Stress testing and financial stability

Mark D. Flood

Department of Finance

University of Maryland

Center for Latin American Monetary Studies (CEMLA)

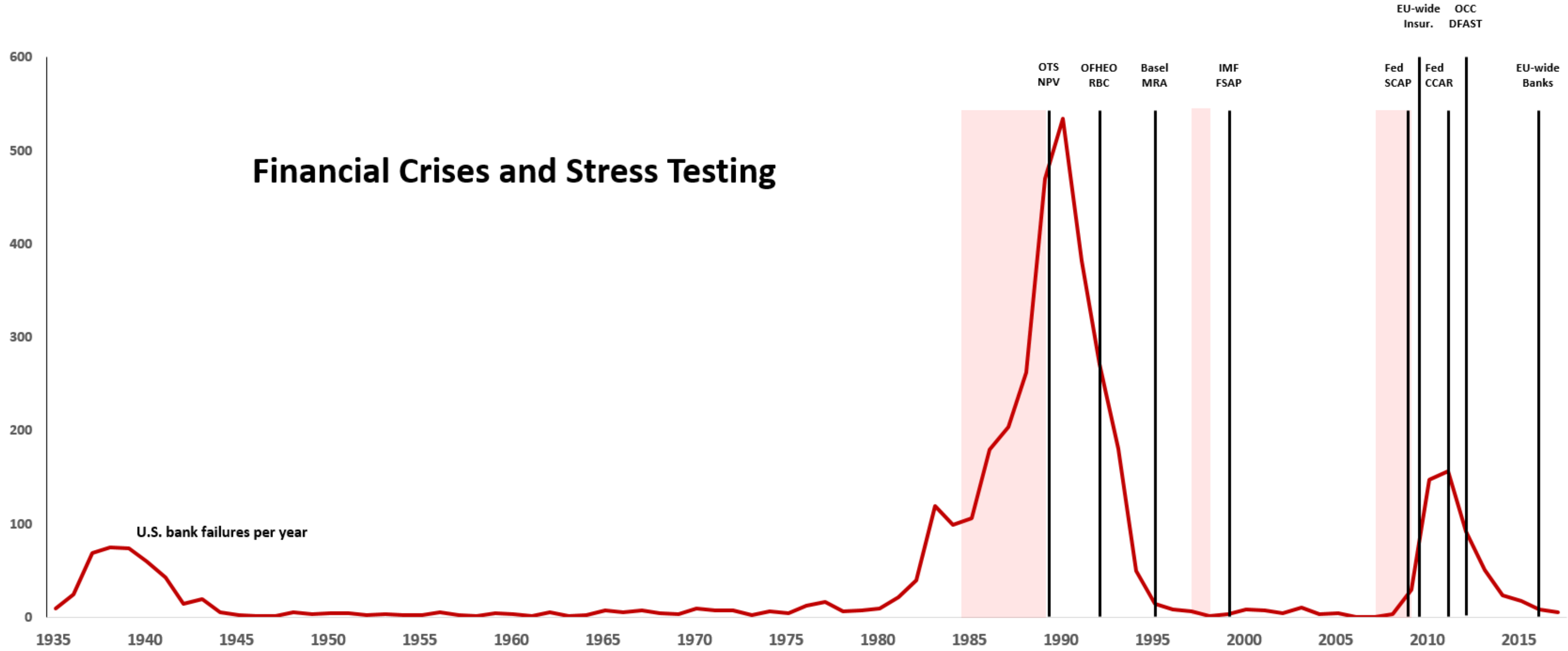
Course on Financial Stability

Mexico City, 19 September 2019

The Blind Side

Financial crises Granger-cause stress testing programs

- 1980s S&L Crisis ➔ OTS NPV, OFHEO RBC, Basel MRA
- 1997 Asian Financial Crisis ➔ IMF / World Bank FSAP
- 2007-09 Financial Crisis ➔ Fed SCAP, Fed CCAR, OCC DFAST, EIOPA EU-wide ST, EBA EU-wide ST



Supervisory Stress Testing v1.0

Some examples

- 1992 OFHEO housing scenario
- 1996 Basel market risk amendment
- 2001 IMF Financial Sector Assessment Programs (FSAPs)

Characteristics

- Microprudential only
- Focus on historical scenarios (“fighting the last war”)
- Scenarios and models inconsistent across firms
- Extrapolating from value-at-risk (VaR)

Supervisory Stress Testing v2.0

Some examples

- 2009 Supervisory Capital Assessment Program (SCAP)
- Comprehensive Capital Assessment and Review (CCAR)
- Dodd-Frank Act Stress Tests (DFAST)
- European Banking Authority (EBA) stress tests

