Government Office for Science

Foresight

Computer trading and systemic risk: a nuclear perspective

Driver Review DR26

Foresight, Government Office for Science

Contents

Computer trading and systemic risk: a nuclear perspective

Robin Bloomfield and Anne Wetherilt

This review has been commissioned as part of the UK Government's Foresight Project, The Future of Computer Trading in Financial Markets. The views expressed do not represent the policy of any Government or organisation.

Summary

Financial markets have evolved to become complex adaptive systems highly reliant on the communication speeds and processing power afforded by digital systems. Their failure could cause severe disruption to the provision of financial services and possibly the wider economy. In this study we consider whether a perspective from the nuclear industry can provide additional insights.

The views expressed in this paper are the personal views of Dr Anne Wetherilt, and not those of the Bank of England.

Extended summary

Introduction

Financial markets have evolved to become complex adaptive systems highly reliant on the communication speeds and processing power afforded by digital systems. Their failure could cause severe disruption to the provision of financial services and possibly the wider economy. In this study we consider whether a perspective from the nuclear industry can provide additional insights.

In fact there are a very wide range of areas where the issues and practices in the nuclear industry might resonate with those raised by the evolution of computer-based trading and in this report we focus on:

- The approaches to systemic risk definition and evaluation;
- The definition of protection system parameters, risk controls and architecture;
- The need for trust in computer-based systems.

Approaches to systemic risk

We begin by examining a number of basic questions: how are the overall risks from a nuclear plant defined and evaluated, and how does this compare with financial markets.

We consider a serious nuclear incident that has the potential for the release of radioactivity with associated plant damage as a “systemic event” and hence make the link to a financial market crash: an event that both damages the market and also potentially impacts the wider financial system and the broader economy.

The development of the nuclear industry approach to safety has been driven by the need to engineer systems that provide social and economic benefits with tolerable risks, to evaluate and explain the nature and extent of these risks and to provide a framework that allows for scrutiny at varying levels of independence ranging from technical experts within the industry as well as pressure groups and those who, quite legitimately, hold very different values and worldviews.

For the nuclear and finance sector we consider:

- The basic concepts of hazard, risk and accident;
- Probabilistic Safety Analysis and the concept of a Design Basis;
- Tolerability of risk and the As Low As Reasonably Practicable (ALARP) principle and
- Numerical risk targets.

Our analysis points to the following similarities and differences:

- When thinking about large-scale risks, both industries employ the concept of systemic risk.
- Both industries use probabilistic concepts of risk and impact.
- The nuclear industry has a clear notion of tolerable level of risks and it can set numerical targets. Hence, it is able to explicitly assess trade-offs between risk reduction and costs (ALARP).
- In its thinking about financial market crashes, the finance industry relies to a large extent on probabilistic methods, using historical data. This is complemented by stress testing, using both historical and theoretical stress scenarios. There is also much emphasis on understanding past events so that potential future problems can be avoided.
- Although these past events have undesirable features, which can be measured precisely, there is no equivalent of the ALARP principle and there are no numerical targets.

In addition, it is useful to note the following particular aspects of the nuclear safety analysis and risk framework:

- The nuclear industry has a formalised approach to defining the classes of consequence, the categories and frequencies of initiating events. It uses using theory, models and experiment to justify the risk analysis.
- This means that the industry can set risk targets for classes of accident and different classes of people, and discusses tolerability and proportionality in reducing them further.
- In doing so, the industry accepts that many things are hard to quantify, but there is nonetheless an emphasis on ranking risks, setting targets for risk reduction, and debating whether both the risks and the targets are accurate and acceptable.
- The nuclear safety analysis framework allows systematic design of protection and mitigation systems that cover not only what they have to do, but also how much they have to be trusted. These systems use diverse mechanisms to ensure that the overall protection works when it is needed.

- The nuclear industry also places greater emphasis on explaining risks to society at large. This in part drives the quantification of risk as there needs to be a basis for comparing different types and sources of risk.

Looking ahead, one could question whether the rapid development of computer-based trading in financial markets requires the adoption of additional risk concepts and tools. Our analysis suggests a number of questions that are worth asking (see conclusions).

Protection parameters and risk controls

Having defined the key risk concepts used in the nuclear and financial markets context, we compare their respective use of risk controls and observe the following:

- Both employ risk controls, based on thresholds beyond which operations need to be halted (or slowed down). These controls are normally subject to regulatory supervision.

- In the nuclear industry, the risk controls are the result of a systematic engineering analysis, summarized in the protection envelope and the Fault and Protection Schedule.

- In financial markets, risk controls typically depend on a smaller and less complex set of parameters (e.g. traded prices or message volumes).

- Unlike the nuclear protection systems, there is no formal mechanism for describing how much the controls themselves have to be trusted (e.g. in terms of probability of failure on demand, probability of spurious activation).

We provide a brief overview of the control and protection of a nuclear plant that raises a number of issues that may be of relevance as financial markets consider how to adapt existing risk controls to the future computerized trading environment:

- Engineering succeeds by making the complex systems controllable and predictable (within limits). Although the underlying processes are complex and complicated the ability to model and design the plant and to have a scientific based understanding of what might happen allows the functional aspects of the protection (controls) to be relatively simple. However there are more onerous requirements on the non-functional aspects (e.g. the probability of failing, the response time) as the systems really do have to operate when needed.

- The ability to engineer a control and protection system relies on observability of the system. The notion that the financial market is an observable system in readily identified states is only partially true. It is clear in our review of market crashes that there are competing theories and perspectives. There is evidence that some crashes appear to just happen and that these are irreducible and so there is no difference between a transition to a systemic loss and an everyday one. Others would argue that indeed there is a difference; it is just that we do not (yet) have the means to identify the states that precede a systemic event. This has implications for the extent to which market controls can be engineered.

- Protection systems have authority to override any other system and force a shut down. If they operate when not needed (e.g. due to internal failures, operator error) they can cause spurious plant disturbances with consequential economic costs and safety implications. There

is a need to define performance measures for spurious activation (e.g. once every 10 years) as well as for the probability of failures per demand.

- There is trade off between economic benefits and having a simple protection envelope. As the understanding of the nuclear plant has developed over the years, protection envelopes have become more complex and computer based, facilitating more efficient operation. There can be considerable off-line data analysis and modelling to derive the parameters for these systems: so that trust is needed in both the protection algorithms and the data.

- Adaptation and learning is very important, but in a nuclear plant they occur in different timebands from the control and protection actions e.g. months for procedures and safety culture, years for updating equipment, decades for design of plant. This is in contrast to the rate of adaptation in markets and computer-based trading in particular.

- There is a need to get the best balance between reliable automation and human analysis and adaptation. This is a sophisticated topic but in brief the design needs to play to humans' strengths of understanding and adaptation.

- In safety engineering there are examples where introducing safety or protection measures can change people's behaviour so that the anticipated safety improvements are not as envisaged. The complex systems nature of markets means that adaptation could be a significant future issue in designing control and protection especially as these might provide unintended opportunities for new forms of regulatory arbitrage or market abuse.

Trust in computer based systems

Computer based systems, and ICT in general, are of course essential for high-frequency trading and for algorithmic trading more broadly: for market participants to process information, design trading strategies, and execute the trades; for operators of trading venues and regulators to collect and process the data for monitoring and market surveillance and for operators to provide protection and intervention via so-called circuit breakers. In all cases, systems will need to be sufficiently trusted: how do we describe that level of trust and how is this evaluated?

In the nuclear sector the reactor protection system is crucially dependent on software and complex electronics. In the UK nuclear industry, the justification is based on two important safety principles: the need for "production excellence" and independent "confidence building". The resulting assurance approach could be summarised as "Do everything and do it at least twice". The specific technical measures that are used to achieve assurance are:

- The use of a very careful development lifecycle with trusted tools and extensive verification and validation;

- The independent static analysis and mathematical proof of the software with respect to its specification and known vulnerability classes;

- The use of statistical testing to simulate live operation;

- The challenging of the system with negative testing designed to abuse and misuse the system and

- The compliance with appropriate standards.

These approaches are deployed on protection systems and adjusted depending on the criticality of the system or component that is being assured.

In general, the functionality and trust required from a protection system depends on the quality of the system that it is protecting and the consequences of failure and spurious activation. In the nuclear example, plant design and siting is used to reduce the exogenous and endogenous hazards. For the latter, redundancy (that is using replicated components) and defence in depth (realised by multiple independent barriers) is used to ensure that single failures, or anticipated frequent failures of components, do not lead to costly challenges on the protection system or to needing higher levels of trust than necessary.

In the context of market protection and computer based trading, similar considerations and trade-offs might apply. At one extreme one could have trading constrained in such a way that there is no need for any additional protection (akin to having intrinsic safety in engineered systems) and at the other, an unconstrained approach where there was fast, trusted and powerful protection that enabled complete freedom for the trading approaches (somewhat analogous to unstable aircraft where they can only be flown with continuous computer based control).

In practice one suspects a balanced strategy would be required and indeed a different strategy for different types of hazard. How to decide on a particular approach is outside the scope of this paper, but it illustrates that there is close coupling between:

- How much trust we need in the trading algorithms and platforms;
- How much trust is needed in any protection mechanisms (whether automated or procedural).

As we are concerned with systemic risk, it is likely that different approaches will be required for:

- Single instances of algorithms;
- Collective behaviour across algorithms/participants and
- Cross platform/venue behaviour.

So we could imagine an approach under which the market and venues should be able to tolerate a single rogue algorithm. In addition some as yet un-designed protection could be deployed against hazardous collective behaviours with measures taken to address correlated and cascade failures across markets/venues. Together, these approaches would be shown to present tolerable systemic risks.

If such an approach was adopted then one could foresee trust requirements being articulated for the computer-based trading and the protection systems. These would differ from the nuclear example in:

- The speed of response and functionality of the protection;
- The trust needed in the protection;

- The nature and assurance of the trading algorithms.

The latter concern the rapid rate of adaptation of the algorithms, the development lifecycles, the emphasis on rapid prototyping and back testing to gain assurance, and the risk management via gradual introduction into service. The dependability properties of the algorithm may be very different from the overall functionality e.g. some high confidence in lack of extreme behavior.

Overall conclusions

The nuclear industry and finance industries may seem worlds apart. A nuclear plant relies on decades of science based engineering, the plant is static, physically identifiable, remotely located, each reactor owned and licensed to a single operator with strong incentives to ensure safety and to ensure the remaining risks are tolerable.

The finance industry relies on centuries-old risk concepts, yet is fluid, innovative, and fast changing. Risk taking is an intrinsic part of its day-to-day functioning. Diversity abounds, both in terms of market participants and infrastructure providers. Competition between participants and infrastructure providers drives both innovation and risk taking. Technology allows participants to be present in multiple venues at once.

Yet this industry too is concerned with safety and systemic risk mitigation as well as its impact on the broader economy. Both market participants and infrastructure providers have incentives to ensure the system is robust and inspires confidence. And as described in [Foresight (2011)], the increase of computer-based trading has created new challenges for the industry. These relate to the understanding of the interaction between human traders and computer algorithms (see also [Foresight (2011), DR13], the implications for systemic risk and the development of new risk controls for use by both market participants and infrastructure providers.

In this paper, we have focused on three areas where the issues and practices in the nuclear industry resonate with those raised by the evolution of computer-based trading in financial markets. These are:

- The approaches to systemic risk definition and evaluation.
- The defintion of protection system parameters, risk controls and architecture.
- The need for trust in computer-based systems.

The paper is written for the Foresight project and is constrained to not develop policy recommendations. However, we have identified a number of key questions that we think capture the findings of this study and that could inform future discussions.

Approaches to systemic risk

Looking ahead, one could question whether the rapid development of computer-based trading in financial markets requires the adoption of additional risk concepts and tools. Our analysis suggests that the following questions are worth asking:

1. Is it possible to have a more precise description of risk categories (e.g. in terms of the type of consequences, who is affected, the initiating events that precipitated them)?

2. Is it possible to define precise tolerability criteria? Can one distinguish between tolerable and broadly acceptable risks?
3. Is it possible to define numerical targets? If not, how does one define 'acceptable' risk?
4. Is it possible to develop the notion of a 'design basis,' which would characterise those adverse endogenous and exogenous events that the system (i.e. the market with its control and protective mechanisms in place) should withstand?

Protection systems

In other Foresight reviews (Foresight (2011) DR4) it is agued that financial markets have become complex adaptive systems, in which extreme events can occur in unexpected ways Moreover, as financial markets have become increasingly interconnected, they can be viewed as 'systems of systems.' This means that a failure in one or more constituent parts (market venues) could have widespread repercussions. It also implies that system-level failure is difficult to predict, not only because both humans and computers can adapt their behaviour over time (and can do so at high speed), but also because of the sheer number of possible interactions between humans and computers, both within and across venues. These complexities make it worth asking whether the concept of a protection or viability envelope would be helpful and at the same time these complexities add enormously to the challenge of designing and validating such an approach. We have identified the following specific questions to help articulate these issues:

1. What would the protection and control envelopes look like?
2. What would be the parameters that need to be measured and what would we infer from them? How are they related to existing controls such as price limits or circuit breakers?
3. What would the availability and reliability requirement be for such a system e.g. the probability of failure on demand, the frequency of spurious activation?
4. What is the balance between automation and operator recovery?
5. What additional understanding (and research) is needed given the complex adaptive systems nature of markets? How would the markets adapt to having such protection?
6. What additional analysis techniques and data are needed to assess the risks arising from correlated failures and to design risk controls to guard against their impact?

The questions raised could be useful in further exploring the challenge of developing viability envelopes and designing protection systems.

Computer assurance

Computer based systems, and ICT in general, are of course essential for high-frequency trading and for algorithmic trading more broadly. In all cases, systems will need to be sufficiently trusted: how do we describe that level of trust and how is this evaluated?

Our comparison with the nuclear sector leads to the following questions:

1. What would be the advantages/disadvantages of having an explicit assessment of the trust needed in computer-based systems and prospective protection and control measures?
2. What are the trade-offs between providing protection mechanisms at a venue level vs those on individual users of algorithms?
3. What different levels of trust for individual, collective and cross-market behaviours are required?

4. What software engineering techniques would be appropriate to assure future algorithmic systems?

Our analysis in this paper underlines the importance of trust in computer-based systems. The questions outlined above may be helpful in exploring this topic in the context of financial markets and to assess whether it would be worthwhile to use some of the nuclear assurance strategies and techniques as a basis for innovative approaches in the financial sector.

And finally, although both industries are so different in terms of the culture, technology, regulation, incentives, geography, history, rate of evolution, and their fundamental purpose, the fact that they both focus on societal significant systemic risks has provided the authors with a stimulating perspective on how risks might be evaluated, controlled and communicated in the future.

