A multilevel factor approach for the analysis of CDS commonality and risk contribution

Carlos Vladimir Rodríguez Caballero[1]
Massimiliano Caporin [2]

[1] Instituto Tecnológico Autónomo de México & CREATES.
[2] University of Padova.

Workshop on Econometrics and Data Science

Motivation

Relevance of Factor Models in the recent literature:

- Common factors capture the most relevant time variation that it is spread across variables in a panel.
- Factor models have been extensively used in finance and macroeconomics due to the relevance and practicality when treating the problem of dimensionality in large data sets.
- The use of factor models for forecasting purposes has become an increasingly widespread tool to forecast macroeconomic variables
- Static and dynamic setups (Bai & Ng(2008) for an extensive review).

Motivation

Relevance of the use of Factor Models in the recent literature:

- Some comments:
  1. Standard factor models assume the common factors to affect all variables of the system (in principle).
  2. As a standard practice, most empirical studies rely on the existence of pervasive factors.
     - ★ A large N not necessarily helps the estimation of common factors, due to the existence of strong cross-sectional correlations. (Boivin& Ng (2006))
     - ★ Particularly, efficiency of the PC estimator may be deteriorated substantially if groups of variables do not provide any information about the factors (loadings = 0) (Boivin& Ng (2006))

Motivation

Multi-level Factor models

The cornerstone of these models is to decompose the common factor structure into different levels, with factors associated to the full cross-section, i.e. pervasive, and factors that impact on and explain only a specific sub-group of variables, the non-pervasive factors.

Caveat:

A separate analysis will mix up non-pervasive and pervasive factors which hampers identification and provoke a severe loss of efficiency.

Motivation

Multi-level Factor models

The cornerstone of these models is to decompose the common factor structure into different levels, with factors associated to the full cross-section, i.e. pervasive, and factors that impact on and explain only a specific sub-group of variables, the non-pervasive factors.

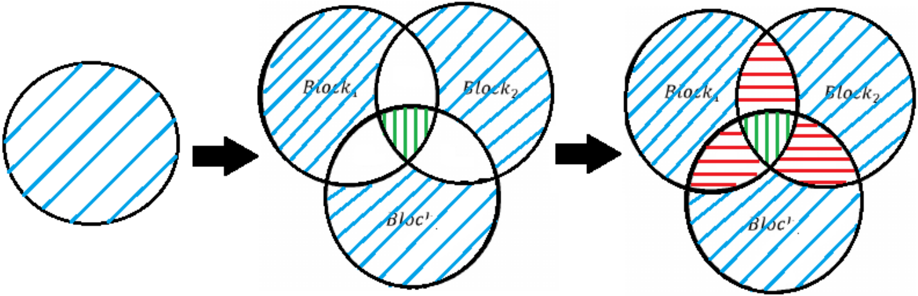
Possible methodologies:

1. Considering zero restrictions in the associated loadings matrix, as discussed in Wang (2010), Choi et al. (2014), Breitung and Eickmeier (2016), and Rodríguez Caballero and Ergemen (2017).
2. Considering a hierarchical approach as discussed Kose, Otrok, and Whiteman (2003), and Moench, Ng, and Potter (2013).

Motivation

In this paper:

- We propose a model that is able to characterize both within-group and between-group variations.



Motivation

In this paper:

- We propose a model that is able to characterize both within-group and between-group variations.
- We accompany the methodological advancement with an empirical analysis on the sovereign Credit Default Swap (CDS) market.
  - The purpose is to identify the sources of commonality and how these impact on the risk of a CDS portfolio.
  - Our empirical study is related to the work of Longstaff et al. (2011) that also discuss the presence and sources of commonality across sovereign CDS spreads, but limiting the analysis to the use of standard factor model.
  - We show that a multilevel model based on country groupings build upon the Debt/GDP ratio or the sovereign rating provide views that differ from those associated with principal components, and that are more easily economically interpreted.

Model setup

The model:

1. We consider a block-structured factor model in that the unobservable common factor structure may be classified in many different levels according to the number of blocks formed by data.
2. These blocks can be formed naturally with some economic sectors or countries as in many macroeconomic studies, or by implementing multivariate statistical procedures as clustering or recursive partitioning.

