Modeling of complex stochastic systems via latent factors

Hedibert Freitas Lopes
Booth School of Business
University of Chicago

Colóquio Interinstitucional
Modelos Estocásticos e Aplicações
Universidade Federal Fluminense
19 de setembro de 2012.

Outline

Factor analysis: early days

Bartholomew (1995)[1] starts his paper by saying that

> *Spearman invented factor analysis but his almost exclusive concern with the notion of a general factor prevented him from realizing its full potential.*

Factor analysis, however, has flourished ever since Spearman's (1904) seminal paper on the American Journal of Psychology entitled "General Inteligente objectively determined and measured".

Factor models are mainly applied in two major situations:

1. Data reduction,
2. Identifying underlying structures.

[1] Spearman and the origin and development of factor analysis, *British Journal of Mathematical and Statistical Psychology*, 48, 211-220.

Basic model

The Gaussian linear factor model relates a m-vector of observables y_t to a k-vector of latent variables f_t via

$$y_t | f_t, \Theta \sim N(\beta f_t, \Sigma),$$

where $\Theta = (\beta, \Sigma)$, $\Sigma = \mathrm{diag}(\sigma_1^2, \cdots, \sigma_m^2)$, and, *a priori*,

$$f_t \Theta \sim N(0, I_k).$$

Conditional variance: The ommon latent factors explain all the dependence structure among the m variables:

$$\mathrm{cov}(y_{it}, y_{jt} | f_t, \Theta) = \begin{cases} \sigma_i^2 & i = j \\ 0 & i \neq j \end{cases}$$

Unconditional variance:

$$V(y_t | \Theta) = \Omega = \beta\beta' + \Sigma$$

Classical literature

- Lawley (1940,1941)
- Anderson and Rubin (1956)
- Jöreskog (1969,1970)
- Rubin and Thayer (1982)
- Bentler and Tanaka (1983)
- Rubin and Thayer (1983)
- Akaike (1987)
- Anderson and Amemiya (1988)
- Amemiya and Anderson (1990)

Bayes pre-MCMC

- Press (1972)
- Martin and McDonald (1975)
- Geweke and Singleton (1980)
- Bartholomew (1981)
- Lee (1981)
- Press and Shigemasu (1989)

Bayes post-MCMC

- Geweke and Zhou (1996)
- Aguilar and West (2000)
- Lopes, Aguilar and West (2000)
- Lopes and Migon (2002)
- West (2003)
- Wang and Wall (2003)
- Lopes and West (2004)
- Quinn (2004)
- Hogan and Tchernis (2004)
- Lopes, Salazar and Gamerman (2008)
- Carvalho *et al.* (2008)
- Chib and Ergashev (2009)
- Frühwirth-Schnatter and Lopes (2009)
- Carvalho, Lopes and Aguilar (2011)
- Bhattacharya and Dunson (2011)
- Lopes *et al.* (2012)
- Hahn and Lopes (2013)

Invariance

The model is invariant under transformations of the form $\tilde{\beta} = \beta P'$ and $\tilde{f}_t = P f_t$, for any orthogonal matrix P:

$$\Omega = \beta\beta' + \Sigma = \tilde{\beta}\tilde{\beta}' + \Sigma$$

Two standard solutions

- Classical approach: $\beta'\Sigma^{-1}\beta = I$.
- Bayesian approach: β is a block lower triangular.

More general solution (Frühwirth-Schnatter and Lopes, 2009): β is generalized block lower triangular.

This last form provides both identification and, often, useful interpretation of the factor model.

Number of parameters

The resulting number of parameters in Ω is

$$m(m+1)/2 - m(k+1) + k(k-1)/2 \geq 0,$$

which provides an upper bound on k.

For example,

- $m = 6$ implies $k \leq 3$,
- $m = 12$ implies $k \leq 7$,
- $m = 20$ implies $k \leq 14$,
- $m = 50$ implies $k \leq 40$,

Even for small m, the bound will often not matter as relevant k values will not be so large.

Full-rank loading matrix

Geweke and Singleton (1980) show that, if β has rank $r < k$ there exists a matrix Q such that $\beta Q = 0$ and $Q'Q = I$ and, for any orthogonal matrix M,

$$\beta\beta' + \Sigma = (\beta + MQ')'(\beta + MQ') + (\Sigma - MM')$$

This translation invariance of Ω under the factor model implies lack of identification and, in application, induces symmetries and potential multimodalities in resulting likelihood functions.

This issue relates intimately to the question of uncertainty of the number of factors.

Ordering of the variables

Alternative orderings are trivially produced via Ay_t for some rotation matrix A.

The new rotation has the same latent factors but transformed loadings matrix $A\beta$.

$$Ay_t = A\beta f + \varepsilon_t$$

This new loadings matrix does not have the lower triangular structure.

However, we can always find an orthonormal matrix P such that $A\beta P'$ is lower triangular, and so simply recover the factor model with the same probability structure for the underlying latent factors Pf_t (Lopes and West, 2004).

The order of the variables in y_t is irrelevant assuming that k is properly chosen.

Prior specification

Loading matrix:

$$\begin{aligned}\beta_{ij} &\sim N(0, C_0) & \text{when} \quad i \neq j, \\ \beta_{ii} &\sim N(0, C_0)1(\beta_{ii} > 0) & \text{when} \quad i = 1, \ldots, k\end{aligned}$$

Idiosyncratic variances

$$\sigma_i^2 \sim IG(\nu/2, \nu s^2/2)$$

where s^2 is the prior mode of each σ_i^2 and ν the prior degrees of freedom hyperparameter.

We eschew the use of standard improper reference priors $p(\sigma_i^2) \propto 1/\sigma_i^2$, since such priors lead to the Bayesian analogue of the so-called *Heywood problem* (Martin and McDonald, 1975, and Ihara and Kano, 1995).

Full conditional distributions

Factor scores

$$f_t \sim N(V_f \beta' \Sigma^{-1} y_t, V_f)$$

where $V_f = (I_k + \beta' \Sigma^{-1} \beta)^{-1}$.

