Climate Change and Risk Management:
Micro-correlations, Fat Tails, and Tail Dependence

Carolyn Kousky[1] and Roger M. Cooke[2]
April 24, 2009

Abstract: Climate change may be altering the frequency and/or magnitude of extreme events, impacting loss distributions. We study public flood and crop insurance claims data from 1980-2008, identifying the presence of micro correlations, fat tailed losses, and tail dependence. These phenomena are easy to overlook, and yet, if unidentified, can undermine traditional risk management approaches.

Introduction

Adapting to climate change will not only require adapting to physical impacts, but also adapting the way we conceptualize, measure, and manage risks. Risk management can involve: *elimination, mitigation,* or *dissipation*. Risk elimination is no longer an option for climate change. Risk mitigation comprises proactive measures to reduce the probability of harmful consequences. Just as a shield dissipates the energy of a hostile blow, risk dissipation distributes risk over many stakeholders. It is commonly achieved through insurance, but also through insurance linked securities or government indemnification, for example.

Climate change presents us with an "Unholy Trinity" of new risk management perils: *micro-correlations*, *fat tails,* and *tail dependence*. These are distinct aspects of loss distributions that challenge our traditional approaches to managing risk. Micro-correlations are negligible correlations which may be individually harmless, but very dangerous in concert. Fat tails apply to losses whose probability declines slowly, relative to their severity. Tail dependence is the propensity of severe losses to happen together. Global changes in climate will be altering loss distributions by way of this trinity. If one does not know how to detect these phenomena, it is easy to overlook them. As the climate changes, it behooves us to start looking.

We draw on two datasets to explore the unholy trinity: flood insurance claims data from the National Flood Insurance Program (NFIP) and crop insurance indemnities paid data from the United States Department of Agriculture's Risk Management Agency.[i] Both datasets are aggregated by county and year for the years 1980 to 2008. The data are in constant year 2000 dollars. Over this time period there has been substantial growth in exposure to flood risk, particularly in coastal counties. To remove the effect of growing exposure, we divide the claims by personal income estimates from the Bureau of Economic Accounts (BEA).[ii] Thus, we study flood claims per dollar income, by county and year. The crop loss claims are not exposure adjusted, as a proxy for exposure is not obvious, and exposure growth is less of a concern.

[1] Resources for the Future, kousky@rff.org

[2] Resources for the Future and Dept. Mathematics, Delft University of Technology, cooke@rff.org

Micro-correlations

Micro-correlations are correlations between variables at or beneath the limit of detection. These tiny correlations are amplified by aggregation, undermining common diversification strategies. The ballooning under aggregation is illustrated by a very simple formula that should be on the first page of every insurance text book, but isn't. Let $X_1, \ldots X_N$ and $Y_1, \ldots Y_N$ be two sets of random variables with the same average variance σ^2 and average covariance C (within and between sets). The correlation of the sums of the X's and the sum of the Y's is easily found to be:

$$\rho(\sum X_i, \sum Y_i) = \frac{N^2 C}{N\sigma^2 + N(N-1)C}. \quad (1)$$

This evidently goes to 1 as N grows, *if* C is non-zero and σ^2 is finite. If all variables are independent, then C = 0, and the correlation in (1) is zero. The variance of ΣX_i is always non-negative; if the σ^2 and C are constant for sufficiently large N, it is easy to see that $C \geq 0$.

The amplification of correlation can be seen most dramatically in the flood insurance claim data. Suppose we randomly draw pairs of US counties and compute their correlation. The green histogram in Figure 1 shows 500 such correlations. The average correlation is 0.04. A few counties have high, positive correlations, but the bulk is around zero. Indeed, based on the sampling distribution for the normal correlation coefficient, correlations less than 0.37 in absolute value would not be statistically distinguishable from zero at the 5% significance level. 91% of these correlations fall into that category.

Instead of looking at the correlations between two randomly chosen counties, consider summing 100 randomly chosen counties, and correlating this with the sum of another, distinct set of 100 randomly chosen counties. If we repeat this 500 times, the blue histogram in Figure 1 results; the average of 500 such correlations-of-100 is 0.23. The red histogram depicts 500 correlations-of-500, their average value is 0.71.

