Jon Danielsson

Does risk forecasting help macroprudential policy makers?

Conference Presentation

Does Risk Forecasting Help Macroprudential Policy Makers?

Jon Danielsson
Systemic Risk Centre
London School of Economics
http://www.RiskResearch.org
http://www.systemicrisk.ac.uk

March 2013

What is macro–prudential policy?

according to the FSB, BIS and IMF

policy that uses primarily prudential tools to limit systemic or system–wide financial risk, thereby limiting the incidence of disruptions in the provision of key financial services that can have serious consequences for the real economy, by

- dampening the build-up of financial imbalances and building defences that contain the speed and sharpness of subsequent downswings and their effects on the economy;
- identifying and addressing common exposures, risk concentrations, linkages and interdependencies that are sources of contagion and spillover risks that may jeopardise the functioning of the system as a whole.

Possible paths

1. Use risk forecasting methods to identify the build up of risk and implement corrective measures
   - Basel II/III is a good example
2. Use supervisory level data to identify vulnerabilities, and upon that implement corrective measures
3. Use methods to prevent vulnerabilities *before* they happen (like LTV, restrictions on capital inflows, etc.)
4. Recognizing that we know little about the nature of systemic risk or financial instability and focus attention on study

I will only focus on the first here

Risk and risk forecasting

- While risk is the probability of something going wrong
- In a financial context, risk forecasting is the application of *formal statistical methods* and *available data* to *predict future* probability of things going wrong
- I want to limit myself to a subset of this,
  - the most common market risk forecasting methods
  - systematic risk forecasting

Key questions

1. Are fancy methods just based on plain old VaR?
   - Is the underlying model reliable?
2. Does every indicator flash at the same time?
3. Is 2007/ 2008 really the right benchmark?
4. Are crises sufficiently similar to be amenable to statistical analysis?
5. How to evaluate the risk forecasts (taking into account the intended application)
6. Paralysis by analysis
7. Is the probability appropriate?
8. What is the objective of micro–prudential policy founded on risk forecasts?
9. Can a formal risk forecast beat the Financial Times?

Some approaches

- Public data
  - Accounting and balance sheet information
  - Interbank and other network linkages
  - Exposure
  - Financial market data
    - equity markets
    - bonds and CDS
- Private data
  - Like supervisory level data

We start with the public

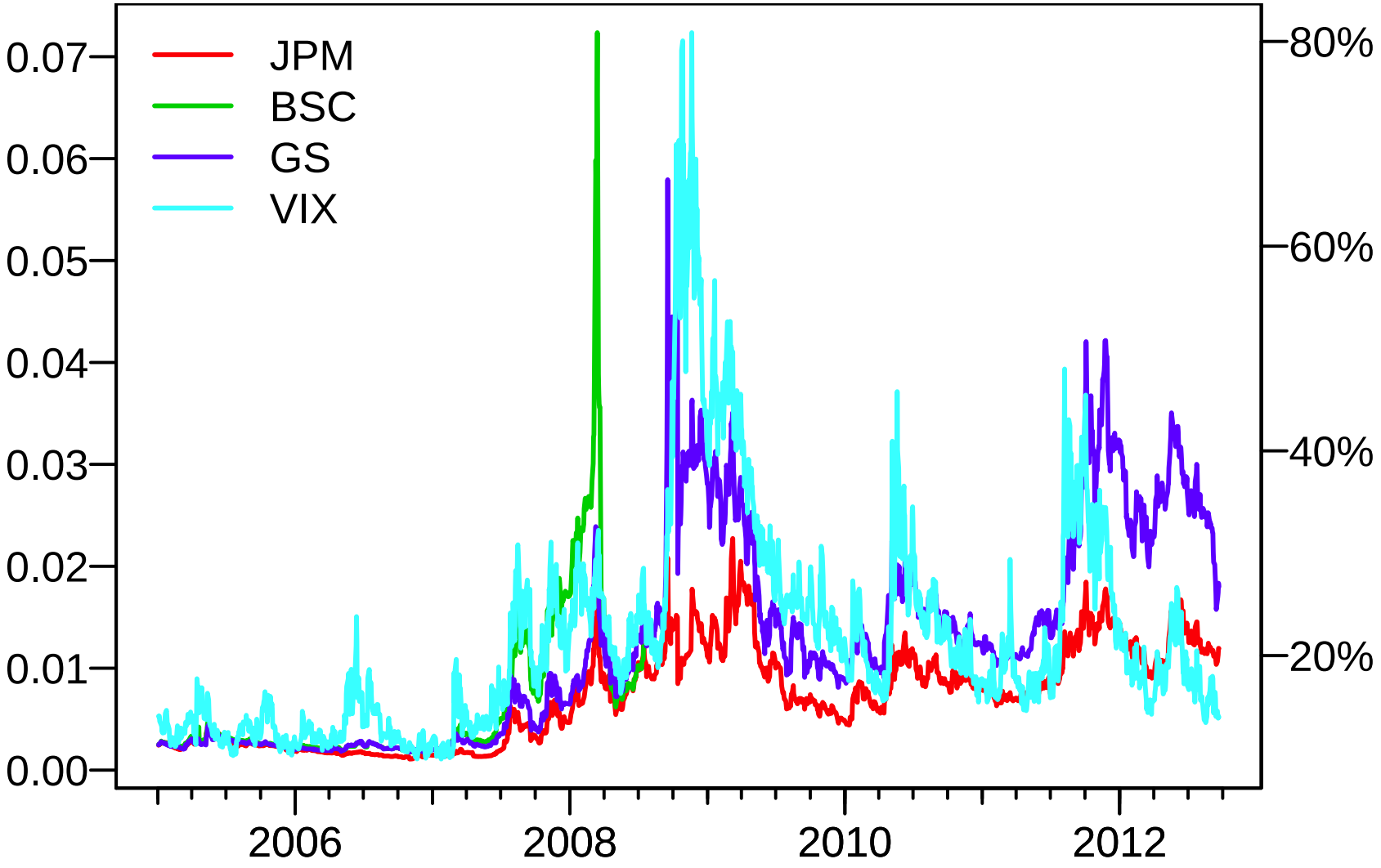
The ECB list

Deviations of credit-to-GDP ratios and (real) house prices from trend, credit growth, loan growth and customer deposits; term and credit spreads, intra-financial credit and credit risk conditions; equity valuations, stock returns and real equity growth; bank efficiency scores, contagion effects and leverage; asset price misalignments related to market sentiment; terms of trade and current account deficit; and measures of fiscal vulnerability. Global credit volume and global credit growth, global GDP growth, global leverage, real equity growth and equity valuations, as well as commodity prices. The decoupling of financial firms' credit risk conditions from the macroeconomic and financial variables that usually explain them.