Idiosyncrasies

$$\sigma_i^2 \sim IG((\nu + T)/2, (\nu s^2 + d_i)/2)$$

where $d_i = (y_i - f \beta_i')'(y_i - f \beta_i')$.

First k rows of β

$$\beta_i \sim N(M_i, C_i) 1(\beta_{ii} > 0)$$

where

$$\begin{aligned} M_i &= C_i \left(C_0^{-1} \mu_0 1_i + \sigma_i^{-2} f_i' y_i \right) \\ C_i^{-1} &= C_0^{-1} I_i + \sigma_i^{-2} f_i' f_i. \end{aligned}$$

Last $m - k$ rows of β

$$\beta_i \sim N(M_i, C_i)$$

where

$$\begin{aligned} M_i &= C_i \left(C_0^{-1} \mu_0 1_k + \sigma_i^{-2} f' y_i \right) \\ C_i^{-1} &= C_0^{-1} I_k + \sigma_i^{-2} f' f. \end{aligned}$$

Example: Lopes and West (2004)

Monthly international exchange rates.

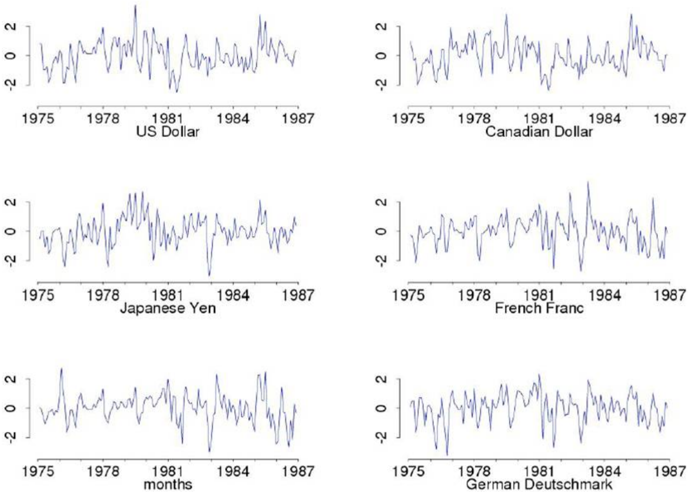
The data span the period from 1/1975 to 12/1986 inclusive.

Time series are the exchange rates in British pounds of

- US dollar (US)
- Canadian dollar (CAN)
- Japanese yen (JAP)
- French franc (FRA)
- Italian lira (ITA)
- (West) German (Deutsch)mark (GER)

Exchange rates

Standardized first differences of monthly log exchange rates



Standardized first differences of monthly observed exchange rates.

Posterior means

1st ordering

$$E(\beta|y) = \begin{pmatrix} \text{US} & 0.99 & 0.00 \\ \text{CAN} & 0.95 & 0.05 \\ \text{JAP} & 0.46 & 0.42 \\ \text{FRA} & 0.39 & 0.91 \\ \text{ITA} & 0.41 & 0.77 \\ \text{GER} & 0.40 & 0.77 \end{pmatrix} \qquad E(\Sigma|y) = \text{diag}\begin{pmatrix} 0.05 \\ 0.13 \\ 0.62 \\ 0.04 \\ 0.25 \\ 0.28 \end{pmatrix}$$

2nd ordering

$$E(\beta|y) = \begin{pmatrix} \text{US} & 0.98 & 0.00 \\ \text{JAP} & 0.45 & 0.42 \\ \text{CAN} & 0.95 & 0.03 \\ \text{FRA} & 0.39 & 0.91 \\ \text{ITA} & 0.41 & 0.77 \\ \text{GER} & 0.40 & 0.77 \end{pmatrix} \qquad E(\Sigma|y) = \text{diag}\begin{pmatrix} 0.06 \\ 0.62 \\ 0.12 \\ 0.04 \\ 0.25 \\ 0.26 \end{pmatrix}$$

More structure

Factor stochastic volatility models

Dynamic stock factor models

Factor-augmented vector autoregressions

Spatial dynamic factor models

Hierarchical factor models

Sparse factor models

Factor SV

The p-vector of time series y_t follows a k-order factor model:

$$
\begin{aligned}
y_t|f_t &\sim N(\beta f_t, \Sigma_t) \qquad \Sigma_t = \text{diag}(\sigma^2_{1t}, \ldots, \sigma^2_{pt}) \\
f_t &\sim N(0, H_t) \qquad H_t = \text{diag}(\sigma^2_{p+1,t}, \ldots, \sigma^2_{p+k,t})
\end{aligned}
$$

where

$$
\begin{aligned}
\eta_{it} = \log(\sigma^2_{it}) &\sim N(\alpha_i + \gamma_i \eta_{i,t-1}, \xi^2_i) \\
\lambda_{jt} = \log(\sigma^2_{it}) &\sim N(\mu_j + \phi_j \lambda_{j,t-1}, \tau^2_j)
\end{aligned}
$$

.

Aguilar and West (2000) introduce contemporaneous covariation in the common factor log-volatilities.

Let $\lambda_t = (\sigma^2_{p+1,t}, \ldots, \sigma^2_{p+k,t})'$, $\mu = (\mu_1, \ldots, \mu_k)$ and $\Phi = \text{diag}(\phi_1, \ldots, \phi_k)$, then

$$\lambda_t \sim N(\alpha + \phi\lambda_{t-1}, U)$$

where U is a full covariance matrix.

Lopes and Carvalho (2007) introduce time-varying loadings, β_t.

The $d = pk - k(k-1)/2$ unconstrained elements of β_t, namely $\beta_{21,t}, \beta_{31,t}, \ldots, \beta_{p,k,t}$, are modeled by simple first order autoregressive models, ie.

$$\beta_{ijt} \sim N(\zeta_{ij} + \Theta_{ij}\beta_{ij,t-1}, \omega^2_{ij})$$

for $i = 2, \ldots, p$ and $j = 1, \ldots, \min(i-1, k)$.