Model setup

For clarity of exposition consider a panel data formed by three different blocks of data, B_1, B_2, and B_3. In this light, the model would include

- top-level pervasive factors that affects all blocks of data,
- pairwise sub-level pervasive factors, and
- block-specific non-pervasive factors.

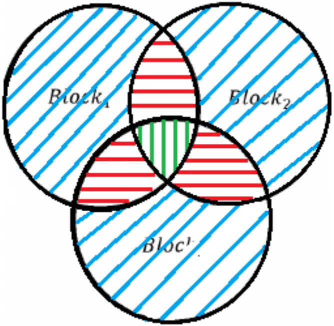


Figure: Factor structure formed by three different blocks of data.

Model setup

For clarity of exposition consider a panel data formed by three different blocks of data, B_1, B_2, and B_3. In this light, the model would include

- top-level pervasive factors that affects all blocks of data,
- pairwise sub-level pervasive factors, and
- block-specific non-pervasive factors.

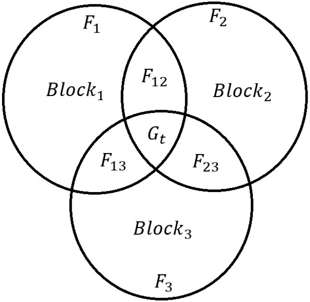


Figure: Factor structure formed by three different blocks of data.

Model setup

Then, the three-block factor model is written as

$$y_{k,it} = \gamma'_{k,i} G_t + \kappa'_{kj,i} F_{kj,t} + \lambda'_{k,i} F_{k,t} + u_{k,it},$$

where

- $k = 1, 2, 3$ indicates the block, (i, t) denotes the usual panel index, and kj means interaction between blocks k and j for $k \neq j$.

The unobservable factor structure is defined by:

- $\mathbf{r}_G \times 1$ vector G_t containing the unobservable global factor,
- the $\mathbf{r}_{F_{kj}} \times 1$ vector $F_{kj,t}$ containing the pairwise block factors that interact only between blocks k and j with $k \neq j$, and
- the $\mathbf{r}_{F_k} \times 1$ vector, $F_{k,t}$, that consists of the $\mathbf{r}_{F_k}$ unobservable block-specific factor of block k.
- Naturally, $\gamma_{k,i}, \kappa_{kj,i}$, and $\lambda_{k,i}$ are the $\mathbf{r}_G, \mathbf{r}_{F_{kj}}$, and $\mathbf{r}_{F_k}$- dimensional factor loadings.

On assumptions

$$y_{k,it} = \gamma'_{k,i} G_t + \kappa'_{kj,i} F_{kj,t} + \lambda'_{k,i} F_{k,t} + u_{k,it},$$

Some comments:

- The idiosyncratic term, $u_{k,it}$, as well as the factor structure satisfy standard assumptions provided by Bai(2003).
- The model may also allows for long-range dependence processes from which the common component and the idiosyncratic components should follow the assumptions provided by Rodríguez Caballero and Ergemen (2017).

Alternative representation

The model implies blocks of zero restrictions on the associated matrix of factor loadings,

$$\begin{pmatrix} y_{1,\cdot t} \\ y_{2,\cdot t} \\ y_{3,\cdot t} \end{pmatrix} = \begin{pmatrix} \gamma_1 & \boldsymbol{\kappa}_{12_1} & \boldsymbol{\kappa}_{13_1} & 0 & \boldsymbol{\lambda}_1 & 0 & 0 \\ \gamma_2 & \boldsymbol{\kappa}_{12_2} & 0 & \boldsymbol{\kappa}_{23_2} & 0 & \boldsymbol{\lambda}_2 & 0 \\ \gamma_3 & 0 & \boldsymbol{\kappa}_{13_3} & \boldsymbol{\kappa}_{23_3} & 0 & 0 & \boldsymbol{\lambda}_3 \end{pmatrix} \begin{pmatrix} G_t \\ F_{12,t} \\ F_{13,t} \\ F_{23,t} \\ F_{1,t} \\ F_{2,t} \\ F_{3,t} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,\cdot t} \\ \epsilon_{2,\cdot t} \\ \epsilon_{3,\cdot t} \end{pmatrix},$$

$$y_t = \Lambda^* F_t^* + \epsilon_t. \tag{1}$$

Estimation

The estimation procedure is based on the sequential approach.

