I. MOTIVATION

• Financial crises are typically associated with countries’ difficulties in repaying their public debt or overall (public and private) external debt

• They are also typically associated with currency depreciations and a disruption in firm and household financing → So, output and employment suffer


• **2000s**: The 2008-09 financial crises in the US, followed by debt crises in some eurozone countries
• Often, a crisis is preceded by some “exuberance” in asset markets, such as housing markets, stock markets, and often also in foreign exchange markets (currency overvaluation).

• The big debate is whether such “exuberance” is “rational” or “irrational”

• For instance, the famous economist Rudi Dornbusch rang a “good alarm” about over-optimism on Mexico months before the crisis; Raghu Rajan (former IMF chief economist) rang another good alarm on financial conditions prior the 2008-09 crisis.

• But in the past, there were also plenty of “false alarms”
• Indeed, Gourinchas, Valdes & Landerretche (2001) document that most financial bonanzas do not end up in crises.

• Yet, the cost of crises like those of the 1990s and 2008-09 has been huge.

• So, a “early warning system” (EWS) capable of identifying main macroeconomic vulnerabilities and maximizing “good alarms” while minimizing “false alarms” is clearly welcome.

• This is a key motivation of the G-20 call in 2009 for the IMF and the FSB to implement Early Warning Exercises (EWE) (the idea being to combine IMF’s macro-financial surveillance strengths with FSB’s expertise on bank regulation and supervision).
II. BUILDING EARLY WARNING SYSTEMS (EWS)

Important to be clear about three questions before “getting hands dirty”:

• **What** (crises) to predict?

• **How** (to choose a method) to predict?

• **When** (in advance) to predict?
WHAT TO PREDICT

• Many “crises” definitions around
• One common denominator is a major recession/depression (i.e., major output and employment losses)
• But triggering factors and other symptoms may be different
• Many output collapses in the past were associated with large exchange rate depreciations and/or widespread run on (and failures of) domestic banks
• Hence, the initial focus of work on EWS was either on “currency crises” or “twin crises” (currency and banking crises)
But currency and banking crises may be just a side-effect of triggers such as un-sustainable fiscal deficits and/or high public debt, given the exposure of banks and currency markets to those fiscal developments

Or simply reflect loss of confidence in the economy due, e.g., to a crisis and lower growth in a neighboring country/trade partner (“contagion”)

In practice, there is positive correlation in these “crisis” symptoms but those correlations are far from perfect

For instance, many external debt crises are not associated with a major currency depreciation (notably as in the eurozone recently)
WHAT TO PREDICT: CLASSIC CRISIS DEFINITIONS

• Some authors focus on other crises in addition to currency crises:

  – Banking Crisis: a systemic banking crisis (as defined in Laeven and Valencia 2008 IMF WP)

  – Growth Crisis: significant slowdown in growth relative to trend

  – External Crisis: default/rescheduling of external debt and/or temporary resort to heavy multilateral financing

  – “Sudden Stops” Crises: Some combination of the above
**WHAT TO PREDICT: CORRELATIONS OF CRISIS DEFINITIONS**

Table 1. Spearman Ranking Correlations Between Distinct Crisis Indicators 1/

<table>
<thead>
<tr>
<th></th>
<th>External Crisis</th>
<th>Currency Crisis</th>
<th>Currency &amp; Growth Crisis</th>
<th>Currency &amp; IMF program</th>
<th>LV Currency Crisis</th>
</tr>
</thead>
<tbody>
<tr>
<td>External Crises</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Currency Crises</td>
<td>0.32</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Currency &amp; Growth Crises</td>
<td>0.32</td>
<td>0.73</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Currency and IMF program</td>
<td>0.57</td>
<td>0.46</td>
<td>0.52</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>LV Currency Crisis</td>
<td>0.33</td>
<td>0.40</td>
<td>0.47</td>
<td>0.42</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*No of obs = 2674*

1/ For specifics of each crisis definition, see Catão and Milesi-Ferretti, "External Liabilities and Crises“, Journal of International Economics, 94, pp. 18-32
Goal of the model is to combine information for a wide set of vulnerability indicators.

Two main choices then involved:

- Indicators to be considered
- How to combine information from indicators into a measure of macroeconomic vulnerability
As mentioned above, the focus of the earlier work was on currency crises (given the prevalence of such crises in the 1990s).

- Frankel and Rose (1996 JIE) use probit regressions to estimate currency crises in emerging markets using annual data.