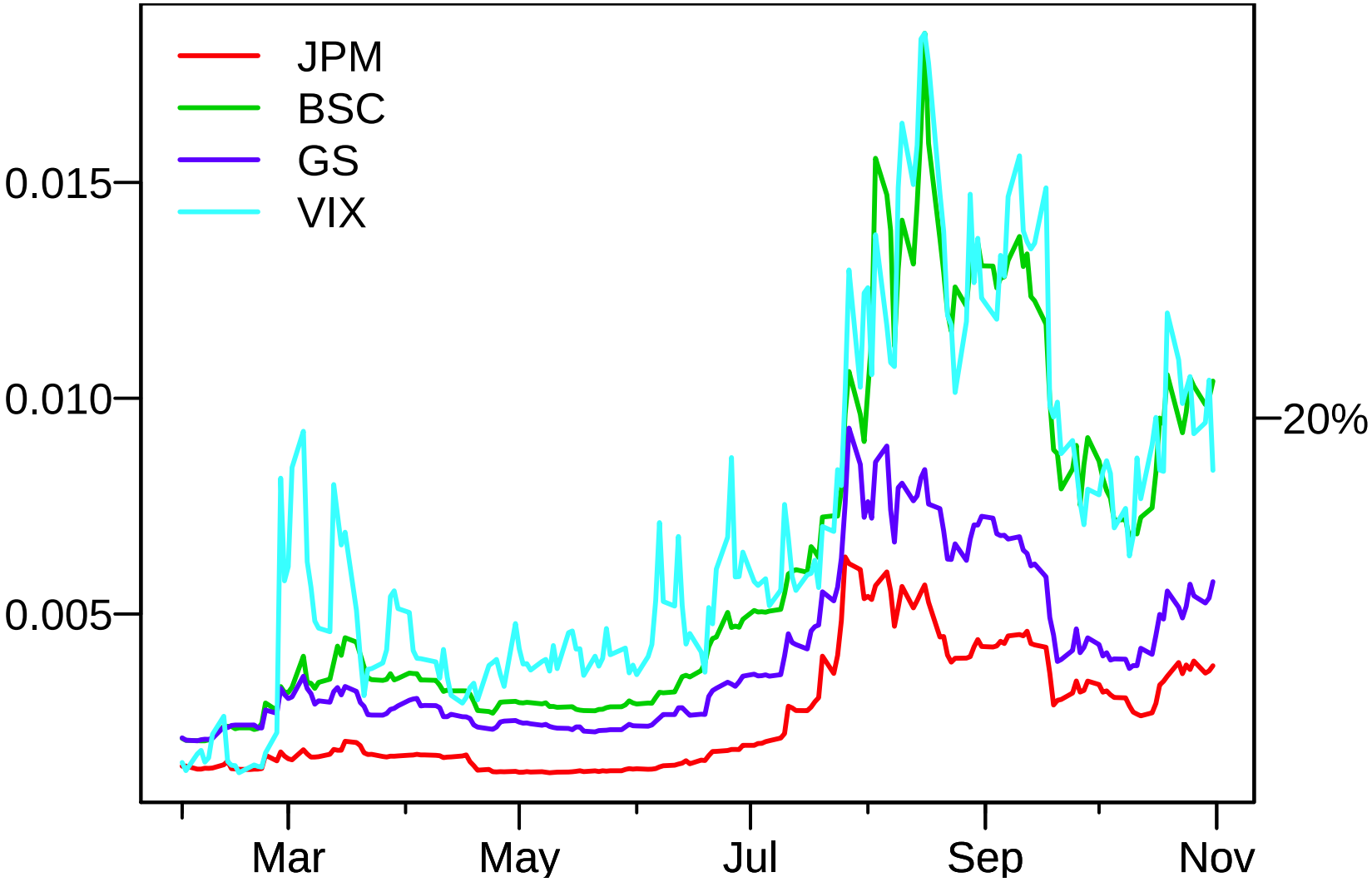
Does every indicator seems to flash at the same time?

- Many authors have proposed indicators/predictors for 2007
- The most obvious are CDS spreads and VIX (next slide)
- The question is, do all the indicators/predictors signal at the same time
- Or can we find one that leads in a *statistically significant* sense
  - taking into account all the *data mining* (hindsight bias)
  - if you try to 20 variables, one will predict significantly at 5%

CDS & VIX


JPM
BSC
GS
VIX
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
80%
60%
40%
20%
2006
2008
2010
2012

CDS & VIX 2007



ft.com/frontpage UK All times are London time

FINANCIAL TIMES

ft.com/frontpage UK All times are London time

FINANCIAL TIMES

Buyers shun Bear Stearns' fire sale

Posted by **Gwen Robinson** on Jul 04 05:10.

ft.com/frontpage UK All times are London time

FINANCIAL TIMES

Buyers shun Bear Stearns' fire sale

Posted by **Gwen Robinson** on Jul 04 05:10.

July 6, 2007

Credit crisis to worsen as banks cut and run

The fallout from the crisis at two Bear Stearns hedge funds is

ft.com/frontpage UK All times are London time

FINANCIAL TIMES

Buyers shun Bear Stearns' fire sale

Posted by **Gwen Robinson** on Jul 04 05:10.

July 6, 2007

Credit crisis to worsen as banks cut and run

FT ALPHAVILLE July 10, 2007

Remember that subprime crisis? Well, it's back

Is 2007/ 2008 really the right benchmark?

- We don't know the nature of the next crisis, but it will be different from this one
- Excessively calibrating models and analysis to 2007/2008 not advisable

How to evaluate the risk forecasts

(taking into account the intended application)

- Basel I/II/II.V/III (and surely IV, V, etc.) mandate backtesting of risk forecasts
- It is only prudent to make the same demand of risk forecasts used for macro–prudential purposes
- This has to explicitly take into account the low probability events and the nature of crisis
- A risk forecast without a proper backtest is like religion, we have to take it on faith
  - Finding that method predicts 2007 is *not nearly enough*

Paralysis by analysis

my skiing instructor once told me that if once on the top of a mountain, one rationally analyzes a decision to ski down, one ends up taking the lift — paralysis by analysis

- Given the fact that there are dozens, hundreds of potential predictors/indicators for macro prudential policy makers
- one of them will at any given time *flash red*
- Therefore, any indicator/predictor needs to be carefully chosen
- without data mining or hindsight bias

Is the probability appropriate?

- Basel I/II/III VaR is 99% daily
  - 2.5 events per year
- Most proposals use similar probabilities
- But crisis are infrequent ($< {}^{1}/_{2500}$ daily)
- So any 99% daily method says nothing about crises probabilities, or the buildup of vulnerabilities
  - Probability shifting (translating from one probability to another) does not work
- One could ask what is the point of the 99% daily in Basel
- (I discuss EVT later)

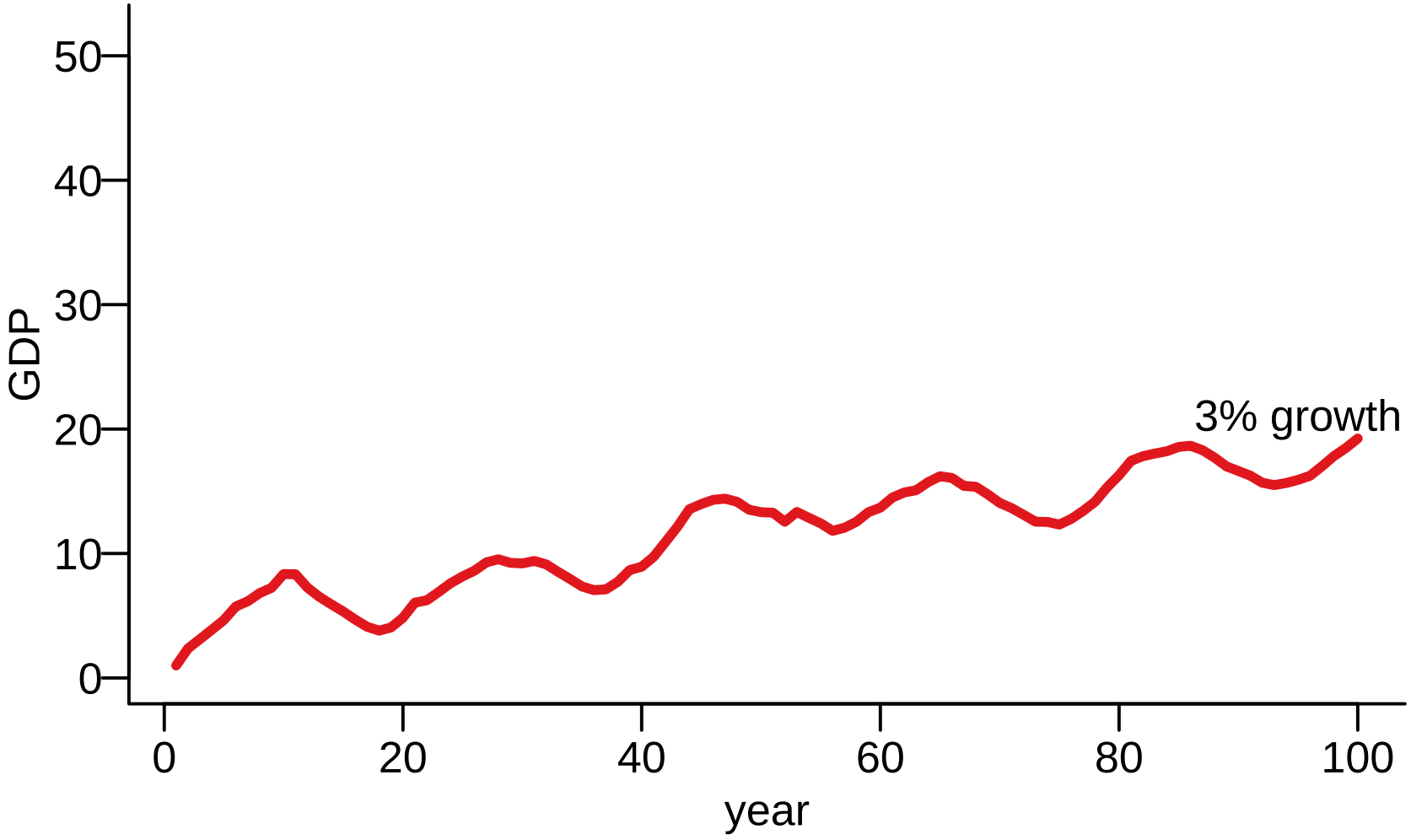
Are crises sufficiently similar to be amenable to statistical analysis?