The goal is to minimize the residual sums of square (RSS) function

$$S\,(F_t^*, \Lambda^*) = \sum_{t=1}^{T} \left(y_t - \Lambda^* F_t^*(\hat{\theta})\right)' (y_t - \Lambda^* F_t^*)$$

by a sequence of two least-squares regressions until RSS achieves a minimum.

Estimation

The algorithm can be executed for the general case of k blocks as follow:

1. Initial step:
   - **a)** Canonical correlation analysis (CCA) on $y_{ki,t}$ to obtain $\hat{G}^{(0)}$.
   - **b)** Regress $y_{k,it}$ on $\hat{G}^{(0)}$ in each block k to filter out the global component. Then we get the residuals, $y^{*(0)}_{ki,t}$.
   - **c)** CCA on $y^{*(0)}_{k,it}$ to obtain the next lower level block factors.
   - **d)** Regress $y^{*(0)}_{k,it}$ on the respective block factors involved and get the residuals.
   - **e)** Steps c) and d) are sequentially executed until getting the initial estimates of the pairwise block factors. $y^{**(0)}_{k,it}$ are the residuals after filtering the pairwise factors on each block k.
   - **e)** Then, run PCA on $y^{**(0)}_{k,it}$ to get $\left(F^{(0)}_{1,it}, \ldots, F^{(0)}_{k,it}\right)$.
   - **f)** The loading factors $\hat{\Lambda}^{*(0)}$ are estimated from time-series regression of $y_{ki,t}$ on the factors involved in each specific block k.

Estimation

The algorithm can be executed for the general case of k blocks as follow:

1. Initial step.
2. The updated estimator for the unobservable common factors are obtained by a sequential procedure as follow:
   a) Run least-square $y_{ki,t}$ on $\hat{\Lambda}^{*(0)}$ to get $G_t^{*(1)}$.
   b) Regress $y_{k,it}$ on $\hat{G}^{(1)}$ in each block k to filter out the global component. Get $y_{k,it}^{*(1)}$.
   c) Run least-square $y_{k,it}^{*(1)}$ on the next lower level block factors.
   d) Repeat the same procedure as before until getting block-specific factors, $\left(F_{1,it}^{(1)}, \ldots, F_{k,it}^{(1)}\right)$.
   e) Next, the updated (and normalized) factors $F_t^{*(1)}$ are used to get the associated updated factor loading matrix $\hat{\Lambda}^{*(1)}$.
3. Step 2 is repeated until RSS converges to a minimum from which F_t^* and $\hat{\Lambda}^*$ are collected.
4. The last step consists on orthogonalize each level of factors estimates.

Finite sample properties

Main findings:

Remarks from MC simulations:

- The methodology proposed in this paper perform well in relatively small samples ($Nk = 20$, $T = 150$).
- Perform very well when sample sizes increase independently of size distortions between N_k and T.
- A low signal-to-noise ratio makes the factors independently of the level a bit less precise although such loss of precision is not dramatic.
- Changes in variances in the factors do not have a considerable impact in the estimation of the factors.
- In cases when a factor structure consists also on some pairwise factors, neglect the existence of such factors, estimation of global and regional factors will be substantially biased even when both dimensions increase.

Economic data and country grouping

- The data set used in this paper collects the 5-years Credit Default Swaps (CDS) premia on government bonds data.
- Our dataset is a balanced panel consisting of 53 countries for each day for the period 1 January 2009 to 11 December 2017, yielding a total of 2,333 daily observations for each country.
- We work with the standardized log-changes of the CDS premia.

We cluster the 53 countries in two different ways;

i) using the median of the debt/gdp ratio of at most 8 years (2009-2017),

ii) by the last credit rating assigned by Standard & Poor's.

Goal:

- We will evaluate the role of both the credit rating and the Debt/GDP ratios in driving the commonality across the countries in our panel.