1. Introduction

Financial markets have evolved to become complex adaptive systems highly reliant on the communication speeds and processing power afforded by digital systems. Their failure could cause severe disruption to the provision of financial services and possibly the wider economy and in this study we consider whether a perspective from the nuclear industry can provide additional insights.

In fact there are a very wide range of areas where the issues and practices in the nuclear industry might resonate with those raised by the evolution of computer-based trading and in this report we focus on:

- The approaches to systemic risk definition and evaluation
- The defintion of protection system parameters, risk controls and architecture
- The need for trust in computer-based systems

The development of the nuclear industry approach to safety has been driven by the need to engineer systems that provide social and economic benefits with tolerable risks, to evaluate and explain the nature and extent of these risks and to provide a framework that allows for scrutiny at varying levels of independence ranging from technical experts within the industry to pressure groups and those who, quite legitimately, hold very different values and worldviews. The Public Inquiry in the 1980s into the Sizewell B PWR provided an unprecedented impetus to articulate and communicate the industry's approach and to provide for public debate of societal risks. More recently the Generic Design Assessment of the proposed new Nuclear Build in the UK has also provided considerable information in the public domain. The regulation and approach of the industry has been strongly shaped by the accidents at Windscale, Three Mile Island and Chernobyl and will be by the events at Fukushima.

The paper is organised in a straightforward manner: the approaches to systemic risk definition and evaluation are addressed in Section 2; the defintion of protection system parameters, risk controls and architecture in Section 3; and the need for trust in computer-based systems in Section 4. Section 5 brings together the key issues and questions that the study raises. We also provide in the Appendices some background material and glossary of terms: they are underlined on first use if they are in the glossary. When referring to the 'nuclear industry' or the 'finance industry', we include authorities, academics and industry participants.

The paper is written for the Foresight project and is constrained to not develop policy recommendations. Furthermore the views expressed in this paper are not those of the Bank of England but the personal views of Dr Anne Wetherilt.

2. Approaches to systemic risk

2.1. Introduction

In this section we examine a number of basic questions: how are the overall risks from a nuclear plant defined and evaluated, and how does this compare with financial markets.

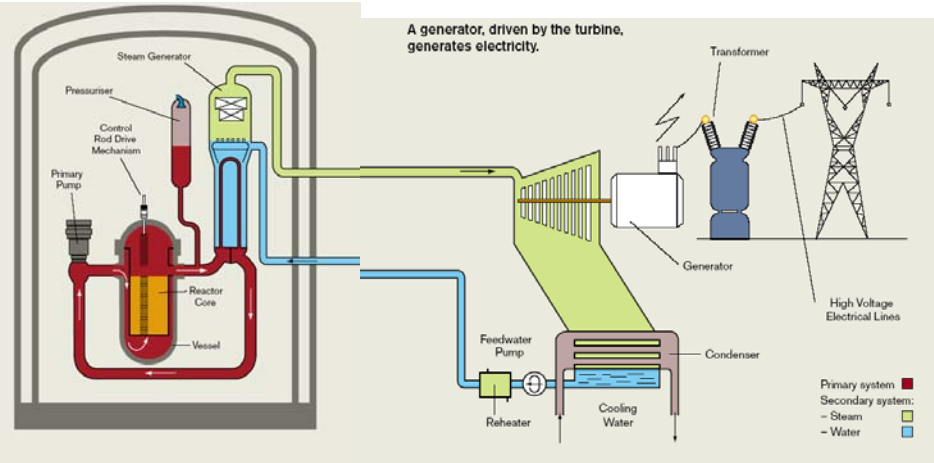
We consider a serious nuclear incident that has the potential for the release of radioactivity with associated plant damage as a “systemic event” and hence make the link to a financial market crash: an event in market that both damages the market and also potentially impacts the wider financial system and the broader economy.

2.2.Nuclear sector

2.2.1 Background

A typical nuclear power station is shown schematically in Figure 1 (Areva 2011). Heat generated from nuclear fission in the reactor core is captured and used to generate steam that drives a turbine - a large piece of rotating machinery. The turbine then drives a generator that produces electricity that feeds into the grid.

Figure 1. Overview of nuclear plant



The nuclear fuel is encased in fuel rods within the reactor core. The reactor core itself is within a thick steel pressure vessel and the overall reactor is isolated from the environment by a number of layers of containment building.

During normal operation automated computer based control systems and operator action maintain the plant in an optimal state. When things go wrong either due to events internal to the plant such as pump failure or operator error or due to external events such as earthquakes the reactor protection systems operate to shut the plant down. The initial part of the shut down would normally involve insertion of control rods into the reactor core to stop the nuclear reaction and associated heat generation. Heat generation can persist for some time and in practice operating the plant into a stable shut down state can be quite a complex operation involving a variety of engineered systems. If the plant is damaged then this can be hard to achieve as was seen at Fukushima. Reactor types vary in how easily they are to control and shut down.

Shutting the plant down and reaching a safe state involves a judicious blend of automation and manual intervention. Where reliable and rapid protective action is required, automatic initiation is required. Where the requirements are less demanding or on a longer timescale, human operator actions or administrative control complement the engineered systems. The UK PWR at Sizewell is designed so that no human intervention is required for the first 30 minutes after the protection system has been activated but that subsequently in an accident situation a range of operator and other expert interaction is anticipated. The balance of automation and control is

much studied and a very important one to get right: on the one hand humans can be the source of error on the other hand their adaptability is essential to deal with unexpected or complex situations, especially in the diagnosis of the problems and the planning of the recovery.

At an abstract level reactor protection systems are quite simple: signals are measured, they are compared to defined limits either directly or after varying degrees of computation, and a decision is made to shut the plant down by dropping rods into the reactor core. Reactor protection systems have to work when needed: they have very high availability and reliability requirements. To achieve this, the architecture incorporates segregation, redundancy and diversity. This is shown schematically in Figure 2.

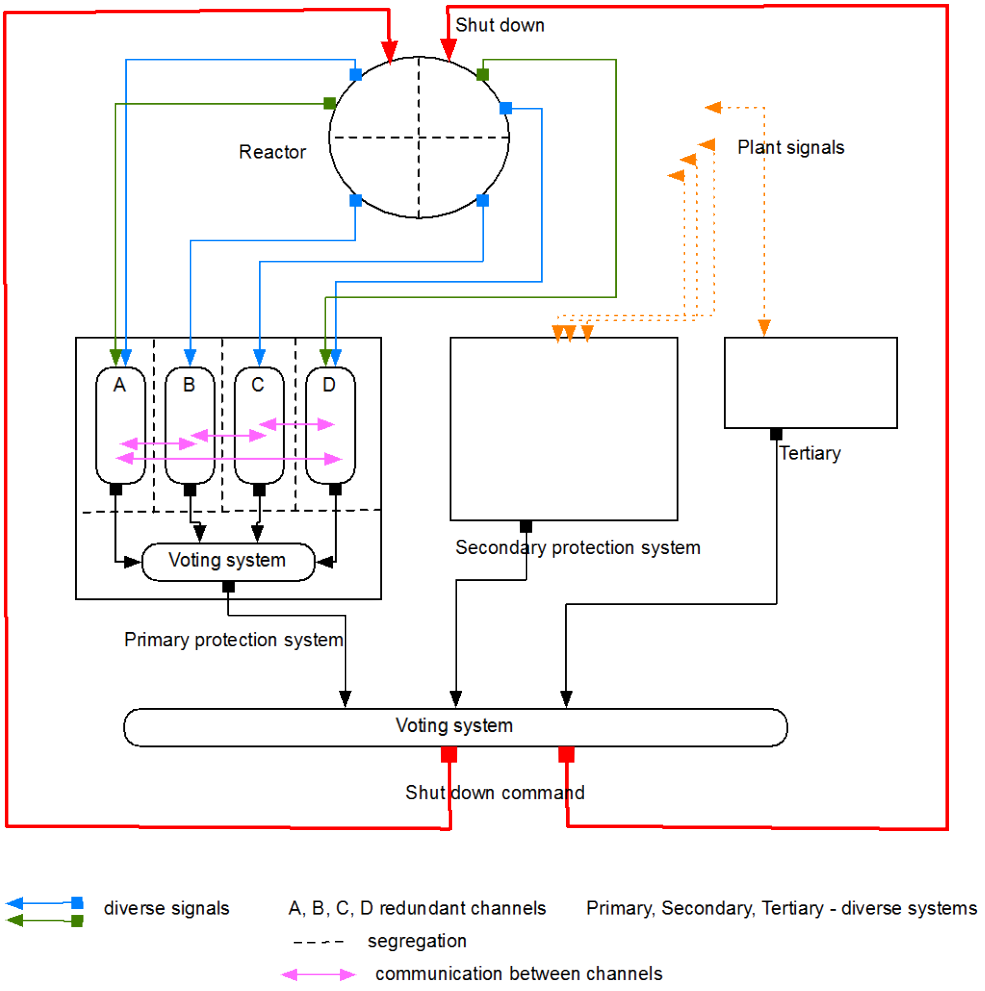


Figure 2. Simplified protection system architecture

For many, the potential for societal consequences of a severe nuclear accident suggests that the requirements for nuclear reactor protection systems are very onerous. However, from the point of view of risk based criteria, the use of diversity, defence in depth and the infrequent demands on the systems means that in the UK the Primary protection systems have had requirements for a probability of failure on demand of 10^{-3} to 10^{-4}. It is difficult to compare these figures directly with devices that have to continuously operate but it is generally accepted that these computer-based probabilistic requirements are modest when compared with those claimed for other safety critical systems (e.g. in avionics and defence applications).

Protection systems are designed to have such authority that they override anything else that is going on in order to force a shutdown of the reactor. This authority means they can also be a source of spurious plant disturbances if they operate when not needed, with consequential economic costs and safety implications. There is a need to define performance measures for spurious activation (e.g. once every 10 years) as well as failures per demand.

The control and protection systems rely on the plant being designed and maintained to support safety. The state of the plant needs to be able to be assessed (e.g. cracks in pressure vessels of critical sizes detected, growth in cracks understood, the idea that systems should leak before break) and again there is much science and engineering devoted to understanding these phenomena sufficiently well to trust the plant e.g. in the integrity of the pressure vessel. Furthermore, the success of the technical systems in preventing accidents, or mitigating those that do occur, is achieved in conjunction with people within a specific organisational and company setting. The wider socio-technical system is responsible for providing leadership, learning from experience and the supporting safety culture.

The nuclear industry in the UK does not claim that nuclear power is zero risk or perfectly safe. Instead there is a framework and supporting evidence that supports a debate on whether the risks associated with the technology are tolerable. Some of this framework is summarised in the following sections in which we explain the basic concepts used in nuclear safety analysis in more detail.

2.2.2 Basic concepts – hazard, risk, accident

Nuclear safety analysis is based on the concepts of hazard, risk and accident.

> "Hazard is the potential for harm from an intrinsic property or disposition of something that can cause detriment..."

So hazard contains the idea of inherent danger, such as in a source of energy or toxicity, as well as a state or "disposition". It is sometimes clearer to qualify the term hazard to clarify this usage i.e. whether it refers to a state of the system or an intrinsic property.

Both endogenous and exogenous events are considered.

External hazards are those natural or man-made hazards to a site and facilities that originate externally to both the site and the process, i.e. the dutyholder may have very little or no control over the initiating event.

> Internal hazards are those hazards to plant and structures that originate within the site boundary and over which the dutyholder has control over the initiating event in some form.

Table 1. Examples of internal and external hazards

External Hazards	Internal Hazards
Earthquake, aircraft impact, extreme weather, electromagnetic interference (off-site cause) and flooding as a result of extreme weather/climate change, terrorist or other malicious acts	Internal flooding, fire, toxic gas release, dropped loads, explosion and resulting debris.

The Health and Safety Executive[1]'s (HSE) publication "Reducing risk, protecting people: HSE's decision making process" (known as R2P2) is an authoritative source for safety concepts and policy. It built on earlier work on "The tolerability of risk from nuclear power stations" (HSE 1992) and it defines risk as

> Risk is the chance that someone or something is adversely affected in a particular manner by a hazard.

The HSE Nuclear Safety Assessment Principles (SAPs) (HSE 2006) use the same definition.

The important point to note is that a risk evaluation has to take into account both the severity and probability of the event. In the context of risks relating to the operation of a nuclear power station, the risks of greatest interest are those associated with radiation; both to individuals and to society:

- *The risk to individuals:* this concerns the risk to the health of any particular individual, worker or member of the public. In this case, the harm may be either in the form of "early effects" or "late effects". Early effects will occur if the radiation dosage is very high, received in a short time and can result in direct death. With regards to late effects, the greatest concern is cancer.

- *The risk to society:* societal risks are those that are wider than those to individuals, having consequences to environment and infrastructure such as loss of electricity and economic loss and were they to materialise, could provoke a socio-political response. For instance, in the case of the Chernobyl disaster, apart from the individuals that died or were affected by the fall-out, there were significant effects to the environment and consequently to the food chain, not only locally but also internationally.

The term accident is also defined and has a meaning close to its general usage

[1] The HSE is the UK's independent health and safety regulator. Its responsibilities for nuclear safety now reside with the Office for Nuclear Regulation (ONR).

> 'accident' includes any undesired circumstances which give rise to ill health or injury; damage to property, plant, products or the environment; production losses or increased liabilities.

Accidents are classified according to an international scale – the International Nuclear and radiological Event Scale (INES) – that considers the off-site effects, on-site effects, and the impact on the defence in depth. There is also the concept in the SAPs of a severe accident in which the highest radiological doses exceed certain targets or unintended relocation of radioactive material within the facility that places a demand on the integrity of the remaining physical barriers. As noted earlier we consider a serious incident that has the potential for the release of radioactivity with associated plant damage a "systemic event", in INES levels this would be a Serious Incident of severity level 3 and above.

To sum up, the nuclear industry safety analysis is based on the concepts of hazard – the potential for harm – as well as the consequences of that potential manifesting itself in an accident. Where possible hazards are removed but if this not possible the decision making, all things being equal, is based not only on the potential consequence but also the likelihood of that consequence. The consequence and probability are combined in the risk.

In this paper we use the term *systemic* to refer to the potential for the release of radioactivity with associated plant damage. Systemic risk, in nuclear terminology, refer both to the consequence and probability of such events.

Having established what we mean by systemic risk the next questions to address are how this is evaluated and how we use this in decision-making.

2.2.3 Probabilistic safety analysis and design basis

The prediction of future risks cannot be extrapolated directly from historic data as there is, thankfully, a lack of data. Instead empirical evidence is combined with imagined scenarios supported with extensive scientific modelling to establish both qualitative descriptions of the risk as well as probabilistic values. Even where there is empirical data (e.g. on large earthquake or flood levels) the risk assessment will rely on models to establish plausible extreme values and their associated probabilities. As with all risk assessments of low probability/high impact events there is considerable use of subjective expert opinion. A major outcome of the risk assessments is an improved understanding of the plant behaviour and the interdependencies and tradeoffs that have been made.