- By a first approximation, we observe roughly one event of high stock market turmoil per decade
- By a first approximation, individual countries observe one crisis per generation
- These tend not to be very similar
- Therefore, it is hard to the point of impossible to rely on statistical analysis of past crises for what might happen in future crises

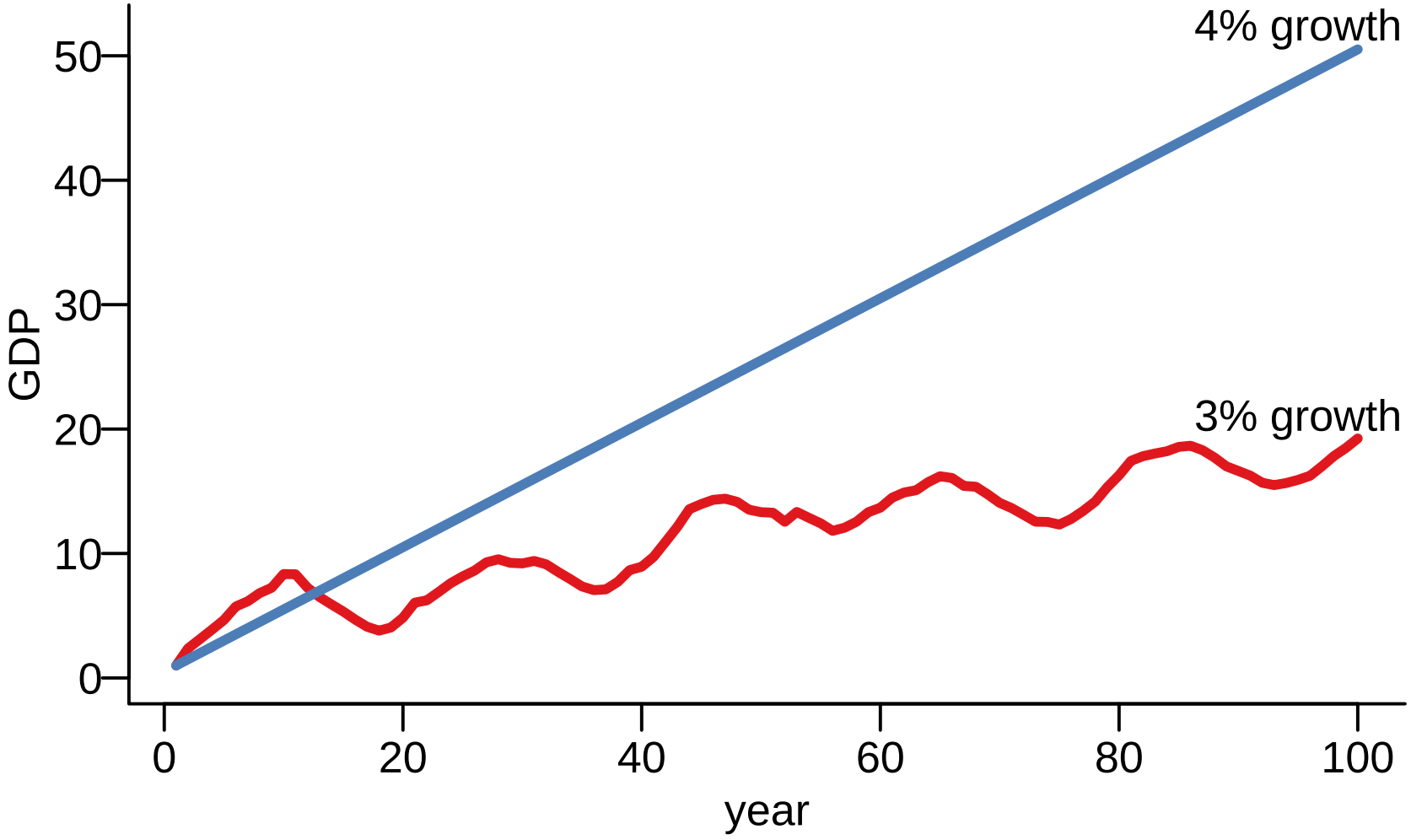
What is the exact objective of micro–prudential policy founded on risk forecasts?

- Understand the vulnerabilities in the system
- Smooth the road
  - Prevent crises altogether
  - Prevent crises that exceed certain severity threshold
- Which leaves the question, is the objective desirable and achievable?