Example: Lopes-Carvalho (2007)

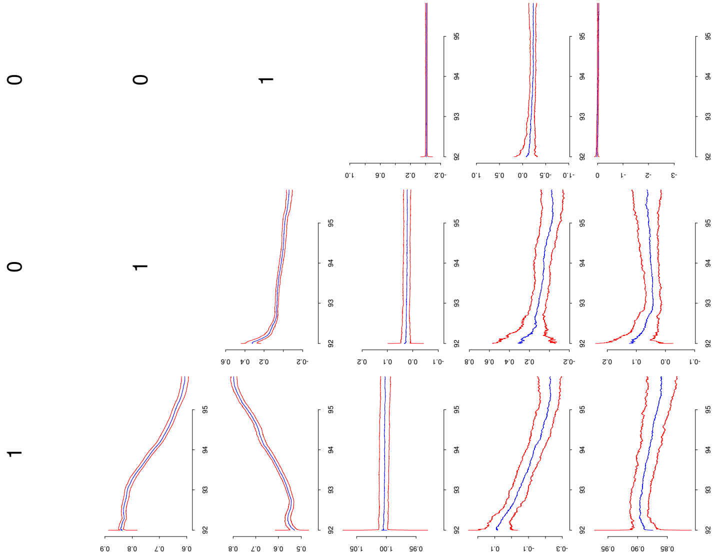
Returns on weekday closing spot prices for six currencies relative to the US dollar.

The data span the period from 1/1/1992 to 10/31/1995.

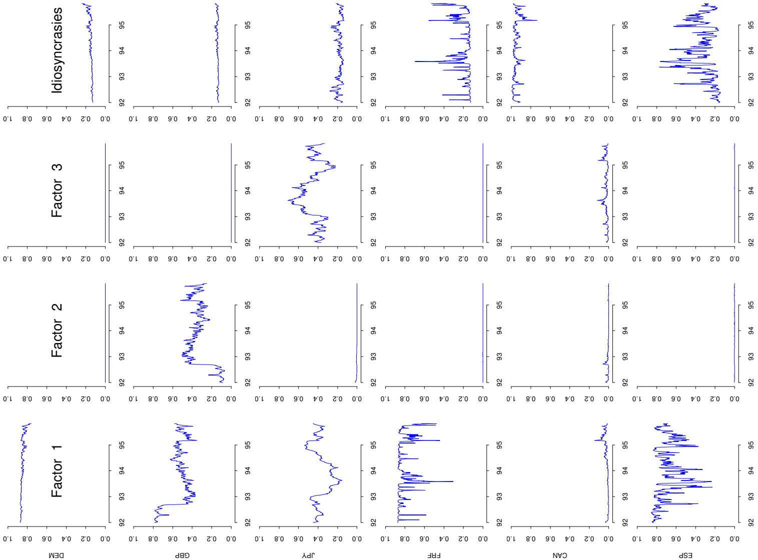
- German Mark(DEM)
- British Pound(GBP)
- Japanese Yen(JPY)
- French Franc(FRF)
- Canadian Dollar(CAD)
- Spanish Peseta(ESP)

A 3-factor stochastic volatility model with time-varying loadings was implemented with relatively vague priors for all model parameters.

Time-varying loadings



Variance decomposition



Spatial dynamic factor models

Lopes, Salazar and Gamerman (2008) introduces the following spatio-temporal model for $y_t = (y_{1t}, \ldots, y_{mT})'$, measurements on m spatial locations and over T time periods:

Dimension reduction:

$$y_t \sim N(\beta f_t, \Sigma)$$

Time series component:

$$f_t \sim N(\Gamma f_{t-1}, \Gamma)$$

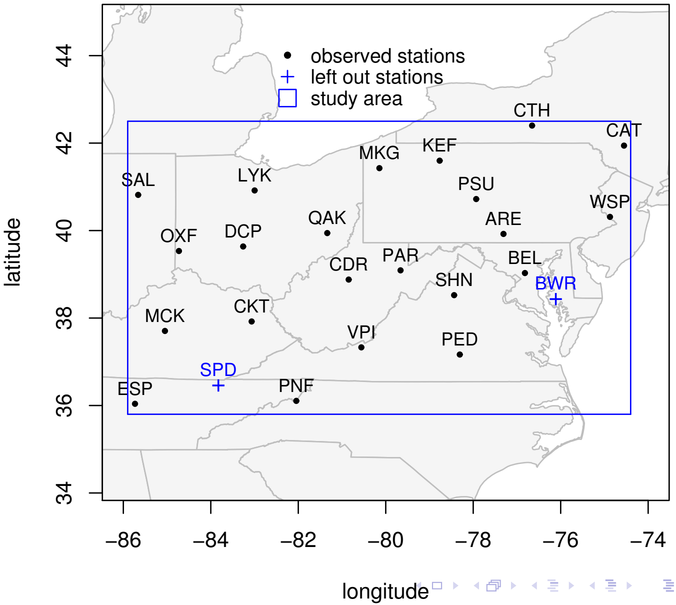
Spatial component:

$$\beta_j \sim GP(\mu_j, \tau_j^2 R_{\phi_j})$$

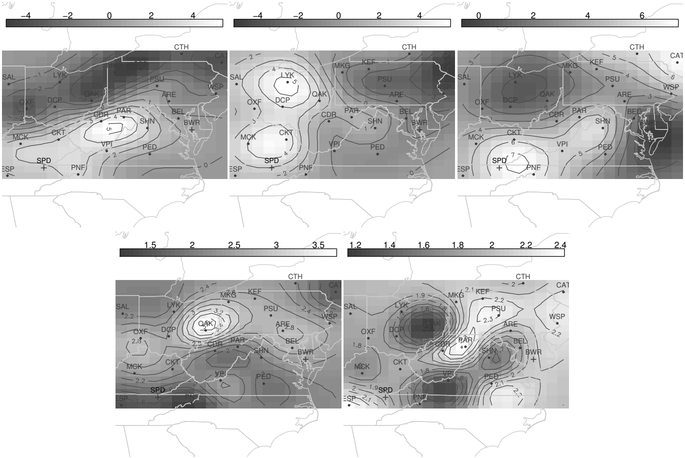
where $\beta = (\beta_1, \ldots, \beta_k)$ and R_{ϕ_j} spatial correlation matrix.

