Entropy and systemic risk measures

Monica Billio  Roberto Casarin
Michele Costola  Andrea Pasqualini

Ca’ Foscari University of Venice

Dealing with Complexity in Society 2015
Systemic Risk measurement

- Given the relevance and the impact of latest financial and sovereign crises, academics and regulators devoted several attention in modeling systemic events.
- Policy makers are looking for analytic tools to better identify, monitor and address risks in the system.
- Systemic risks involve the financial system, a complex and strongly interrelated system where the interconnectedness among financial institutions in period of financial distress may result in a rapid propagation of illiquidity, insolvency, and losses.
- We define systemic risk as “any set of circumstances that threatens the stability of or public confidence in the financial system” (Billio et al., 2012).
Aim of the paper

- We analyze the time evolution of systemic risk in Europe and construct an early warning indicator for banking crises based on different entropy measures (Shannon, Tsallis and Rényi).
- The analysis is based on the cross-sectional distribution of systemic risk measures by considering two classes of these measures:
  1) Tails of the financial returns that captures the co-dependence between financial institutions and the market (ΔCoVaR and MES).
     ⇒ Theoretical models showing that shocks to volatility or to tail risk provoke common fluctuations across firms (Acemoglu et al., 2011).
  2) Network linkages among financial institutions (IO network degree).
     ⇒ Skewness and fat tails suggest the presence of heterogeneity in the linkages among institutions: a large majority of financial institutions have low degree, but a small number (HUBS-SIFI) have a high number of linkages (Acemoglu, 2012).
The considered systemic risk measures are at micro-level for the financial industry in the Euro area (Cross-sectional distribution).

Structural changes in the proximity of a systemic event: the relevant or frail financial institutions are probably be the first to react and thus to provoke a structural change in the cross-sectional distribution.

We aim to exploit the ability of the entropy indicator to detect this heterogeneity and time variations in the financial system.

We do not impose any assumption on the cross-sectional distribution of the risk measures (non parametric approach).
Entropy in finance

- Jiang et al. (2014) propose an entropy measure for asymmetrical dependency in asset returns.
- Chabi-Yo and Colacito (2013) propose an entropy-based correlation measure to assess the performance of international asset pricing models.
- Bera and Park (2008) use cross-entropy measure as shrinkage rule to overcome extreme portfolio weights in the Mean-Variance estimation framework.
- Gao and Hu (2013) study the income structures of different sectors of an economy and provide an early warning indicator measuring losses in term of quarterly negative income where exposure networks are modelled by the Omori-law-like distribution.
- Albares-Ramirez et al. (2012) apply approximated entropy measures at univariate level to study the dynamic of the market efficiency from an informational perspective.
Our Research approach

- We sequentially apply over time entropy to systemic risk measures.
- We test the predictive ability of entropy indicators for banking crisis (Alessi and Detken, 2011).
- These entropy indicators are used as covariates in a logit model to build an early warning system (Davis and Karim, 2008).
- We show that entropy can be used to build a quasi-real time early warning indicator (nowcasting)
  - entropy of the network characterizes the complexity of the system.
  - entropy of loss distributions is useful to describe the systemic risk level of the system.
- The empirical analysis involves the Euro area: A Recent sovereign debt crisis mainly due to a frail financial and banking system (Lane, 2012).
Entropy Measures

Entropy is used in a variety of fields to characterize the complexity of a system and to summarize the information content of a distribution.

- Let $\pi_t = (\pi_{1t}, \ldots, \pi_{mt}), t = 1, \ldots, T$, with $\pi_{jt} \geq 0$, $\sum_j \pi_{jt} = 1$, be a sequence of probability vectors.

- In our case, these vectors represent a sequence of cross-sectional distributions of a given systemic risk measure for a set of financial assets available at time $t$ in the market.