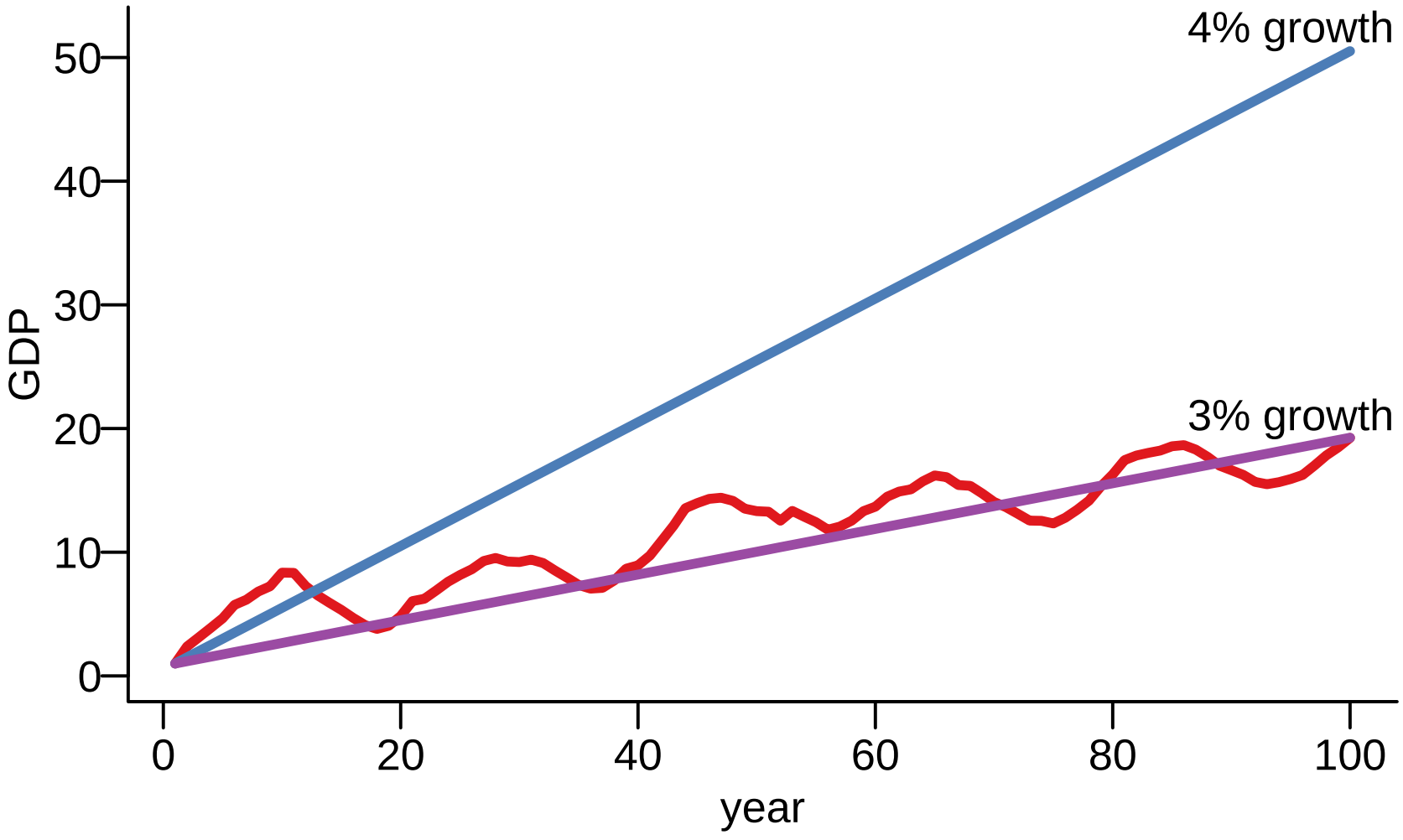
GDP over a century



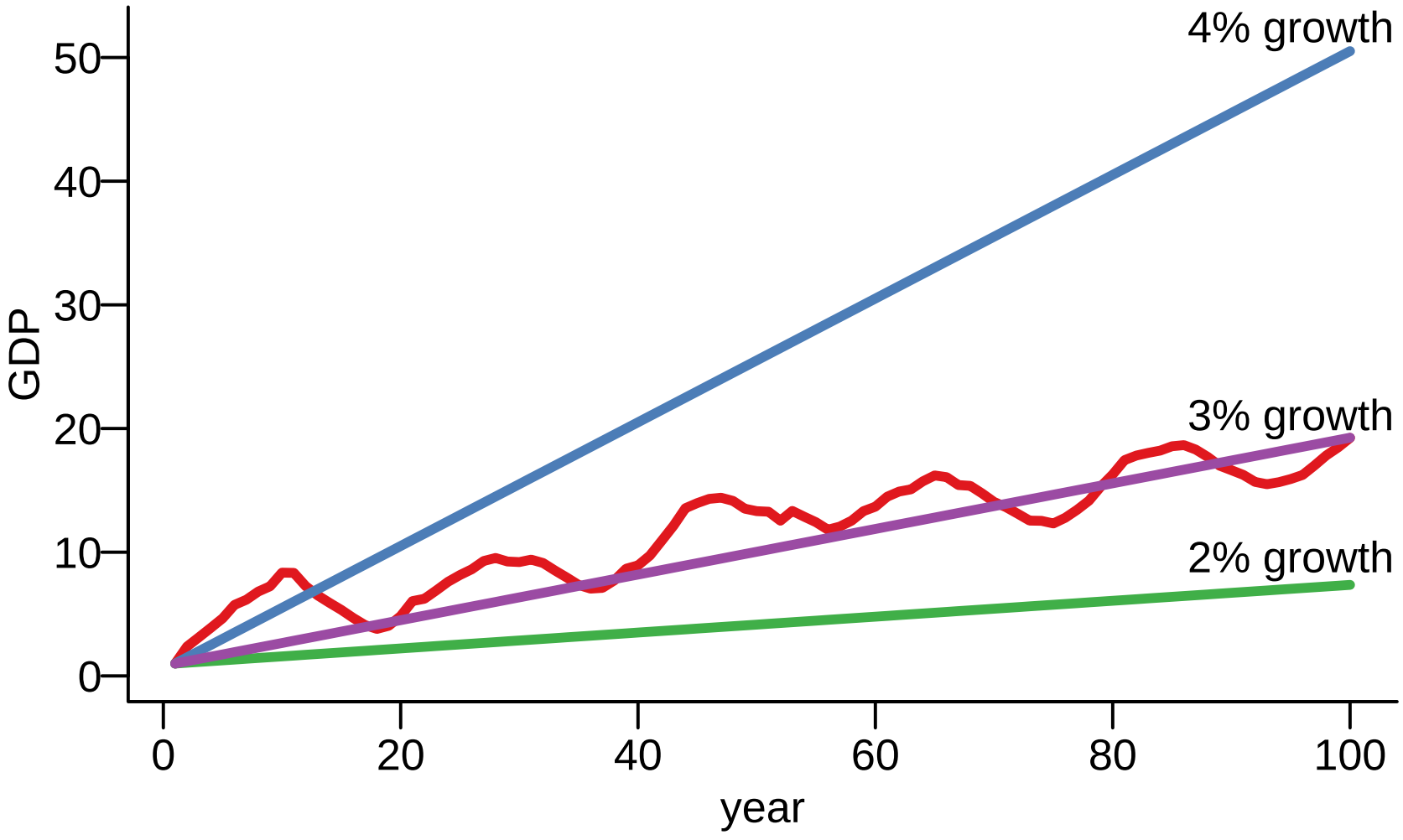
GDP over a century



GDP over a century



GDP over a century



Basic problem

- Supposedly a successful in smoothing out the cycles (at least the most extreme shocks)
- That increases appetite for risk
- And endogenously increases risk taking
- Thereby undermining the smoothness success

Forest fires in the US

- Forest fires are endemic in the Southwest US
- Historically, they would flare every few years, burn the undergrowth, but spare the big trees
- Fires were frequent and small
- Then people moved into forests — all fires fought
- Successful for a generation
- Then when a fire starts, there is so much dry undergrowth that fire becomes out of control and burns the big trees
- The US fire authorities opted for what was in effect lowering volatility and fattening the tails

Macro prudential policy founded on extant market risk methodologies

some might say they failed before the crisis, but...

Systemic risk from market risk models

- R_i is risky outcomes of institution i
- R_S is outcomes from the entire financial system
- Joint distribution is:

$$f\left(R_i, R_S\right)$$

- *Marginal* density is $f(R_i)$, and the two *conditional* densities are $f\left(R_i|R_S\right)$ and $f\left(R_S|R_i\right)$
- *VaR*, (where Q is a quantile)

$$\text{pr}[R_i \leq Q_i] = p$$

Common measures

these things are much more similar than often maintained

Marginal risk measure	Condition on system	Condition on institution
	MVaR	*CoVaR*
VaR	$\text{pr}[R_i \leq Q_i \mid R_S \leq Q_S] = p$	$\text{pr}[R_S \leq Q_S \mid R_i \leq Q_i] = p$
	MES	CoES
ES	$\text{E}[R_i \mid R_S \leq Q_S]$	$\text{E}[R_S \mid R_i \leq Q_i]$

- Other measures like Shapley fit into this
- *All depend on daily VaR* and have more model risk than VaR

Experiment

From one of my papers "Model Risk of Systemic Risk Models"

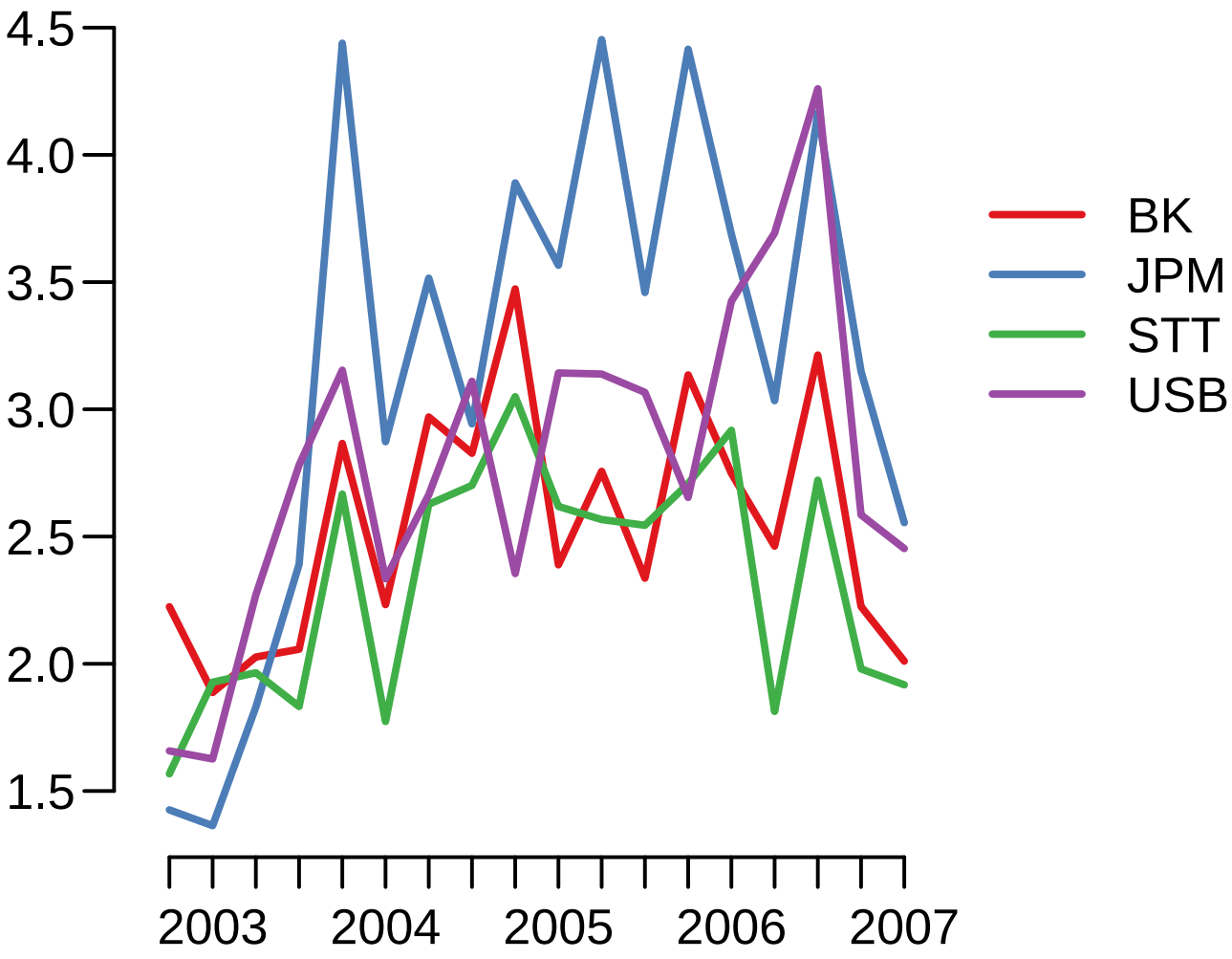
- Daily total returns January 1997–December 2010
- 92 largest US financial institutions
- 99% daily VaR with the most widely used state–of–the–art methods
  - HS, MW, student-t MW, EWMA, GARCH and student-t GARCH
  - Range of estimation windows (500, 1000, 1500)
  - \$1000 portfolio
  - End of quarter results
  - Probability is 1%