A RJMCMC is proposed to select k.

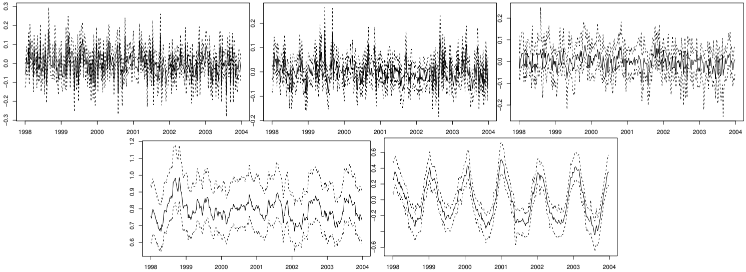
Example: SO_2 in Eastern US



Spatial loadings



Dynamic factors



Early days

Basic model

Literature

Classical literature

Bayes pre-MCMC

Bayes post-MCMC

Basic model (cont.)

More structure

Factor SV

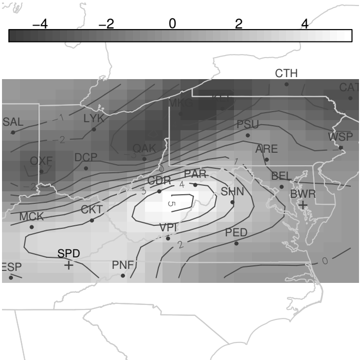
SDFM

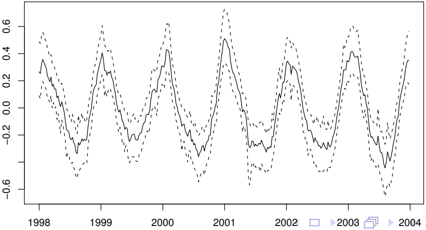
Sparse FA

SHFM

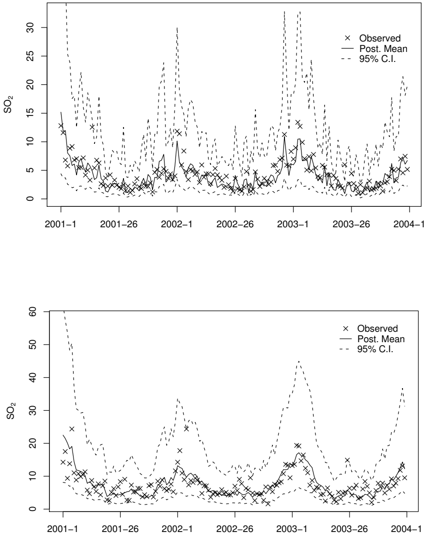
Final remarks

Seasonal factor



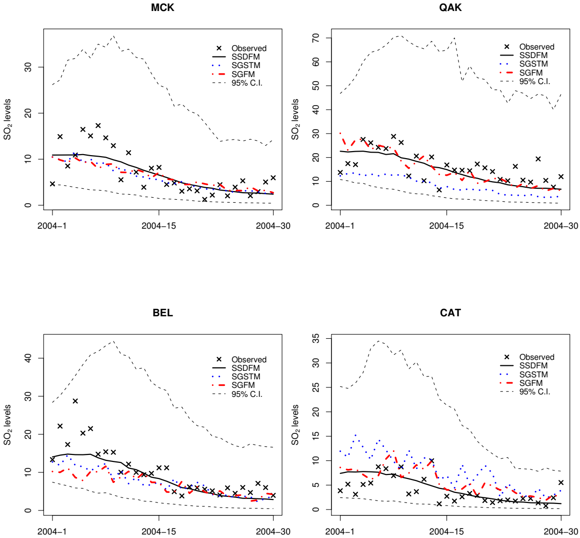


Spatial interpolation



Interpolated values at stations SPD and BWR.

Forecasting



Sparse FA[2]

Application to the 1970 British Cohort Study to analyze

the effect of child cognition, mental/physical health

on

educational choices and adult economic and health outcomes.

[2] Conti, Heckman, Lopes and Piatek (2011).

The British Cohort Study

A survey of all babies born (alive or dead) after the 24th week of gestation from 00.01 hours on Sunday, 5th April to 24.00 hours on Saturday, 11 April, 1970 in England, Scotland, Wales and Northern Ireland.

Follow-ups (so far): 1975, 1980, 1986, 1996, 2000, 2004, 2008.

Background characteristics:

- Cognitive, mental, physical health measurements (age 10)
- Education and adult outcomes (age 30)

Sample size: 5,105 women and 5,420 men.

Outcomes

Schooling outcomes (D)

- O-level
- A-level
- Higher Education

Post-schooling outcomes (Y)

- Health outcomes
  - poor health
  - obesity
  - daily smoking
- Labor market outcome
  - log hourly wage

Measurement system (M)

The measurement system includes more than one hundred and thirty indicators of child

- cognitive traits,
- mental health traits,
- physical health traits

all collected at age ten.

Cognition

- Picture Language Comprehension Test (PLCT):
  vocabulary, sequence, sentence comprehension.

- Friendly Math Test (FMT):
  arithmetic, fractions, algebra, geometry, statistics.

- Shortened Edinburgh Reading Test (SERT):
  vocabulary, syntax, sequencing, comprehension, retention.

- British Ability Scales (BAS):
  similar to IQ: two verbal and two non-verbal scales.

Mental health

- Rutter Parental ‘A’ Scale of Behavioral Disorder (19 items)
  Administered to the mother.

- The Conners Hyperactivity Scale (19 items)
  Also administered to the mother.

- The Child Developmental Scale (53 items)
  Answered by a teacher with knowledge of the child.

- The Locus of Control Scale (16 items)
  Measures the child’s perceived achievement control.
  Administered by the teacher and completed by the child.