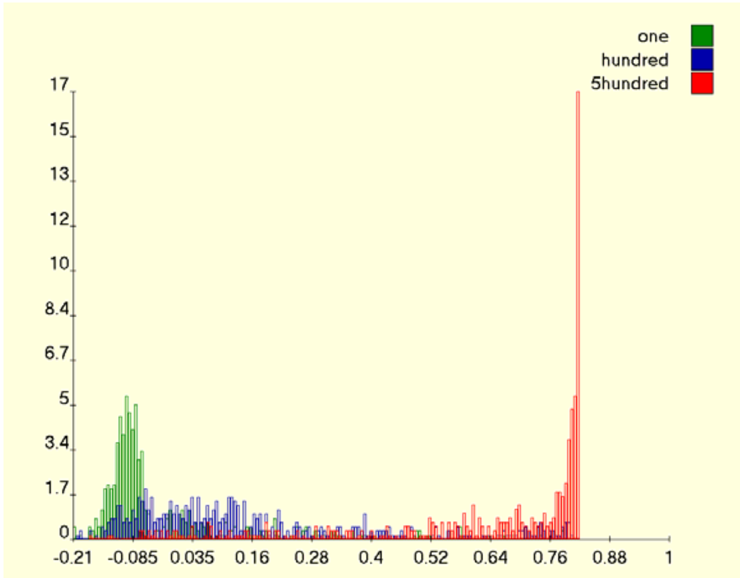

Figure 1: US flood claims, correlations-of-1 (green), correlations-of-100 (blue), and correlations-of-500 (red).

This dramatic increase in correlation is a result of the micro-correlations between the individual variables. Compare Figure 1 with Figure 2, in which each county is assigned an independent uniform variable, for each of 29 years. The correlations-of-1 and correlations-of-500 are effectively the same.

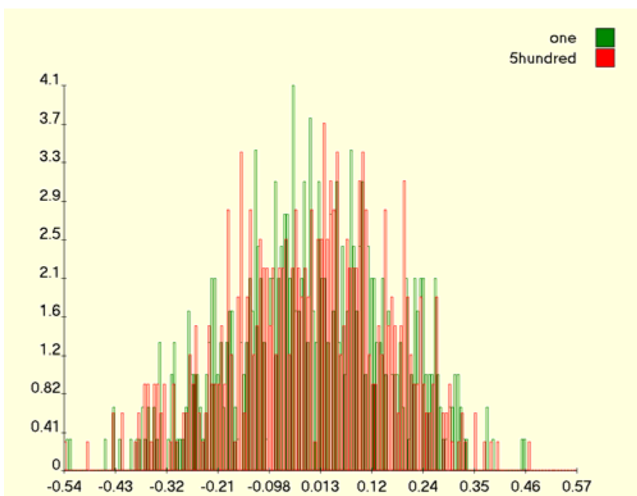

Figure 2: Correlations-of-1 and correlations-of-500, for 30 realizations of independent uniform variables

The difficulty with micro-correlations is that they could so easily go undetected. One might not readily assume that fires in Australia and floods in California are correlated, for example, but El Niño events induce exactly this coupling.

Fat Tails

Fat tails were introduced in mathematical finance in 1963 by Benoit Mandelbrot to describe cotton price changes (*1*), and are characteristic of many loss distributions. There is growing recognition among economists that, when it comes to climate change, "the tails matter." There is already evidence that damages often follow distributions with fat tails (e.g., *2*) and climate change may be directly fattening the tails of the distributions of many extreme events (e.g., *3*). The uncertainty surrounding climate change may also generate fat tails, as in Weitzman's analysis, where updating a non-informative prior yields a fat-tailed posterior damage distribution (*4*).

The precise mathematical definition of tail obesity is rather subtle (*5*), but a working notion is that damage variable X has a fat tail if, for sufficiently large values x, the probability that X exceeds x is ax^{-k}, for some constants $a,k > 0$. It is easy to see that the *m-th* moment is infinite if $m \geq k$. If $k \leq 1$, the tail is "Super Fat" and the mean or first moment is infinite. Of course on N samples from such a distribution, the average of the N sample values will be finite, but it increases with N. "Really Fat" tails, $1 < k \leq 2$, have infinite variance. The sample mean also has infinite variance, no matter how many samples we draw.