- We apply to $\pi_t$ three definitions of entropy: Shannon, Tsallis and Rényi.
Entropy indicators

- **Shannon entropy** (Shannon, 1948),
  \[
  H_S(\pi_t) = -\sum_{j=1}^{m} \pi_{jt} \log \pi_{jt} \text{ where } m < \infty, \tag{1}
  \]

- **Tsallis’ entropy** (Tsallis, 1988),
  \[
  H_T(\pi) = \frac{1}{\alpha - 1} \left(1 - \sum_{i=1}^{m} \pi_{it}^{\alpha} \right), \tag{2}
  \]

- **Rényi’s entropy** (Rényi, 1960),
  \[
  H_R(\pi) = \frac{1}{1 - \alpha} \log \left(\sum_{i=1}^{m} \pi_{it}^{\alpha} \right). \tag{3}
  \]
The extensive index $\alpha$

Shannon and Tsallis entropy allow for power tail behaviour in function of $\alpha$.

- $\alpha$ allows the researcher to identify the relevance of the tails in the crises prediction...

- ...by assigning more or less weight to the tails of the distribution.
  - For the Renéy’s entropy, the higher the parameter $\alpha$, the less the entropy for distributions is far from the uniform (the tails of the distribution are penalized).
  - For the Tsallis’ entropy, the higher the parameter $\alpha$, the less the entropy is sensitive to changes in the probabilities associated to common events.

- It is useful in our application since the weight of the tails in the cross-sectional distribution of the systemic risk measures may change dramatically during periods of financial distress.
Marginal Expected Shortfall (MES)

\( MES_{it} \) is defined as the expected value of \( r_{it} \) when the market \( r_{mt} \) is below a given quantile \( q_k \) (Acharya et al., 2010).

\[
MES_{it} = \mathbb{E} (r_{it} | r_{mt} < q_{5\%}) ,
\]

where \( r_{it}, t = 1, \ldots, T \) denotes the series of asset returns for the asset \( i \).
$\Delta \text{CoVaR}$

$\Delta \text{CoVaR}$ (Adrian and Brunnermeier, 2011) is defined as the difference between the CoVaR conditional on an institution being under distress and the CoVaR in the median of the institution, that is

$$\Delta \text{CoVaR}_{mit,q} = \text{CoVaR}_{mit,q} - \text{CoVaR}_{mit,0.5}, \quad (5)$$

where $r_{it}$ is the asset return value of the institution $i$ and $r_{mt}$ represents the system. $\text{CoVaR}_{mit,0.5}$ represents the VaR of the system at time $t$ when returns of asset $i$ are at $50^{th}$ percentile.

- Both $MES$ and $\Delta \text{CoVaR}$ aim to measure the contribution of a given institution to systemic risk.
Network based risk measures

- A network is defined as a set of nodes $V_t = \{1, 2, \ldots, n_t\}$ and directed arcs (edges) between nodes.
- The network can be represented through an $n_t$-dimensional adjacency matrix $A_t$, with the element $a_{ijt} = 1$ if there is an edge from $i$ directed to $j$ with $i, j \in V_t$ and 0 otherwise.
- The matrix $A_t$ is estimated by using a pairwise Granger causality approach to detect the direction and propagation of the relationships between the institutions (Billio et al., 2012).
Granger Causality Network

In order to test the causality direction the following model is estimated,

\[
\begin{align*}
    r_{it} &= \sum_{l=1}^{m} b_{11l} r_{it-l} + \sum_{l=1}^{m} b_{12l} r_{jt-l} + \epsilon_{it} \\
    r_{jt} &= \sum_{l=1}^{m} b_{21l} r_{it-l} + \sum_{l=1}^{m} b_{22l} r_{jt-l} + \epsilon_{jt}
\end{align*}
\]

with \(i \neq j, \forall i, j = 1, \ldots, n_t\). \(m\) is the max lag (selected according a BIC criteria) and \(\epsilon_{it}\) and \(\epsilon_{jt}\) are uncorrelated white noise processes. The definition of causality implies that,

- if \(b_{12l} \neq 0\) and \(b_{21l} = 0\), \(r_{jt}\) causes \(r_{it}\) and \(a_{ijt} = 1\).
- if \(b_{12l} = 0\) and \(b_{21l} \neq 0\), \(r_{it}\) causes \(r_{jt}\) and \(a_{ijt} = 1\).
- if \(b_{12l} \neq 0\) and \(b_{21l} \neq 0\), there is a feedback relationship among \(r_{it}\) and \(r_{jt}\) and \(a_{ijt} = a_{jit} = 1\).
IO network degree