Characteristics

- Detailed, consistent data collection – e.g., FRB Y-14
- Detailed analytics – supervisors augment firms' models
- Public disclosure – more than a compliance exercise
- Still largely microprudential

Possibilities for Stress Testing v2.1

Enhanced scenario selection

- Enhanced scenario design
- Increased scenario counts
- Reverse stress testing

Selective resolution

- Coarse stress tests for typical high-level assessment
- Detailed (granular) analysis for critical cases

Alignment with internal risk management

Stressing liquidity and solvency jointly

- Liquidity stress is likely to accompany capital stress

Next Generation Stress Testing – v3.0

Modeling Systemic Effects

- Systemically important institutions
- Correlated exposures
- Feedback dynamics (e.g., fire sales and funding runs)

Incorporating Reaction Functions

- Firms' reactions
- Policymakers' reactions

Shifting Landscape

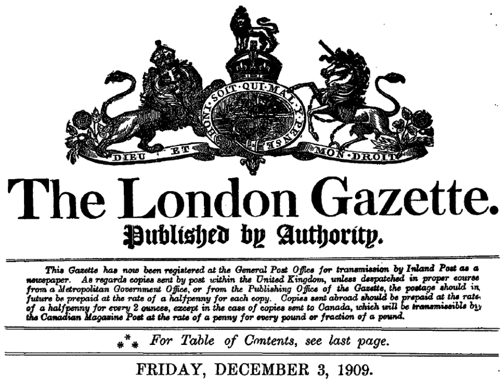
- New institutions (not just large BHCs)
- New risks and asset classes

Agent-based Modeling

- A possible methodology for Stress Testing v3.0

Applied Economic Epistemology

Economist's view of the world


DIEU ET MON DROIT

The London Gazette.

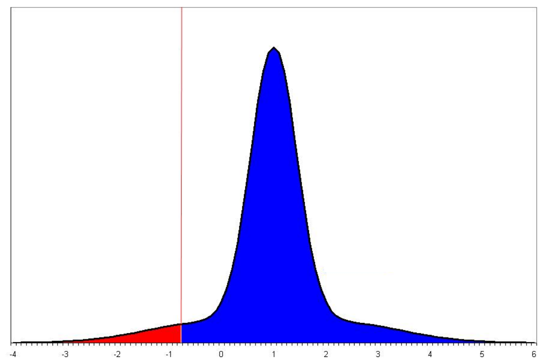
Published by Authority.

This Gazette has now been registered at the General Post Office for transmission by Inland Post as a newspaper. As regards copies sent by post within the United Kingdom, unless despatched in proper course from a Metropolitan Government Office, or from the Publishing Office of the Gazette, the postage should in future be prepaid at the rate of a halfpenny for each copy. Copies sent abroad should be prepaid at the rate of a halfpenny for every 2 ounces, except in the case of copies sent to Canada, which will be transmissible by the Canadian Magazine Post at the rate of a penny for every pound or fraction of a pound.

*** *For Table of Contents, see last page.*

FRIDAY, DECEMBER 3, 1909.

Ex-post published facts



Ex-ante measurable risk

Knightian uncertainty

But also

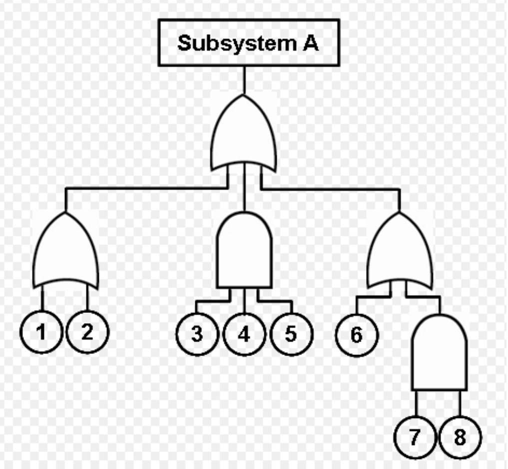
- **Model risk and ambiguity**
- **Asymmetric information**
- **Moral hazard and incentives**

Risk measurement without a "measure"

Financial context

- **Stress testing**
- **Stress scenario selection**
  - Severe, yet "plausible"
  - Plausibility wars

Functional hazard identification and fault tree analysis



Engineering context

- **Uncertainty quantification**
- **Maximum permissible probability of failure**
  - 10^{-9} aviation industry (catastrophic event per flight hour)
  - 0.00 nuclear power plants (seismic design)
  - 0.05 surface mining (collapse of soil embankments)
- **Worst case scenario analysis**

Optimal Uncertainty Quantification (OUQ)