- The easiest risk forecast scenario possible
- *If this looks bad, surely everything else will be worse*

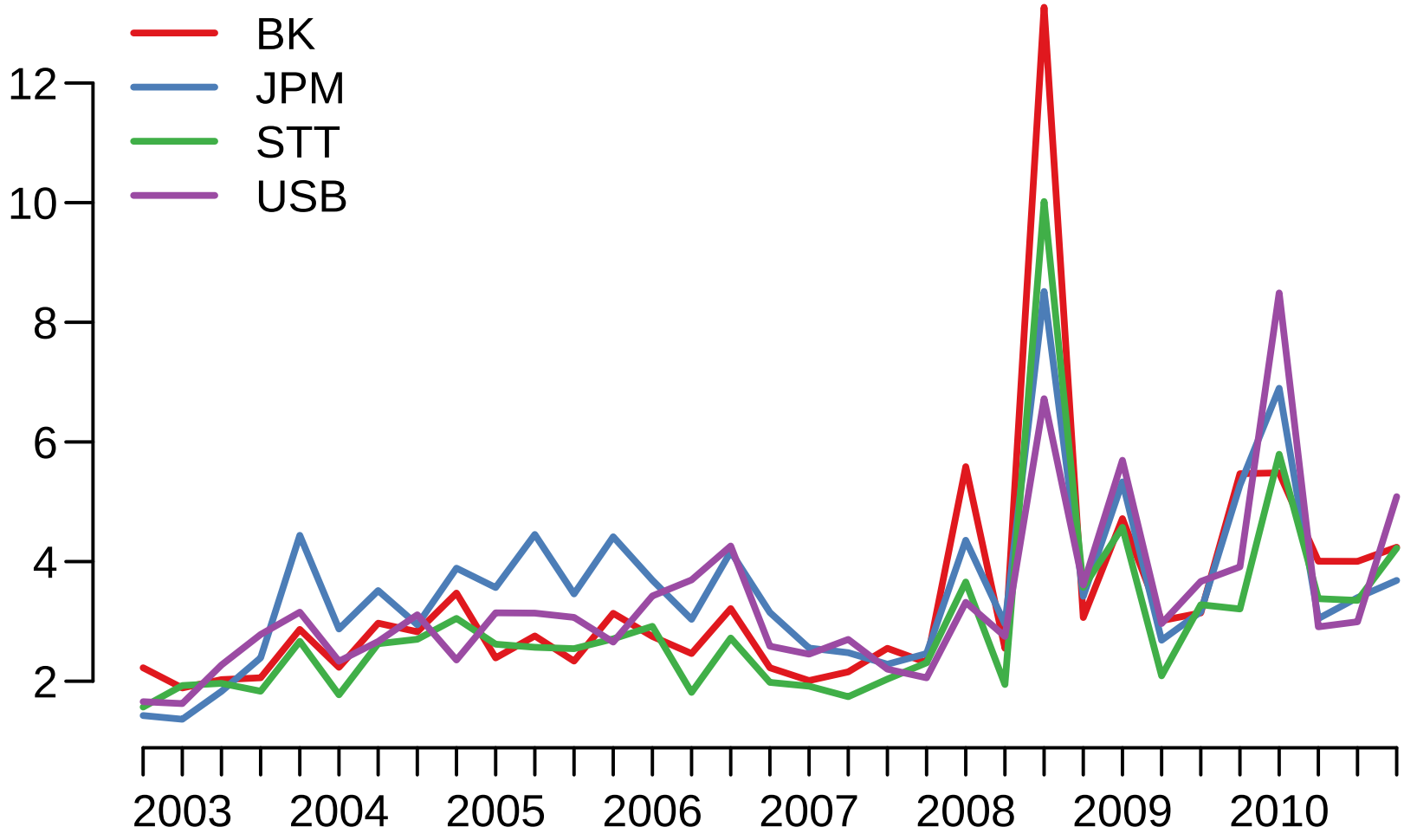
Q4 2002

Bank	$\frac{\text{max VaR}}{\text{min VaR}}$	Method at min	Method at max
BK	2.22	MA	t–GARCH
JPM	1.43	MA	t–GARCH
STT	1.57	GARCH	t–MA
USB	1.66	EWMA	t–GARCH

