The Mathematics of Highly Optimized Tolerance*

Alexander W. Salter
Rawls College of Business
Texas Tech University
alexander.w.salter@ttu.edu

Vlad Tarko
Economics Department
Dickinson College
tarkov@dickinson.edu

Consider event space X – the set of all locations or organizations that could be affected by a shock. For simple natural disaster applications, X has corresponds to the geographical landscape, but, in general, X is a set of organizations that are part of a complex network. The analysis here is independent of the actual topology of X.

Let s_x denote the spillover effect, i.e. a subset of X affected by the shock occurring at location x, where "affected" means that each member of s_x suffers a cost above some conventional threshold. For example, s_x may be an area destroyed by a forest fire that started at x. In our banking example, it will be the set of all banks and other financial institutions affected by a shock (e.g. a run on bank x or the bankruptcy of financial institution x). Let $c(s_x)$ be the

* Appendix to "Governing the Banking System: An Assessment of Resilience Based on Elinor Ostrom's Design Principles", *Journal of Institutional Economics*

total cost suffered by the affected group of actors.[1] The expected cost over the entire system over all possible shocks is:

$$C = \sum_{x \in X} p_x \, c(s_x) \tag{1}$$

where p_x is the probability that a shock will occur at x.[2]

Agents who have a vested interest in the overall resilience of the system will try to set up prevention mechanisms to minimize the expected cost. Let r_x be the distribution of resources spent at various possible locations, x, in the attempt to prevent the problem. For example, these might be resources used to pay for financial audits and enforcement of various banking rules, as well as financial guarantees, interests paid on reserves, subsidized insurance rates, etc. These resources are subjected to a budgetary constraint,

$$B = \sum_{x \in X} r_x \tag{2}$$

where B is the available budget for promoting resilience.

We assume that using more resources in location x diminishes the spillover cost:

$$c(s_x) = c_0(s_x) - r_x^{-\beta} \tag{3}$$

where $c_0(s_x)$ is the cost in the absence of prevention, and β is a measure of how interconnected the system is.[3] The nature of the network affects both the size of spillovers and the economies of scale in prevention. The exponential form reflects the fact that interconnectivity generates a combinatorial problem. The more interconnected the system is (higher β) the more difficult it is

---

[1] In principle, the shock can be either positive or negative. For positive spillovers the cost $c(s_x)$ will be negative, i.e. a benefit.
[2] For a continuous geographical space, all the sums in this appendix become integrals.
[3] For convenience, the units of measure of $r^{-\beta}$ are chosen to match c.

to stop the spillovers. (As we shall see, the flipside of this is that economies of scale are present.)