The Certification Problem

- Guarantee that

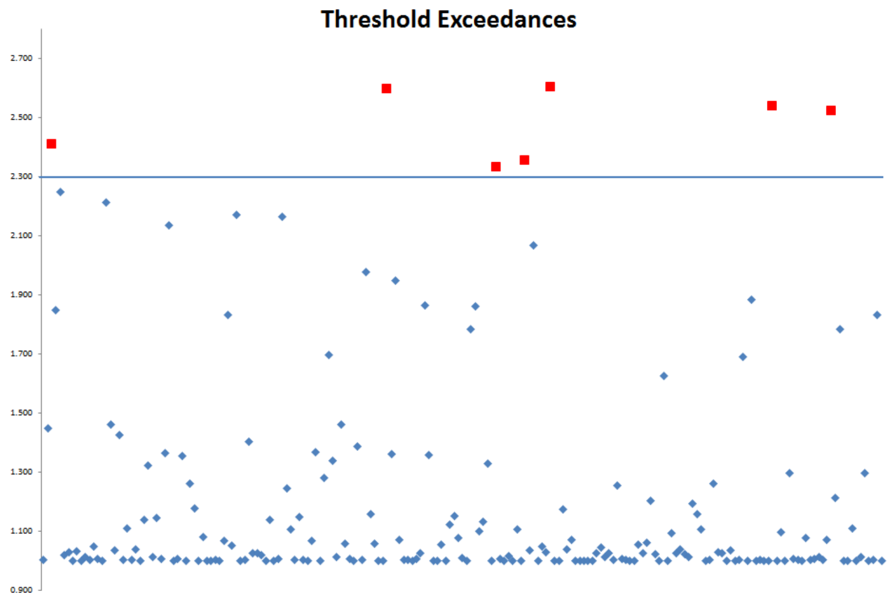
$$P\{G(X) \geq \alpha\} \leq \varepsilon$$

Where

- X is a risky or uncertain scenario
- P is a probability measure
- $G(X)$ is a system response (the "quantity of interest")
- $G(X) \geq \alpha$ is some event (typically undesirable)

But

- P is unknown or partially known
- G is unknown or partially known



Optimal Uncertainty Quantification (OUQ)

SCAP as Certification

Ben Bernanke (2013)

Stress testing banks: What have we learned?

“In retrospect, the SCAP stands out for me as one of the critical turning points in the financial crisis. It provided anxious investors with something they craved: credible information about prospective losses at banks. Supervisors' public disclosure of the stress test results helped restore confidence in the banking system and enabled its successful recapitalization.”

Concentration inequalities

Chebyshev's Inequality

- Let X be an integrable random variable with finite mean, μ, and finite (non-zero) variance, σ^2.
- Then

$$P\{\ |X - \mu| \geq k\alpha\ \} \leq 1/k^2$$

McDiarmid has two key assumptions

Concentration inequalities bound the difference between an RV and its mean by limiting the extent of possible variation in the RV.

E.g., a finite diameter restriction.

McDiarmid's Inequality

In bounding $P\{\ G(X) \geq \alpha\ \}$, if:

- The components of X are statistically independent, and
- The component-wise oscillations of $G(X)$ have finite diameter,
- Then

$$P\{\ G(X) \geq E[G(X)] + \varepsilon\ \} \leq \exp[-2\varepsilon^2/\Delta^2]$$

- Where Δ^2 is the "wiggle room" in $G(X)$:

$\Delta^2 \equiv \sum_m \delta_m^2$ for the component-wise oscillation bounds, δ_m

Application to financial stress testing

A laddered portfolio of U.S. Treasuries

- **Response function defined by profit or loss:**

$$G(X) \equiv \mathrm{E}[L(X)]$$

Where

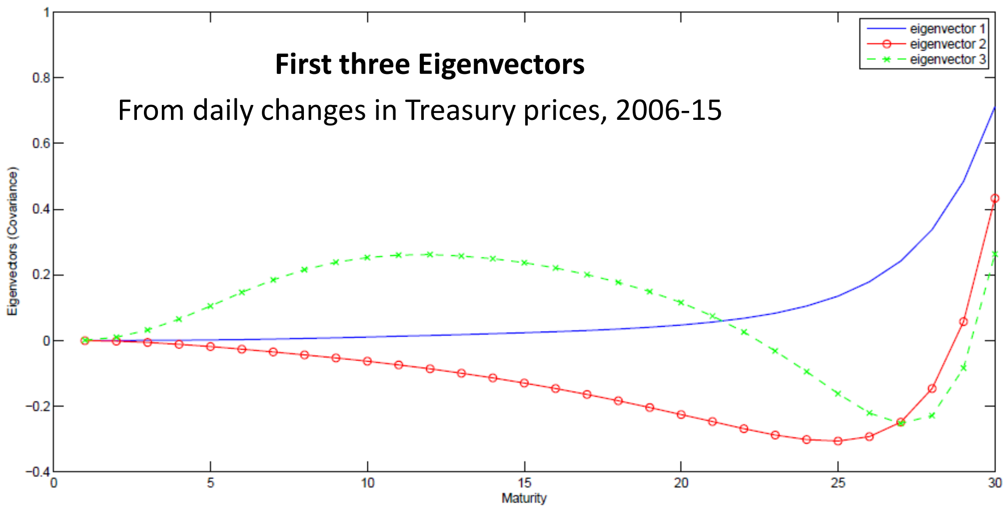
- $X \in \mathrm{R}^D$ **is embedded in the yield curve**
- **E[●], is w.r.t. an unknown dist'n**