- The Self-Esteem Scale (12 items)
  Measure the child’s self-esteem with reference to teachers, peers and parents. It was administered by the teacher and completed by the child.

Physical health

- height
- head circumference
- weight
- diastolic blood pressure
- systolic blood pressure

Control variables (X)

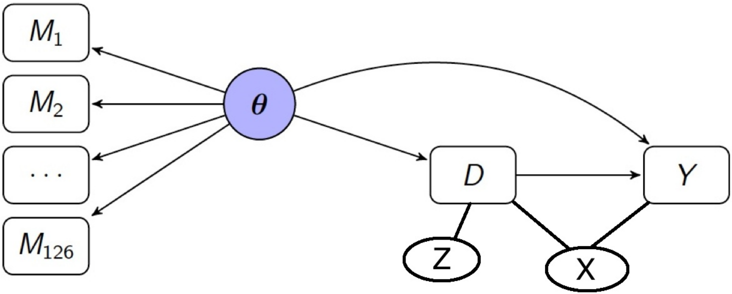
- mother's age at birth
- mother's education at birth
- father's high social class at birth
- total gross family income at age 10
- an indicator for broken family
- the number of previous livebirths
- the number of children in the family at age 10

Exclusion variables (Z)

Gender-specific, county-level deviation from long-run average.

- unemployment rate
- gross weekly wage

British study



Variables		Definitions
Education	D	Observed: achieving A-level or higher
Outcomes	Y	Observed in one state only! Poor health, Obesity, Smoking, Wage
Measurements:	M_j	Observed: 126 items (binary and cont.)
Cognitive skills Personality	θ	**Unobserved dimensions**

Education choice

Education outcome D is related to

- Latent factors θ (via measurements M)
- Observed characteristics X
- Exclusion restrictions Z

via the continuous latent utility $D^\star$:

$$D^\star = \alpha'_D X + \alpha'_Z Z + \beta'_D \theta + \varepsilon_D$$

where $D = 1$ if $D^\star > 0$, and zero otherwise.

Potential outcome

Let $(Y_1, Y_2, \ldots, Y_S)$ be health and labor market outcomes

The measured outcome Y_s can thus be expressed as:

$$Y_s = D\, Y_{1s} + (1 - D)\, Y_{0s}.$$

We assume that each potential outcome Y_{ds} is generated by a latent outcome $Y^{\star}_{ds}$, for $d = 0, 1$, through the following linear-in-parameter model:

$$Y^{\star}_{ds} = \alpha'_{ds} X + \beta'_{ds} \theta + \varepsilon_{ds}$$

Latent traits

We assume that each observed measurement is determined by an underlying latent variable $M_q^\star$ that linearly depends on the observed characteristics X and on the latent factors θ:

$$M_q^\star = \alpha_q' X + \beta_{M_q}' \theta + \varepsilon_{M_q}$$

Overall model

$$\begin{pmatrix} M_1^\star \\ \vdots \\ M_Q^\star \\ D^\star \\ Y_{01}^\star \\ Y_{11}^\star \\ \vdots \\ Y_{0S}^\star \\ Y_{1S}^\star \end{pmatrix} = \begin{pmatrix} \alpha_1' & 0 \\ \vdots & \vdots \\ \alpha_Q' & 0 \\ \alpha_D' & \alpha_Z' \\ \alpha_{01}' & 0 \\ \alpha_{11}' & 0 \\ \vdots & \vdots \\ \alpha_{0S}' & 0 \\ \alpha_{1S}' & 0 \end{pmatrix} \begin{pmatrix} X \\ Z \end{pmatrix} + \begin{pmatrix} \beta_{M_1}' \\ \vdots \\ \beta_{M_Q}' \\ \beta_D' \\ \beta_{01}' \\ \beta_{11}' \\ \vdots \\ \beta_{0S}' \\ \beta_{1S}' \end{pmatrix} \theta + \begin{pmatrix} \varepsilon_{M_1} \\ \vdots \\ \varepsilon_{M_Q} \\ \varepsilon_D \\ \varepsilon_{01} \\ \varepsilon_{11} \\ \vdots \\ \varepsilon_{0S} \\ \varepsilon_{1S} \end{pmatrix},$$

Or, more compactly,

$$y = \alpha W + \beta\theta + \varepsilon$$

Parsimonious BFA

Frühwirth-Schnatter and Lopes (2009)

- Lay down a new and general set of identifiability conditions that handles the ordering problem present in most of the current literature,
- Introduce a new strategy for searching the space of parsimonious/sparse factor loading matrices,
- Designed a highly computationally efficient MCMC scheme for posterior inference which makes several improvements over the existing alternatives,

for the important class of Gaussian factor models:

$$y = \beta\theta + \varepsilon$$

where $\varepsilon \sim N(0, \Sigma)$.

Identification issues

- Block lower triangular
  Generalized lower triangular alternative

- Rank deficiency
  If β is rank-deficient, then $\exists Q$ such that

$$\beta\beta' = (\beta + MQ')(\beta + MQ')' + (\Sigma - MM').$$

for some orthogonal M with $\beta Q = 0$ and $Q'Q = I$.
We use this "deficiency" in our model search strategy.

Generalized lower triangular

$$\begin{pmatrix} \beta_{11} & 0 & 0 & 0 \\ \beta_{21} & \beta_{22} & 0 & 0 \\ \beta_{31} & \beta_{32} & \beta_{33} & 0 \\ \beta_{41} & \beta_{42} & \beta_{43} & \beta_{44} \\ \beta_{51} & \beta_{52} & \beta_{53} & \beta_{54} \\ \beta_{61} & \beta_{62} & \beta_{63} & \beta_{64} \\ \beta_{71} & \beta_{72} & \beta_{73} & \beta_{74} \end{pmatrix} \Longrightarrow \begin{pmatrix} \beta_{11} & 0 & 0 & 0 \\ \beta_{21} & 0 & 0 & 0 \\ \beta_{31} & \beta_{32} & 0 & 0 \\ \beta_{41} & \beta_{42} & 0 & 0 \\ \beta_{51} & \beta_{52} & 0 & 0 \\ \beta_{61} & \beta_{62} & 0 & \beta_{64} \\ \beta_{71} & \beta_{72} & 0 & \beta_{74} \end{pmatrix}$$

Birth/death of loadings
Birht/death of columns.