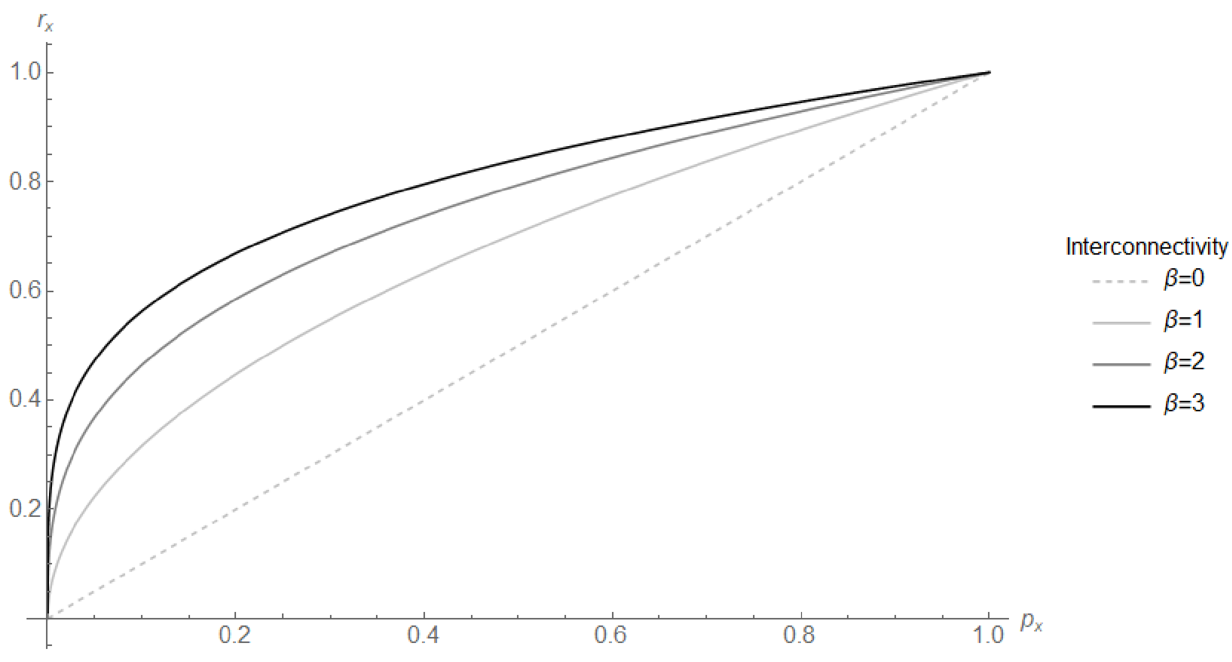
To determine the most efficient allocation of resources, we minimize the total expected cost subjected to the budgetary constraint, by varying the possible distributions r_x:

$$\frac{\delta}{\delta r_x}\left[\sum_{x\in X} p_x\left(c_0(s_x) - {r_x}^{-\beta}\right) + \lambda\left(B - \sum_{x\in X} r_x\right)\right] = 0 \tag{4}$$

where λ is the Lagrange multiplier. The resources for prevention should, thus, be allocated in the locations and proportions r_x proportional to (fig. A1):

$$r_x \sim p_x^{\frac{1}{1+\beta}} \tag{5}$$

Figure A1: The allocation of preventive resources based on estimated probability



Note: To help interpretation, the figure illustrates the proportionality relation between r_x and p_x. For the complete result, each curve must be shifted downwards to take the budget constraint into account, i.e. the areas under the curves need to remain constant and equal to B.

For example, if the interconnectivity is minimal ($\beta = 0$), the most efficient allocation of resources is $r \sim p$, i.e. locations more likely to be affected by a shock should receive proportionally more resources. If the interconnectivity is greater than zero, the resources should be spread more broadly, even to lower probability locations (fig. A1). For very high interconnectivity ($\beta \gg 1$), it is best to spread protective resources almost equally to all locations, including those that are highly secure. Intuitively, this is because, in an interconnected system, a shock originating in an unlikely location will still spread widely.

This is the efficient solution in light of one's estimates of risk, p_x, and of the interconnectivity, β. But whether allocating resources this way is indeed a good idea (i.e. a robust and "wise" solution) is still debatable. The critical question about resilience concerns the probability that, despite the prevention efforts, a large scale (even system-wide) problem will still occur. The expected cost of a system-wide catastrophe is:

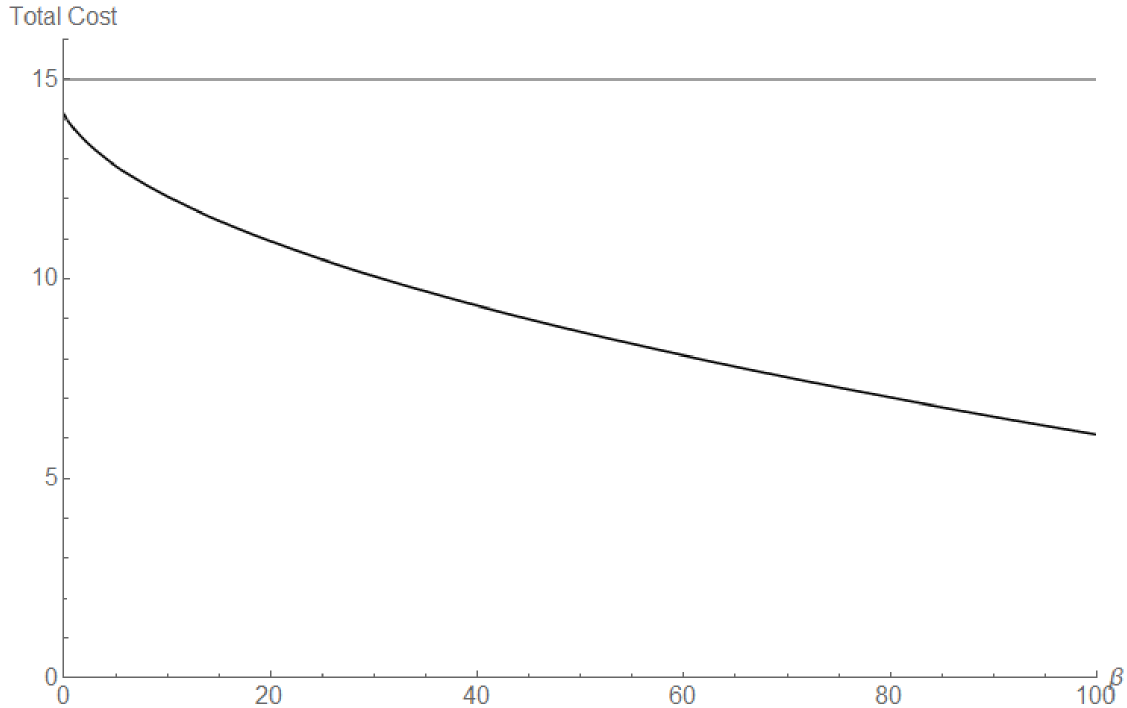
$$\overline{C} = \sum_{\substack{x \in X \\ s_x = X}} \left[p_x \, c_0(s_x) - p_x \, r_x^{-\beta} \right] \tag{6}$$

Substituting r from eq. 5, we obtain the expected cost of a system-wide catastrophe for the most efficient allocation of preventive resources:

$$\overline{C} = C_0 - \sum_{x \in X} p_x^{\frac{1}{1+\beta}} \tag{7}$$

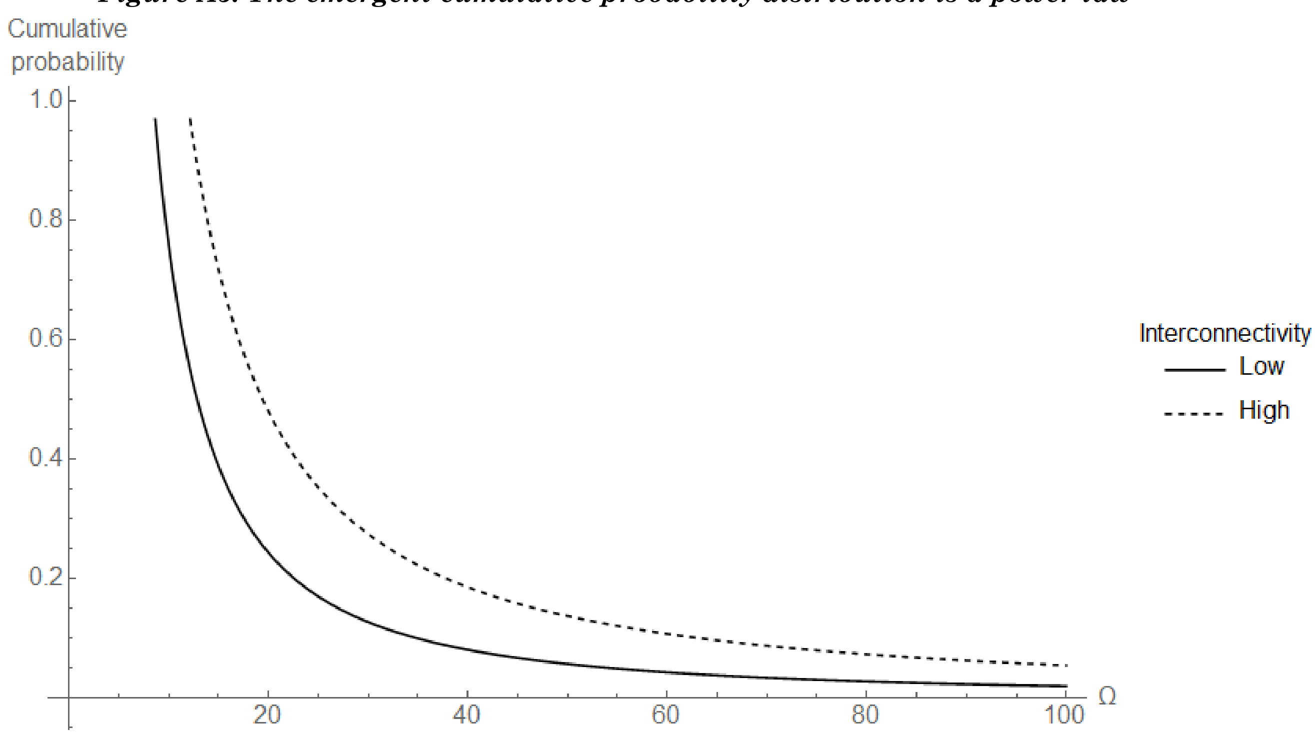
where C_0 is the cost in the absence of preventive actions.

