Market Based Risk Mitigation:
Risk Management vs. Risk Avoidance

Shmuel Oren

University of California at Berkeley

Presented at a White House OSTP/NSF

Workshop on Critical Infrastructure Interdependencies

June 14-15, Washington DC

1. Introduction:

Civil and critical infrastructure systems such as transportation, communication, power, and financial systems have provided the foundation for modern society. Not surprisingly, much of the research work in many areas of engineering was directed over the years at the advancement and application of scientific principles to the design, maintenance and improvement of the critical infrastructures in our society. Risk assessment and systematic consideration of risk in the design and operation of infrastructure is a relatively new phenomenon that emerged over the last thirty years. While many engineering disciplines still do not consider probability and statistics as essential basic knowledge for engineers, like physics for instance, consideration of risk has penetrated all engineering disciplines. In most cases, however, such considerations take the form of setting thresholds and safety margins so as to avoid “unacceptable” risk, where acceptability levels are typically determined by experts. In power systems for instance, the common wisdom has traditionaly been to build sufficient capacity so that the system will fail to meet demand no more than one day in ten years. Similarly, power system operation is governed to a large extent by an “N-1 security criterion” which requires that the system as a whole can sustain failure of any one element (e.g. generator, transmission line, transformer etc.).

The last two decades brought about two revolutionary changes in critical infrastructures that will have a lasting impact on infrastructure related research and development. These changes present new challenges for academic research and training of professionals in infrastructure related fields. In particular, the way we think about risk and risk mitigation in the context of critical infrastructure is at the verge of a revolutionary transformation that will empower more personal choice and accountability.

The first change is the IT revolution. Electronic sensing, telecommunication, controllers and computation have proliferated into every infrastructure system and revolutionized their operational paradigms. The new information technologies also led to new critical infrastructures, such as the Internet, and to interdependencies and complexities that create new vulnerabilities and risks.

The second change is the massive deregulation of infrastructure industries such as the airline industry, natural gas industry, trucking, telecommunications and electric power. The deregulation of these industries in the US, and around the world, has challenged traditional beliefs that infrastructure industries should be publicly owned or be run by vertically integrated regulated monopolies, and that consumers should have little or no say in choices that concern tradeoffs between physical and financial risks vs. cost. Deregulation of infrastructure industries has brought about decentralization of the decision-making, and creation of markets that enable and incentivizes customer choice. By empowering consumers to make informed choices, the deregulation movement has triggered a process that is revolutionizing the concepts of reliability and risk. For example in an infrastructure industry governed by an "obligation to serve", service curtailment (e.g. being bumped off a flight due to overbooking) is considered a failure that counts toward reliability metrics. By contrast, in a deregulated environment governed by "obligation to serve at a price" service curtailment becomes voluntary (in exchange for compensation or a cheaper fare) and no longer counts as a failure. Airline passengers would probably feel more tolerant to schedule delays as well given the proper financial compensation.

2. Risk Mitigation Paradigms

We can identify three basic approaches to risk mitigation:

a. Risk avoidance

This has been the traditional approach adopted by engineering design and regulatory bodies. Typically it consist of screening alternative courses of action by performing a risk assessment, and enforcing a threshold criterion for acceptable risk. Such criteria are set on the basis of expert opinion and policy consideration including political compromise. Alternatives that fail to meet the set criterion are rejected. Typical applications of this approach include for instance FDA approval of drugs, EPA approval of emission levels or toxic waste levels, OSHA safety standards, and NERC [North American Electric Reliability Council] operating guidelines for the electricity grid.

b. Decision-making under uncertainty

This approach, known as Decision Analysis, evaluates alternative courses of action by modeling the uncertain outcomes of each alternative, and taking into consideration subjective and objective information including assignment of values to potential outcomes and consideration of the decision maker's risk preference. The approach aims to identify the "best alternative" among the choices under consideration (including the alternative to seek more information), given current information and the decision maker's preferences. It also provides a variety of sensitivity analyses. Decision Analysis is a well-established discipline rooted in statistical decision theory and system analysis. This approach has been applied successfully to many decision problems in the private and public sector, in the context of decisions such as whether to seed hurricanes, decisions concerning earthquake reinforcement, product development decisions, investment decisions etc.

c. Risk management

This approach aims to control financial consequences of uncertain outcomes through economic mitigation. Such mitigation often takes the form of trading contingent claims whose financial settlement depends on realizations of uncertain state variables. Contractual arrangements such as insurance contracts are examples of such contingent claims. Risk sharing agreements and syndication is another form of economic risk mitigation. These are mechanisms for "spreading the risk" or more precisely the financial consequences of the risk. Selling shares in a company through a public offering is an example of a risk sharing arrangement. Hedging is a third form of economic risk mitigation in which a risk bearer reduces its exposure by creating a portfolio of ventures whose outcomes happen to be correlated so as to reduce total variability. A farmer for instance can hedge the price risk of its crop by purchasing a financial contract that pays the difference between some set price and the market price of the crop. If the price of the crop goes down the payoff from the contract goes up so the gains from the financial contract offsets the losses due to the decline in crop prices. Hedging can be financial or physical (operational). For instance, a PC manufacturer can hedge the risk associated with a long term supply contract due to fluctuating prices of some critical input (e.g. memory chips) by merging with the company producing that critical input or accumulating large inventories of that input.