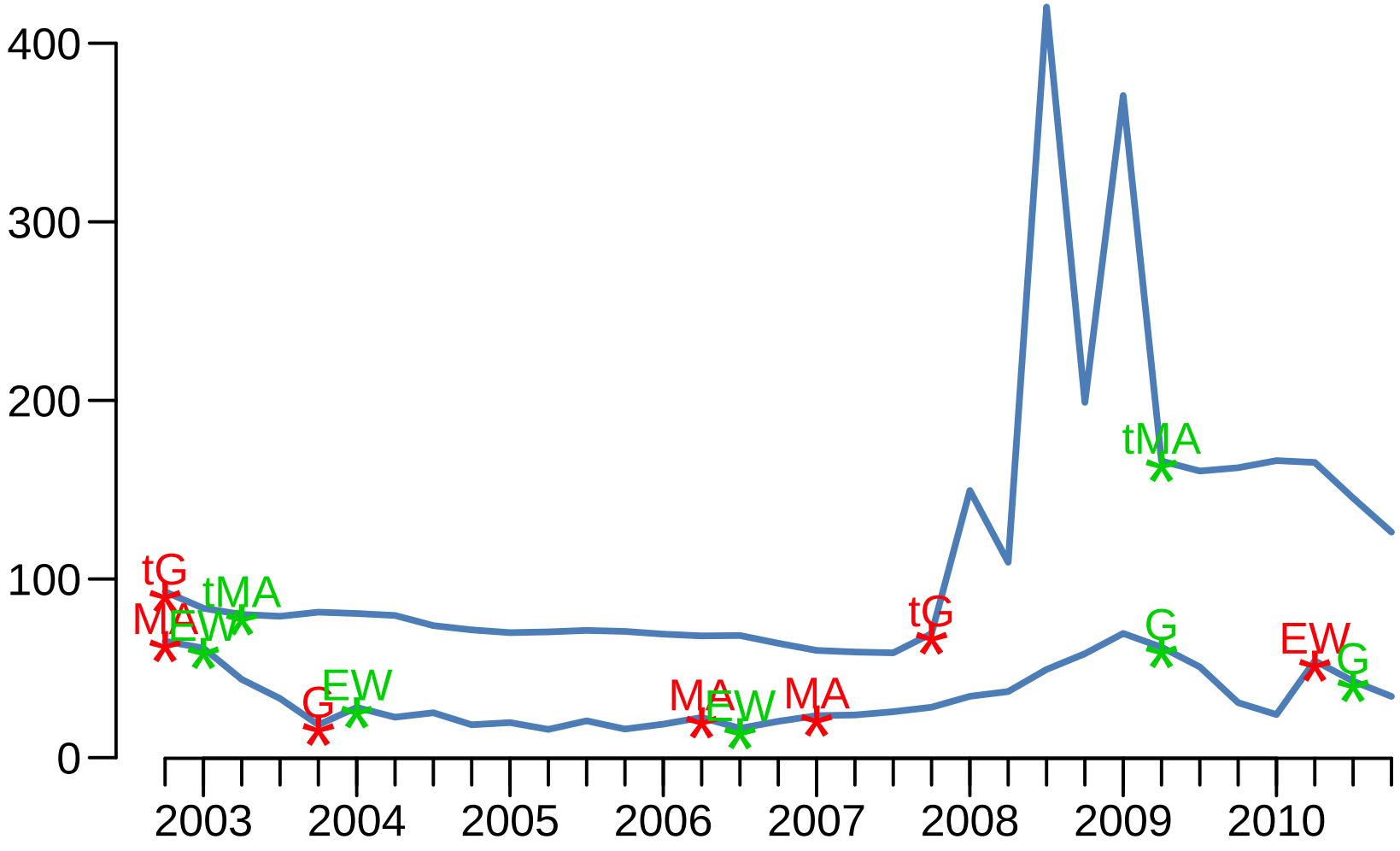
Ratio of highest to lowest



Ratio of highest to lowest



JP Morgan highest and lowest VaR



Results

- Best case scenario
- Only model risk, not parameter risk
- No multivariate issues clouding the results
- With 3 different industry–standard estimation methodologies can make

VaR=\$25 or VaR=\$60 or VaR=\$160

- We can make VaR be anything

More realistically

The problem gets worse with

- Aggregation
- Smaller markets
- Manipulation
- More model driven pricing
- More mechanistic application of models

Two explanations

1. Failure to comply with the underlying statistical assumptions
    - I am not the only who is non–compliant, nobody is
2. Endogenous risk

Statistical assumptions

Most risk models are based on snapshot measurements (maybe sequence of daily observations)

1. One stochastic model covering all states of the world
2. Modeller has never seen data, and runs model once
3. Backtester
   - has never seen data
   - is fully independent of modeller
   - does not provide any feedback to model design

and these are like always obeyed

Snooping

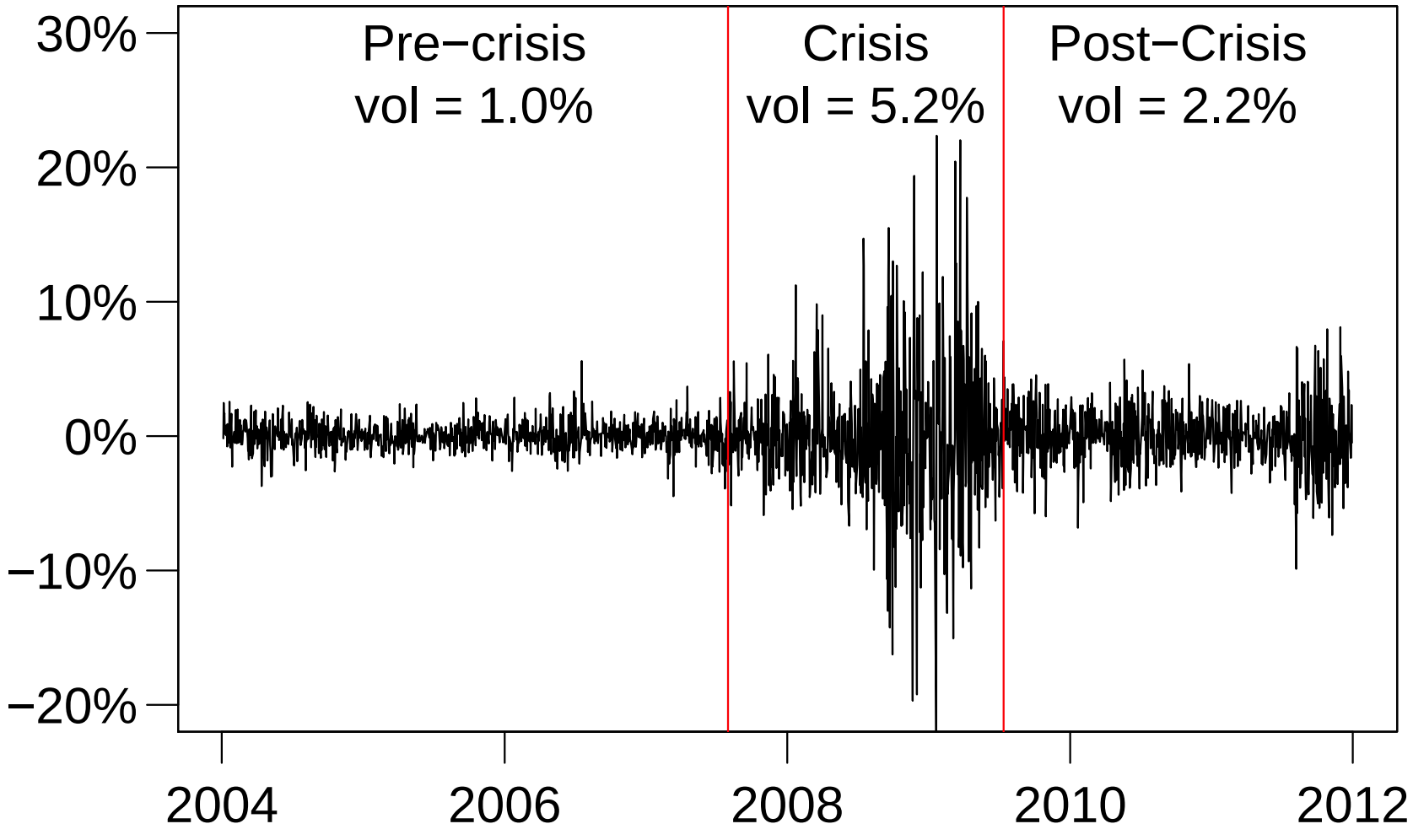
- If we know what data looks like
  - (e.g. how a model performs in back testing)
- We will cheat (unless a saint, and even then)
- It is inevitable that modellers *know the data*
- And inevitable that this will *bias* their analysis
- This means that the actual confidence bounds on risk model outputs are much larger than indicated
  - (in the few cases anybody pays attention to confidence)

Backtesting

- Statistically analyze violations of risk forecasts
  - occurrence, clustering, magnitude, ...
  - Bernoulli hit sequences
- Small sample problems
- Very hard to identify clustering
- Very hard to analyze magnitudes
- Whilst at the same time the backtester is not the person who designed the model and does not feed back to the model design

Backtesting is unreliable

Three states of the world: JPM



Structural break modelling

- Most models use a single stochastic process to capture *all states of the world*
- And hence miss the structural breaks (like in JPM)
- Models generally do not entertain *structural breaks*
- Because it is almost impossible to do so
  - (yes I know Markov switching, and no, it doesn't work here)
  - the uncertainty around the magnitude and location of the switching point is too high
- So we have to assume that the JPM regimes are really one

Immediate conclusion

- State–of–the–art market risk models are highly inaccurate
- Just by tweaking the model I can make VaR be \$100, \$200 or \$300, using only models accepted by the supervisors
- Model risk much higher during extreme turmoil
- Model risk much higher at systematically important probabilities
- To me this suggests market risk approaches are not all that reliable for macroprudential purposes

Nature of risk

Endogenous risks vs. Exogenous risks

- *Endogenous risk*: the risk from shocks that are generated and amplified *within* the system
- *Exogenous risk*: shocks that arrive from *outside* the system
- Analogies
  - A financial hedge (futures contract) vs. a weather hedge (umbrella)
  - Poker vs. Roulette
- Essentially situations where an agent affects outcomes
  vs.
  situations where the agent cannot

Assumptions behind almost every risk and pricing model known

- Risk is *exogenous* (we are strictly price takers)
- Market prices are the best reflection of value
- Today's price has most information

If true, consequently

- The best way to forecast risk (even prices) is to combine a *historical sample* of prices
  - With a model like EWMA, HS, IV, etc. ...
- Best to *down weigh history*
- Price dynamics in a crisis belong to the *same stochastic process* as price dynamics outside of crisis

and VaR works great

However . . .