Other contributions

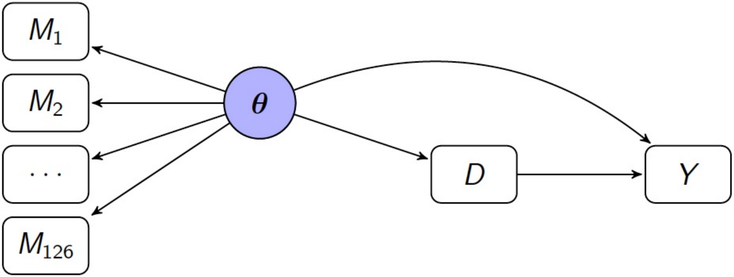
- Our approach provides a principled way for inference on the number of factors, as opposed to previous work (Carvalho et al., 2008; Bhattacharya and Dunson, 2009).

- Our prior specification on Σ properly addresses Heywood problems

- Our fractional-like prior on β is more robust than the existing ones (Lopes and West, 2004, Ghosh and Dunson, 2009)

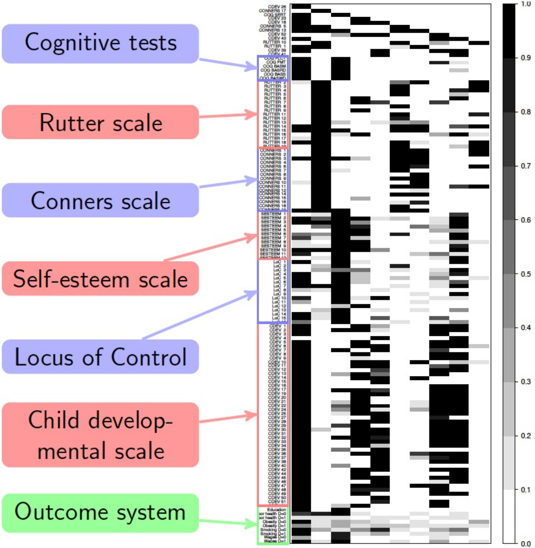
- Efficient (and correct) parameter expansion where the prior is unchanged (as opposed to GD2009).

British study



Variables		Definitions
Education	D	Observed: achieving A-level or higher
Outcomes	Y	Observed in one state only!
		Poor health, Obesity, Smoking, Wage
Measurements:	M_j	Observed: 126 items (binary and cont.)
Cognitive skills	θ	**Unobserved dimensions**
Personality		

Females — Factor loadings posterior probabilities



Cognitive tests

Rutter scale

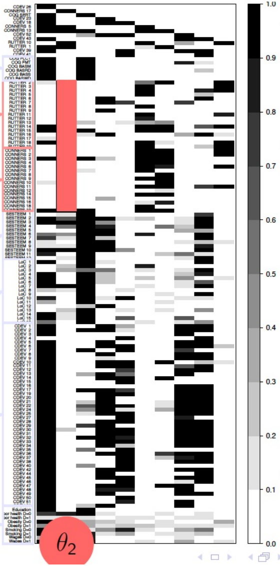
Conners scale

Self-esteem scale

Locus of Control

Child develop-mental scale

Outcome system



Females — Factor loadings posterior probabilities

Factor 2
Significantly loaded by items from the Rutter and and the Conners scales associated with:
'Anxiety Disorders'

Cognitive tests

Rutter scale

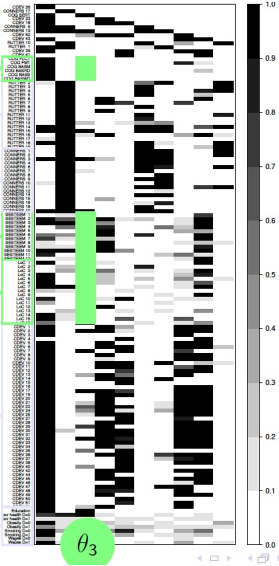
Conners scale

Self-esteem scale

Locus of Control

Child developmental scale

Outcome system



Females — Factor loadings posterior probabilities

Factor 3

Significantly loaded by items from the cognitive tests and locus of control items: **Cognitive factor**

Vulnerability index for Uruguay

Uruguay has an area of 176,215 km^2 and roughly 3.3 million inhabitants, half of which live in the capital, Montevideo. Around 93% of the population lives in urban areas.

Census tracts per capital

Capital	Census tracts	Capital	Census tracts
Bella Unión	11	Durazno	35
Canelones	20	Maldonado	36
Colonia	21	Tacuarembó	38
Fray Bentos	22	Mercedes	39
Trinidad	27	Melo	43
Rocha	28	Rivera	45
Treinta y Tres	29	Paysandú	72
Florida	31	Salto	84
Minas	33	Montevideo	1031
San José	34		

Main goals

To characterize the vulnerability of the population of Uruguay to diseases transmitted through vectors (e.g. Dengue Fever, Malaria, etc.);

To help prioritizing the allocation of fundings;

We have information on $p = 11$ variables per census tracts of the $I = 19$ Departamental Capitals of the country.

Source: Census 1996 (latest Census in Uruguay)

Table : Description of the $p = 11$ variables, observed in the census tract level of the departmental capitals, to build the vulnerability index of the population of Uruguay to vector-borne diseases.