The applicability of the various approaches in a specific context largely depends on the type of risk we are facing and the perception of that risk. In particular, it is useful to classify types of risks based on the following categories:

- Voluntary vs. Involuntary
- Private vs. Public
- Diversification options (risk sharing, portfolio approaches)
- Tradability (insurance, hedging)
- Interdependency (ability to provide differential protection)

It is important to note, however, that in some instances these categories depend on the framing of a situation. For instance, it is well known that individuals' risk tolerance is higher for voluntary risk then for involuntary risk. However, the distinction between what is voluntary and what is involuntary is vague and can shift depending on framing and incentives. Similarly, it is often possible to localize risk or convert public risk to private risk through the use of technology and information that can localize the consequences of risky actions.

3. Electric Reliability Example

The electric power system is undoubtedly one of the most important and complex critical infrastructures, which is currently undergoing massive restructuring in the US and around the world from a vertically integrated regulated monopoly structure to a deregulated market based industry. The reliability standards for the power industry, however, continue to be regulated by NERC (the North American Electric Reliability Council). The electric power industry serves as an excellent example of how technological and regulatory changes in a critical infrastructure can affect the risk mitigation paradigm.

NERC defines reliability as: “the degree to which the performance of the elements of the electrical system results in power being delivered to consumers within accepted standards and in the amount desired” This definition of Reliability encompasses two concepts:

- *Security:* “the ability of the system to withstand sudden disturbances.” This aspect concerns short-term operations and is addressed by ancillary services, which include: Voltage support, Congestion relief, Regulation (AGC) capacity, Spinning reserves, Non-spinning reserves, Replacement reserves.
- *Adequacy:* “the ability of the system to supply the aggregate electric power and energy requirements of the consumers at all times”. This aspect concerns planning and investment and is addressed by Planning reserves, Installed capacity, Operable capacity or Available capacity.

Security and Adequacy are clearly related since it is easier to keep a system secure when there is ample excess generation capacity. However, a system with limited adequacy can also be operated securely by simply maintaining sufficient reserves on hand even if that means curtailing customer loads (that approach has been used in California during Stage 3 alerts when reserves dropped below 2%).

From an economic perspective, security and adequacy differ in the sense that security is a *public good* (like fire protection, national defense or clean air) while adequacy is (or can be) treated as a *private good.* With existing technology, it is not practical to offer differential security to customers, offering for instance exclusionary protection against system collapse. This creates externalities and potential “free rider” effects. Thus, while the resources that are needed to meet security needs can be procured through a competitive process the decisions concerning the amounts of ancillary services that are needed, how they should be dispatched and how should their cost be allocated to the various parties need to be centralized. By contrast, generation adequacy decisions can be decentralized and left to the market. Inadequate supply will result in high prices, which in turn encourage new capacity. Customers can be allowed to decide how much reserve capacity they want to pay for, whereas suppliers can decide how much to invest in new capacity. These are individual economic and risk management decisions. Unfortunately, in most of the restructured systems the independent system operators (often influenced by political considerations) continue to operate under the old "obligation to serve" paradigm which leads to the treatment of generation adequacy as a public good, and to undermining the foundation of a market based system.

The remainder of this discussion will focus on the provision of generation adequacy. There are three basic approaches that have been adopted in deregulated electricity markets for ensuring generation adequacy. The first two can be classified as risk avoidance methods while the third approach is market based.

Risk avoidance methods for ensuring generation adequacy:

a. A central agency (independent system operator or regulator) specifies requirements for planning reserves based on traditional planning tools. ICAP (Installed Capacity) markets allow supplier to trade reserves and efficiently reallocate the reserves requirements (PJM, New York, New England). The drawback of this approach is that the capacity markets and energy market may not be in equilibrium, i.e. market prices for energy and capacity do not reflect the fact that capacity has no intrinsic value other than providing an option to produce energy. In particular, short run supply and demand curves governing the ICAP market are inelastic (vertical) so ICAP prices are either very low (when supply exceeds demand) or very high (when demand exceeds supply).

b. Generators receive capacity payments based on availability in order to induce investment and availability and to provide a base cash flow for generation capacity sitting on the sidelines as planning reserves (old UK system, Argentina, Spain). Unfortunately, while capacity payments support incumbent generators they suppress energy prices and particularly scarcity rents that provide price signal for new investment. Furthermore, suppression of energy prices due to capacity payments reduces demand response and leads to excess capacity.

Market approach to ensuring generation adequacy:

This approach relies fully on energy markets to compensate generators for fixed investment costs through inframarginal payments embedded in market clearing prices (that may include scarcity rents) and to provide price signals that will attract new investment in generation capacity. Consumers and suppliers interact through unrestricted energy spot markets. Energy spot and future energy prices provide price signals and compensation for capacity investment. Technology mix and generation capacity are determined by entry and exit of suppliers and by customer choice of desired price risk (California, Nordpool) .