Note

- **The profit-loss function, $L(X)$, is bounded, both above and below**
- **To apply McDiarmid, we must show the risky inputs, X, are independent**

Principal components analysis

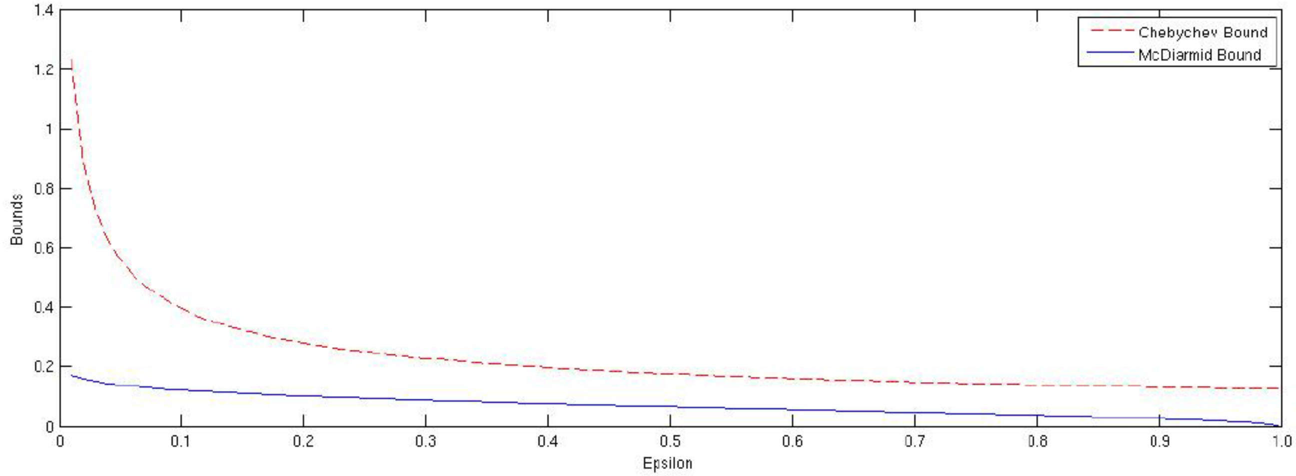
- **Extracted from time series of daily bond price *changes*, 2006-15**
- **First 3 components explain 99.9977%**
- **First 2 components explain 99.9733%**



Results

Result #1 — Proof of Concept

- We can implement OUQ for a simple financial stress test
- McDiarmid's distance allows for formal certification guarantees
- McDiarmid is indeed stronger than Chebyshev