- Berg and Patillo (IMF SP 1999) uses monthly data and a regression framework to predict currency crises.
The performance of earlier EWM models is reviewed in Berg et al (IMF OP 186, 1999) and others

- Out-of-sample performance can be reasonable, e.g. probability of crisis given an alarm of 59 percent (11 percent without alarm), with 41 percent alarms proving false.

- However, not widely recognized as very good: still too many missed crises, too many false alarms.
How and When to Predict: The Signaling Approach

• Predicting (the timing of) crises is very hard

• Even when vulnerabilities identified, it may take several years for a crisis to materialize
  – Signal “fatigue”
  – Crises are rare (e.g. 5% of sample)—if proba<10% not “flag” it

• Thus, instead of trying to predict timing of crises, more promising to focus on identifying vulnerabilities

• Detect crisis symptoms early so as to allow time for adopting preemptive measures

• The basic idea is that economic indicators systematically behave differently before a crisis
Stylized Facts on External Crises

Source: Catão and Milesi-Ferretti (2014)
Stylized Facts on External Crises cont.

Source: Catão and Milesi-Ferretti (2014)
• In the crisis risk models we seek to identify particular vulnerabilities that have been associated with crises episodes in the past

• The approach similar to predicting crises, but interpretation of results different

• The approach is also more lenient on false alarms than on missing crises

• This will be the base for the criterion of risk threshold selection we will implement in this afternoon’s workshop
## IMF VE Indicators

<table>
<thead>
<tr>
<th>External Sector</th>
<th>EMs</th>
<th>Financial Sector</th>
<th>EMs</th>
<th>Corporate/Real Sector</th>
<th>EMs</th>
<th>Public Sector</th>
<th>EMs</th>
<th>Public Sector</th>
<th>EMs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current Account/GDP</td>
<td>✓ ✓</td>
<td>Capital Adequacy Ratio (Banks)</td>
<td>✓ ✓</td>
<td>Return on Assets (Corporate)</td>
<td>✓ ✓</td>
<td>Primary Fiscal Gap</td>
<td>✓ ✓</td>
<td>Public Debt Exposed to Rollover Risk</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>External Debt/Exports of Goods and Services</td>
<td>✓</td>
<td>Nonperforming loans (% of total)</td>
<td>✓</td>
<td>Interest Coverage ratio</td>
<td>✓</td>
<td>Real export growth</td>
<td>✓</td>
<td>Public Debt Exposed to Rollover Risk</td>
<td>✓</td>
</tr>
<tr>
<td>Real export growth</td>
<td>✓</td>
<td>Foreign liability (% of domestic credit)</td>
<td>✓✓</td>
<td>Real GDP Growth</td>
<td>✓✓</td>
<td>Real Effective Exchange Rate Overvaluation</td>
<td>✓</td>
<td>Loan/deposits ratio</td>
<td>✓</td>
</tr>
<tr>
<td>CGER Current Account Norm Deviation</td>
<td>✓✓</td>
<td>Loan/deposits ratio</td>
<td>✓✓</td>
<td>Real GDP Growth</td>
<td>✓✓</td>
<td>Private sector external debt</td>
<td>✓</td>
<td>Loan/deposits ratio</td>
<td>✓</td>
</tr>
<tr>
<td>Public Sector</td>
<td>EMs</td>
<td>Financial Sector</td>
<td>EMs</td>
<td>Corporate/Real Sector</td>
<td>EMs</td>
<td>Public Sector</td>
<td>EMs</td>
<td>Public Sector</td>
<td>EMs</td>
</tr>
</tbody>
</table>
## IMF VE Indicators, Cont’d

### Other (Medium Term) Indicators

<table>
<thead>
<tr>
<th>Indicator</th>
<th>AEs</th>
<th>EMs</th>
</tr>
</thead>
<tbody>
<tr>
<td>House Prices</td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td>Stock Prices</td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td>Private Credit growth</td>
<td>✅</td>
<td>✅</td>
</tr>
<tr>
<td>Construction sector contribution to GDP growth</td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td>Financial sector contribution to GDP growth</td>
<td>✅</td>
<td></td>
</tr>
</tbody>
</table>