Figure A2: The benefits of prevention as a function of interconnectivity



Note: The straight line at 15 indicates the total cost in the absence of preventive measures.

Figure A3: The emergent cumulative probability distribution is a power law



To make the strongest possible case for the importance of this HOT phenomenon, i.e. about the vulnerabilities created by optimizing based on the known risks, let us assume the best-case scenario in which we have a good idea about the possible distribution of shocks, and that this distribution is normally distributed,[4] $p_x \sim \mathrm{Exp}(-x^2)$. Substituting this in eq. 7, we obtain the following expected cost of a system-wide catastrophic shock:

$$\overline{C} \approx C_0 - \int_0^\infty [\mathrm{Exp}(-x^2)]^{\frac{1}{1+\beta}} = C_0 - \frac{\sqrt{(1+\beta)\pi}}{2} \qquad (8)$$

As illustrated in fig. A2 for $C_0 = 15$, the benefits of prevention are greater if interconnectivity is higher – prevention efforts have significant economies of scale. However, even perfectly efficient prevention will not decrease the expected cost to zero. As illustrated in fig A3, the cumulative probability that the cost of the shock is greater than some overwhelming cost, Ω, is a power law with a fat tail (see Carlson and Doyle 1999, Table 1):

$$p_{\mathrm{cum}}(c(s_x) > \Omega) \sim \frac{\Omega^{-\left(1+\frac{1}{\beta}\right)}}{\sqrt{\mathrm{Log}\,\Omega}} \qquad (9)$$

While Taleb argues that fragility occurs because of Knightian uncertainty and the misuse of the normal distribution, the HOT result is even more pessimistic. Even when we know the sources of danger, and the underlining distribution is Gaussian, we can still end up with a system vulnerable to large-scale destructive events, as a result of its interconnectivities. The probability that the *entire* system is negatively affected is non-negligible.[5]

The pair of figures A2 and A3 highlights the fundamental problem: Figure A2 shows that

[4] In the Bayesian model selection approach (Bretthorst 1996; Jaynes 2003), this corresponds to us knowing both the average and the variance, and assuming that they don't change in time. The case when there are several possible shocks at different locations is easily generalized as a sum of normal distributions.

[5] The power law of the cumulative probability distribution emerges because of eq. 3.

economies of scale are present, and, hence, according to considerations of *efficiency*, all the parts of the system should be combined into a single large system. This way we get the largest bang-for-the-buck out of the resources we have for prevention. By contrast, figure A3 shows that, the greater the interconnectivity, the more likely it is for the system to fall prey to a system-wide failure, even when preventive resources are perfectly allocated. Creating a unified single system is a bad idea from the perspective of *resilience*. We thus have a fundamental trade-off between efficiency and resilience.