- $IO_{it}$ is defined as

$$IO_{it} = \sum_{j=1}^{n_t} a_{ijt} + \sum_{j=1}^{n_t} a_{jit}. \quad (7)$$

- We also consider the DCI (in Rob. Checks),

$$DCI_t = \binom{n_t}{2}^{-1} \sum_{i=1}^{n_t} \sum_{j=1}^{n_t} a_{ijt}. \quad (8)$$

If $(DCI_t - DCI_{t-1}) > 0$, there is an increase of system interconnectedness.
The European financial sector

The dataset constitutes of the European financial institutions according to the ICB (code:8000).

<table>
<thead>
<tr>
<th>Market</th>
<th>$n_T$</th>
<th>Market</th>
<th>$n_T$</th>
<th>Market</th>
<th>$n_T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>43</td>
<td>Belgium</td>
<td>73</td>
<td>Denmark</td>
<td>179</td>
</tr>
<tr>
<td>Finland</td>
<td>30</td>
<td>France</td>
<td>285</td>
<td>Germany</td>
<td>344</td>
</tr>
<tr>
<td>Greece</td>
<td>82</td>
<td>Hungary</td>
<td>16</td>
<td>Ireland</td>
<td>30</td>
</tr>
<tr>
<td>Italy</td>
<td>139</td>
<td>Latvia</td>
<td>1</td>
<td>Lithuania</td>
<td>5</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>40</td>
<td>The Netherlands</td>
<td>87</td>
<td>Norway</td>
<td>78</td>
</tr>
<tr>
<td>Portugal</td>
<td>29</td>
<td>Spain</td>
<td>84</td>
<td>Sweden</td>
<td>113</td>
</tr>
<tr>
<td>Switzerland</td>
<td>149</td>
<td>United Kingdom</td>
<td>1310</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- daily closing price time series from Jan-1985 to May-2014
- MSCI EU as proxy for the market
- Use of the rolling window approach (252 obs.)
Estimated Cross-sectional MES
Estimated Cross-Sectional $\Delta\text{CoVaR}$

![Graph showing Estimated Cross-Sectional $\Delta\text{CoVaR}$]
Estimated Cross-Sectional IO network degree
Estimated entropies for MES

**Figure:** Shannon entropy (solid black line), Tsallis entropy (dashed blue line) and Rényi (dotted red line).
Estimated entropies for $\Delta$-CoVaR

**Figure:** Shannon entropy (solid black line), Tsallis entropy (dashed blue line) and Rényi (dotted red line).
Estimated entropies for IO network degree

Figure: Shannon entropy (solid black line), Tsallis entropy (dashed blue line) and Rényi (dotted red line).
Comments on the estimated entropies

All entropy measures exhibit long-term upward trend (magenta dashed line) and short-term oscillations about the trend.

- Tail risk measures (MES and ΔCoVaR) exhibit increasing and persistently high entropy during periods of financial turmoil.
- According to information theory, the increasing trend of the IO degree distribution indicates that the number of future possible network configurations is growing and thus the system becomes less predictable (Cover and Thomas, 2012).
  ⇒ A high level of entropy suggests that the network is heterogeneous with a degree distribution that is close to the uniform.
- Our findings also show that during tranquil periods the entropy is low, which indicates a more robust financial system.
- The observed persistence of the entropy dynamics makes this indicator suitable for forecasting financial crises.
Entropy as early warning signal

- An Early Warning System (EWS) is defined as a system that will issue a signal in case the likelihood of a crisis crosses a specified threshold (Davis and Karim, 2008).
- Literature has focused on the development of EWS for currency crisis, banking crisis and debt crises using different methodological frameworks (signaling approach, discrete choice models, and regression trees models).
- Our analysis implements entropy on systemic risk measures as an early warning indicator to signal banking crisis.
- This indicator represents also one of the target variables monitored by European Systemic Risk Board (ESRB).
Entropy as early warning indicator