The nuclear safety analysis uses a formal framework for describing the assumptions and scenarios that are considered. Initiating faults are identified that may challenge the safety via a consequential fault sequence. These consider both internal failures such as disintegration of turbine blades or operator error as well as external events such as earthquakes. The analysis is simplified by the consideration of bounding cases that represent extreme consequences and the simplification of the number of accidents that need to be considered by the use of accident classes. The safety analyses of the design justifies that taking a conservative view of the initiating faults (in terms of frequency and size), together with bounding scenarios for the propagation of the event into a class of accidents, provides an overall conservative result. The range of conditions and events that the plant is explicitly designed to withstand without exceeding safety limits are known as the Design Basis.

The safety analyses are supported by Probabilistic Safety Analyses (PSA) that seek to establish the probability of a certain undesirable events which, when combined with knowledge of the potential consequences, allows the risks of nuclear accidents to be calculated. It proceeds by a mixture of top down analysis (e.g. using a Fault Tree to analyse “how can this pipe break”) as well as bottom up that takes the occurrence of a hazard and seeks to trace the consequences (e.g. using Event Tree analysis to assess what happens if “this control system fails”). It requires a comprehensive analysis of the overall plant design and operation together with detailed evaluation of failure modes of components and their probabilities. It covers all initiating faults that are in the design basis. It is performed using best-estimate methods and data to demonstrate the acceptability, or otherwise, of the plant risk. The PSA also seeks to demonstrate that a balanced design has been achieved. This means that no particular class of accident of the plant makes a disproportionate contribution to the overall risk e.g. of the order of one tenth or greater. The PSA provides information on the reliability, maintenance and testing requirements for the safety systems and the rigour required of operating procedures.

The PSA provides a means of claiming that the risk is lower than the intolerable region established by the SAPs (see also Figure 3). The PSA can then be used to demonstrate that risks are As Low As Reasonably Practicable by investigating the effect on plant risk of modifying the plant safety provisions. Another important role of the PSA is to establish how much trust needs to be placed in the functions that the various safety systems perform (e.g. in terms of the Safety Category and the probability of failure on demand) and their allocation to the different safety systems.

The output from the safety analyses is summarised in a Fault and Protection Schedule that details for each fault sequence in the Design Basis the impact on the plant and how it will be protected and mitigated. A simplified example is provided in Box 1.

Box 1. Nuclear fault and protection schedule

Simplified Fault and Protection Schedule (excerpt)								
		Functional Level			Equipment		Equipment	
Fault description	Frequency	Safety Function	Plant Level	Cat	Primary Protection System	Class	Secondary (diverse) Protection System	Class
Partial loss of coolant flow	10^{-3} to 10^{-4}	Reactivity Control	Shut down and maintain core sub-criticality	A	Reactor trip on low flow rate in one loop	1	High pressure Boron injection on ATWS signal	2
		Heat Removal	Transfer heat from the reactor coolant to the ultimate heat sink	A	Steam generator and emergency feedwater control	1	Heat removal by Emergency Core Cooling System (ECCS)	2

The above table provides an example of a summary Fault and Protection Schedule. The first column describes the type of accident that in this example is a partial loss of coolant flow. This requires two types of safety function, the first to control the reactivity of the core and the amount of heat being produced and the second a safety function to remove the heat from the core. These functions are classed as Category A – the highest safety category. The next part of the table summarises the equipment that will implement these Safety Functions. There is a Primary Protection System that will drop the rods in to the core and control the turbine and feedwater so as to continue to extract heat. If this Primary System fails, the diverse Secondary Protection System detects this condition (known as ATWS – Anticipated Transient Without Scram) and uses a diverse method of controlling the reactivity via injecting boron. It also activates a different Emergency Core Cooling System. The Secondary System has a lower safety class as it is only needed when the Primary System fails.

In an actual Fault Schedule there will be additional details of the protection functions and cross-references to the safety analysis and other faults as well as

details of the plant state when the accident occurs.

However, there are some aspects of safety that the SAPs recognise as not readily amenable to simple quantification of failure. The role of human factors (at an individual, group and organisational level) in achieving safety and initiating accidents is hard to quantify meaningfully, especially when knowledge based activities are concerned. Similarly the contribution of good management practices is hard to assess, although research and some progress has been made in this area. Other areas identified that are difficult to quantify, are common mode failures and failures due to design faults or specification omissions including software faults.

In dealing with the necessary uncertainties and incompleteness in nuclear system analyses it is useful to distinguish between epistemic and aleatory uncertainties (Oberkampf (2004), Littlewood (2010)). Epistemic uncertainty concerns incompleteness in our knowledge about the world e.g. in the models and reasoning that we use to estimate and predict reliability. Epistemic uncertainty is in principle reducible, e.g. by collecting more and better evidence concerning the subject of the uncertainty. There is also uncertainty associated with what we see as random processes in the world: quite when will a component fail, when will it rain. This is "uncertainty in the world" and is irreducible2. One consequence of the recognition of epistemic uncertainties is in the use of claim limits to prevent unreasonable low figures being claimed for the probability of failure of sub-systems (IAEA (2000), HSE (2006)). While it is very hard to justify the limits chosen the concept is appealing.

The existence of a formal framework for nuclear safety analysis does not of course guarantee that it will be applied appropriately or that epistemic uncertainties will be properly addressed. Although there are significant differences in how safety is assessed in the UK and Japan, the recent accident at Fukushima should reinforce our scepticism and emphasise further the need for challenge and continuous review of safety analyses. The appreciation of the risks from tsunamis and earthquakes is deeply embedded in the culture (and art) of Japan. Yet, the apparent shortcomings in Fukushima's risk controls remind us that constant vigilance is required in order to ensure that there are no significant 'blind spots' in an industry's thinking about systemic risk[3].

2.2.4 Tolerability of risk and ALARP[4]

The overall approach to safety and risk evaluation in the nuclear industry in the UK is described in the HSE Safety Assessment Principles (SAPs) (HSE 2006). These are the primary principles that define the overall approach to be followed and are consistent with international guidance from the International Atomic Energy Agency and are supported by more detailed Technical Assessment Guides (TAGs) and other guidance documents. The SAPs are based on eight

[2] The term comes from aleator – a dice player.

[3] See also possible problems from normalization of deviance as discussed in Foresight (2011).

[4] This section is based on a report for the UK Nuclear Safety Advisory Committee (Littlewood (1998)) The HSE provides additional detail on the definition and interpretation of ALARP in (HSE (2010)).

fundamental principles that address, responsibility, leadership, understanding, fairness, inter-generational issues, protection, recovery and mitigation.

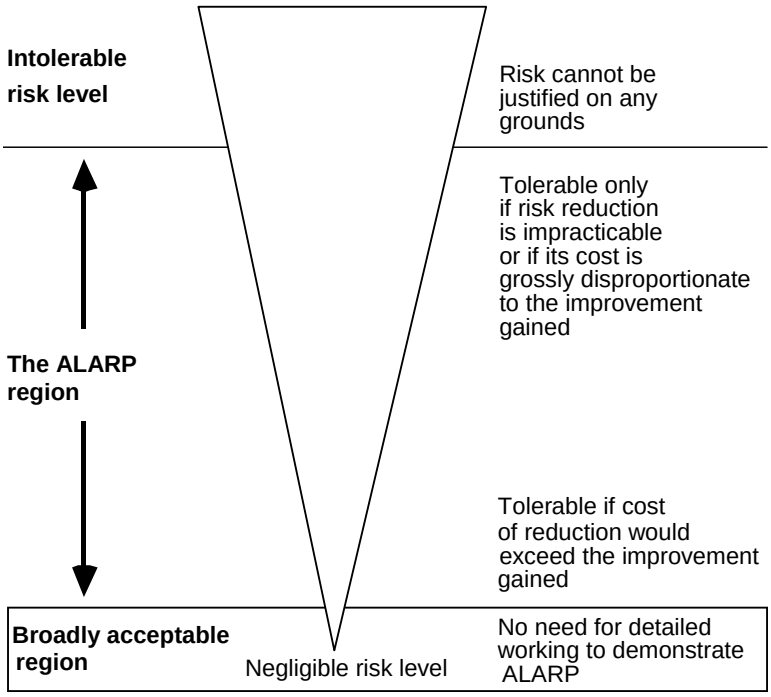
It is important to note that the Principles explain *what* has to be done, but do not become prescriptive as to *how* these requirements will be met: they reflect a goal-based approach to regulation. They provide flexibility to the duty holder to achieve what is expected by the SAPs. It is then a matter for the safety case to demonstrate adherence to the SAPs. In Section 4 we discuss the application of these principles to computer-based system.

The risks to workers and the population at large are considered both in normal operation of nuclear installations and in accident conditions and have to be reduced to a level of "tolerability". Tolerability does not mean "acceptability"—it refers to a willingness to live with a risk to yield certain benefits so long as there is confidence that it is properly controlled. The design of nuclear installations and their supporting safety systems focuses on minimizing and controlling risks. Cost and rigour of activities must be proportionate to those risks. Calculations of risk, taking into account severity and likelihood, have to then demonstrate that the risk is appropriately mitigated to a tolerable level.

The "ALARP principle": the principle that certain nuclear risks have to be demonstrated to be "As Low As Reasonably Practicable" is key to discussing risks and the stopping rules associated with additional design and operational measures. The ALARP principle is based on the assumption that it is possible to compare *marginal improvements in safety* (marginal risk decreases) with the *marginal costs of the risk reduction measures*. Nuclear risks may offer this possibility when they are quantified (i.e. in terms of event probability and of radiation releases), and when the failure rate improvements of the systems controlling the relevant events can be evaluated. Note that the application of the ALARP concept does not necessarily need a quantification of risk reduction. For example, the simple addition of a further safety feature, which costs relatively little, may be obviously worthwhile—qualitative judgements of this nature can often be readily made. Also, marginal does not mean one just considers incremental or small perturbations to the design: sometimes creative design changes (e.g. substitution of hazardous materials with benign ones) are needed to justify that the risks have been reduced to ALARP.

ALARP found its expression in the well known 'carrot diagram' (see Figure 3 below), which has become the standard means for the exposition of the principle. There are two significant boundaries: the upper one beyond which risks are not acceptable at all and cannot be justified on any grounds and a lower one beyond which risks are considered negligible and no detailed assessment is required. Regulators would not usually require further action to reduce risks unless reasonably practicable measures are available. Within these two boundaries is the ALARP region. At the upper, more risky end, of the ALARP region the risks are only tolerable if costs are judged *grossly* disproportionate to the risk reduction gained.

Figure 3. The ALARP principle: levels of risk are divided into three bands. Width of wedge represents level of risk.



A key part of assessing tolerability is the effective communication of safety and risk. The nuclear industry, in keeping with other high hazard sectors, uses the concept of a safety case to facilitate this.

2.2.5 Safety cases

The nuclear safety case is defined by the HSE as

> "…the totality of documented information and arguments which substantiates the safety of the plant, activity, operation or modification in question. It provides a written demonstration that relevant standards have been met and that risks have been reduced as low as reasonably practicable (ALARP)."

The licensee is legally responsible for the safety case. Given the magnitude and complexity of the legislative and technical requirements that have to be met, safety cases have to be structured in a logical manner and be demonstrably adequate. The safety case has to support an argument that the requirements placed upon it are met. As such, the safety case contains *claims* about the properties of the system and, following a systematic approach, demonstrates that these claims are substantiated by *evidence* (see Box 2). Safety cases can be seen to support the following (Bishop 2010):

- *Reasoning and argumentation.* A safety case can be seen as an over-arching framework that allows us to argue whether the claims are substantiated by the evidence. The case might be mainly narrative, using prose to explain the connections between claims and evidence. However cases deal with highly technical subjects and hence they might use specialist notations from the particular discipline concerned (e.g. from fluid mechanics, computer science). The case will then integrate a selection of technical analyses and other evidence

using a formal or graphical notation to show whether the claims have been met; how the evidence is integrated; and the overall structure of the case and the thrust of the argument.

- *Negotiation, communication, trust.* The safety case represents a boundary object between the different stakeholders who have to agree (or not) the claims being made about the system. To this end it has to be detailed and rigorous enough to effectively communicate the case and allow challenges and the subsequent deepening of the case.

Box 2. Safety and assurance cases: claims, arguments, evidence

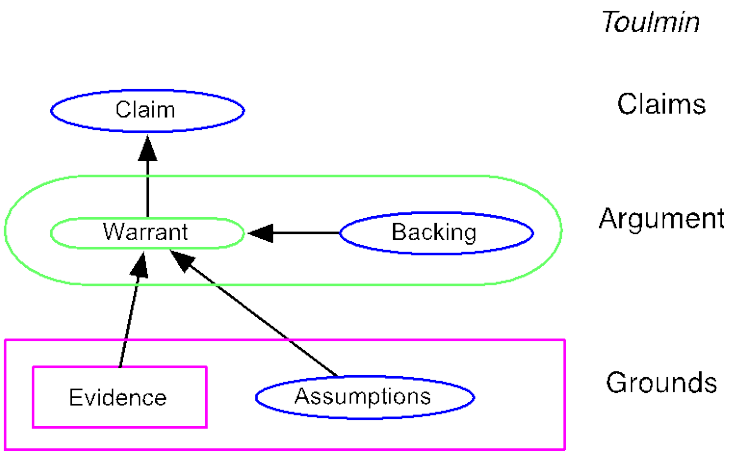
A system safety case is now a requirement in many safety standards and regulations. Explicit safety cases are required for military systems, the off shore oil industry, rail transport and the nuclear industry. An early definition of a safety case (Bloomfield 1998) was

*"*a documented body of evidence that provides a convincing and valid argument that a system is adequately safe for a given application in a given environment*".*

More recent definitions make explicit the concept of structured argumentation

"A structured *argument*, supported by a body of *evidence*, that provides a compelling, comprehensible and valid case that a *system is safe* for a given application in a given environment"

Current safety case practice makes use of the basic approach that can be related to the concepts developed by (Toulmin 1958) where claims are supported by evidence and a "warrant" or argument that links the evidence to the claim. There are variants of this basic approach that present the claim structure graphically such as Goal Structuring Notation (GSN) (Kelly 2004) or Claims-Argument-Evidence (CAE) (Bloomfield 1998).



As noted in Section 2.2.1 defence in depth is an important principle so that there is not over reliance on any single system. There is an analogous approach within the safety case in which diverse arguments and evidence are used to support key claims. These arguments are sometimes called the 'legs' of the safety case and are based on different evidence. Another important feature of the safety case process is independent assessment: both within the organisation responsible for it as well as by the regulator. Safety case reports are produced that provide sufficient detail of the claims, arguments and evidence to enable this independent review. A nuclear plant has many subsystems and the safety documentation will be a series of linked cases for different aspects of the plant and its operation. The objective of independent assessment is to ensure that more than one person or team sees the evidence so as to overcome possible conflicts of interest and blinkered views that may arise from a single assessment.