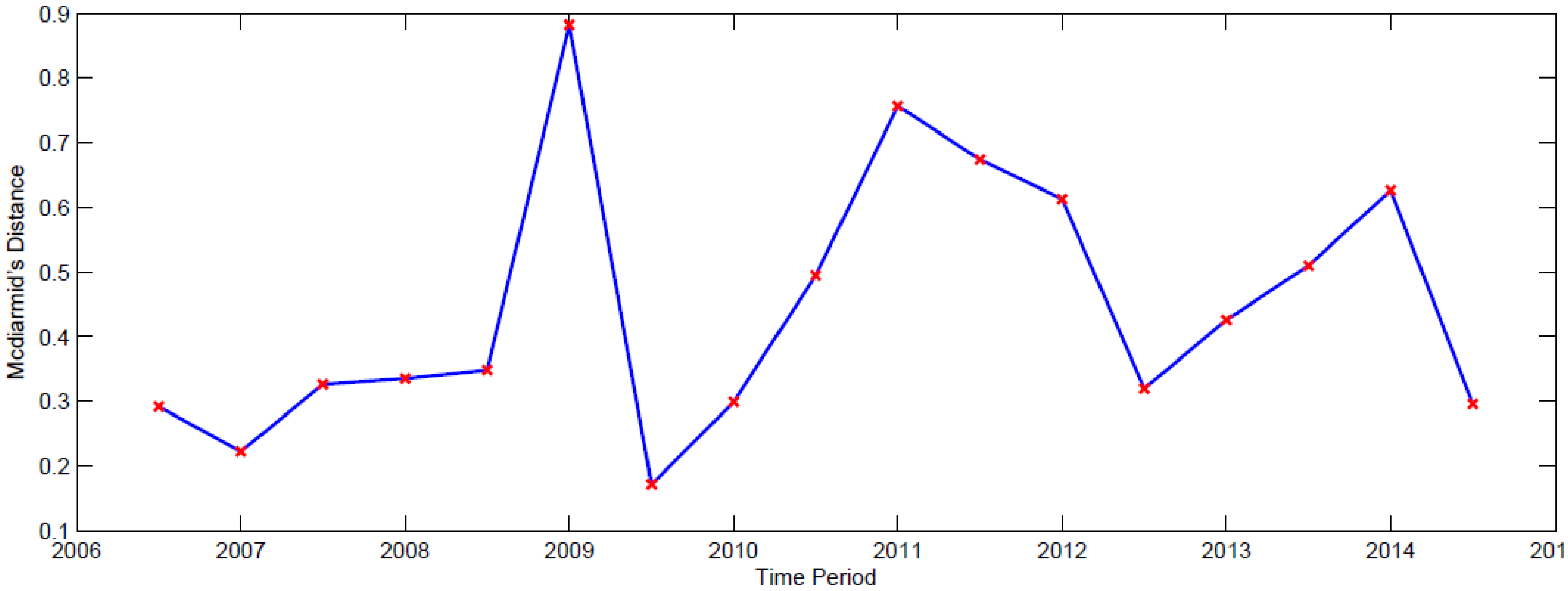
But this a limited case study

- Static stress test, no policy response or human factors
- Simple long-only portfolio, no optionality
- Exploited a well-understood principal component analysis decomposition

Results

Result #2 – Formal measure of macroeconomic uncertainty

- **McDiarmid's distance extracted from yield curve**
- **Minimal assumptions required**
- **Significant intertemporal variation**
- **Peaks in 2009 (just when certification would be most valuable…)**



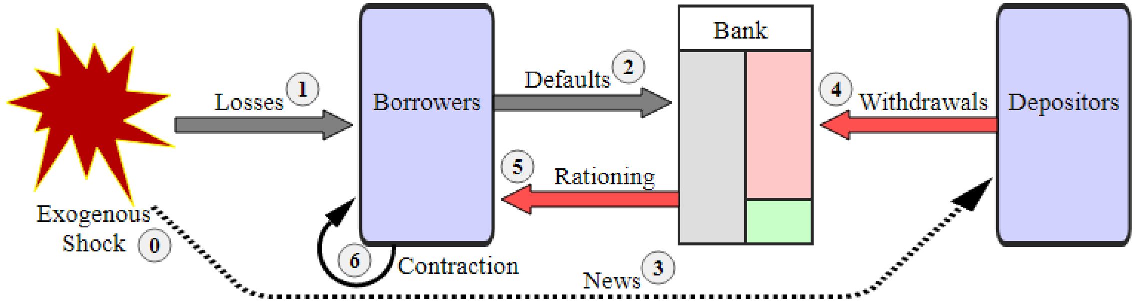
Heterogeneity in macroprudential stress testing

What is a supervisory stress-test and what are its goals?

- **Stress tests of individual FIs in isolation are microprudential stress tests**
- **Microprudential tests examine an FI's viability**
  - **In several dimensions (capital, liquidity, etc.)**
  - **When the FI faces several general stress scenarios**
  - **And for institution-specific scenarios for the FI's vulnerabilities**
- **A macroprudential stress test accounts explicitly for the systemic aspect and connection to the rest of the economy**
- **The macroprudential approach focuses on stability of the whole system**

Heterogeneity in macroprudential stress tests

A microprudential stress test considers a firm in isolation:



Fallacy of composition:

- **Each individual FI (or sector) is robust to a shock**
- **Together, the FIs compose the full financial system**
- **Therefore, the full system is also robust to the shock too (*right*?)**
- **Except – interactions among the FIs matter too!**

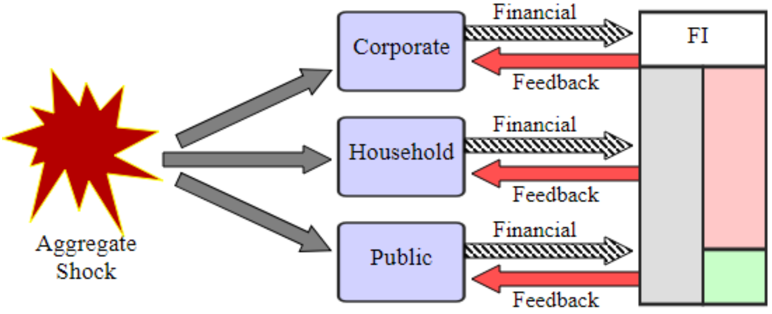
Responding to the fallacy:

- **Common scenario(s) for all FIs simultaneously**
- **General equilibrium approach, with feedback and propagation**

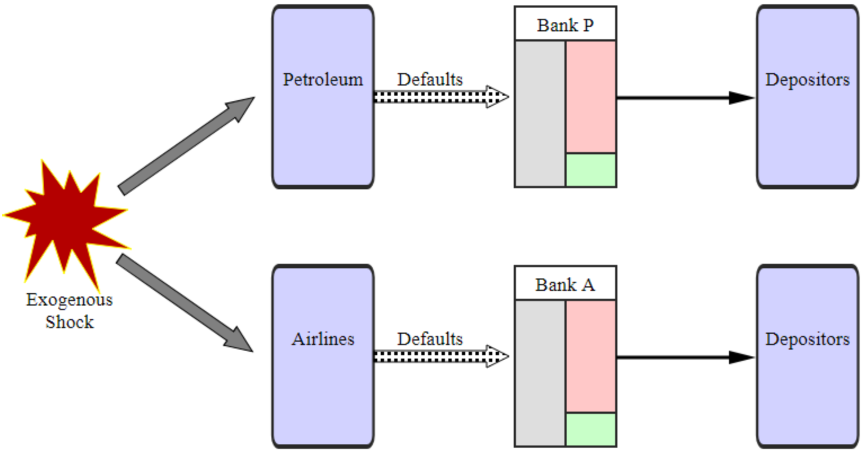
Heterogeneity in macroprudential stress tests

Macroprudential stress tests must consider

- **Multiple transmission channels**



- **Multiple firms/sectors**



Importance of modeling heterogeneity

Example

- 200 Banks
- Systemic Risk Objective (*SRO*):
  - Probability < 5% that more than 10% of banks default

Microprudential approach achieves the macroprudential objective

- Case 1
  - Highly heterogeneous banks – bank defaults independent
  - *SRO* achieved if each bank is capitalized so P(default) = .07155
- Case 2
  - No heterogeneity – bank defaults perfectly correlated
  - *SRO* achieved if each bank is capitalized so P(default) = .05000

Importance of modeling heterogeneity

Case 3 - Moderate heterogeneity:

- **Groups of FIs have similar risks**
  - **100 FIs lend primarily to airlines (default if oil prices are high)**
  - **100 FIs lend primarily to oil companies (default if oil prices are low)**

- ***SRO* <u>not</u> achieved if each FI is capitalized as in Case 1 or Case 2**

$$\text{Num. defaults} = \begin{cases} 100 & \text{Prob} = .07155 & P_{oil}\ \text{high} \\ 100 & \text{Prob} = .07155 & P_{oil}\ \text{low} \\ 0 & \text{Prob} = 1\text{–}2(.07155) & P_{oil}\ \text{moderate} \end{cases}$$

- **Instead, capitalize FIs so *P*(default) = .025 for high and low oil prices**

Lessons

1. **Must account for heterogeneity to achieve the *SRO***
2. **Multiple scenarios may be needed to achieve the objective**
3. **Extension should address hedging, feedback, and counterparty risk**