A good way to gauge tail obesity is by "mean excess" plots (*6*). At each damage level x, we consider all damages greater than x, and plot the average by which these exceed x. For fat tailed distributions, the mean excess plot is increasing. If the slope is greater than one, the tail is Really Fat, less than one, but greater than zero, only Meso Fat. The distinction between Really Fat and Meso Fat turns out to be very important. There is not enough data to study tail behavior county-wise. We therefore pool all counties and consider each county-year as a realization of a single loss variable, yielding about 75,000 loss events for crop losses and for flood claims. Figure 3 shows mean excess plots for crop (above, left) and exposure adjusted flood insurance claims (below, left). The unit slope lines have been added.

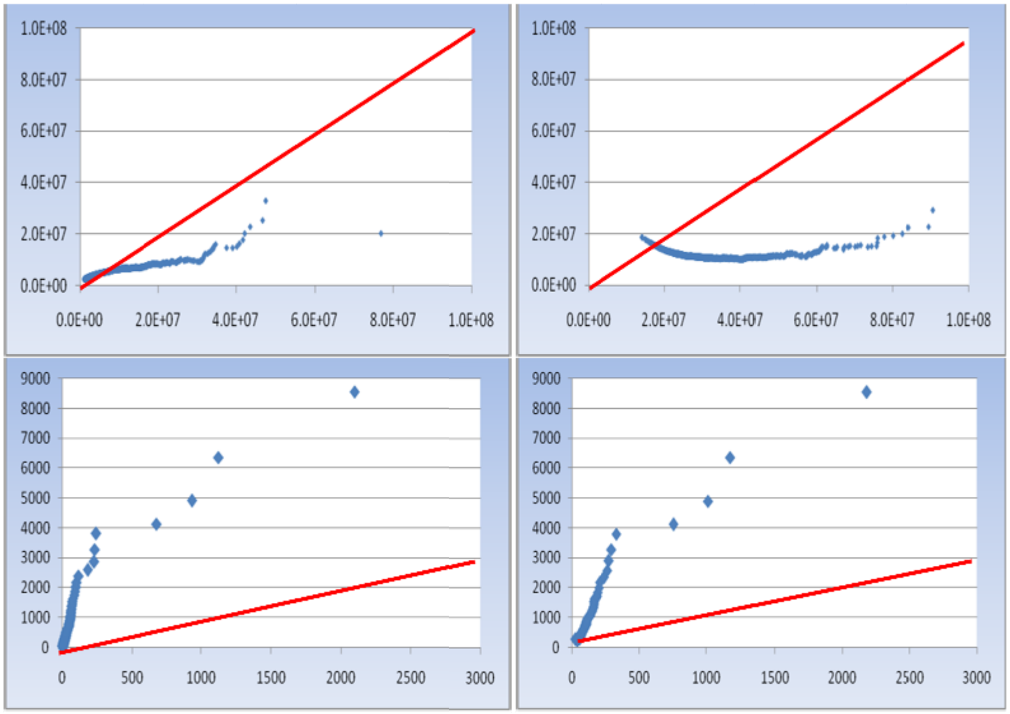

Figure 3: Mean excess plots for US crop loss (above) and exposure corrected US flood claims (below). The vertical axis gives mean excess loss, given loss at least as large as the horizontal axis. The upper right picture shows crop loss mean excess for random aggregations of ten counties. The lower right picture shows flood claims mean excess for random aggregations of 50 counties.

Whereas mean excess crop losses show a slope less than 1, mean excess exposure adjusted flood claims have a slope greater than one, indicative of a Really Fat Tail. The implication of this slope can be profound when we aggregate losses. Suppose we form packages of crop losses by randomly grouping the county-years ten at a time. Since the variance exists, the sum will approach a normal (thin tailed) distribution. This is reflected in the flattened mean excess plot of a random aggregation-by-ten of crop losses in Figure 3 (upper right). Compare this with Figure 3 (lower right) showing the mean excess of a random aggregation-by-fifty of exposure adjusted flood losses. Aggregation does not thin the tail significantly. Meso fat tailed loss data converge to normal distributions as we form random aggregations, but really fat tailed distributions converge to so-called "stable laws" which retain the tail behavior of the original variables. When insurance companies hold really fat tailed policies, the aggregated portfolio is also really fat, and is poorly estimated by historical averages.

Tail Dependence

Tail dependence is the most unwholesome and least well understood member of our unholy trinity. It refers to the tendency of dependence between two random variables to concentrate in the extreme high values (known as upper tail dependence, UTD). For loss distributions we are interested in UTD of non-negative variables. Technically, upper tail dependence of variables *X* and *Y* is defined as the limit (if it exists) of the probability that *X* exceeds its *r-percentile*, given that *Y* exceeds its *r-percentile,* as *r* goes to *100*.