Finally, one can wonder how the total cost changes if the probability was misestimated. Suppose we allocated protective resources based on probability distribution p_x but the correct estimate was, p'_x, and $\delta p_x = p'_x - p_x$ is the error. Substituting p'_x in eq. 6, and r_x from eq. 5 as before, we obtain that the total cost changes by:

$$\delta\overline{C} = \delta C_0 + \sum_{x\in X} \delta p_x \, p_x^{-\frac{\beta}{1+\beta}} \tag{9}$$

Figure A4 illustrates the contribution to the total cost error from a single location for a probability error of $\delta p_x = 0.1$. As we can see, low probability locations create the biggest problems. If an error occurs at a high probability location, the impact is minimal because plenty of resources are spent there anyway. By contrast, even a relatively small estimate error for a low probability location has a significant impact on the amount of resources spent, and, hence, on the total cost. Furthermore, as expected, higher interconnectivity amplifies the problem. Figure A5 shows the effect of different probability errors from a location that was deemed to have a $p_x = 0.2$ probability to be the origin of a shock. A greater error leads to a proportionally greater cost, and once again, interconnectivity makes matters worse.

Figure A4: Contribution to total cost error from one single location

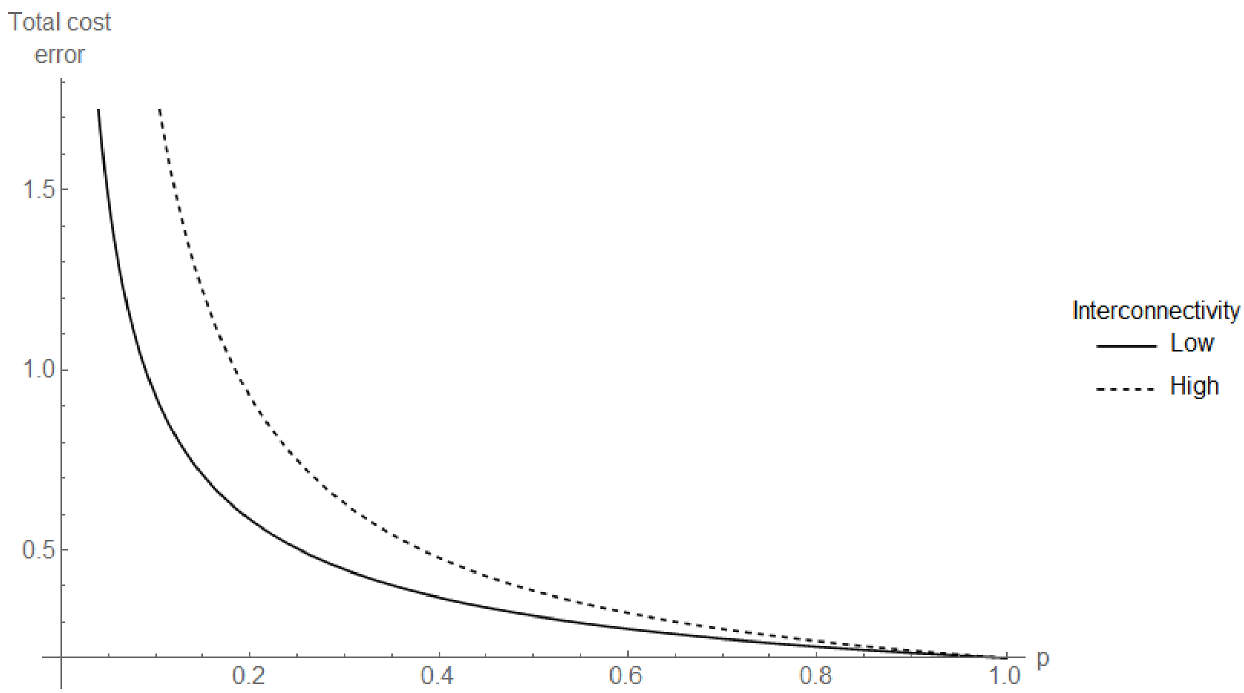
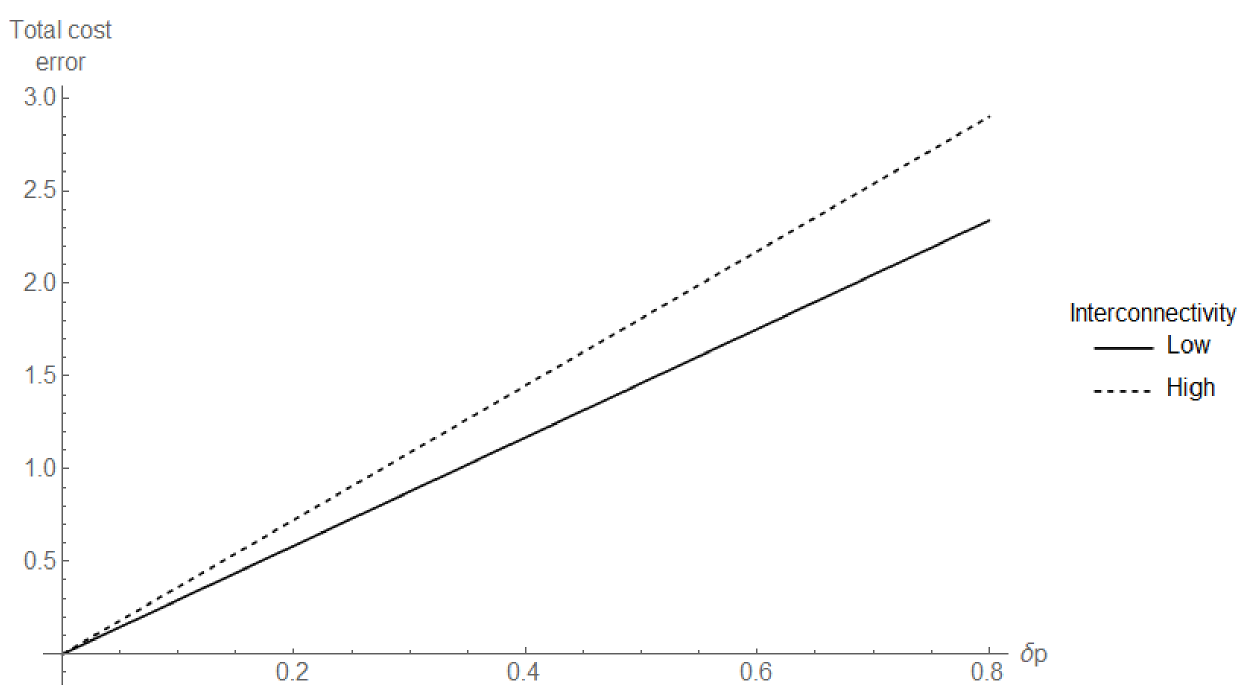


Figure A5: The effect of a probability estimate error at a low probability location



A few consequences follow. First, given the possibility of probability estimates errors, to promote resilience one needs to distribute prevention resources more equally across locations. But, *how much* more equally is hard to tell as, by definition, we don't know the size of our error. Secondly, it's important to have institutional mechanisms for accurate knowledge aggregation, i.e. for minimizing δp_x for all x. A top down policy is problematic for several reasons: (1) It will be based on a single specific probability distribution estimate, and whatever error this estimate involves, it will be imposed system-wide. (2) Hayekian distributed knowledge concerns contribute to the problem. (3) Top-down regulatory systems incentivize the discovery of rules evasions, hence, *endogenously* increasing the size of δp_x. By contrast, a polycentric system, especially when market prices are available, can both (a) better enable "wisdom of crowds" mechanisms for discovering the correct probability distribution (e.g. divergent probability estimates will be embedded in the prices for insurance and in the interest rates), and (b) limit the scope of errors.