}_{\substack{\text{Ber}(\pi) \\ \text{on each link}}} \underbrace{P(\mathbf{\Theta})}_{\text{Normal}}$$

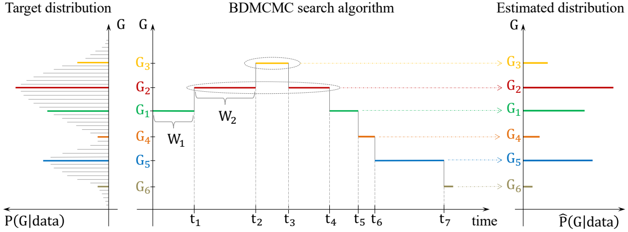
# Bayesian inference: MCMC scheme

1. Gibbs sampling (probit regression with offset)
   - $\boldsymbol{\alpha}|\boldsymbol{\beta}, \mathbf{c}, \{G^{(k)}\}_k, \mathbf{w}$
   - $\boldsymbol{\beta}|\boldsymbol{\alpha}, \mathbf{c}, \{G^{(k)}\}_k, \mathbf{w}$
   - $\mathbf{c}_k|\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{c}_{-k}, \{G^{(k)}\}_k, \mathbf{w}$
2. Gibbs sampling ${z_{ij}}^{(k)} \mid \boldsymbol{\Omega}^{(k)}, {\mathbf{z}_{i,-j}}^{(k)}, {\boldsymbol{y}_i}^{(k)}$, truncated on $\mathcal{I}_{\hat{F}_j}(y_{ij}|\mathbf{x}_i)$
3. Gibbs sampling $\boldsymbol{\Omega}^{(k)} \mid G^{(k)}, \mathbf{z}^{(k)}$ (G-Wishart)
4. Continuous time birth-death MCMC for

$$\left(G^{(k)}\right)^{\pm e} \mid \boldsymbol{\Omega}^{(k)}, \mathbf{z}^{(k)}, G^{(k)}, \boldsymbol{\Theta}, \mathbf{w}$$

$\rightsquigarrow$ Posterior distributions of $G^{(k)}$, $\boldsymbol{\Omega}^{(k)}$, $\boldsymbol{\Theta}$

# Bayesian structural learning: graph uncertainty




From the graphs posterior, **posterior edge inclusion probabilities**:

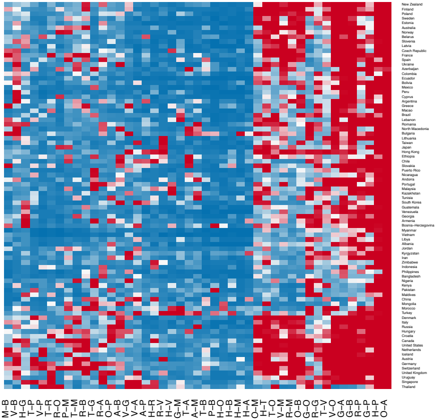
$$\pi_e{}^{(k)} = P(e \in E \mid \boldsymbol{Y}, k) = \frac{\sum_{t=1}^{N} 1(e \in G_t{}^{(k)}) W(\boldsymbol{\Omega}_t{}^{(k)}, \boldsymbol{\Theta})}{\sum_{t=1}^{N} W(\boldsymbol{\Omega}_t{}^{(k)}, \boldsymbol{\Theta})},$$

with $E$: set of edges, $N$: MCMC iterations and $W(\boldsymbol{\Omega}_t{}^{(k)}, \boldsymbol{\Theta})$ waiting time for graph $G_t{}^{(k)}$ with precision matrix $\boldsymbol{\Omega}_t{}^{(k)}$

# EWVS study: graph uncertainty

Heatmap of posterior edge probabilities for each country:

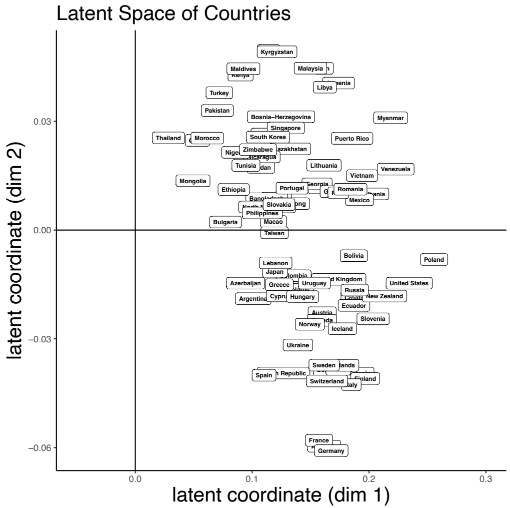
**blue** $\approx 0 \rightarrow$ **red** $\approx 1$




- High sparsity across countries ($\rightsquigarrow$ low $\alpha_k$ values)
- High structural similarity with heterogeneity