If *X* and *Y* are independent, their tail dependence is zero. If their tail dependence is positive, then when one variable takes on an extreme value, it is more likely the other variable will as well. Note that UTD does not depend on the marginal distributions of *X* and *Y*; if we apply any 1-to-1 transformation to *X*, say $X^* = X^{1/N}$ (which will thin *X*'s tail), then $UTD(X^*,Y) = UTD(X,Y)$. UTD has no simple relation to the standard Pearson correlation coefficient used in eq. (1). For example, normal variables with any correlation ρ strictly between -1 and 1, have zero tail dependence (*6*).

Correlations between sums always grow with aggregation, if the average covariance is positive. Under certain conditions, tail dependence can also grow (*7*), further foiling diversification. Little is known about general conditions under which aggregation amplifies tail dependence, but we can see it in data. We examine monthly flood loss data from Florida (the only state for which we have monthly observations), per county, as it gives more observations than yearly data. If we consider two random groups of five different counties, and make a scatter plot of the percentiles of their monthly losses, the left plot of Figure 4 emerges. The points along the axes correspond to months in which no losses were reported in these counties. We may discern a weak tendency for points to cluster in the upper right corner. This tendency grows appreciably stronger if we take two random groups of 30 different counties, as in the right plot.

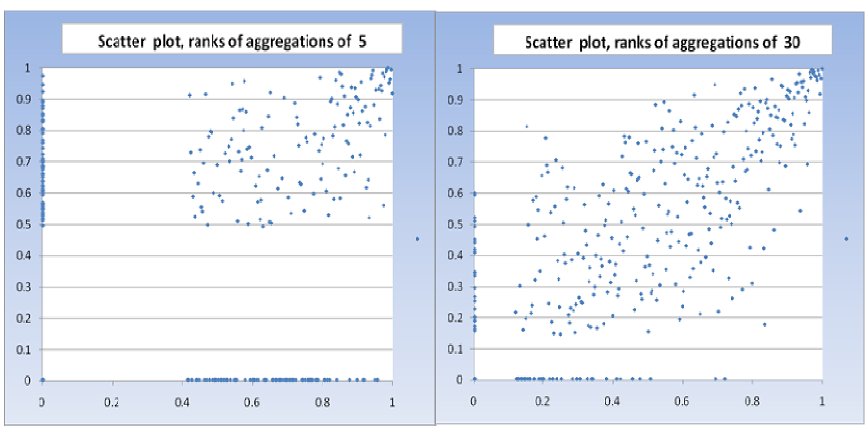

Figure 4: Percentile scatter plots of random aggregations of 5 (left) and 30 (right) Florida counties, monthly flood losses.

After Hurricane Katrina, heretofore independent lines of insurance saw many large claims, among them property, cargo, inland marine and recreational watercraft, floating casinos, onshore energy, automobile, worker's comp, health, and life insurance (*8*), pointing to tail dependence across these lines of business. Similar tail dependence, specifically between property and motor vehicle claims, was also observed in France (*9*). Such tail dependence has the potential to wipe out insurance companies.

Managing Risk

We don't need an unholy trinity to tell us that we as a society manage catastrophe risks badly, but it may help us fix what's broken. Since Mandelbrot's 1963 study of cotton prices, we know that fat tails require different methods than thin tails; the sample mean from a Really Fat tailed loss distribution has infinite variance—but risk management strategies are still based on historical averages. The NFIP not only uses an "historical average loss year" in setting risk based capital requirements for flood insurers, they also give a 1% weight to 2005 (with hurricanes Katrina, Rita and Wilma) in "an attempt to reflect the events of 2005 without allowing them to overwhelm the pre-Katrina experience" (*10*). NFIP is itself overwhelmed with Katrina's claims, and on its own admission is unlikely to be able to cover the interest on its debt to the treasury.

Ask someone from St. Tammany county in Louisiana:

'After Katrina, flood loss claims in your county totaled $240 per dollar income (2000 dollars); in the next hurricane at least as bad as Katrina, what do you expect your (2000) dollar loss per dollar income to be?"