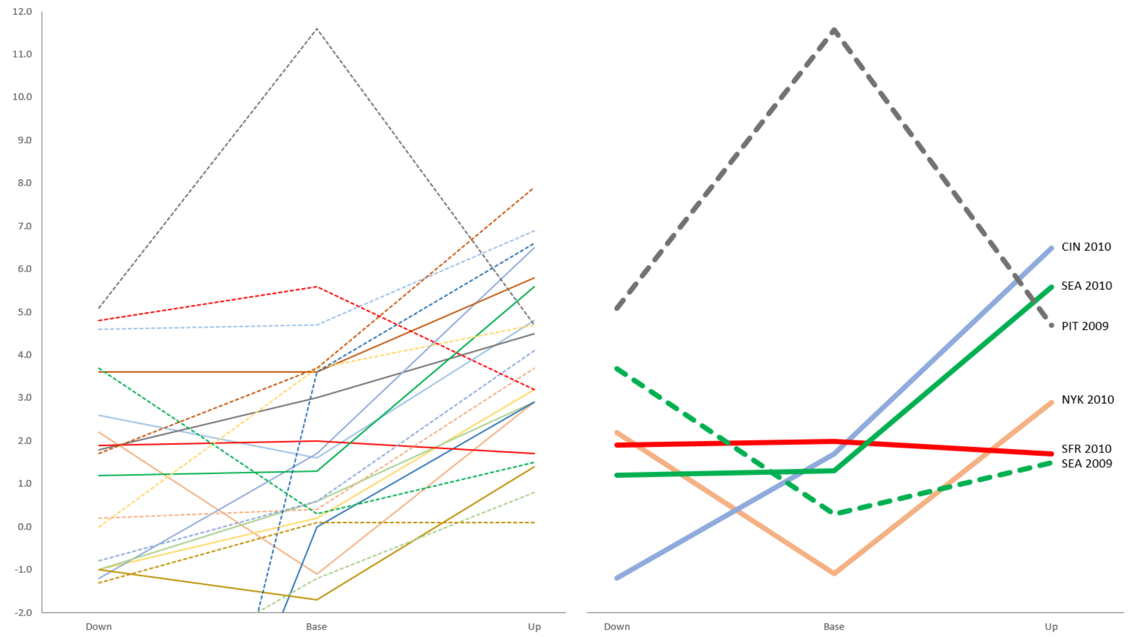
Heterogeneity of stress responses

Example – Diverse portfolio responses to interest-rate shocks

- Federal Home Loan Banks – identical mission: liquidity for mortgage lenders
  - 12 institutions, regional scope
  - 2009 and 2010
- Duration of equity = $(D_A - D_L) / V_e$
- Three parallel yield-curve shocks:
  - – 200 bp (but ZLB)
  - Base case
  - + 200bp

Mostly upward-sloping

- Except …
  - Seattle 2010
  - Pittsburgh 2009
  - New York 2010
  - San Francisco 2010
  - Seattle 2009



A Game of Battleship

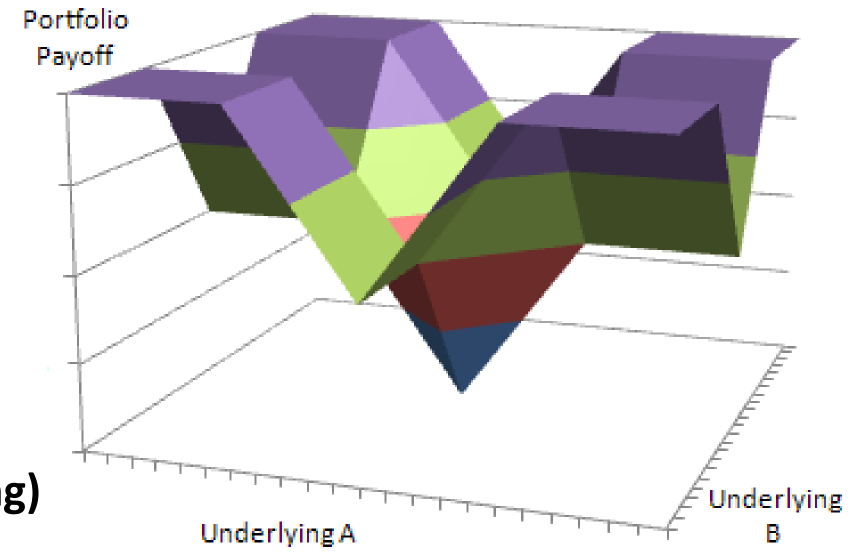
Forward stress test – McNeil and Smith (2010)

$$x_{LSLE} \equiv \arg\min \{g(x) : x \in S\} \text{ for } S \subset \Re^d$$

- where *LSLE* = least solvent likely event (i.e., among $x \in S$)

CCAR / DFAST has three “likely events” (scenarios):

- Baseline
- Adverse
- Severely adverse



Is three enough?

- Non-monotonicity of payoffs
- Anisotropy of payoffs
- Model risk
- Data limitations
- Strategic behavior (e.g., window dressing)

Inverting the question

Reverse stress test – McNeil and Smith (2010), again

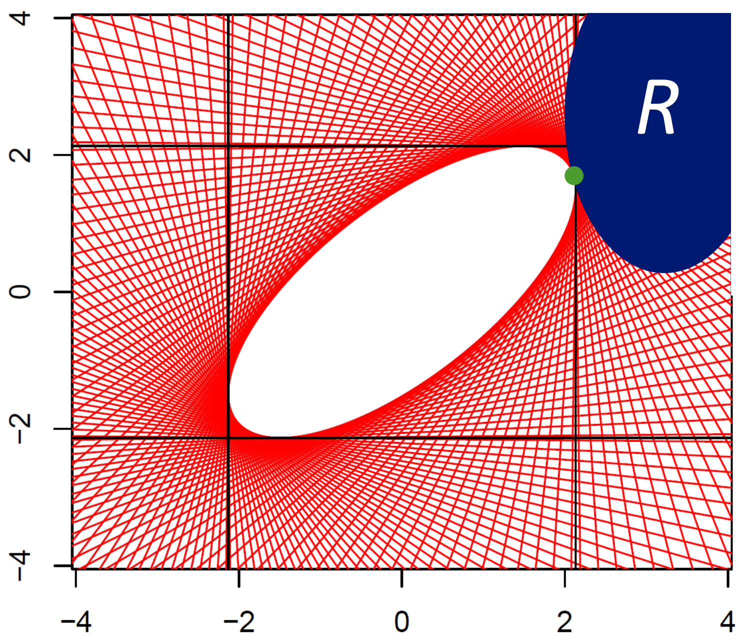
$$x_{MLRE} \equiv \arg\max \{\text{depth}(x) : x \in R\} \quad \text{for } R \subset \Re^d$$