### Other Household Sector Indicators

<table>
<thead>
<tr>
<th>Indicator</th>
<th>AEs</th>
<th>EMs</th>
</tr>
</thead>
<tbody>
<tr>
<td>House Price Acceleration</td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td>Stock Price Acceleration</td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td>Household Liabilities/GDP</td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td>Interaction (Household Liabilities)*(Medium-Term House Price Increase)</td>
<td></td>
<td>✅</td>
</tr>
<tr>
<td>Interaction (Household Liabilities)*(House Price Acceleration)</td>
<td></td>
<td>✅</td>
</tr>
</tbody>
</table>
**ESTIMATION OF EARLY WARNING**

- Probit/Logit regression techniques widely used

- One problem is to accommodate a large number of explanatory variables in one regression

- And the choice of the form of the probability distribution function/functional form (linear regression, probit, logit)

- An alternative we will discuss in the workshop is a non-parametric **signal extraction model**
THE NON-PARAMETRIC THRESHOLD APPROACH

• This approach allows us to consider a wide array of variables
• Thresholds themselves are robust to “over-fitting”
• For example, threshold for current account/GDP is not affected by any of the other variables considered
• Threshold flags a risk without making parametric probabilistic assumptions:
  
  — E.g. a 10 percent current account deficit is not assumed to be twice as risky as a 5 percent one. Instead both will lead to a same “flag” being raised.

• But easy also to see disadvantages to the latter
**WHEN TO: DEFINING THE SIGNALING HORIZON (KLR)**

- Time period within which the indicators are expected to have an ability to anticipate crises.

- Policymakers (central banks, treasuries, int. orgs.) interested in early assessment of crisis likelihood to allow for time to implement pre-emptive policies.

- Usual focus is **the 24 months prior** to the crisis.

- **Good signal**: signal followed by crisis within 24 months, otherwise ‘false signal’ (noise).

- If we only have good data on annual basis (as in many EMs), then a one year lag is standard too (today’s workshop).
If a variable crosses a “threshold,” sending a “signal” of a future crisis, there are two possible scenarios:

– The crisis happens within one or two years after the signal “1” was given (the signal was accurate)
– The crisis does not occur within one or two year (the signal is labeled a “false alarm”)

**Threshold levels must strike a balance between**
- the **risk of missing many crises**: ‘false negative’ or “no alarm”
- the **risk of having many false signals**: “false positives”
The "Confusion Matrix"

Actual realization

\[
\text{Noise} = \frac{B}{B+D} \quad \text{Accuracy} = \frac{A+D}{A+B+C+D} \quad \text{Sum of errors} = \frac{C}{A+C} + \frac{B}{B+D}
\]

THE SIGNAL

Positive ("1")

Crisis

Missed Crises

Negative ("0")

No-Crisis

False alarms

Model classification

\[
\begin{array}{|c|c|}
\hline
\text{Positive ("1")} & \text{Negative ("0")} \\
\hline
\text{Crisis} & \text{No-Crisis} \\
\hline
A & B \\
C & D \\
\hline
\end{array}
\]

Noise-to-Signal: \( \frac{B}{B+D} \)  
Accuracy: \( \frac{A+D}{A+B+C+D} \)  
Sum of errors: \( \frac{C}{A+C} + \frac{B}{B+D} \)
THE SIGNAL cont: “TYPE I” VS. “TYPE II” ERRORS

• If the null hypothesis is that there will be a crisis, predicting a non-crisis (“tranquility period”) when there happens to be one, would be a “type I error”

• Using the same convention (H0: there will be crisis), when the model predicts a crisis and there is none, then this would be a “type II error”

• In short, if the null hypothesis is there will be a crisis, then
type I error = missed crisis
type II error = false alarm

• This is the convention we will use for the workshop

Note: Notice that if the opposite convention used, i.e. the null is no crisis. In this case, the above error typology is flipped.
Example of abnormal behavior of an indicator (signal) and a normal behavior (no signal)?

(Likely) Abnormal: a current account deficit of 20 percent of GDP

(Likely) Normal: a current account deficit of 1 percent of GDP

The question is then where to choose a threshold below which the variable (e.g. CA) is behaving “normally” (i.e., consistent with tranquil period) of “abnormally” (as in the verge of a crisis)
Another Example of abnormal behavior of an indicator (signal) and a normal behavior (no signal)?