- Risk is really *endogenous*
- Prices reflect *constraints* (margins, capital, politics, etc.)
- These effects are *stronger during crises*
- Forces driving prices and risk are different in a crisis than out of crisis
- The underlying economic process may be the same, but we are talking statistics

Risk in crises

The monster gets quietly stronger under the bed

- Estimating a model on *non–crisis* data is unlikely to say very much about risk *during* times of stress
- Very little data available
- Risk building up in quiet times

We can not get from the failure process in normal times to the failure process in crisis times

Minsky

- Minsky (1992) argued that economies have either *stable or unstable* financial regimes. Even if the economy starts out stable, continued prosperity paves the way for an unstable system
- *Stability is destabilizing* because financial institutions have a tendency to extrapolate stability into infinity, investing in ever more risky debt structures, followed an abrupt correction
- Like before 2007s when all were blind to the hidden risk during the "*great moderation*"

When risk is created

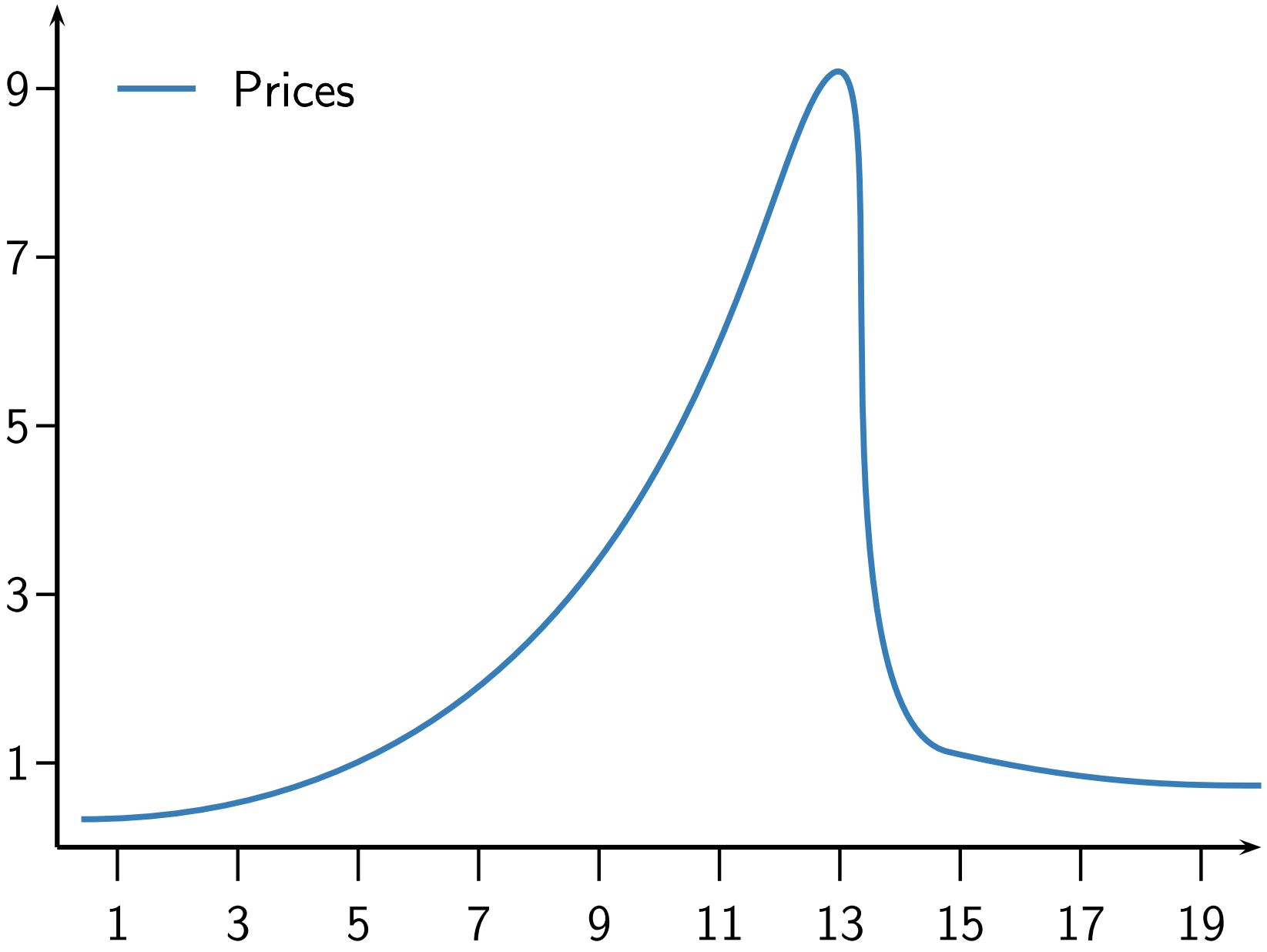
Former head of the BIS, Andrew Crockett in 2000

"The received wisdom is that risk increases in recessions and falls in booms. In contrast, it may be more helpful to think of risk as increasing during upswings, as financial imbalances build up, and materialising in recessions."

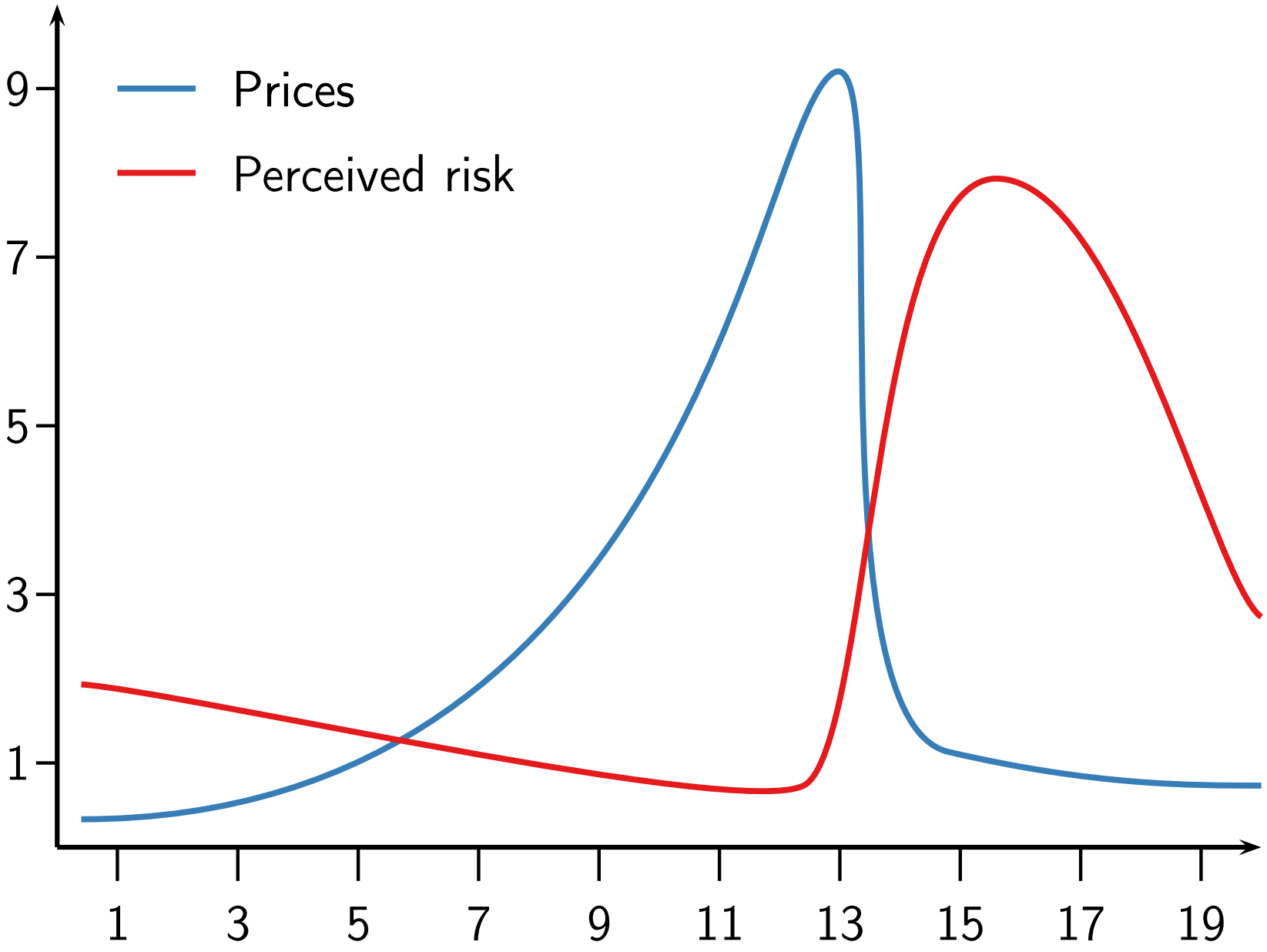
Actual and perceived risk

- The monster under the bed is actual risk
- Perceived risk is what the models tell us
- These two tend to be negatively correlated