# EWVS study: random graph model (latent space)

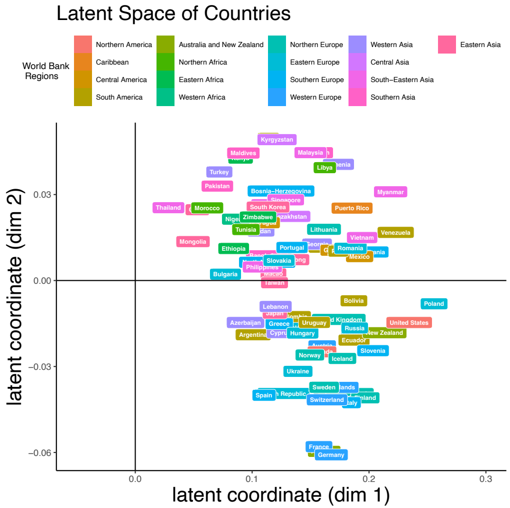
$$P({G_{j_1 j_2}}^{(k)} = 1 | G^{(-k)}) = \Phi\Big(\alpha_k + \mathbf{c}_k^t \sum_{k' \neq k} \mathbf{c}_{k'} (1_{\{{G_{j_1,j_2}}^{(k')}=1\}} - 1_{\{{G_{j_1,j_2}}^{(k')}=0\}})\Big)$$




Some heterogeneity but also some noise ...

# EWVS study: random graph model (latent space)

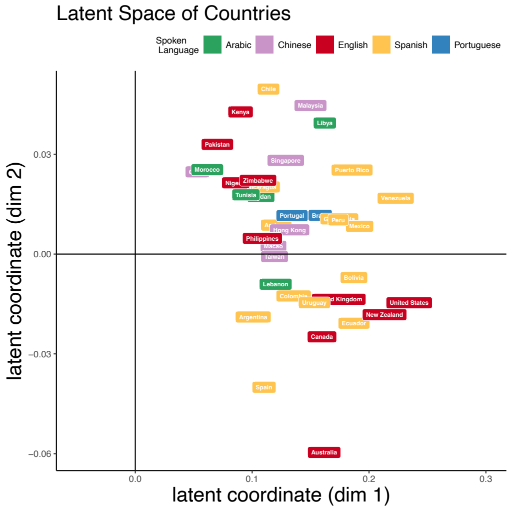
$$P({G_{j_1 j_2}}^{(k)} = 1 | G^{(-k)}) = \Phi\Big(\alpha_k + \mathbf{c}_k^t \sum_{k' \neq k} \mathbf{c}_{k'} (1_{\{G_{j_1, j_2}{}^{(k')} = 1\}} - 1_{\{G_{j_1, j_2}{}^{(k')} = 0\}})\Big)$$




Some heterogeneity but also possible drivers ...

# EWVS study: random graph model (latent space)

$$P({G_{j_1 j_2}}^{(k)} = 1 | G^{(-k)}) = \Phi\Big(\alpha_k + \mathbf{c}_k^t \sum_{k' \neq k} \mathbf{c}_{k'} \big(1_{\{{G_{j_1, j_2}}^{(k')}=1\}} - 1_{\{{G_{j_1, j_2}}^{(k')}=0\}}\big)\Big)$$

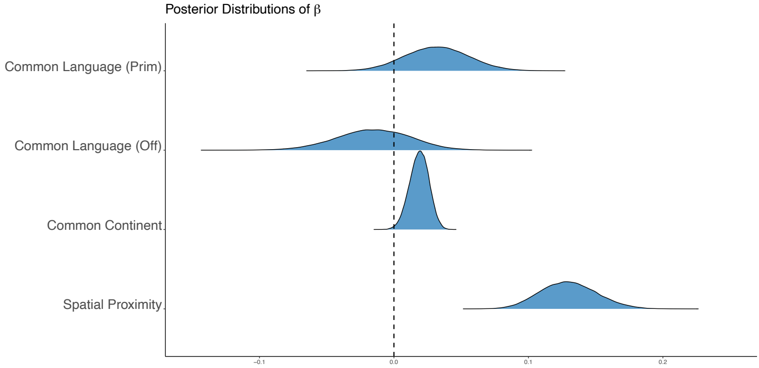



Some heterogeneity but also possible drivers ...