Economic data and country grouping

Debt to GDP Ratio (Median -2009-2017)					
High		Medium		Low	
Country	Ratio	Country	Ratio	Country	Ratio
Japan	186.27	Germany	68.30	Sweden	41.60
Italy	132.60	Colombia	67.28	Latvia	40.10
Portugal	130.40	Netherland	65.62	Panama	39.20
Slovak	112.00	Abu Dhabi	65.52	Korea	38.39
Cyprus	107.80	Dominican	64.87	Denmark	37.80
Belgium	105.90	Finland	63.60	Romania	37.60
Iceland	105.73	Vietnam	62.40	Czech	37.31
Ireland	100.43	Costa rica	62.00	Turkey	36.10
UK	96.78	Israel	61.90	Bahrain	35.94
France	96.00	Brazil	60.58	Norway	35.60
USA	94.41	Poland	54.10	Thailand	31.61
Egypt	92.30	Malaysia	52.16	Venezuela	28.20
Hungary	92.17	South africa	51.60	Indonesia	27.78
Spain	91.54	Philipines	51.00	Guatemala	24.27
Austria	84.60	El salvador	49.52	Peru	21.68
Slovenia	78.50	Mexico	47.90	Chile	21.30
		Dubai	47.60	Kazakhstan	11.74
		Qatar	47.60	Russia	9.10
		China	46.20		

Table: Countries clustered with respect to the median of Debt/GDP ratio

Economic data and country grouping

COUNTRY RATING S&P 2017							
Highest		High-Medium		Medium-Low		Lowest	
Country	Rating	Country	Rating	Country	Rating	Country	Rating
Austria	AA+	Chile	A+	Colombia	BBB-	Brazil	BB-
Belgium	AA	China	A+	Hungary	BBB-	Costa rica	BB-
Czech	AA-	Iceland	A	Indonesia	BBB-	Cyprus	BB+
Denmark	AAA	Ireland	A+	Italy	BBB	Dominica	BB-
Dubai	AA	Israel	A+	Kazakhstan	BBB-	Guatemala	BB-
Abu Dhabi	AA	Japan	A+	Mexico	BBB+	Russia	BB+
Finland	AA+	Latvia	A-	Panama	BBB	South africa	BB
France	AA	Malaysia	A-	Peru	BBB+	Turkey	BB
Germany	AAA	Slovak	A+	Philipines	BBB	Vietnam	BB-
Korea	AA	Slovenia	A+	Poland	BBB+	Bahrain	B+
Netherland	AAA			Portugal	BBB-	Egypt	B-
Norway	AAA			Romania	BBB-	El Salvador	CCC+
Qatar	AA-			Spain	BBB+	Venezuela	SD
Sweden	AAA			Thailand	BBB+		
UK	AA						
USA	AA+						

Table: Countries clustered with respect to the last credit rating provided by S&P

Global and economic-driven factors in CDS

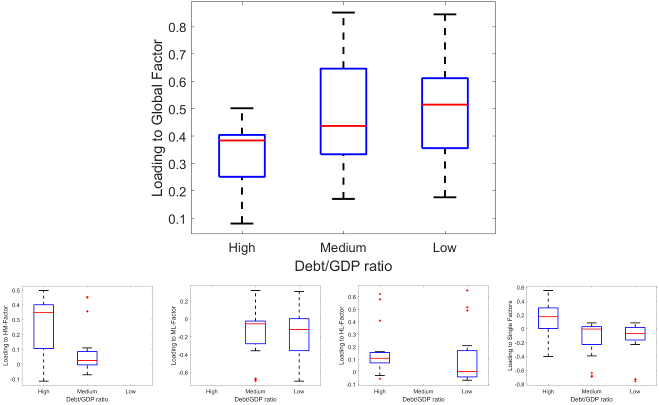


Figure: Box plots of estimated coefficients (loadings) for country groups based on the Debt/GDP ratio (High, Medium and Low levels). By row, from top to bottom, left to right: loadings to the global factor; loadings to the HM factor; loadings to the ML factor; loadings to the HL factor; loadings to the single group factor.