Although the nuclear industry was instrumental in establishing the need for safety cases, the incorporation of software safety case requirements in the UK defence standards helped drive interest in structured safety cases, and other forms of assurance case. The UK Civil Aviation Authority has a goal based software regulation and more recently the FDA in the US has required Assurance Cases to be developed for medical infusion pumps. There is an international standard ISO/IEC 50126 on the basic concepts and work within the OMG on standardisation of interchange formats.

2.2.6 Numerical targets

A significant aspect of the nuclear safety approach is the role of numerical, probabilistic, targets. The individual risk of death levels in "Reducing risk, protecting people: HSE's decision making process" (HSE 2001) cover risks to workers and to members of the public from activities on the site and are summarised in Table 2.

Table 2. Numerical risk targets for workers and the public

Boundary	
Boundary between the 'tolerable' and 'unacceptable' regions for risk entailing fatality	Worker: 1 in 1,000 pa Member of the public: 1 in 10,000 pa
Boundary between the 'broadly acceptable' and 'tolerable' regions for risk entailing fatality	Worker: 1 in 1,000,000 pa Member of the public: 1 in 1,000,000 pa".

There is also a consideration in the safety analysis of societal risks arising from severe but very unlikely accidents. The SAPs note that the nature of radioactive release from a major accident at a nuclear site will mean that long term, large distance stochastic effects are important. As a measure of the societal concerns that would result from a major accident, a representative target has been defined based on an accident leading to an immediate or eventual 100 or more fatalities. The target does not cover all the factors related to societal concerns so that in making an ALARP argument other societal effects must also be considered.

The distinction between societal effects, risk and concerns is explained in Table 3.

Table 3. Concerns, effects and risks

	Definition
Societal concerns	Societal concerns are the risks or threats from hazards which impact on society and which, if realised, could have adverse repercussions for the institutions responsible for putting in place the provisions and arrangements for protecting people.
Societal effects	A term used to describe those societal concerns that are capable of quantitative prediction such as numbers of deaths or injuries, numbers of people evacuated, area of land contaminated and general economic loss.
Societal risk	The risk of an accident causing the death of a specified number of people in a single event from a single major industrial activity, i.e. an activity from which risk is assessed as a whole and is under the control of one company in one location, or within a site boundary.

Table 4. Target for total risk of 100 or more fatalities

The targets for the total risk of 100 or more fatalities, either immediate or eventual, from on-site accidents
Basic Safety Level (BSL): 1 x 10^{-5} pa (level that all new plant should reach) Basic Safety Objective (BSO): 1 x 10^{-7} pa (boundary of broadly acceptable region)

The consequence of normal operation and accidents are studied in detail and the targets are set for events that within or on the boundaries of the plant e.g. targets for the frequency of core damage due to internal hazards of 10^{-6}p.a. (HSE 2011).

2.3.Financial markets

2.3.1 Background

Financial markets are a key component of the wider financial system. The OECD defines the latter as follows:

> The financial system consists of institutional units and markets that interact, typically in a complex manner, for the purpose of mobilizing funds for investment, and providing facilities, including payment systems, for the financing of commercial activity (OECD, glossary of statistical terms)

Financial markets themselves are varied, both in terms of their economic function, and their organisational characteristics.

In line with the overall aim of the Foresight project, this paper focuses on those markets that are referred to as continuous auctions (Foresight (2011)), with most examples taken from stock exchanges. It is in these markets that computerised trading is most prevalent and the widest range of algorithms have been developed. Algorithmic trading is growing, however, in other auction-type markets, such as FX trading platforms (BIS (2011)).

2.3.2 Basic concepts: market crashes and systemic risk

This Section starts by defining a number of key concepts. First, in common with the nuclear, industry, the finance industry employs key risk concepts:

> Operational risk is defined as the risk that shortcomings in information systems or internal processes (including human error) could result in unexpected losses (CPSS (2003)).
>
> Market risk can be defined as the risk of a change in the value of a portfolio of assets arising from changes in the value of the underlying assets.
>
> Systemic risk is broadly defined as the risks affecting the functioning of the financial system as a whole (see Box 3).

A market crash is an extreme event, whereby prices fall by "very large" amounts. Although there is no commonly accepted numerical definition of a market crash, studies of historical market crashes reveal the following characteristics:

- Market crashes constitute large price falls. For example, on Monday October 19, 1987, the Dow Jones Industrial Average fell by 24%.

- They are a market-wide phenomenon (a large price fall in a single stock does not constitute a market crash).

- They are often triggered by relatively small events. In a much quoted study, Cuttler, Poterba and Summers (1989) show that extreme stock price movements rarely coincide with significant economic news. In fact, many of the largest price movements in their sample occurred on days with no notable news events.

- They may be relatively short-lived events (e.g. 1987) or they may play out over a longer time period (e.g. 1929).

Very large price falls result in a significant reduction of economic wealth (e.g. the assets held by pension funds). But for such large price falls to be considered a systemic event (see box 3), a number of characteristics must apply:

1. The large price fall undermines confidence in the market, so investors withdraw, possibly for an extended period. Reduced participation may affect the ability of a financial market to fulfil its essential economic functions of price discovery and risk sharing.
2. The large price fall 'spills over' into other markets via so-called contagion effects. Hence, a wide range of markets may be impaired in their essential functions.
3. The large price fall affects the funding decisions of systemically important financial institutions.5 Institutions may find it difficult to raise new funding. Hence, the contagion spreads from markets to institutions.

Note that there is a link between institutional funding needs and market functioning. A financial entity needing to sell assets when prices are falling may find itself in a positive feedback loop: the more assets it needs to sell, the greater its contribution to the overall selling pressure, and the larger the price fall, the more assets it needs to sell in turn (Brunnermeier and Pedersen (2009)). This downward spiral is one of the possible characteristics of a systemic price fall.

Note also that while operational and systemic risk are distinct concepts they are not independent classes of risk. For example, a technical outage which lead to a halt of trading would not be considered a systemic event, unless of course it was accompanied by or led to the conditions described in (i) – (iii). At the same time, if market participants had concerns about the controls in place to manage operational risk (either at the level of a venue or its participants), then they may be less inclined to trade, thus reducing liquidity in a given venue, and affecting its ability to function as noted above.

[5] FSB (2009) defines a systemically-important institution as financial institutions which provide critical functions in the financial system (p. 6).

Box 3. Defining financial stability and systemic risk

Definitions of financial stability vary: some refer in broad terms to safeguarding the core functions of the financial system; others emphasise the occurrence of specific financial crisis events. As an example of the former, Haldane et al (2004) define financial instability as 'the deviation from the optimal saving-investment plan of an economy that is due to imperfections in the financial sector.' Tucker (2011) states 'financial stability prevails where the financial system is sufficiently resilient that worries about bad states of the world do not affect confidence in the ability of the system to deliver its core services to the rest of the economy.'

The Financial Stability Board (FSB,(2009)) defines a systemic event as a 'disruption to the flow of financial services that is (i) caused by an impairment of all or parts of the financial system and (ii) has the potential to have serious negative consequences for the real economy.' The FSB further outlines three broad criteria to assess systemic importance (which apply both to financial institutions and financial markets):

- Size (which captures the amount of financial services provided);
- Lack of substitutability;
- Interconnectedness (which captures both direct and indirect interlinkages)

When applying these criteria to financial markets, FSB (2009) notes that it is difficult to assess their systemic importance independently, as this depends on the systemic importance of institutions that participate in these markets. Moreover, financial markets are often characterised by network effects, which effectively concentrate trading, thereby increasing the size of particular markets and reducing possibilities for substitutability.

A related concept is systemic risk. The Office of Financial Research (Bsiais et al (2012)) defines systemic risk as 'any set of circumstances that threatens the stability of or public confidence in the financial system.' IMF (2010) defines systemic risk as 'the large losses to other financial institutions induced by the failure of a particular institution due to its interconnectedness.'

A more detailed examination of the literature reveals a long list of factors contributing to systemic risk, including contagion, asset bubbles, information disruption, feedback behaviours, negative externalities etc. (Bsiais et al (2012)). Hence, many studies highlight the difficulty of measuring systemic risk: 'it is there when we see it' (IMF (2009). Building on the substantial literature in this area, Bsiais et al (2012) propose a total of 31 quantitative measures of systemic risk. These include measures of financial interlinkages, indicators of credit risk, and results of stress testing, to name but a few.

Apart from concerns about 'fuzzy measurement,' as termed by Borio and Drehman (2009), systemic risk metrics often fail to detect financial distress. Borio and Drehman also point out that periods of instability may persist for a long time, without a financial distress event occurring.

When thinking about the sources of systemic risk, one can distinguish between three broad scenarios:

(i) a wide range of institutions and markets being exposed to a common shock;

(ii) an initially idiosyncratic shock which subsequently spreads to a wider range of institutions and markets and

(iii) the gradual build up of an imbalance such as an asset price or a credit bubble, which subsequently unravels and affects a broad range of institutions and markets.[6]

A further distinction can be made between exogenous and endogenous sources of systemic risk. The latter can be defined as the risks attributable to the actions of risk-averse market participants.[7] For example, market participants may respond to a change in exogenous risk by selling assets. But this in turn may lead to sharp asset price falls, thus increasing the perceived risk of holding such assets, thereby further increasing endogenous risk.

Market crashes are not a new phenomenon, as noted above. But there are concerns that the widespread use of automated trading strategies may have made financial markets more vulnerable to such crashes. In that sense, the May 6 'Flash Crash' is viewed by some as a signal of increased market fragility (Foresight (2011)). A key development is that market crashes can no longer be attributed to the behaviour of human traders alone. Indeed, it is not inconceivable that a market crash could be triggered by a computer algorithm, propagated by high-frequency trading, and eventually leading to a series of automated feedback loops (Foresight (2011)).

So far the discussion has focused on the concept of 'systemic risk' in general, and market crashes in particular. Related concepts include market integrity and market efficiency. These are defined as follows by IOSCO:

- 'Market integrity is the extent to which a market operates in a manner that is, and is perceived, to be fair and orderly, and where effective rules are in place and enforced by regulators so that confidence and participation in the market is fostered' (IOSCO (2011a), p.8).

- 'Market efficiency refers to the ability of market participants to transact business easily and at a price that reflects all available information. Factors considered when determining if a market is efficient include liquidity, price discovery and transparency' (IOSCO (2011a), p8).

Another much-used term is 'orderly' market,[8] or alternatively, the need to avoid a 'disorderly market.'[9] These terms are rarely defined, but generally refer to the absence of large order

[6] See e.g. Schwaab et al (2011).

[7] See e.g. Danielson et al (2009).

[8] See e.g. FSA (2010), The FSA's markets regulatory agenda, May. See also ESMA (2011).

[9] See e.g. MiFID, recitals 46 and 48.

imbalances, erroneous orders, system overloads, large price swings, or other events that may disrupt the trading process.

Traditionally, securities markets regulators have focused mostly on market integrity. Recent amendments to the international guidelines for securities markets regulation, however, acknowledge the importance of systemic risk (see Box 4).

Box 4. Regulation of financial markets

The international guidelines for securities markets regulation are contained in IOSCO's *Objectives and Principles of Securities Regulation* (IOSCO (2010)). Their main aim is to ensure the integrity of trading by:

- Requiring that exchanges and trading systems are subject to authorization and oversight;
- Maintaining fair and equitable rules;
- Promoting transparency;
- Detecting and deterring market manipulation;
- Seeking to ensure the proper management of large exposures;

In July 2010, IOSCO adopted a new principle that acknowledges the need for securities regulators to monitor, mitigate and manage systemic risks arising in securities markets, in addition to their traditional micro-prudential objectives (listed in the previous paragraph).

This systemic risk principle covers a wide range of issues that may affect systemic risk, including financial innovation, product complexity and interconnectedness (between institutions and markets)

Liquidity features in IOSCO's earlier-mentioned definition of market efficiency, but is also related to systemic risk. A market crash is an instance of extreme illiquidity (see Box 5 for a definition and common measures of liquidity).

Box 5. Defining and measuring market liquidity

Definition

Market liquidity is defined as the ability of a market to process buy and sell orders with minimal delay and price impact. In a liquid market:

(i) buy and sell orders can be matched with minimal delay;

(ii) buy and sell orders can be matched without requiring significant price concessions from either party;

(iii) buy and sell orders of varying sizes can be processed.

The concept of market liquidity is often associated with the absence of market frictions. In an illiquid market, participant interest may be unbalanced (more sellers than buyers), or it may only be possible to process relatively small trade sizes.

Liquidity is conventionally measured by the bid-ask spread, the difference between the price at which a market participant (e.g. a market maker) is willing to buy and the price at which he is willing to sell. Bid-ask spreads are narrow when the market is liquid. They widen when markets become less liquid, e.g. when there are more sellers than buyers.

Metrics

Financial economists have developed a wide range of liquidity metrics that can be used to describe a market under stress. Each captures different aspects of market liquidity. The choice of metric is often determined by data availability (some use transaction-level data, some use data on both prices and trade sizes).

A few illustrative examples:

(i) Price impact: based on Kyle (1985), this popular measure is obtained by regressing transaction prices on transaction sizes. The price impact coefficient measures the transaction volume required to move the price by one unit of price (say one dollar).

(ii) Order flow toxicity: developed by Easley et al (2011), this measures the presence of informed traders in a market. The intuition behind this measure is that uninformed investors typically lose money when trading with more informed counterparties, and will reduce their trading activity, thus resulting in lower liquidity.

(iii) Returns on market making: Khandany and Lo (2011) document significant falls in the returns on simulated portfolios associated with market making during a period of reduced market liquidity.

Empirical evidence shows that these metrics significantly deteriorate during periods of market stress. For a recent summary of such metrics, see Bsiais et al (2012) and Billio et al (2010).

Two final terms of use are operational resilience, which refers to the operational capacity of a trading venue (e.g. whether its systems can cope with sudden surges in transaction volumes), and liquidity resilience, which refers to its ability to remain liquid during periods of market stress. Both are important objectives in their own right, but as will be discussed in more detail in Section 3, are also important mitigants in reducing systemic risk.

The increased reliance of computer-based systems in financial markets has led some to worry abut operational resilience, in particular the massive increase in message traffic, and the widespread use of order cancellations. There are also questions about the reliability and impact of algorithms (often untested in their interaction with others), and the greater potential for erroneous algorithms to disrupt the market (Foresight (2011)).

Worries about liquidity resilience focus on the specific role of high-frequency traders; whether they contribute to liquidity provision, both during normal and stressful market conditions, and whether they discourage others from providing liquidity. Some are also concerned about the impact of algorithmic trading more broadly on market liquidity (see e.g. Kay (2012)).