Levels	Variables
Personal characteristic	Illiteracy rate (ILL)
	Population with access to public health care (PHC)
	Male without formal jobs (UQW)
Household characteristic	Owed houses (OWH)
	Households headed by a woman (WHF)
	Households without sewage system (AHS)
	Average number of persons per household (APH)
	Households with more than two persons per room (OVC)
	Households without access to drinkable water (ADW)
	Households with air conditioner (ACO)
	Households poorly built (HOQ)

Sample correlations

	ILL	PHC	OVC	UQW	AHS	ADW	APH
PHC	0.85						
OVC	0.78	0.79					
UQW	0.67	0.65	0.68				
AHS	0.64	0.59	0.67	0.60			
ADW	0.60	0.47	0.49	0.51	0.62		
APH	0.53	0.52	0.54	0.38	0.32	0.26	
HOQ	0.45	0.36	0.43	0.40	0.63	0.57	0.23

The sample correlations between OWH or WHF or ACO and any one of the attributes are below 18% (in absolute value).

Model structure

Observational Level:

$$y_{ijk} = \mu_k + \beta_k f_{ij} + \sigma_k \varepsilon_{ijk} \quad k = 1, \cdots, p,$$

where μ_k represents the overall grand mean.

Modeling f_{ij}:

$$f_{ij} = \theta_i + \tilde{f}_{ij} + \sqrt{\omega_i} u_{ij}$$

where θ_i is the common factor for capital i.

Spatial variation within capitals:

$$\tilde{f}_i \sim N(0, \tau_i^2 P_i)$$

where $P_i = (I_{n_i} + \phi M_i)^{-1}$, $M_i = D_i - W_i$, with w_{ijl}, the (j, l) component of W_i, given by $w_{ijl} = 1/d_{jl}$ if j and l are neighbors (denoted here by $j \sim l$) and zero otherwise, $d_{jl} = ||s_j - s_l||$ is the Euclidean distance between centroids of capitals j and l, $D_i = \text{diag}(w_{i1+}, \ldots, w_{in_i+})$ and $w_{ij+} = \sum_{l \sim j} w_{ijl}$.

Model structure (cont.)

Spatial variation between capitals:

$$\theta \sim N\left(1_I\theta_0, \delta^2 H(\lambda)\right),$$

where $\theta = (\theta_1, \cdots, \theta_I)$.

Although each capital i has its own vulnerability factor, the above model allows borrowing-strength across neighboring regions.

Table : *Comparing SHFM and UHFM:* Comparing the unstructured hierarchical factor (UHFM) and spatial hierarchical factor models (SHFM) for different values of ϕ. Best models appear in italic. DIC: deviance information criterion, EPD: expected posterior deviation, CRPS: continuous ranked probability score, MSE: mean square error and MAE: mean absolute error. CRPS are in tens of thousands.

	UHFM		SHFM		
Criterion	$\theta = 0$	unknown θ	$\phi = 1$	$\phi = 5$	$\phi = 7$
DIC	-21445.4	-21493.3	-21785.8	*-21827.4*	-21827.0
EPD	2557.4	2510.9	2453.1	2433.6	*2432.6*
CRPS	1030.7	1024.2	1014.2	*1010.3*	1010.3
MAE	2397.0	2381.8	2374.5	*2367.9*	2369.1
MSE	1222.3	1200.1	1177.2	1169.2	*1168.9*

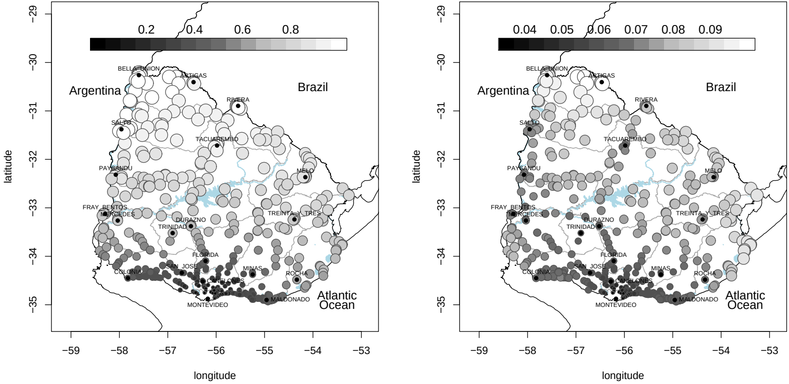


Figure : Posterior mean of θ_i and standard deviations (second column) for observed and unobserved cities under the SHFM when $\phi = 5$.

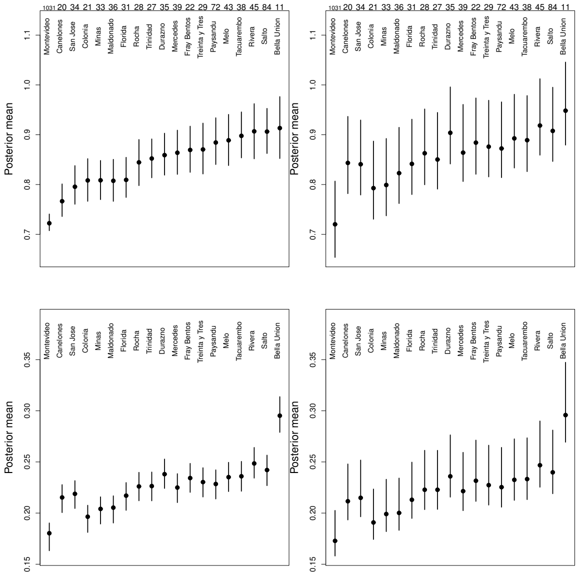


Figure : Posterior means of the θ_i and 95% CI. *Top row:* SHFM with $\phi = 5$ (left) and UHFM (right). *Bottom row:* ASFM (left) and AFM (right).