(Likely) Abnormal: FX reserve coverage of external repayment obligations during the year of < 100% (Guidotti-Greenspan rule)

(Likely) Normal: FX reserve cover > 100%

→ Note that **unlike the CA deficit**, the higher the safer! So a risk signal will be issued for lower (rather than higher) values of that variable
CHOOSING THE THRESHOLD: TRADE-OFFS

• **Too loose a threshold**: the indicator will catch all the crises, but will also give lots of false alarms (noise):
  
  – Current Account Deficit of 1 percent of GDP

• **Too tight a threshold**: the indicator will never issue a false signal, but it will miss many crises:
  
  – Current Account Deficit of 20 percent of GDP

• The idea is then some balancing between these extremes
A PERFECT INDICATOR?

• If existed, would only produce observations in cells A and D

• But in practice, no indicator(s) produce(s) that!

• So, some indicators are better than others depending on the objective function of the classifier.

• IMF VE criterion: minimize the absolute number of errors, i.e:

  * Min. \[ \text{Missed crises/Total crises} \] (Type I error, cell C) \[ + \] \[ \text{False Alarms/Total non-crises} \] (Type II error, cell B)

→ Because crises are much rarer (typically 5% or less of time), this implies giving a much higher weight to avoiding missing a crisis
Putting the VE in Practice

• With this conceptual discussion in the bag, you are now in a position to start “dirtying your hands” by implementing an EWS exercise

• In the remaining of this lecture and in the workshop we will focus on the IMF VE, but the idea is the same as in widely used implementations (like in Kaminsky-Lizondo-Reinhart and the Kaminsky-Reinhart papers – see reading list)

• Before we turn to specific steps, let’s just take a quick look at how this is institutionally implemented at the IMF under the so-called IMF/FSB VEE approach
Assessment of underlying vulnerabilities and crisis risks in Emerging Market (EM) economies

Emphasis on financial account ("sudden stop") crises

Began 2001, periodically updated

Sample of 51 EM countries, starting from 1993, annual data

Semi-annual exercise

- Collaboration between functional and area departments
- Increased interaction with market participants.
Like choosing an insurance policy, you must first be clear about which kind of macroeconomic crisis you want to be protected of.

Since the correlation between different variety of crises is far from one (as we saw above), this is important.

In the IMF VE for Emerging markets (the so-called “VEE”), the emphasis is on external financial crises under the broad umbrella or “Sudden Stop” crisis.

We will stick to that in the workshop.

Then the question is how to identify sudden stops for various countries/years, but here we will take that as coming from given data.
HOW TO CHOOSE A RISK THRESHOLD

• Different criteria have been used by distinct analysts

• Here, let’s stick to the criterion of minimizing type 1+ type 2 errors (that used in the IMF VEE), which is quite sensible

• The next step is then to pick the (macro) variables we think behave very differently between crisis and non-crisis (we already discussed a list)

• Then find a “threshold” for each of these variables (e.g. reserve cover)

• Think of each country/year observation for the reserve cover as coming from two possible cumulative distributions: the “non-crisis” distribution and the other is the crisis distribution. Plot them. E.g.:
OPTIMAL THRESHOLD SELECTION

Threshold maximizes the distance between the two CDFs.

Higher reserve cover --> low crisis risk

Type I error → Threshold

Type II error ←
**SIGNALING AND AGGREGATION OF INDIVIDUAL SIGNALS**

- **Signaling**: A binary variable is assigned for each indicator taking value of:
  - 1 if in crisis-prone side of threshold (“flag”)
  - 0 otherwise.

- One possibility is to compute the final index $I$ weighing all indicators $S_j$ equally

\[
I_t = \sum_{j=1}^{n} S_t^j
\]

Another is to weigh acc. to their relative performance in minimizing the sum of type 1 and 2 errors:

\[
I_t = \sum_{j=1}^{n} \frac{S_t^j}{\omega_j}
\]
AGGREGATION OF INDIVIDUAL SIGNALS, CONT’D

• But you can also use economic intuition and the specificities of your country in choosing the weights

• In the IMF VE exercise for advanced countries (VEA), for instance, the weights adjusted to offset impact of including highly correlated indicators (“pruning the model”).

• Or we could group different indicators by sectors (external, fiscal, financial, and real), and distribute weights within sectors and across sectors based on your and other policy makers’ experiences

• Then you can check which combination gives a better in-sample prediction (high correlation) with actual crises

• The last approach is what we will implement in the workshop
SUMMARY OF STEPS

• **Challenge:** transform continuous values of many variables into an index of overall underlying vulnerability

• **First Step:** identify thresholds for individual variables

• **Second step:** transform indicator values into scores
  – One if indicator value is in the crisis-prone side of threshold
  – Zero, otherwise.