2.3.3 Probabilistic safety analysis – the finance approach

This and the next Section explain how the finance industry measures and manages the risk arising from very large price changes. In this Section we summarise work describing past market crashes and how this is used to infer the probability of such crashes occurring in the future. We then look, in the following Section, at the models used to measure these risks.

Much like the nuclear industry, the finance profession has developed a quantitative approach towards the measurement of market and systemic risk in general, and market crashes in particular. By and large, the profession relies on historical data to measure both the probability and the impact of the very large price changes that may constitute a market crash.

The literature on market crashes is closely related to that on asset price bubbles. A bubble is defined as an instance where asset prices deviate for a prolonged period of time from the value justified by underlying fundamentals (such as the growth prospects of a company). An asset price bust occurs when prices rapidly return to their fundamental value (the bubble bursts), or in some cases even fall below levels justified by fundamentals. The consensus in the literature is that asset price bubbles and busts are difficult to predict, even though some long-term indicators (such as rises in credit availability) are in some instances associated with subsequent asset price bursts.[10]

Some of the work describing historical market crashes is descriptive and looks at the big picture – how often crashes occur and what their economic impact is. Bordo (2003) describes 20 stock market crashes in the US and 17 in the UK, between 1800 and 2000. Barro and Urusa (2009) use long-term data (in some cases going back to the 1930s) ending in 2006 and report 232 stock market crashes in 30 countries.

These studies typically measure the extent of the market crash using low-frequency return data. For example, Barro and Ursua (2009) consider the peak-to-trough movement in stock returns, with a multi-year return of -25% or lower classified as a market crash. Bordo (2003)

[10] See e.g. Kannan et al (2009).

uses 20% as his cut-off point. Including war periods, Barro and Urusa (2009) report that in 124 out of 232 stock market crashes the total price fall was greater than 40%, with 79 of those greater than 50%.

When measuring impact, these studies focus on broad economic measures. Indeed, a significant number of financial market crises were associated with economic recessions. Hence, the fall in economic output (measured by GDP) is frequently used as a measure of the cost of a financial crisis. Increased public debt is another measure. Barro and Urusa (2009) use the decline in per capita consumer expenditure as their measure of economic recession, and find that 71 of the above-mentioned crashes coincided with or were adjacent to economic recessions.

Other studies describe the statistical properties of large price changes at a higher frequency (say daily or intraday). A key question is whether price dynamics observed during a market crash are fundamentally different from those in normal times (see Box 6). As explained in the Box, one approach is to measure the statistical frequency of very large price changes. This critically relies on the choice of the frequency distribution for price changes. Alternatively, one can develop a non-parametric approach, which does not require such a choice.

Box 6. Understanding the statistical properties of extreme price movements

Empirical studies of stock price changes (or returns) show that these are not well described by the normal distribution (see e.g. Campbell et al (1996)). In particular, a number of stylised facts appear at a range of frequencies (e.g. daily, weekly, monthly, or even intraday):

(i) Volatility of returns varies over time;

(ii) Volatility of returns is clustered, i.e. prolonged periods of higher than average volatility are observed, as well as periods of lower than average volatility.

(iii) Return distributions display heavy tails, so very large price changes are more frequent than indicated by the normal distribution;

(iv) Very large price changes tend to cluster.

Hence, it is commonly accepted that stock price returns can be described more accurately by a power law distribution, at least for the earlier-mentioned frequencies. This implies that:

(i) Both small and large price changes can be expected. In other words, very large price changes are not exceptional.

(ii) Both small and large price changes are drawn from the same distribution. Hence, *ex ante* it is not possible to distinguish between these two types of events. In other words, large price changes cannot be predicted.

But some studies view very large price changes as extreme occurrences or outliers, which lie beyond the fat tail of a power distribution (see e.g. Sornette (2003)). Moreover, these studies show that extreme price changes are drawn from a different distribution than those describing smaller price changes. It then follows that the extreme events could be more frequent than suggested by a power law distribution.

As an example of non-parametric work, Johanssen and Sornette (2001) measure the accumulated loss over a period of consecutive daily price falls. Thus they capture the difference between the local maximum (the highest observed price fall) and the local minimum (the lowest observed price fall). They refer to this metric as 'drawdown.' In other words, the drawdown metric quantifies the loss an investor would make if he had bought at the most recent maximum price and sold at the next minimum price. The authors find that historical market crashes have all been preceded by such a drawdown. They also confirm that drawdowns cannot be explained within the power-law framework.

Sornette's work further shows that stock market crashes are not isolated events that occur very rarely and are entirely unpredictable. Instead, he and others argue that these large price falls mark the end of a (sometimes very long) period of instability during which prices rose to unsustainable levels (see also Sornette and Johanssen (2001)).

A separate question is whether the power law distribution adequately describes price changes at very high frequencies (e.g. measured over seconds or milliseconds, rather than minutes). In recent work, Johnson et al (2012) examine extreme price falls that occur over very short time intervals (less than 25 milliseconds) for the period 2006-2011. They find that the power-law distribution no longer describes price changes at these frequencies: price changes that last about 1 second can be described by a power law, but as the duration falls, and arguably human traders can no longer process the data, the distribution ceases to be a power law.

In sum, there is considerable academic evidence that some extreme price movements exhibit distinctive statistical properties. Whether and how these might be used to assess future risks is very much a research challenge.

2.3.4 Tolerability of risk, numerical targets & the design basis – the finance approach

This Section starts by discussing how participants in financial markets measure the risks arising from large asset price falls. This is done by looking at methods to measure market risk in investor portfolios and the regulatory framework applying to firms participating in financial markets. The Section will then look at the approaches towards measuring and evaluating large stock price falls from the perspective of the market as a whole.

Financial market participants use probabilistic methods to quantify the risks arising from large, unexpected price falls. This involves assessing the probability of large price falls, measuring their impact on asset portfolios and setting risk tolerance levels. In part, individual firms' risk tolerance levels will be determined by regulatory capital requirements. Indeed, for financial institutions subject to the Basel Committee's capital framework, this means setting aside regulatory capital that reflects market risk, amongst other risk factors.

A much-used modelling approach is Value at Risk (VAR). This measures the maximum potential loss in value of a portfolio of assets over a given period of time for a given confidence level. For example, a ten-day 99% VAR measures the maximum expected loss on a portfolio, considering 99% of possible losses, as measured over a two-week period. In other words, VAR describes a quantile of the loss distribution. VAR methodology is an essential part of firms' risk management and underpins the Basel II capital framework for market risk.[11]

But VAR is associated with 'normal' price falls. Indeed, traditional VAR models rely on the normal distribution. Hence, they are less suited to capture the extreme price falls that constitute market crashes.[12] More recent risk measurement models recognise the need to incorporate extreme price falls. These alternative approaches include:

(i) Stressed VAR: This is a VAR measure which is based on alternative assumptions underpinning the VAR calculation (e.g. changing the correlations between the prices of the assets comprising a portfolio).

Under its revised framework for market risk, the BCBS requires firms to supplement their standard VAR with a stressed VAR, which accounts for price movements observed over a continuous 12-month period of significant market stress (BIS (2009)).

(ii) Expected Shortfall, defined as the expected loss arising from price moves beyond the VAR threshold.[13]

Additional insights are provided by research in Extreme Value Theory (EVT) which considers a wider range of theoretical distributions with fat tails.[14] EVT also offers empirical methods to estimate the tails. Subsequent VAR models have built on these insights and replace the normal distribution with a particular extreme value distribution.[15]

Financial market participants also use stress testing, to complement their VAR calculations. Stress testing involves running a variety of extreme scenarios, some based on historical events (e.g. 1987), and others on simulated scenarios. As such stress testing allows firms to measure the impact of a range of large price changes, all beyond the VAR threshold.

Stress testing is also part of banks' regulatory requirements under the Basel framework (BIS (2009)). In their stress testing, banks are required to consider a range of stress events, including large stock price falls and severe reductions in market liquidity. Some stress

---

[11] See Basel Committee on Banking Supervision (BCBS) Revisions to the Basel II market risk framework (BIS (2009)).

[12] See e.g. BCBS (2011) for an overview of academic and other critiques of the VAR method. Shortcomings of the VAR method highlighted include: (i) the length of the horizon; (ii) liquidity is not taken into account; (iii) portfolios are assumed to be unchanged See also Flannery et al (2012)

[13] The Basel Committee (BCBS, 2012) is currently consulting on whether to consider Expected Shortfall as an alternative risk metric.

[14] See e.g. McNeil et al (2005).

[15] See e.g. Longin (2000).

scenarios may be prescribed by the authorities, others may be considered by the individual firm as particularly relevant for its business.

Stress tests are often run independently of VAR exercises. Recent research advocates that firms incorporate the results of stress test scenarios into their VAR analysis. This would require firms to assign probabilities to their stress test scenarios, which is inevitably a subjective exercise. [16] This integrated stress testing approach is currently not part of the revised Basel II framework.[17]

Common to all the approaches described so far is that very large price falls are deemed undesirable. This is particularly true when considering the behaviour of prices within a single trading day. Section 3 will discuss the controls in place on trading venues (price limits and circuit breakers) to mitigate such intraday price falls.

In summary, financial market participants use a combination of probabilistic analysis and stress testing to measure and mitigate the impact of very large price changes on their portfolios. Historical data and events play an important part in this process, but simulations based on theoretical stress scenarios are used too.

However, when it comes to measuring the impact of a large price fall on the market as a whole, there is no formal framework. Unlike the nuclear industry, the economics profession does not categorise the risk of market crashes, or label them 'tolerable' or 'acceptable.' And there is no commonly accepted numerical target for market crashes (e.g. 'one market crash every x decades is acceptable' or 'a price fall of x% is acceptable but y% is not'). Instead, market practitioners and financial economists have taken a more descriptive approach aimed at understanding (i) under which conditions market crashes occur and (ii) whether there are any common characteristics. Meanwhile market regulators focus on the need to avoid 'disorderly' markets, which is a qualitative notion, rather than a precise quantitative target (see Section 2.3.2).

2.4. Commentary

Our analysis suggests the following similarities and differences:

- When thinking about large-scale risks, both industries' practices can be interpreted in terms of systemic risk
- Both industries use probabilistic concepts of risk and impact.
- The nuclear industry has a clear notion of tolerable level of risks and it can set numerical targets. Hence, it is able to rigorously assess trade-offs between risk reduction and costs (ALARP)
- In its thinking about financial market crashes, the finance industry relies to a large extent on probabilistic methods, using historical data. This is complemented by stress testing, using

[16] See e.g. BIS (2011).

[17] See BIS (2011).

both historical and theoretical stress scenarios. There is also much emphasis on understanding past events so that potential future problems can be avoided.

- Although these past events have undesirable features, which can be measured precisely, there is no equivalent of the ALARP principle and there are no numerical targets.

In addition, it is useful to note the following particular aspects of the nuclear safety analysis and risk frameworks:

- The nuclear industry has a formalised approach to defining the classes of consequence, the categories and frequencies of initiating events. It uses theory, models and experiment to justify the risk analysis.

- This means that industry can set risk targets for classes of accident and different classes of people, and discusses tolerability and proportionality in reducing them further.

- In doing so, the industry accepts that many things are hard to quantify, but there is nonetheless an emphasis on ranking risks, setting targets for risk reduction, and debating whether both the risks and the targets are accurate and acceptable.

- The nuclear safety analysis framework allows systematic design of protection and mitigation systems that cover not only what they have to do, but also how much they have to be trusted. These systems use diverse mechanisms to ensure that the overall protection works when it is needed.

- The nuclear industry also places greater emphasis on explaining risks to society at large. This in part drives the quantification of risk as there needs to be basis for comparing different types and sources of risk.

2.5. Questions and issues

The nuclear industry has developed an approach that reflects the need to engineer a complex safe system and to explain the safety of what is proposed and implemented to regulators and society at large.

The finance industry has developed a framework for analysing the risks associated with the extreme price falls that constitute a market crash. In doing so, the profession has traditionally viewed asset prices as being the outcome of human decisions. Hence, they have focused on understanding (i) the behaviour of human traders participating in the markets and (ii) how these decisions affect the dynamics of asset prices. The risk concepts and tools developed by the industry are also largely descriptive: historical data are used to describe extreme events in precise statistical terms. These are complemented with scenario-based stress testing. But the industry has so far not defined numerical risk targets or tolerability levels for the risk of market crashes, that could be used to guide new systemic risk reduction initiatives.

Looking ahead, one could question whether the rapid development of computer-based trading in financial markets requires the adoption of additional risk concepts and tools. When viewing a financial market as a complex adaptive system (Foresight 2011), where a large proportion of interactions are the result of trades executed by algorithms rather than humans, concepts and tools developed in other disciplines may become much more relevant. In particular, the discussion in this Section suggests that the following questions are worth asking:

1. Is it possible to have a more precise description of risk categories (e.g. in terms of the type of consequences, who is affected, the initiating events that precipitated them)?
2. Is it possible to define precise tolerability criteria? Can one distinguish between tolerable and broadly acceptable risks?
3. Is it possible to define numerical targets? If not, how does one define 'acceptable' risk?
4. Is it possible to develop the notion of a 'design basis,' which would characterise those adverse endogenous and exogenous events that the system (i.e. the market with its control and protective mechanisms in place) should withstand?

3. Protection parameters and risk controls

Having defined the key risk concepts used in the nuclear and financial markets context, this Section explores in more detail how risk controls are designed and the overall protection system architecture is defined.

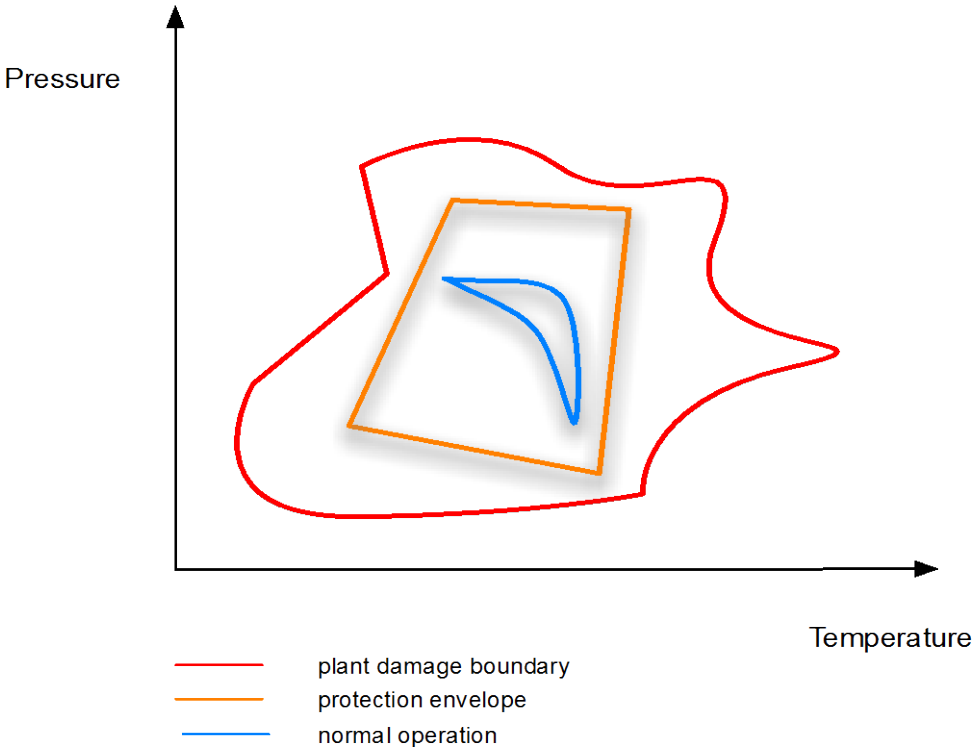
3.1. Nuclear protection and control

3.1.1 Protection and control envelopes

In an engineered complex system such as a nuclear power plant the protection system is designed so that it is possible to infer the state of the plant, (e.g. the condition of the core) from a combination of direct and indirect measurements. In Section 2.2.1 we have summarised how a systematic safety analysis and design approach leads to a definition of the safety systems needed to protect the plant and the parameters that need to be measured.