- We use the dataset of Alessi and Detken (2014), which includes banking crisis at country level for the Euro area. We define a European global variable,
  \[ C_t = \begin{cases} 
  1 & \text{if more than one country is in crisis at time } t \\
  0 & \text{otherwise.} 
\end{cases} \quad (9) \]

- We denote with \( E_t^i \), the entropy \( i \) filtered out by the dataset dimension \( n_t \) to disentangle changes due to the variation in sample size (trend component),
  \[ H_{kt} = \gamma n_t + E_{kt}, \quad k = S, R, T. \quad (10) \]

- We specify a logistic model of the form
  \[ \mathbb{P}(C_t = 1|E_{kt}) = \Phi(\beta_0 + \beta_1 E_{kt}), \quad (11) \]
  where \( \Phi(x) = 1/(1 + \exp(-x)) \) is the logistic function.
### Regression results - Shannon

<table>
<thead>
<tr>
<th></th>
<th>Shannon ((k = S))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-5.3340 ***</td>
</tr>
<tr>
<td></td>
<td>(-5.3911) ***</td>
</tr>
<tr>
<td></td>
<td>(-6.4137) ***</td>
</tr>
<tr>
<td></td>
<td>(0.1996)</td>
</tr>
<tr>
<td></td>
<td>(0.1641)</td>
</tr>
<tr>
<td></td>
<td>(0.1851)</td>
</tr>
<tr>
<td>(E_k(MES))</td>
<td>10.9669 ***</td>
</tr>
<tr>
<td></td>
<td>(0.4151)</td>
</tr>
<tr>
<td>(E_k(\Delta CoVaR))</td>
<td>15.5536 ***</td>
</tr>
<tr>
<td></td>
<td>(0.4001)</td>
</tr>
<tr>
<td>(E_k(IO))</td>
<td>20.0670 ***</td>
</tr>
<tr>
<td></td>
<td>(0.5817)</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.1140 ()</td>
</tr>
<tr>
<td>Adj-(R^2)</td>
<td>0.1139 ()</td>
</tr>
<tr>
<td>Log-Lik</td>
<td>-4455.92</td>
</tr>
<tr>
<td>LLR</td>
<td>-3790.80</td>
</tr>
<tr>
<td>AIC</td>
<td>8915.84</td>
</tr>
<tr>
<td>BIC</td>
<td>8929.56</td>
</tr>
</tbody>
</table>

**Note:** \(***\) denotes statistical significance at the 0.01 level.
## Regression results - Tsallis

<table>
<thead>
<tr>
<th></th>
<th>Tsallis ($k = T$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>$-85.4113^{***}$</td>
</tr>
<tr>
<td></td>
<td>(5.0148)</td>
</tr>
<tr>
<td>$E_k(MES)$</td>
<td>$86.7045^{***}$</td>
</tr>
<tr>
<td></td>
<td>(5.0961)</td>
</tr>
<tr>
<td>$E_k(\Delta CoVaR)$</td>
<td>$29.0981^{***}$</td>
</tr>
<tr>
<td></td>
<td>(1.0367)</td>
</tr>
<tr>
<td>$E_k(IO)$</td>
<td>$25.0981^{***}$</td>
</tr>
<tr>
<td></td>
<td>(0.8479)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.0457</td>
</tr>
<tr>
<td>Adj-$R^2$</td>
<td>0.0456</td>
</tr>
<tr>
<td>Log-Lik</td>
<td>-4706.41</td>
</tr>
<tr>
<td>LLR</td>
<td>0.0340</td>
</tr>
<tr>
<td>AIC</td>
<td>9416.83</td>
</tr>
<tr>
<td>BIC</td>
<td>9416.83</td>
</tr>
</tbody>
</table>
### Regression results - Rényi