Look at Figure 3 (lower left): Given a loss at least as great as $240 per dollar income, the expected loss in excess of $240 is $4,000. Expect the next one at least as bad as Katrina to be much worse than Katrina. It sounds like a contradiction, but it isn't, it's just the logic of Really Fat tails.

Banks, insurance companies, and mutual funds are all risk dissipaters. Mortgages, corporate bonds, and stocks carry risk, but you cannot lose more than everything you invest. Fat tails are not a problem here—but micro-correlations are. Suppose you buy a bond for $1,000 that pays $1,100 in one year, unless it defaults, in which case you get nothing. If the probability of default is 0.01, your expectation is μ = $89, but the standard deviation of your return is σ = $109.4 . Your "comfort ratio" μ/σ tells you how by many standard units your mean exceeds zero. $89/$109.4 = 0.81 is not a good comfort ratio. If you buy N identical *independent* bonds, your comfort ratio becomes $\sqrt{N}\mu/\sigma$. The more bonds you buy, the more comfortable you become. After 100 purchases, your mean is already 8 standard units from zero. Now suppose that all these bonds have a statistically insignificant, but positive correlation ρ; then your comfort ratio does *not* keep going up but levels off to $\mu/(\sigma\sqrt{\rho})$. That may not be so bad, *if* you know it…but if you overestimate your comfort, you might just do something foolish, like leveraging yourself at 30 to 1.

The last person in the unholy trinity, tail dependence, implies that when the chips are down, we're all in it together. The possibility for simultaneous losses across multiple insurance lines or locations may suggest the need for government intervention to protect against insolvency and cover extremely high loss layers. Unfortunately, many such attempts, such as solvency guarantee funds, have the perverse effect of discouraging insurance companies from managing catastrophe risk. This is the shadow side of risk dissipation, *moral hazard.* Government reinsurance with risk based prices could be designed to cover tail dependent regions of loss distributions. Whether policymakers can maintain risk-based pricing to encourage risk mitigating activities in the face of political pressure is uncertain at best. About 40% of the total premiums paid to the NFIP are subsidized at about 35% of the fair price. The reason? "It was anticipated that very high premiums would cause great resistance to insurance purchase. However, with reasonable [sic] premiums, property owners purchasing insurance at less than full-risk rates would still be funding at least part of their recovery from flood damage." (*10*). The "full-risk rates" are based on 1% if Katrina, Rita and Wilma. Dealing with moral hazards will require more than mathematics.

References

(1) B. Mandelbrot, *The (Mis)Behavoir of Markets* (Basic Books, New York, NY, 2004).
(2) B. D. Malamud, *Journal of Hydrology*. **322**, 168-180 (2006).
(3) U.S. Climate Change Science Program, *Weather and Climate Extremes in a Changing Climate* (U.S. Climate Change Science Program, Washington, D.C., 2008).
(4) M. Weitzman, *Review of Economics and Statistics*. **Forthcoming,** (2008).
(5) S. Resnick, *Heavy Tailed Phenomena, Probabilistic and Statistical Modeling* (Springer, New York, NY, 2007).
(6) A. J. McNiel, R. Frey, P. Embrechts, *Quantitative Risk Management: Concepts, Techniques, and Tools* (Princeton University Press, Princeton, NJ, 2005).
(7) C.Kousky, R. Cooke, *Climate Change and Risk Management: Challenges for Insurance, Adaptation, and Loss Estimation; RFF Discussion Paper 09-03-REV* (Resources for the Future, Washington, D.C., 2009).
(8) RMS, *Hurricane Katrina: Profile of a Super Cat: Lessons and Implications for Catastrophe Risk Management* (Risk Management Solumtions, Newark, CA, 2005).
(9) L. Lescourret, C. Y. Robert, *Scandinavian Actuarial Journal*. **4**, 203-225 (2006).
(10) T. L.Hayes, and D. R. Spafford, *Actuarial Rate Review: In Support of the May 1, 2008, Rate and Rule Changes* (Federal Emergency Management Agency, Washington, D.C., 2008).

[i] We would like to thank Ed Pasterick, Tim Scoville, and Barbara Carter for providing us with this data.

[ii] The exposure adjusted data are from 1980 to 2006, the last two years are dropped from the flood claims data. Income data was not avilable for Guam, Puerto Rico, or St. Croix, so these are dropped from our dataset. Further, the income data for some counties in Virigina was for aggregations of counties. These are also dropped as they cannot match cleanly with our flood claims data.