A simplified, conceptual, view of the results of this analysis is shown in Figure 4. The red line indicates the boundaries for the parameters – in this case indicated as pressure and temperature – beyond which plant damage and accidents might occur. The actual boundaries could be very complex. In typical commercial reactors there might be 20-30 different physical parameters that are measured and used to infer the state of the plant. Furthermore as we have seen in Section 2.2.3, the actions to be taken depend on the state of the plant (e.g. is it starting up, at power, shutting down, a failure of a protection system detected, in a degraded state) so that the boundaries will be a more complex shape than the two dimensions of Figure 4 suggests.

Figure 4. Protection envelope



In the design of the protection system, a variety of uncertainties are taken into account to simplify the region into an area that is easy to measure and considered as conservative. This region is shown in orange in the diagram: the idea is that shutting the plant down when these boundaries are reached will lead to a safe state. Within this protection envelope there is an economically optimum operating region and this is shown in blue. For example, there might be a limiting pressure at which a pipe is weakened to such an extent that there is an unacceptable probability that it might burst. Operating just below this pressure may be tolerably safe but would reduce the lifetime of the plant and lead to unacceptable cost of replacement and inspections. Instead an operating pressure would be chosen that would lead to the pipe being de-stressed and would have a long operating life.

The design of the reactor control system operating envelope draws on classical linear control theory and computer simulation of the plant to validate the stability of the overall system. There is a long-standing engineering discipline of control theory and modeling that is used to design a variety of feedback and feed forward loops in a plant. This could provide a framework for assessing some of the feedback loops discussed in (Foresight 2011). There is also more ecologically oriented work that is relevant: viability domains describe a similar concept to protection envelopes and their mathematical underpinnings may provide some further insights (Deffuant (2011)). It is likely that the non-linear aspect and performative nature of the underlying modeling would need some novel approaches to take into account adaptation to the viability envelope and possible new failure modes

In designing the plant's measurement strategy, a number of issues have to be taken into account. It is not possible to measure at all points in the plant, so properties have to be inferred from a few measurements (e.g. by exploiting the symmetry of the plant). In addition, measurement errors and the differences between the values measured and the abstract

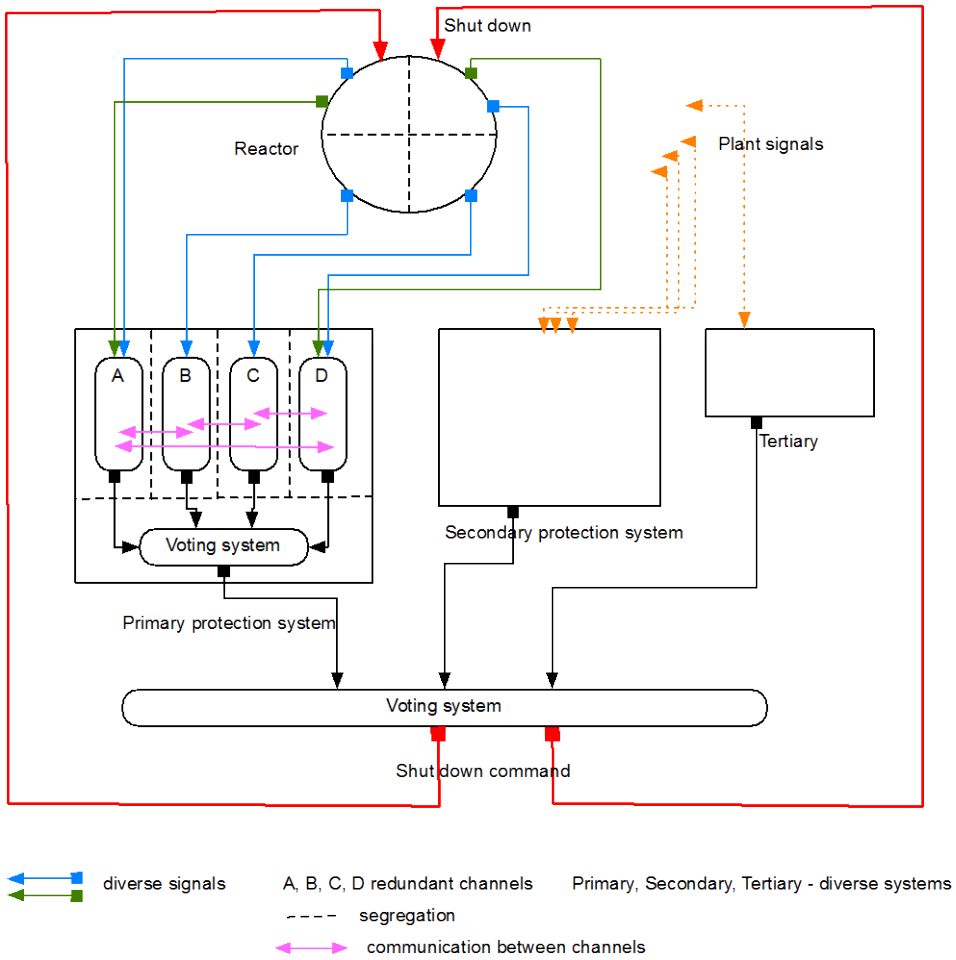
physics based values that are used to define the safe envelopes need to be addressed. Redundant sensors are often used to provide measured values that are more trusted (e.g. by taking a “vote” and making a decision to act when three out of four values agree).

There is a need to ensure that spurious trips, that is shutting down the plant when it is not necessary, are avoided as these carry both a safety risk and commercial cost. This can be due to both an inference problem (i.e. misdiagnosis) or from the failure of components within the protection systems. Considerable design effort is used to ensure that there are suitable levels of diversity and redundancy so that single points of failures are tolerated.

3.1.2 System architecture

Reactor protection systems have evolved a long way since the Manhattan project when Enrico Fermi deployed a woodsman who was to cut a heavy rope holding a “rod” above the core and also three men with buckets of cadmium sulphate, all controlled by Fermi’s hand signals. Yet the essential elements were there: diversity and the need for the protection to work when it was needed. Modern reactor protection systems have of course also to work when needed: they therefore have very high availability and reliability requirements. To achieve this, the architecture incorporates segregation, redundancy and diversity as shown schematically in Figure 5.

Figure 5. Simplified protection system architecture



Although the protection functions may be conceptually simple, such as shutting the plant down when the threshold on parameter is reached, it can lead to a variety of complex systems, with

hundreds of computers running many hundreds of thousands of lines of code in part because of the need to satisfy the non-functional requirements for the system (e.g. the availability requirements). There are many tradeoffs to be made and it is fair to say that, while the principles of diversity and defence in depth are accepted globally, their implementation in system architectures is not internationally agreed. For example, the I&C architectures of the proposed US, UK EPRs and the plants being built in Finland, China and France are all different.

Diversity[18] is a key concept and issue in the protection system architectures. In the simplified architecture diagram of Figure 5, there are three diverse protection systems – that is different in design and implementation technologies - that can each shut down the reactor. The primary is generally more complex and computer based with the tertiary being limited to only a few functions and possibly implemented with a simpler technology (e.g. dynamic logic or with FPGAs).

The Primary System in Figure 5 is shown as having a number of redundant channels which are implemented with the same design of equipment and software. The outputs of these channels are voted upon with typically two out of the four channels needed to command a trip. In case of channel failure the logic adapts to 2 out of 3 and then 1 out of 2. The redundant channels are physically segregated so as to protect against possible common cause and cascade failures such as fire. Nevertheless there is often information exchange between channels and it is necessary to show that this does not introduce any covert channels and so nullify the benefits of the segregation and redundancy.

The redundancy is necessary to provide the high availability that is needed and particularly protects against random failure of hardware. The four channels are needed to keep the spurious trip rate low and allow for equipment being out of service and maintained while the plant is at power. The diverse redundancy is needed to counter common cause failures and epistemic uncertainties e.g. due to common software faults in the common software in the four protection channels.

3.1.3 Failure correlation

The issue of correlation of failures between protection systems is important and has been much studied, particularly in the UK. It is known that claims for independent failure are unrealistic (so it is not possible to just multiply the probability of failures of two subsystems to get the system probability of failure) yet the degree of correlation is very hard to assess.

Some experimental results on correlation are shown in Figure 6 and 7. Figure 6 is from a seminal Nasa funded experiment (data from Knight (1986)) that shows the improvement in the probability of failure of missile detection algorithm as the mean performance improves. The other (Figure 7) is from a software competition with many thousands of entrants and shows the reliability improvement of a diverse pair, relative to a single version (from Meulen (2008)). The horizontal axis shows the average probability of failure on demand of the pool from which both

[18] Or diverse redundancy, “The presence of two or more systems or components to perform an identified function, where the systems or components have different attributes so as to reduce the possibility of common cause failure, including common mode failure”. Diversity could result in different development lifecycles, different organisations, and different implementation technologies. The term “redundancy” denotes replicated, sometimes identical, systems or structures e.g. in protecting against fire by having identical systems located in different places.

programs are selected. The vertical axis shows the reliability improvement from having a second algorithm.

The main message from these experiments is that on average one only gets one or two orders of magnitude improvement in the probability of failure on demand by deploying diverse systems. One explanation for this is that independent designers and developers make similar mistakes because of the inherent difficulty of the problem that the algorithm is solving. The presence of these correlations and the non-independence of failures is a robust result, replicated across experiments sponsored by Nasa, the nuclear industry and others. It may be an area where there could be some cross-over to the assessments of diverse trading algorithms (e.g. see the Foresight Driver Review DR13, (Foresight 2011b))

Figure 6. Experimental results on diversity

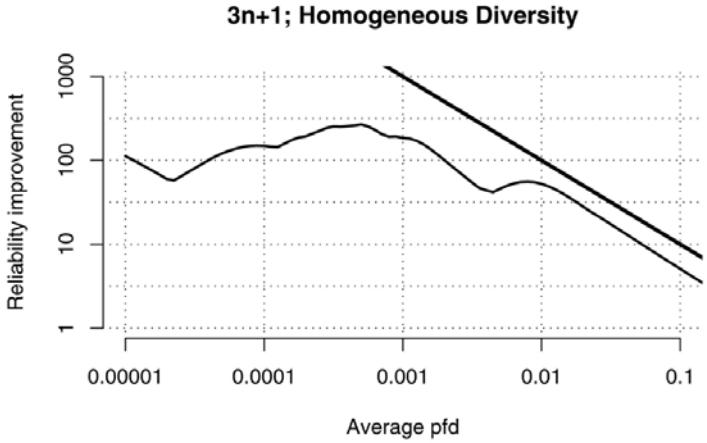
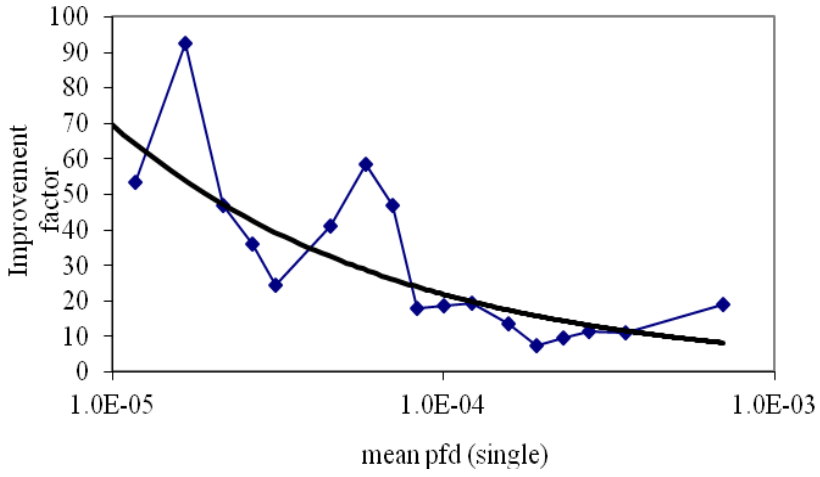


Figure 7. Experimental results on diversity



In summary, the functions of the protection systems in a nuclear plant can be conceptualised as enforcing a safe conservative envelope on plant parameters (e.g. temperature, flow, pressure). To implement this abstract view requires a detailed approach to measurement accuracy and trustworthiness, as further explained in Section 4. In addition, the need to implement the non-functional aspects of the protection (e.g. to operate when required with a very low probability of failure) leads to the provision of multiple diverse systems.

3.2.Financial markets

In Section 3.3., we consider whether the concepts defined in Section 3.1can be applied to understand and mitigate risks in financial markets. Before doing so, Section 3.2. describes the existing processes in financial markets to prevent and halt large intraday price falls. As before, the focus is on continuous-auction type equity markets. We also provide a very brief overview of the regulatory framework.

The controls described in this Section are aimed at ensuring that trading venues have both the operational and liquidity resilience that is required for their smooth daily functioning, including their essential price discovery role (see Section 2.3.2). Hence, one of their main objectives is controlling intraday price movements. As the history of market crashes shows, they may not be able to prevent all large price falls, particularly those that happen over an extended period of time and/or are the result of a sudden shift in investor sentiment. But together, these controls are aimed at increasing the robustness of trading venues.

3.2.1 Objectives

Most trading venues have a number of processes in place aimed at avoiding extreme price movements. These include price limits, trading halts (also referred to as circuit breakers) as well as controls to detect erroneous messages and to limit the number of messages (order submissions and/or cancellations) that market participants can send to the trading venue. In addition, market participants themselves have operational controls in place, e.g. to limit the risk of sending erroneous messages to the trading venue.

Both trading venues and market participants have incentives to introduce such controls: trading venues aim to avoid disorderly market conditions (See Section 2.3.2 for a definition), while market participants wish to avoid trading losses that may arise from erroneous trades. In addition, both trading venues and market participants are subject to regulatory requirements (see box 7).