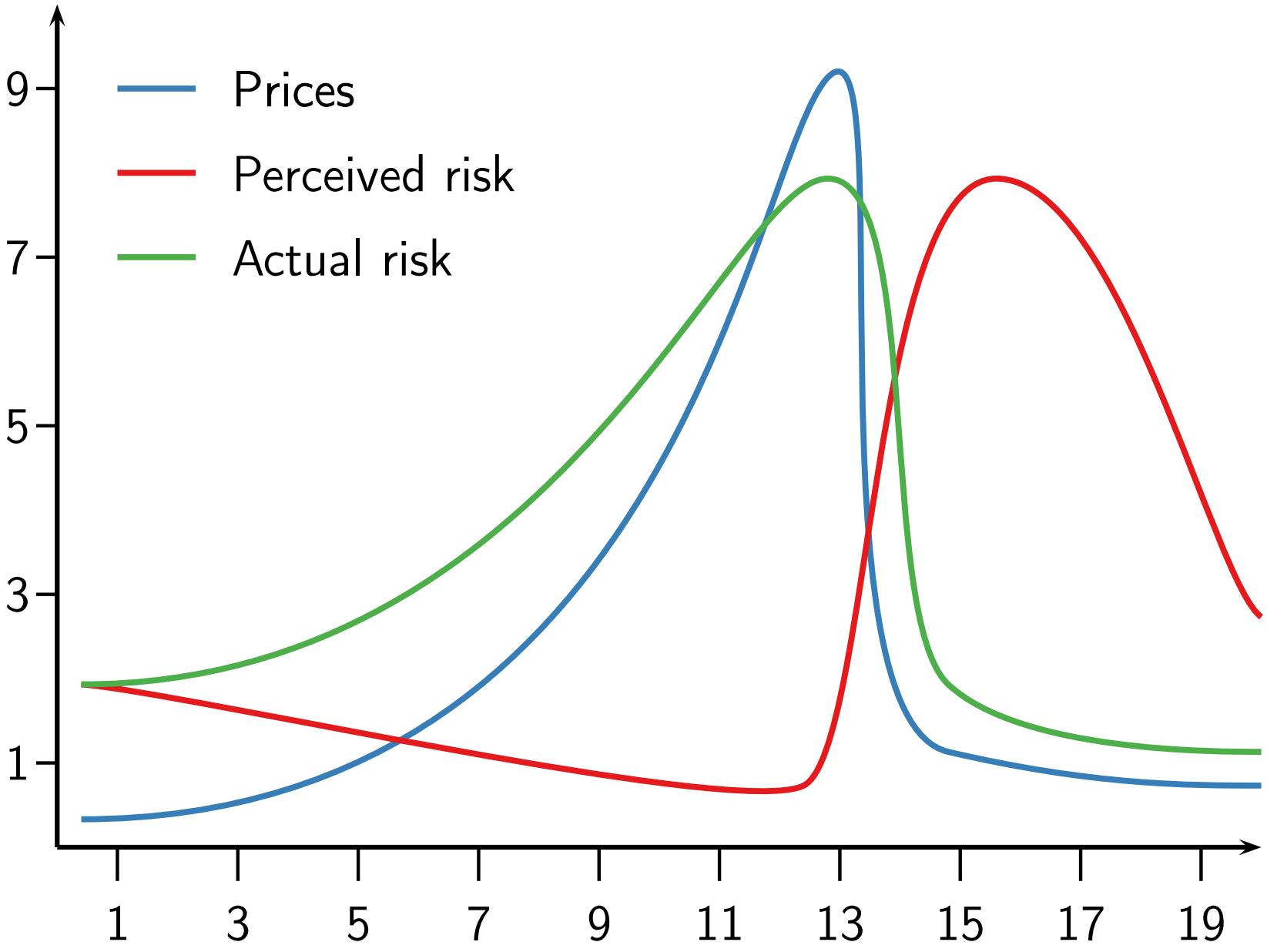
Endogenous bubble



Endogenous bubble



Endogenous bubble



Dual role of prices

- They are a passive reflection of the underlying economic fundamentals, an aggregation of all available information but on the other
- Also an imperative to action
- Implications (see next slide)

Role of prices

- Market prices during periods of calm are a poor input into forecast models
- They are not informative about the distribution of prices that follow after a crisis is triggered
- Price dynamics during one crisis may be quite different in the next, limiting the ability to draw inference from crisis events

Risk models underestimate risk during calm times and overestimate risk during crisis — they get it wrong in all states of the world

Systemic risk forecasting

- Market variables as indication of the risk of future systemic event
- Systemic risk is concerned with events that happen during crisis conditions, looking far into the tails of distributions
  - Little relevant data
  - Over the last fifty or so years we have observed less than a dozen episodes of extreme international market turmoil, all unique
  - Models that are fed with inputs from calm periods will perform much less well during periods of stress

What about tail methods

like extreme value theory (EVT)

- They are fundamentally based on a constant stochastic process
- The only way EVT can be reliable in this application is if
- Financial crisis are regular (and similar) events that can be modeled
- However, they are infrequent and unique
- Therefore, I don't think tail methods are much use in this case

Quality control

Quality control for systemic risk measures

1. Point forecasts are not sufficient: need *confidence intervals* incorporating both estimation risk and model risk
2. Data should be *predictive* and not reactive
3. Statistical method needs to include *backtesting*
4. Event probabilities need to correspond with the probability of systemic events
   - If such events happen once every 10 years, 99% probabilities (2.5 times a year) are of little relevance
   - One *can not* map failure probabilities from less extreme to more extreme. (estimate at 99% use for 99.9%)

Is a bad systemic risk measure better than none?

- Current systemic risk measures are quite bad, perhaps indistinguishable from random noise or at best weakly better in prediction
- High cost of using an incorrect method
- A bad systemic risk measure should not be acceptable for policy purposes, it should be of a proven quality
- *Type 2 errors are very costly* (falsely finding high sysrisk)
- Avoid the fallacy of requiring a number for decision–making regardless of the number quality

Private data
Supervisory level data

Problem with public market data

- It reflects expectations
- And this suggests every indicator flashes at the same time
- It however only reacts to perceived and not actual risk

Supervisory data

- Supervisory data is not affected by this problem as much
- And therefore might be used to develop reliable systemic risk indicators/predictors
- For example exposure data with named counterparties
- Can help with network models
- Or data on chances in trading activities

Stress tests for individual institutions

- I don't think stress tests applied to individual banks are very informative about systemic risk
  1. Impossible to assign probabilities to stress events and there there is an infinite number of potential stress events
  2. Endogenous risk

Industry stress tests

- Individual banks are not hit by stress and isolation
- They will react by trading into a general distressed environment
- Therefore interaction between banks matter
- This means that we would need that industrywide stress test

Does Risk Forecasting Help Macroprudential Policy Makers?

Does Risk Forecasting Help Macroprudential Policy Makers?

Not really

Origins of methods

- While many techniques are fancily dressed up
- Underneath is usually a standard market risk model
- These did not cover themselves in glory prior to 2007
- And the macro prudential problem expects more of these methods than most internal applications in banks
- Meanwhile, the quality control in the macro prudential space is lower than in the market risk space
- Therefore, I don't think such approaches can be of much use

Why

- The underlying models are highly unreliable
- Very hard to model actual risk
- Mistakes are very costly
  - Type 1
  - Type 2