Global and economic-driven factors in CDS

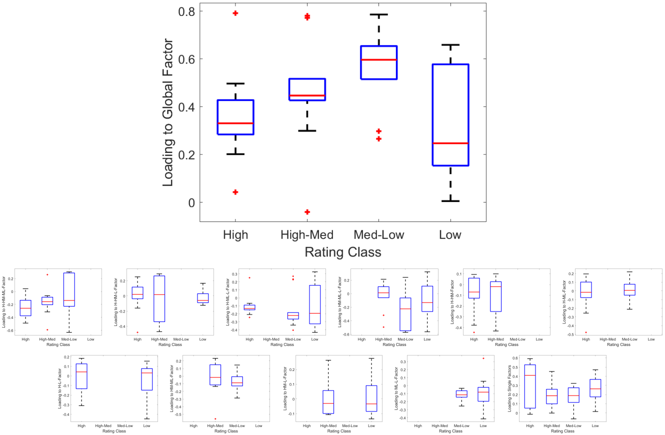


Figure: Box plots of estimated coefficients (loadings) for country groups based on the sovereing debt rating (High, High-Medium, Medium-Low, and Low levels). By row, from top to bottom, left to right loadings to the factors: global; H-HM-ML, H-ML-L, HM-ML-L; H-HM; H-ML; H-L; HM-ML; HM-L; ML-L; single.

Global and economic-driven factors in CDS

Main findings for the standard case (PCA).

- We observe that the first PC has the largest and more stable loadings across all the CDS, without remarkable differences across groups.
- Principal components have a less clear connection with the country groups based either on the Debt/GDP ratio or the sovereign rating.

Global and economic-driven factors in CDS

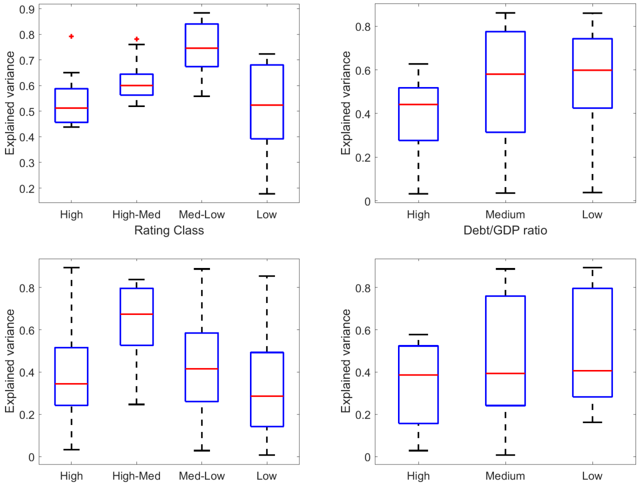


Figure: Box plots of the fraction of variance explained by the various methods. Top row, left plot, multilevel factor model with country rating groups, right plot, multilevel factor model with debt/GDP groups. Bottom row, PCA model evaluated on the same groups of the top plots.

Global and economic-driven factors in CDS

Figure: Global Factor estimated by a multilevel factor model from the Debt/GDP ratio group.

A risk contribution analysis

Main goal:

Evaluate the role of the latent factors in driving the risk of the CDS over country groups.

- We follow a recent technical financial literature, (Fabozzi et al., 2016 for extraction of volatility, a recent risk decomposition proposed by Roncalli and Weisang (2016), and the risk contribution decomposition technique proposed by Roncalli and Weisang (2016)).

A risk contribution analysis

Main Findings. Rating case:

- In the rating case, we note that the global factor is the most relevant risk contributor for all groups.
- Group-specific factors assume a relevant role, in particular for higher rating groups.
- Finally, the residual risk is very low, indicating that, from a risk perspective, the identified factors capture most of the risk in the country groups based on rating.

Main Findings. Debt/GDP case:

- We note that all factors have now some impact on the various country portfolios.
- The residual risk is higher than in the rating case, even if it remains at sensibly lower levels compared to the PCA risk contribution.

PCA:

- This is completely different in comparison with the fraction of risk not explained by the principal components.

Further analysis not presented here

1. Identification of the Global and economic-driven factors in CDS.
2. The relevance of multilevel factors during and after financial crises.

Concluding remarks

- We propose a model that allows for a rich interaction among different blocks of data.
- The risk contribution analysis highlights the role of latent factors for monitoring the risk of country groupings, in particular those based on ratings.
- The role of the global factor remains predominant but we also highlight a relevant contribution from semi-pervasive factors associated with specific sets of country groups.
- The adoption of economically-based country grouping criteria in combination with a multilevel model could provide relevant insights in the analysis of the risk drivers of CDS spreads.