Box 7. Regulation of financial markets – recent developments

IOSCO regularly reviews its regulatory standards in the light of market developments. A recent study assesses the impact of technological change on market integrity. In this report, IOSCO (2011) sets out a number of high-level recommendations for market participants using computerised trading:

- Pre-trade controls that are suitable for high-speed markets;
- Stress testing of algorithms before they are used
- While trading venues should have the following in place:
- Order entry controls to detect and stop anomalous order entry
- Trade cancellation process
- Stress testing (ability of platform to cope with unusually high order numbers)

- A testing environment for market participants to test new algorithms
- Price limits or circuit breakers

Efforts are also underway to improve real-time monitoring of trading. For example, in the US, regulators have recommended the establishment of a consolidated audit trail which would allow monitoring of the flow of orders executed in different venues. In response to the May 2010 Flash crash US regulators have also introduced new pre-trade controls and are piloting new circuit breakers.

In Europe, the Markets in Financial Instruments Directive (MiFID) is being reviewed, and is expected to include a range of measures to be adopted by both trading venues and market participants using computerised trading, including possibly order cancellation measures; market maker commitments and circuit breakers.

Recent guidelines published by ESMA, the European Securities and Markets Authority, include enhanced operational requirements (ESMA (2011)). These apply both to platform operators and to investment firms using computerised trading. They include amongst other things:

For trading platforms:

- Effective arrangements to ensure trading systems are resilient, have sufficient capacity and are able to ensure fair and orderly trading.
- Effective arrangements to prevent excessive flooding of the order book (e.g. through participant limits) and to ensure capacity limits are not being breached.
- Measures to constrain or halt trading in individual or multiple securities, when necessary.
- Business continuity arrangements to deal with unforeseen failure of trading systems.
- Effective arrangements to monitor market activity as close to real time as possible.
- Appropriate risk controls for members providing direct access, and for members who are not credit institutions.
- Record keeping requirements.

For investment firms:

- Effective systems and risk controls to ensure trading systems are resilient and have sufficient capacity, are subject to appropriate thresholds and limits and prevent the sending of erroneous orders
- Effective arrangements to block or cancel orders that do not meet pre-trade controls.
- Business continuity arrangements to deal with unforeseen failure of trading systems.
- Testing requirements for new trading algorithms.
- Effective arrangements to monitor trading activity as close to real time as possible, including control of messaging traffic to individual platforms.
- Appropriate risk controls for direct access arrangements, including pre-set trading and credit thresholds.
- Record keeping requirements.

3.2.2 Risk controls - slowing down or halting trading

Section 2 explained that extreme price changes are a manifestation of extreme illiquidity. In some cases, this happens because market participants choose to coordinate their actions (all become sellers and/or withdraw from the market). In other cases, illiquidity may be the result of an algorithm generating a very large order imbalance. This may in turn cause human traders to react to the observed order imbalance, resulting in further illiquidity.

A first set of risk controls is aimed directly at slowing down the intraday price changes that may result from sudden changes in liquidity. This may be done in the following ways:

- Price limits constrain trading within a pre-specified band: trades within the band are processed as normal; trades outside the band are rejected. Sometimes these are referred to as price-collars.
- A trading halt stops all trading for a (sometimes) pre-specified period. On some markets, trading is initially resumed via a batch auction. After this, normal processes continue.
- Price limits and trading halts can be static or dynamic: in the latter case, the triggering point is updated as market conditions develop. For example, a price fall that reflects a change in underlying fundamentals can be ‘managed’ in an orderly manner.

These price-based limits differ from procedures in place to avoid price changes that would result from erroneous trades (so-called fat-finger mistakes). Both trading platforms and market

participants are required to have operational controls in place to avoid such errors disrupting the market (ESMA (2011) and box 7).

A key challenge for financial market operators (and their regulators) is how to define a price limit:

- Whether the limit applies to a single stock or to the wider market;
- Whether the limit applies to a single trading venue, or is set in coordination with other trading venues (e.g. other equity markets or derivatives markets) (see also 3.2.4);
- How to set the trigger points;
- Whether to choose static or dynamic limits;
- Whether to halt trading or simply slow down price changes.

The principal trade-off for market operators is between avoiding large intraday price changes on the one hand, and disrupting trading on the other hand. The empirical literature on this topic is largely inconclusive, with some, but not all, finding evidence that price limits and trading halts are effective in reducing price volatility. Proponents of trading halts further argue that they offer a useful pause, allowing market participants to assess their exposures, before returning to the market. Critics, however, argue that price limits and trading halts may contribute to price volatility, e.g. by acting as a 'magnet.' [19]

Likewise, there is some debate about coordinating trading halts across venues, often referred to as cross-market circuit breakers. Some argue that this is essential to avoid cross-market contagion (SEC (2011)), while others point out that it is very difficult to determine the relevant parameters given the diversity of trading venues.

A second set of controls, often used in conjunction with price limits, is aimed at controlling the flow of order message traffic, rather than the price changes that may result from those orders. These include:

- Measures to control the number of messages sent to the exchanges;
- Measures to control order cancellations.

Practical challenges for the market operator are:

- Whether to control all messages or order cancellations only;
- Setting critical thresholds (e.g. cancellations up to a point may be allowed);
- Choosing between hard limits (e.g. a fee on orders or cancellations above a certain threshold) or soft limits;

[19] Foresight is undertaking more work on the effectiveness of price limits.

- Whether to target orders generated by automated strategies, or all orders.

Most trading venues have measures in place that allow them to monitor and control message traffic, and this is typically a regulatory requirement to avoid capacity problems (ESMA (2011) and box 7). Measures to control or limit order cancellations are more controversial, with some arguing that their increased prevalence is the direct result of computerised trading and may result in an 'illusion' of liquidity. Others point out that order cancellations are a useful tool in managing trading exposures. The academic literature in this area remains sparse.

It is an open question as to whether the very high-frequency price changes documented in recent academic work (see Box 6) warrant controls that operate at these higher frequencies. Moreover, such controls would need to act very quickly, to match the speed of the algorithms behind the price movements. Furthermore, it is as yet unclear how venues would define such controls, how they would set the trigger levels, or how they would measure their effectiveness.

3.2.3 Risk controls – ensuring resilience and diversity

The controls in Section 3.2.2 either aim to control the overall volume of transactions, or the resulting price changes. Alternatively, trading venues could take measures to ensure that liquidity is resilient so there are sufficient buyers and sellers in the market throughout the trading day. These measures include:

- Market maker schemes: these are formal arrangements, which require a select group of market participants to be active in the market for a pre-specified portion of the trading day. Market makers will typically be required to quote continuous prices, keep prices within a pre-specified band, and quote above a pre-specified minimum size. In return, they may receive fee rebates or bonus payments.

- Liquidity incentive schemes: these are less formal arrangements, which encourage market participants to bring their trades to the market, typically in return for a fee rebate. Some encourage the posting of limit orders. Others offer more attractive fee rebates for orders posted during periods of increased market stress (so-called peak load pricing).

Again, the key challenge for financial market operators is to define the precise scheme:

- Whether to choose a formal or informal scheme;

- How to set the minimum requirements;

- How to set the incentives;

- Whether to put in place special arrangements for times of market stress, and how to define the thresholds for 'market stress.'

When working, such arrangements result in greater liquidity resilience (i.e. liquidity providers will stay in the market for longer), delaying the point at which price limits or trading halts may need to be invoked. Again, the academic evidence is mixed. Some studies show that the introduction of market maker arrangements does indeed lower intraday price volatility and improves liquidity, e.g. for less actively traded stocks. But there is no academic evidence on their effectiveness in periods of market stress. Indeed, a key issue for debate is whether

market maker arrangements can be effective in periods of market stress, or whether dedicated market makers would simply withdraw (like other market participants) in order to limit their exposure when prices are falling rapidly.

Market maker and liquidity provision schemes may help liquidity resilience by encouraging more traders to come to the market. A further question is whether one can ensure that the limit order book is sufficiently rich and deep at all times (i.e. the book contains buy and sell orders for a wide range of prices and sizes), so incoming market orders do not cause the limit order book to become unbalanced or empty. In other words, would encouraging diversity in trading lead to greater liquidity resilience in the short term, and greater market confidence in the longer term? Diversity in trading may arise from having participants with different trading motives and horizons (e.g. short-term market makers versus long-term buy-and-hold investors), different views (e.g. on whether a stock is over or underpriced) or different trade execution preferences (e.g. whether to use market or limit orders). Diversity may also be the result of having competing trading venues with different trading models. More work is needed to understand how to promote diversity of trading in financial markets, and whether measures to increase such diversity would have a discernible impact on liquidity resilience.[20]

3.2.4 Risk controls – correlated failures

In the nuclear industry, correlated failures are an important topic of concern. These can arise for a variety of reasons: there can be common mechanisms such as flood or security vulnerabilities that nullify several risk control barriers at once, there can be common design flaws in similar barriers and there can be correlated failures among diverse systems due to coupling in the problem domain (see Figure 6). In addition there can be cascade failures where failure of one component or system can impact another. All these issues need to be addressed in the design and evaluation of the systems and their risk controls.

In the context of financial markets, correlated failures may refer to instances where problems in one trading venue "spill-over" into another venue. Such spill-overs may take the following forms:

- Capacity constraints on one venue may cause orders to be re-routed (sometimes automatically) to other, related venues;
- Delays in processing orders on one venue may affect the availability of up-to date prices, both in the affected and in related trading venues;
- Controls applied in one venue may cause orders to be re-routed to other, related venues; in turn contributing to price volatility, and possibly to the triggering of controls in these venues.

Hence, an important question to ask is: how can one venue protect itself within a correlated environment? What are the implications of controls that are applied in a non-coordinated fashion (e.g. circuit breakers triggered at different price levels)? How would co-ordinated circuit breakers work instead? These questions are currently being explored by both European and US regulators.

[20] Foresight is undertaking more work on the effectiveness of market making and liquidity incentives schemes.

Further insights may be gained from the literature on complex systems. In Foresight (2011) DR4), it is agued that financial markets have become complex adaptive systems, in which extreme changes can happen in unexpected ways (Foresight (2011) DR4). Moreover, as financial markets have become increasingly interconnected, they can be viewed as 'systems of systems.' (DR4). Even if trading venues operate independent systems, they may be linked by algorithmic trading systems operating in more than one venue and/or conditioning their trading strategies on prices in more than one venue. Moreover, such connections are made at very high speeds. It follows that a failure in one or more constituents (market venues) could have widespread repercussions. It also implies that system-level failure is difficult to predict, not only because both humans and computers can adapt their behavior over time (and can do so at high speed), but also because of the sheer number of possible interactions between humans and computers, both within and across venues.

3.3. Commentary

Comparing the use of risk controls in the nuclear industry and financial markets reveals the following:

- Both employ risk controls, based on thresholds beyond which operations need to be halted (or slowed down). These controls are set by the operators, and are typically subject to regulatory supervision.

- In the nuclear industry, this is the result of a systematic engineering analysis, summarised in the protection envelope and the Fault and Protection Schedule.

- In financial markets, these controls typically depend on a smaller and less complex set of parameters (e.g. traded prices or message volumes).

- Unlike the nuclear protection systems, there is no formal mechanism for describing how much the controls themselves have to be trusted (e.g. in terms of probability of failure on demand, probability of spurious activation).

- There is considerable debate on the effectiveness of some of the risk controls, partly because of the lack of evidence on how they might perform during periods of market stress.

The brief overview of the control and protection of a nuclear plant raises a number of issues that may be of relevance as financial markets consider how to adapt to existing risk controls to the new computerised trading environment:

- Engineering succeeds by making the complex systems controllable and predictable (within limits). Although the underlying processes are complex and complicated the ability to model and design the plant and to have a scientific based understanding of what might happen allows the functional aspects of the protection (controls) to be relatively simple. However there are more onerous requirements on the non-functional aspects as the systems really do need to operate when needed.

- The ability to engineer a control and protection system relies on observability of the system. The notion that the financial market is an observable system in readily identified states is only partially true. It is clear in our review of market crashes that there are competing theories and perspectives. There is evidence that some crashes appear to just happen and that these are irreducible and so there is no difference between a transition to a systemic loss and an

everyday one. Others would argue that indeed there is a difference; it is just that we do not (yet) have the means to identify the hidden states that precede a systemic event (Foresight (2011)). This has implications for the extent to which market controls can be engineered.

- Protection systems have authority to override any other system and force a shut down. If they operate when not needed (e.g. due to internal failures, operator error) they can cause spurious plant disturbances with consequential economic costs and safety implications. There is a need to define performance measures for spurious activation (e.g. once every 10 years) as well as for the probability of failures per demand. (The impact of this on architectures and assurance is discussed in Section 4)

- There is trade off between economic benefits and having a simple protection envelope. As the understanding of the nuclear plant has developed over the years, protection envelopes have become more complex and computer based. There can be considerable off-line data analysis and modelling to derive the parameters for these systems: so that trust is needed in both the protection algorithms and the data.

- Adaptation and learning is very important, but in a nuclear plant occurs in different timeband from the control and protection actions e.g. months for procedures and safety culture, years for updating equipment, decades for design of plant.

- There is a need to get the best mix of between reliable automation and human analysis and adaptation. This is a sophisticated topic, but in brief, the design needs to play to humans' strengths of understanding and adaptation.

- In safety engineering there are examples that introducing safety or protection measures can change people's behaviour so safety improvements are only temporary as people adapt. The complex adaptive nature of markets means that this could be a significant future issue in designing control and protection especially as these might provide unintended opportunities for new forms of regulatory arbitrage or market abuse.

- The nuclear industry uses defence in depth and diversity to achieve effective risk controls and these require independently developed solutions to avoid systemic failures. However there is experimental data and associated theories that show that despite this independence, failures are correlated. This may be an area where there could be some crossover to the assessments of diverse trading algorithms and risk controls (e.g. see the Foresight Driver Review DR13, (Foresight (2011b))).

3.4. Questions and issues

As financial markets are developing into more complex computer-based systems, it is worth asking whether the concept of a protection envelope would be helpful.

1. What would the protection and control envelopes look like?
2. What would be the parameters that need to be measured and what would we infer from them? How are they related to existing controls such as price limits or circuit breakers?
3. What would the availability and reliability requirement be for such a system e.g. the probability of failure on demand, the frequency of spurious activation?
4. What is the balance between automation and operator recovery?
5. What additional understanding (and research) is needed given the complex adaptive systems nature of markets? How would the markets adapt to having such protection?

6. What additional analysis techniques and data are needed to assess the risks arising from correlated failures and to design risk controls to guard against their impact?

4. Trust in computer based systems

Computer based systems, and ICT in general, are of course essential for high-frequency trading and for algorithmic trading more broadly: for market participants to process information, design trading strategies, and execute the trades; for operators of trading venues and regulators to collect and process the data for monitoring and market surveillance and for operators to provide protection and intervention via so-called "circuit breakers". In all cases, systems will need to be sufficiently trusted: how do we describe that level of trust and how is this evaluated?