# What are the drivers of cultural heterogeneity?

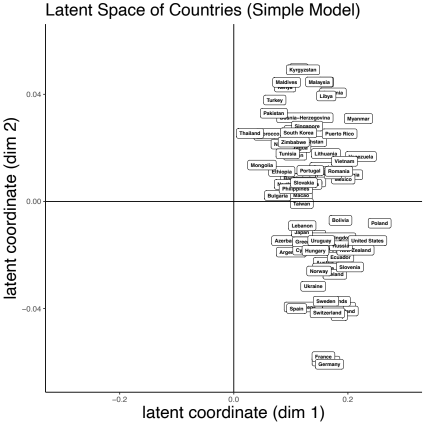
$$P(G_{j_1 j_2}{}^{(k)} = 1 | G^{(-k)}) = \Phi\Big(\alpha_k + \beta^t \sum_{k' \neq k} \underbrace{\textbf{sim}_{k,k'}}_{\substack{\text{similarity between} \\ \text{country } k \text{ and } k'}} \Big(1_{\{G_{j_1 j_2}{}^{(k')}=1\}} - 1_{\{G_{j_1 j_2}{}^{(k')}=0\}}\Big)$$

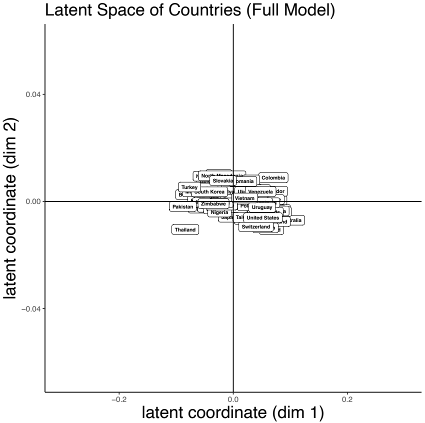
$$+ \mathbf{c}_k^t \sum_{k' \neq k} \mathbf{c}_{k'} (1_{\{G_{j_1, j_2}{}^{(k')}=1\}} - 1_{\{G_{j_1, j_2}{}^{(k')}=0\}})\Big)$$




Structural similarities driven by geographical and historical factors

# Latent space less important ...



# Latent space statistically not needed

Deviance Information Criterion comparison:

| random graph model | parameters | DIC |
|---|---|---|
| country-specific intercepts | $\alpha_k$ | 3,116,506 |
| country-specific intercepts + latent space | $\alpha_k$, $\mathbf{c}_k$ | 3,115,201 |
| country-specific intercepts + proximity measures | $\alpha_k$, $\beta$ | 3,095,245 |
| country-specific intercepts + proximity measures + latent space | $\alpha_k$, $\beta$, $\mathbf{c}_k$ | 3,100,461 |

Best model:

$$P({G_{j_1 j_2}}^{(k)} = 1 | G^{(-k)}) = \Phi\Big(\alpha_k + \beta^t \sum_{k' \neq k} \mathbf{sim}_{k,k'} \Big(1_{\{{G_{j_1 j_2}}^{(k')}=1\}} - 1_{\{{G_{j_1 j_2}}^{(k')}=0\}}\Big)\Big)$$

# References I


Alesina, A. and P. Giuliano (2015). Culture and Institutions. *Journal of Economic Literature 53*(4), 898–944.

Hofstede, G. (1980). *Culture's consequences: International differences in work-related values*. Sage.

Inglehart, R. and C. Welzel (2005). *Modernization, cultural change, and democracy: The human development sequence*. Cambridge University Press.

Jahoda, G. (2012). Critical reflections on some recent definitions of "culture". *Culture & Psychology 18*(3), 289–303.

Kashima, Y., P. G. Bain, and A. Perfors (2019). The psychology of cultural dynamics: What is it, what do we know, and what is yet to be known? *Annual Review of Psychology 70*, 499–529.

Kroeber, A. L. and C. Kluckhohn (1952). Culture: A critical review of concepts and definitions. *Papers. Peabody Museum of Archaeology & Ethnology, Harvard University*.

# References II


Lane, J.-E. and S. Ersson (2016). *Culture and politics: A comparative approach*. Routledge.

Ruck, D. J., R. A. Bentley, and D. J. Lawson (2020). Cultural prerequisites of socioeconomic development. *Royal Society Open Science 7*(2), 190725.

Schwartz, S. H. (1994). Beyond individualism/collectivism: New cultural dimensions of values. *SAGE publications, Inc*.

Schwartz, S. H. (2008). Cultural value orientations: Nature and implications of national differences. *Moscow: Publishing house of SU HSE*.

Taras, V., J. Rowney, and P. Steel (2009). Half a century of measuring culture: Review of approaches, challenges, and limitations based on the analysis of 121 instruments for quantifying culture. *Journal of International Management 15*(4), 357–373.

# References III

Taylor, E. B. (1871). Primitive culture: Researches into the development of mythology, philosophy, religion, language, art and custom. *NY, US: Henry Holt and Company. xii*.

UNESCO (2001). Universal declaration on cultural diversity. Adopted by the 31st Session of the General Conference of UNESCO, Paris, 2 November.

Yeganeh, H. and Z. Su (2006). Conceptual foundations of cultural management research. *International Journal of Cross Cultural Management 6*(3), 361–376.