- where *MLRE* = most likely ruin event (i.e., among $x \in R$)

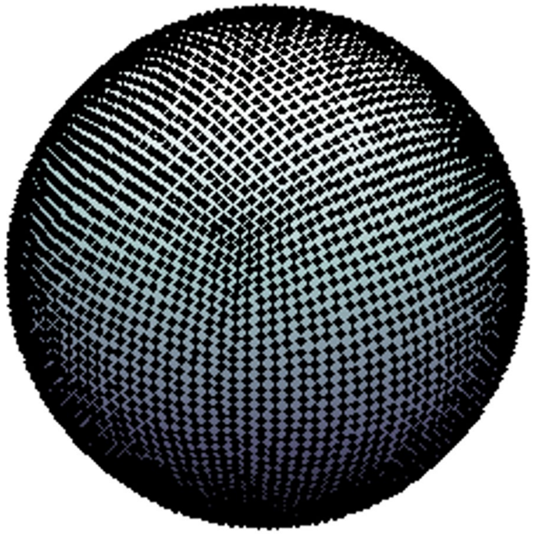
Finding the portfolio "hot spots"

Identify the set of ruin events, *R*

- **Pick the most likely $x \in R$**
- **Payoff surface involved directly**
- **Idiosyncratic scenarios**
  - **Helps reveal cross-sectional exposure concentrations**
  - **Challenge for public disclosure and accountability**



Applied reverse stress testing



Many dimensions of heterogeneity

- Portfolio exposures (a.k.a. “business lines”)
- Transmission channels
    - Feedback
    - Propagation
- Diverse scenarios
    - Including behavioral challenges

Scenario design approach

- Grid search to find the hot spots
    - Arbitrary number of scenarios to cover possible “hot spots”
    - Focus on macroprudential hot spots
- Capitalize to minimize systemic risk

Reading Suggestions

- V. Acharya, R. Engle and D. Pierret (2014), "Testing macroprudential stress tests: The risk of regulatory risk weights," *J. of Monetary Economics*, 65, 36-53.
- C. Aymanns, J. D. Farmer, A. Kleinnijenhuis and T. Wetzer (2018), "Models of Financial Stability and Their Application in Stress Tests," Ch. 6 in Hommes and LeBaron (eds.), *Handbook of Computational Economics, Vol. 4: Heterogeneous Agent Modeling*, North-Holland, 329-391.
- R. Bookstaber, J. Cetina, G. Feldberg, M. Flood and Glasserman (2014), "Stress Tests to Promote Financial Stability," *J. of Risk Management in Financial Institutions*, 7(1), 16-25.
- J. Chen & M. Flood and R. Sowers (2017), "Measuring the unmeasurable: an application of uncertainty quantification to Treasury bond portfolios," *Quantitative Finance*, 17(10), 1491-1507.
- M. Flannery, B. Hirtle and A. Kovner (2017), "Evaluating the information in the Federal Reserve stress tests," *J. of Financial Intermediation*, 29, 1-18.
- M. Flood, J. Jones, M. Pritsker and A. Siddique (2019), "The Role of Heterogeneity in Macroprudential Stress Testing," in: Farmer, Kleinnijenhuis, Schuermann and Wetzer (eds.), *Handbook of Financial Stress Testing*, forthcoming.
- M. Flood and G. Korenko (2015), "Systematic scenario selection: stress testing and the nature of uncertainty," *Quantitative Finance*, 15(1), 43-59.
- B. Hirtle and A. Lehnert (2015), "Supervisory Stress Tests," *Annual Review of Financial Economics*, 7, 339-355.
- P. Kapinos, C. Martin and O. Mitnik (2015), "Stress testing banks: Whence and whither?" *J. of Financial Perspectives*, forthcoming.
- A. McNeil and A. Smith (2012), "Multivariate stress scenarios and solvency," *Insurance: Mathematics and Economics*, 50(3), 299-308.
- T. Schuermann (2014), "Stress testing banks," *International J. of Forecasting*, 30(3), 717-728.

Thanks!