<table>
<thead>
<tr>
<th></th>
<th>Rényi ((k = R))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-4.4896***</td>
</tr>
<tr>
<td></td>
<td>(0.1647)</td>
</tr>
<tr>
<td>(E_k(MES))</td>
<td>11.2202***</td>
</tr>
<tr>
<td></td>
<td>(0.4159)</td>
</tr>
<tr>
<td>(E_k(\Delta CoVaR))</td>
<td>14.5302***</td>
</tr>
<tr>
<td></td>
<td>(0.3791)</td>
</tr>
<tr>
<td>(E_k(IO))</td>
<td>20.0556***</td>
</tr>
<tr>
<td></td>
<td>(0.5803)</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.1196 0.2887 0.1965</td>
</tr>
<tr>
<td>(\text{Adj-}R^2)</td>
<td>0.1195 0.2886 0.1963</td>
</tr>
<tr>
<td>Log-Lik</td>
<td>-4438.51 -3828.46 -4103.57</td>
</tr>
<tr>
<td>LLR</td>
<td>0.0889 0.2142 0.1577</td>
</tr>
<tr>
<td>AIC</td>
<td>8881.03 7660.92 8211.14</td>
</tr>
<tr>
<td>BIC</td>
<td>8894.75 7674.64 8224.86</td>
</tr>
</tbody>
</table>
Estimated Response variables

Figure: Overview of the actual (horizontal red lines) and the estimated response variable over time of MES (solid black line), ΔCoVaR (dashed blue line) and IO network degree (dotted red line).
Empirical Analysis

Percentage of correctly predicted indicators

Considering $\hat{\Phi}_t$ as the predicted probability of crisis returned by the logit model, we define a binary variable $\hat{C}_t$ such that:

$$\hat{C}_t = \begin{cases} 1 & \text{if } \hat{\Phi}_t \geq 0.50, \\ 0 & \text{otherwise}. \end{cases}$$

The Percentage of correctly predicted indicators is defined as the number of times where $C_t = \hat{C}_t$.

<table>
<thead>
<tr>
<th>% corrected predicted</th>
<th>MES</th>
<th>$\Delta CoVaR$</th>
<th>IO degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>65.32%</td>
<td>77.51%</td>
<td>69.79%</td>
<td></td>
</tr>
</tbody>
</table>
The extensive index $\alpha$ allows the researcher to identify the relevance of the tails of the distributions in the crises prediction.

Follow the information theory perspective, we minimize the ratio between the extensive entropy and the maximum entropy, we choose the $\alpha$ which minimizes the AIC criterion of the logit model,

$$\min_{\alpha \in (0, +\infty)} \text{AIC}(\alpha).$$

where $\text{AIC}(\alpha) = 2k - 2 \ln(L(\alpha))$, with $L$ the likelihood function associated to the logit specification.
**AIC(α) for MES**

![Graph showing AIC as function of α for Tsallis (solid) and Rényi (dashed) entropies in the logit model.](image)

**Figure:** AIC as function of $\alpha$ for Tsallis (solid) and Rényi (dashed) entropies in the logit model.
AIC($\alpha$) for $\Delta$-CoVaR

**Figure:** AIC as function of $\alpha$ for Tsallis (solid) and Rényi (dashed) entropies in the logit model.
AIC(α) for IO network degree

Figure: AIC as function of α for Tsallis (solid) and Rényi (dashed) entropies in the logit model.
Robustness Checks

We perform two types of Robustness Checks:

- Alternative indicators: Entropy measures show their superior explanatory power with respect to the DCI and alternative specifications for cross-sectional systemic risk measures (mean and volatility).

- Results are robust also when we formulate alternative banking crisis definition by changing the number of countries required to define an European banking crisis ($N > 2$ and $N > 3$).
Conclusions

- We proposed a new approach based on the cross-sectional entropy of systemic risk measures by exploiting their ability to detect systemic events.
- Entropy measures have estimated considering alternative definitions.
- Given the persistence showed by entropy indicators, we built an EWS for banking crises using entropy measures.
- Our findings highlighted the goodness of entropy indicators in measuring these crises.
- Findings also show that the values of the extensive index ($\alpha$) are greater than one: entropy changes due to changes in the tail probability of the loss distribution are important in order to detect periods of high systemic risk level.
- Further investigation could be performed using other risk measures and target variables.