• **Third step:** Group variables into sectors and choose within-sector weights

• **Fourth step:** Sector indices = weighted average of the indicator scores

• **Fifth step:** overall vulnerability index = weighted average of the four sector indices
## EXAMPLE: INDICATORS BY SECTOR

<table>
<thead>
<tr>
<th>Direction to be safe</th>
<th>External Sector</th>
<th>Financial Sector</th>
<th>Real Sector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reserve coverage GIR/ (ED+CA)</td>
<td>&gt;</td>
<td>Capital adequacy ratio (%)</td>
<td>&gt;</td>
</tr>
<tr>
<td>Current Account ( % of GDP)</td>
<td>&gt;</td>
<td>Return on assets (%)</td>
<td>&gt;</td>
</tr>
<tr>
<td>External Debt ( % of Exports)</td>
<td>&lt;</td>
<td>Loan-to-deposit ratio (%)**</td>
<td>&lt;</td>
</tr>
<tr>
<td>REER overvaluation</td>
<td>&lt;</td>
<td>3-year cumulative change in credit-to-GDP (%)**</td>
<td>&lt;</td>
</tr>
<tr>
<td>Private sector external debt (%)**</td>
<td>&lt;</td>
<td>Foreign liability ( % domestic credit)**</td>
<td>&lt;</td>
</tr>
<tr>
<td>Public Sector</td>
<td>&gt;</td>
<td>Black-Scholes-Merton probability of default ( %)</td>
<td>&lt;</td>
</tr>
<tr>
<td>Fiscal balance ( % of GDP)</td>
<td>&gt;</td>
<td>Interest coverage ratio</td>
<td>&gt;</td>
</tr>
<tr>
<td>Primary Gap ( in % of GDP)</td>
<td>&gt;</td>
<td>Valuation (earnings-to-price ratio, adj. for CPI inflation)</td>
<td>&gt;</td>
</tr>
<tr>
<td>Public Debt ( % of GDP)</td>
<td>&lt;</td>
<td>Real GDP growth (%)**</td>
<td>&gt;</td>
</tr>
<tr>
<td>Public debt exposed to FX risk ( % of GDP)</td>
<td>&lt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public debt exposed to rollover risk ( % of GDP)</td>
<td>&lt;</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: IMF (2010)
Overall Vulnerability Index

External sector Index

Public sector index

Financial sector index

Real/Corporate sector index

Reserves in percent of ST debt + CA deficit

CA in % of GDP

External debt in % of GDP

External debt in % of exports

REER misalignment

Bottom-up Approach
Now that we have values for the overall vulnerability index for country/year, the question is to assess higher vs. lower risk.

One criterion is if the number of sector flags is more than one stdev, or within one stdev above the mean, respectively.

The other is simply to compute a broad cross-country mean of the final index and see which country/year is above (and how much above).

We will follow this latter approach in the workshop.
LAST STEP: VERY FINAL ROUND

Vulnerability Index (VI)

Indicator-based vulnerability rating*
(High, Medium, Low)

Role of Qualitative/Policy Factors

Final Ratings

Judgment
**APPLICATION IN MULTILATERAL SURVEILLANCE: MAP OF COUNTRY RISK**

<table>
<thead>
<tr>
<th>Countries with high underlying vulnerabilities overall</th>
<th>Countries with medium underlying vulnerabilities overall</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Country</strong></td>
<td><strong>Overall</strong></td>
</tr>
<tr>
<td>----------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Country 1</td>
<td></td>
</tr>
<tr>
<td>Country 2</td>
<td></td>
</tr>
<tr>
<td>Country 3</td>
<td></td>
</tr>
<tr>
<td>Country 4</td>
<td></td>
</tr>
<tr>
<td>Country 5</td>
<td></td>
</tr>
<tr>
<td>Country 6</td>
<td></td>
</tr>
<tr>
<td>Country 8</td>
<td></td>
</tr>
<tr>
<td>Country 9</td>
<td></td>
</tr>
<tr>
<td>Country 10</td>
<td></td>
</tr>
<tr>
<td>Country 11</td>
<td></td>
</tr>
<tr>
<td>Country 13</td>
<td></td>
</tr>
<tr>
<td>Country 14</td>
<td></td>
</tr>
<tr>
<td>Country 15</td>
<td></td>
</tr>
</tbody>
</table>