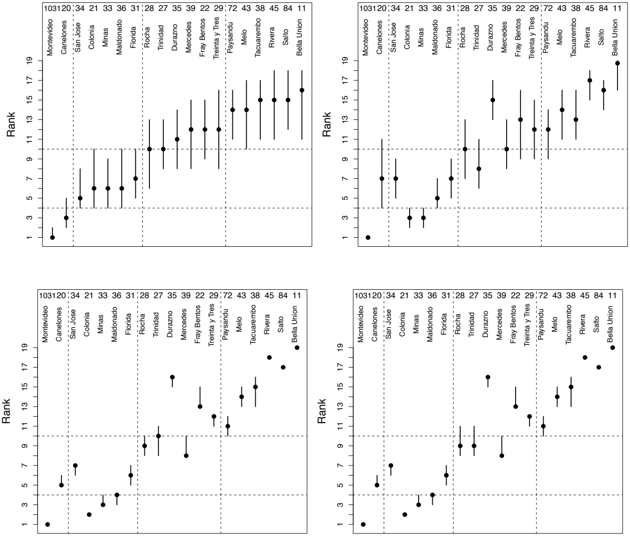
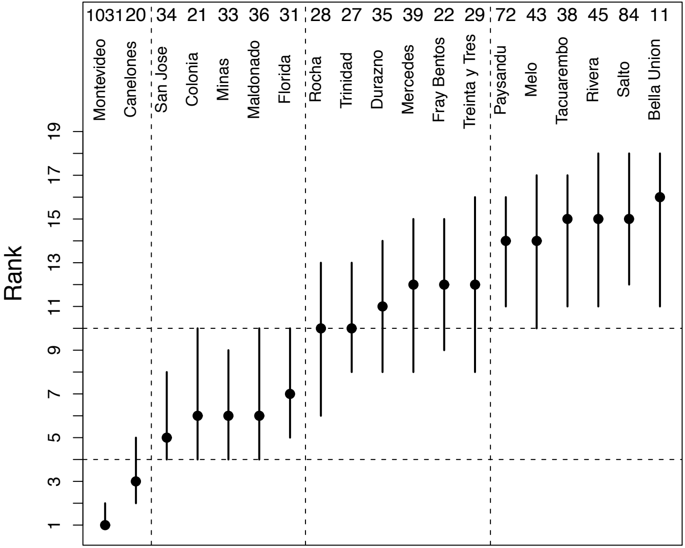


Figure : Posterior rankings of the capitals. *Top row:* SHFM with $\phi = 5$ (left) and UHFM (right). *Bottom row:* ASFM (left) and AFM (right).



Rank
103
120
34
21
33
36
31
28
27
35
39
22
29
72
43
38
45
84
11
Montevideo
Canelones
San Jose
Colonia
Minas
Maldonado
Florida
Rocha
Trinidad
Durazno
Mercedes
Fray Bentos
Treinta y Tres
Paysandu
Melo
Tacuarembo
Rivera
Salto
Bella Union
1
3
5
7
9
11
13
15
17
19

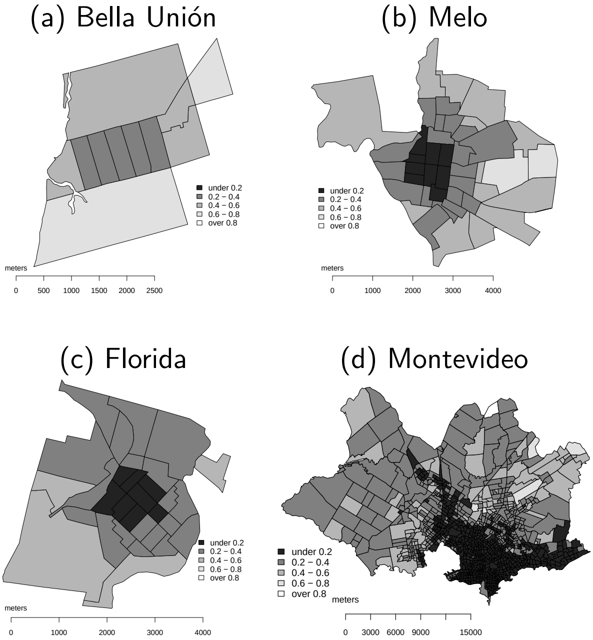
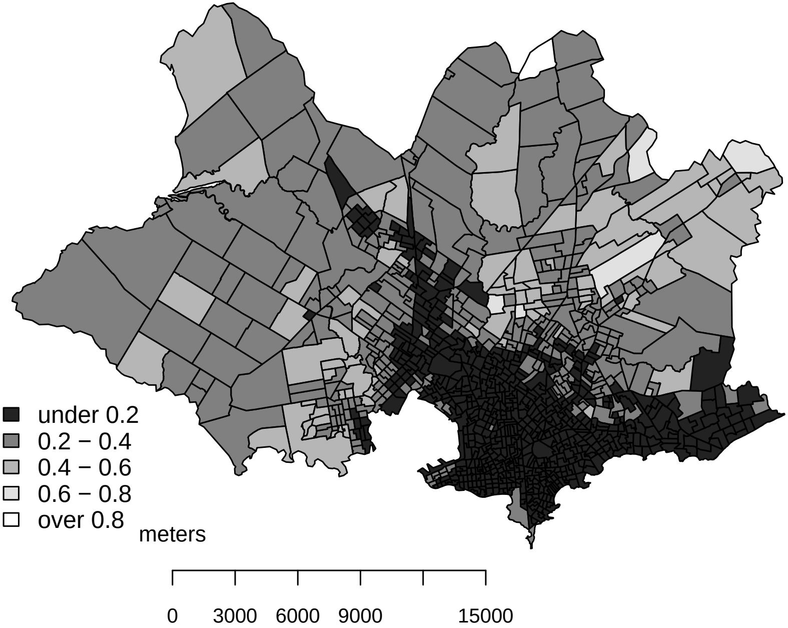


Figure : Within-city posterior vulnerability index per census tract.


under 0.2
0.2 − 0.4
0.4 − 0.6
0.6 − 0.8
over 0.8
meters
0
3000
6000
9000
15000

Final remarks

Massive datasets
GWAS, high-frequency econometrics, climatology

Factor-augmented VAR
(Ahmadi and Uhlig, 2009)

Many weak instruments
(Hahn and Lopes, 2012)

Sparse loadings via regularization
(Polson and Scott, 2011,2012)

Text document modeling via independent factor topic models
Latent Dirichlet allocation and correlated topic model
Putthividhya, Attias and Nagarajan (2012)