While the latter approach is the most desirable from an economic theory perspective it has the undesirable feature that prices can reach, for short durations (or even long ones, as we have witnessed in California during the summer of 2000 and winter of 2001) very high levels. Such price spikes are exacerbated by the fact that for most practical purposes electricity is nonstorable (with the exception of hydro) and by the shape of the electricity supply curve that has a "hockey stick" shape with a steep rise at the end (so that when demand is near the peak, small fluctuations in demand due to a rise in temperature, for instance, can cause large price swings). Price spikes trigger customer discontent and consequently political reactions as we have seen recently in California. While in theory forward contracts would provide the appropriate market instruments that enable market participants to mitigate the financial consequences of price spikes, such mechanisms are not well developed (and often suppressed due to regulatory interference in the market) especially in young markets. Consequently, practical considerations may require some form of adequacy risk mitigation that goes beyond the ideal economic solution that relies on energy only markets. Nevertheless, in developing such supplemental measures for adequacy risk mitigation it should be recognized that generation capacity reserves beyond security needs are no more than a hedge against high prices. Hence, customers and suppliers should be free to choose levels of exposure to price risk through risk management and contractual agreements.

As indicated above, forward markets and hedging instruments can provide a competitive market alternative to capacity payments or ICAP requirements. When such markets function properly, new generation gets built if the market value of capacity (as reflected by long term contracts) exceeds cost of new generation. Unfortunately, regulatory uncertainty and failure of capital markets (to provide investment capital at competitive interest rates) could result in under-investment and involuntary exposure of consumers to price volatility. Again, the California electricity crisis of 2000 and 2001 provides a good example of such market failure.

The following is a sketch of a market based risk management system that retains the essential elements of a market-based approach to ensuring generation adequacy while providing safeguards that protect customers (who wish to be protected) from undesirable price volatility.

A blueprint for market-based provision of generation adequacy

Under the following scheme, the role of regulatory agencies is reduced to ensuring that load serving entities and generators have the resources to meet contractual obligations. Public risk is limited by requiring regulated load serving entities to provide hedges in the form of forward contracts and call options (with strike prices set by the regulator) up to, say X% above their annual peak load (planning reserves). The key difference between this method and the installed capacity requirement approach is that it provides load serving entities with the flexibility to choose how to hedge their capacity risk. This can be done through physical hedging by contracting with generators who have idle capacity, by contracting with loads that are willing to be interrupted in case of a shortage or by providing a financial security that ensure their ability to purchase power at high spot prices (in case of a shortage) and absorb the price spikes for their retail customers. This flexibility provides price elasticity that will discipline the prices for hedges. Furthermore, unlike ICAP contracts, a physical capacity hedge entails an obligation by the generator to supply power at a pre-specified strike price (from its own generation or from other sources). By contrast to the capacity payment approach where payments for capacity are uniform, and centrally determined, generators offering physical hedges to load serving entities must compete and the resulting payments for reserve capacity are market based and vary from one deal to the next.

Let us go back to the insurance framework that may help us understand the implication of the risk management approach described above. As discussed earlier, the objective of generation adequacy is to provide price insurance to customers. While in theory risk averse customers will obtain insurance on their own, some do not, either because of lack of information or due to some form of market failure. The social and political consequences of extreme price spike are so severe that regulators may wish to impose some level of mandatory insurance much like the mandatory insurance imposed on automobile users. The capacity payment approach would be equivalent in that context to imposing a uniform level of coverage at a regulated fixed price whereas the approach proposed above may be viewed as allowing parties to negotiate the prices for the level of coverage that will meet their obligation or even (as in the case of automobile insurance) to self-insure by posting a bond that will allow them to cover their liability if needed.

3. Implementation, Research and Educational Challenges

Economic-based risk mitigation, whenever appropriate, poses methodological and implementation challenges that require new research as well as curricular innovation and training of a new breed of engineers that are versed in technology, statistics, economics and finance. Such multidisciplinary training might be practical only at the graduate level.

As far as a research and development agenda, following is a partial list of topics that span theoretical research, application-oriented subjects and public policy.

- Development of market instruments, markets and institutions for risk trading and risk sharing. Such development should be viewed as “Market Engineering” which builds on the “physics” of markets explored by social sciences (including economics) but focuses on the harnessing of market forces and human behavior to achieve a desired outcome.
- Create mechanisms for transforming involuntary private risks to voluntary risk so as to empower individual choice. Without individual choice market based risk mitigation is not possible.
- Remove regulatory and institutional impediments to the exercise of individual risk preferences.
- Develop decision analytic methodologies that integrate expert risk assessment with market based pricing of risk (e.g. decision theory, real option theory, financial engineering methods)
- Develop pricing and portfolio analysis methods for valuation and optimization of risk mitigation options