4.1. Software assurance in the nuclear industry

In this Section we consider the principles, strategy and techniques used to assure the software in reactor protection systems.

4.1.1 Excellence of production and confidence building

The reactor protection system is crucially dependent on software and complex electronics. In the UK nuclear industry the justification is based on two important safety principles that come from the SAPs. These are

> "Where the system reliability is significantly dependent upon the performance of computer software, the establishment of and compliance with appropriate standards and practices throughout the software development life-cycle should be made, commensurate with the level of reliability required, by a demonstration of 'production excellence' and 'confidence-building' measures." (ESS27 from HSE (2006))

This approach of "confidence building" and "production excellence" is known in the nuclear safety domain as the "two-legged approach" and is discussed below in more detail.

4.1.2 Production excellence

"Production excellence", as the name suggest, requires a demonstration of excellence in all aspects of production, covering initial specification through to the finally commissioned system. It expects compliance with standards and implementation of a QA programme as well as a comprehensive testing and analysis to check every system function, including:

- The verification of all phases of the system production process and the validation of the integrated system against its requirements specification by persons not involved in the specification and design activities;
- A demonstration that the safety system, in conjunction with the plant, performs to requirements, this demonstration being devised by persons other than the system specifiers, designers or manufacturers; and
- A programme of dynamic testing, applied to the complete system that is capable of demonstrating that the system meets its reliability requirements.

This latter point is important: it introduces the requirement for statistical testing in which simulated operation of many thousands of tests are used to provide a numerical indication of the reliability (subject of course to many assumptions and uncertainties).

Box 8. Software reliability

The use of probabilities to characterize software dependability is sometimes challenged. It is argued that because software failures are systematic – if it fails in certain circumstances, it will always fail when those circumstances are repeated – there is no role for probability. The answer to this apparent paradox is that there is uncertainty about *which* inputs will cause failure, and about *when* these will be executed. So the software failures that occur form a stochastic process and the uncertainty associated with this process is *aleatory*, i.e. it is "uncertainty in the world" and is irreducible.

In addition to this aleatory uncertainty, there is also *epistemic* uncertainty concerning the models and reasoning that are used to estimate and predict reliability. For example, in operational testing, there may be uncertainty about the correctness of the test oracle that is used to asses whether a test is successful or not, and about the representativeness with respect to real-life operation of the test case selections. Epistemic uncertainty is in principle reducible, e.g. by collecting more and better evidence concerning the subject of the uncertainty.

If weaknesses are found in the production process, compensating measures must be applied which are targeted at the specific weaknesses: these may be designed to rectify or mitigate the weaknesses.

4.1.3 Independent confidence building measures

Independent confidence building measures (ICBMs) are required to provide a thorough and challenging assessment of fitness for purpose, but should be reasonably practicable (see Section 2.2.5). They will involve:

- Complete and preferably diverse checking of the software (after validation has been completed) by a team independent of the suppliers;
- Independent assessment of the full test programme, covering the full scope of the testing activities.

4.1.4 Assessment strategy

In a reactor protection system there will be many different software components with different supply chains and provenance. In a typical modern reactor protection system there will be hundreds of processors and associated software for smart sensors, communication, voting, signal processing, data fusion, decision-making, self-testing, interface to maintenance equipment and provision of information. Practice and research in the nuclear industry and elsewhere has found it helpful to use the concepts of claims, argument and evidence (introduced in Box 2) in discussing the assurance of systems and to view the overall assessment strategy in terms of the Assessment Triangle (as in Figure 8) that balances: a rule based compliance approach (the blue rectangle), the goal-based demonstration of positive

safety properties (the yellow rectangle), and the risk-informed demonstration of absence of negative vulnerabilities (the red rectangle).

Figure 8. Assessment triangle

It is perhaps not surprising that for a critical system utmost care is required in its development; that claims made about the system should be challenged and that high quality evidence is required to show that it is fit for purpose. As outlined above in the principles (Section 4.1.1) and in the more detailed supporting standards and guidelines, independence is considered essential and appears both in the initial verification and validation of the system; in the challenge and confidence building activities and in the layers of internal and external regulatory oversight.

4.1.5 Assurance techniques

The assessment triangle and associated Claims-Argument-Evidence structures can be used to guide the selection of development and assessment techniques. The approach might be guided by goal-based assessments of attributes (e.g. demonstration of accuracy, reliability, etc.) including assessment of potential vulnerabilities in the implementation of the system/component and formal demonstration of certain integrity properties (e.g. deadlock, run time exceptions caused by divide by zero).

In terms of the critical evidence that is required, two aspects should be emphasised from the Safety Assessment Principles. The first is the use of (modern) static analysis techniques and the second the need for statistical testing. Both of these activities can require significant resources or elapsed time. While no estimates are available for New Build an informed guess might put the analysis activity, even with modern computer science techniques of proof and model checking to be 10s if not 100s of person years of effort, and the high-fidelity testing to last at least a year. This is for systems of several hundreds of thousands of line of code and some hundreds of processors.

Advances in computer science and the engineering of the associated tools has been dramatic in the past five years. The types of modern techniques that can now be deployed on the

analysis and testing of nuclear software is varied (Guerra (2010)). These include formal mathematical proof supported by model checking and SAT solvers; concurrency analyses based on process algebra tools and abstract interpretation to find runtime errors (such as divide by zero). However despite these advances, assessment is still a specialist activity and considerable applied research and innovation is needed to scale these techniques to forthcoming protection systems such as envisaged on New Build.

In addition, more traditional techniques will also be deployed such as: testing against requirements; negative testing designed to abuse and misuse the system; the analysis of test coverage of the code; manual expert review; checking compliance with standards; demonstrating that the software tools used do not compromise the integrity of the code. This may involve source code to binary comparisons or a justification of code generators either in general or for a specific system that has been developed.

It is also noteworthy that apart from the simulated experience of statistical testing, there is limited scope for using experience to justify reliability claims particularly as confidence is needed before these systems are put into operation. Even in France after many years of computer based protection systems they still have probably responded to only 500 trips. However, experience can still provide important evidence that can be used to help focus the assessment and should be scrutinised for counter-evidence to the confidence building results.

4.1.6 Summary

One could sum up the assurance approach for reactor protection systems as “Do everything and do it at least twice”. Specific measures that are used to achieve assurance are:

- The use of a very careful development lifecycle with trusted tools and extensive verification and validation;
- The independent static analysis and mathematical proof of the software with respect to its specification and known vulnerability classes;
- The use of statistical testing to simulate live operation;
- The challenging of the system with negative testing designed to abuse and misuse the system;
- The compliance with appropriate standards.

Thee approaches are deployed on protection systems and adjusted depending on the criticality of the system or component that is being assured.

4.2. Assurance of computer based trading in financial markets- some observations

In Section 2.3.2, we explained in which circumstances an operational outage ceases to be ‘just another computer glitch.’ Section 3.2 described the set of risk controls employed by operators of trading venues aimed at avoiding extreme price movements. In this Section, we return to the issue of operational errors and outages, and focus on those problems that may arise from computer malfunctioning, either at the level of the user (market participant) or the trading venue. This allows us to introduce the notion of trust in trading algorithms.

Table 5 below provides a simple classification of problems caused by faulty algorithms. The first set of incidents concerns individual trades that have gone wrong – often dubbed “fat fingers” to capture the idea that someone has hit the wrong button or punched the wrong key. Some of these are quite straightforward and may be easily attributable and understandable. Others, however, may be more difficult to detect or assess, particularly if the fault lies buried inside a complicated algorithm.

The second set of operational problems concern trading platforms. An operational outage may involve a disruption in the data feed to or from the trading venue; it may involve a failure of the computer system matching orders; or it may be result of the erroneous deletion of trades entered in the venue’s system. In some instances, an operational fault in one venue may affect a range of related venues.

The third set of problems arises from the interaction between algorithms, either within a single venue, or across multiple venues. These incidents are the hardest to understand and assess. Anecdotal evidence suggests that when a price falls very rapidly (say to zero) and then recovers equally fast, this might be the result of a new algorithm being tested in the market.

Table 5. Problem classification

	Type of problem	Possible impacts
Problems with individual algorithms	'Fat finger' trades i.e. a faulty trade arising from a keyboard error.	Can lead to a trade which exceeds a firm's capital. Can lead to further price falls if undetected.
	Faulty algorithm	Algorithm triggering erroneous sell orders may result in significant price falls across range of stocks Algorithm sending very high number of erroneous messages may overwhelm trading venue's systems
Problems at trading venues	Trade matching engine not available Data feed not available Erroneous cancellation of orders entered into the venue's system	Impact may be mild if restored swiftly and technical nature of problem rapidly known Longer-lasting outage at primary exchange may affect price discovery in related markets. May lead to fall in confidence and lower liquidity in short/medium term
Problems arising from the interaction between algorithms	Atypical and short-lived price movements	Difficult to attribute and assess impact. Explanations offered include algorithm being tested/ withdrawn

ESMA (2011) outlines detailed guidelines for trading venues operating electronic trading systems and for market participants using electronic trading systems and trading algorithms. Some of these guidelines can be said to deal with the issue of trust, as described in Section 4.1. Specifically:

- Governance: ESMA requires clear lines of accountability for the sign-off of development, initial deployment, updates and resolution of problems in relation to electronic trading systems and trading algorithms (ESMA (2011), p. 32 & 35);

- Testing: ESMA requires clearly-delineated development and testing methodologies, including performance simulations, back testing and offline testing. These tests need to ensure that electronic trading systems or trading algorithms can work in stressed market conditions (ESMA (2011), p. 33 & 36);

- Record keeping: ESMA requires trading venues and market participants to keep records regarding key decisions, testing methodologies, test results and periodic reviews (ESMA (2011), p. 34 & 37);

In sum, in Section 3, we explained how the functionality and trust required from a protection system depends on the quality of the system that it is protecting and the consequences of failure and spurious activation. In the nuclear example plant design and siting is used to reduce the exogenous and endogenous hazards. For the latter, redundancy and defence in depth is used to ensure that single failures, or anticipated frequent failures of components, do not lead to costly challenges on the protection system or to higher levels of trust than necessary.

In the context of trading venues and computer-based trading, similar considerations and trade-offs apply. At one extreme one could have trading constrained in such a way that there is no need for any additional protection (akin to having intrinsic safety in engineered systems) and at the other an unconstrained approach where there was fast, trusted and powerful protection that enabled complete freedom for the trading approaches (somewhat analogous to unstable aircraft where they can only be flown with continuous computer based control).

In practice one suspects a balanced strategy would be required and indeed a different strategy for different types of hazard. How to decide on a particular approach is outside the scope of this paper, but it illustrates that there is close coupling between:

- How much trust we need in the trading algorithms and platforms;

- How much trust is needed in any protection mechanisms (whether automated or procedural)

As we are concerned about systemic risk it is likely that different approaches will be required for (see also Table 5 earlier in this Section):

- Single users of algorithms

- Collective/correlated behaviour across algorithms/participants within a platform/venue;

- Collective/correlated behaviour across platform/venue

So we could imagine an approach under which the market and venues should be able to tolerate a single rogue algorithm. In addition some as yet un-designed protection could be deployed against hazardous collective behaviours with measures taken to address correlated and cascade failures across markets/venues. Together, these approaches would be shown to present tolerable systemic risks.

If such an approach was adopted, then one could foresee trust requirements being articulated for the computer-based trading and the protection systems. These would differ from the nuclear example in:

- The speed of response and functionality of the protection
- The trusted needed in the protection
- The nature and assurance of the trading algorithms.

The latter concern the rapid rate of adaptation of the algorithms, the development lifecycles, the emphasis on rapid prototyping and back testing to gain assurance, and the risk management via gradual introduction into service. The safety properties of the algorithm may be very different from the overall functionality e.g. the need for high confidence in lack of extreme behavior.

4.3. Questions and issues

The above discussion leads to the following questions:

1. What would be the advantages/disadvantages of having an explicit assessment of the trust needed in computer-based systems and prospective protection and control measures?
2. What are the trade-offs between providing protection mechanisms at a venue level vs. those on individual users of algorithms?
3. What different levels of trust for individual, collective and cross-market behaviours are required?
4. What software engineering techniques would be appropriate to assure future algorithmic systems?

5. Conclusions

The nuclear industry and finance industries may seem worlds apart. A nuclear plant relies on decades of science based engineering, the plant is static, physically identifiable, remotely located, each reactor owned and licensed to a single operator with strong incentives to ensure safety and to ensure the remaining risks are tolerable.

The finance industry relies on centuries-old risk concepts, yet is fluid, innovative, and fast changing. Risk taking is an intrinsic part of its day-to-day functioning. Diversity abounds, both in terms of market participants and infrastructure providers. Competition between participants and infrastructure providers drives both innovation and risk taking. Technology allows participants to be present in multiple venues at once.

Yet this industry too is concerned with safety and systemic risk mitigation as well as its impact on the broader economy. Both market participants and infrastructure providers have incentives to ensure the system is robust and inspires confidence. As described in [Foresight (2011)], the increase of computer-based trading has created new challenges for the industry. These relate to the understanding of the interaction between human traders and computer algorithms (see also [Foresight (2011), DR13], the implications for systemic risk and the development of new risk controls for use by both market participants and infrastructure providers.

In this paper, we have focused on three areas where the issues and practices in the nuclear industry resonate with those raised by the evolution of computer-based trading in financial markets. These are:

- The approaches to systemic risk definition and evaluation.
- The defintion of protection system parameters, risk controls and architecture.
- The need for trust in computer-based systems.

The paper is written for the Foresight project and is constrained to not develop policy recommendations. However, we have identified a number of key questions that we think capture the findings of this study and that could inform future discussions.

5.1. Approaches to systemic risk

Looking ahead, one could question whether the rapid development of computer-based trading in financial markets requires the adoption of additional risk concepts and tools. Our analysis suggests that the following questions are worth asking:

1. Is it possible to have a more precise description of risk categories (e.g. in terms of the type of consequences, who is affected, the initiating events that precipitated them)?
2. Is it possible to define precise tolerability criteria? Can one distinguish between tolerable and broadly acceptable risks?
3. Is it possible to define numerical targets? If not, how does one define 'acceptable' risk?
4. Is it possible to develop the notion of a 'design basis,' which would characterise those adverse endogenous and exogenous events that the system (i.e. the market with its control and protective mechanisms in place) should withstand?

Our analysis showed that systemic risk is a multi-faceted concept, which is difficult to measure and monitor. Nonetheless, the comparison with the nuclear industry outlines useful questions for future work.

5.2. Protection systems

In other Foresight reviews (Foresight (2011) DR4) it is agued that financial markets have become complex adaptive systems, in which extreme events can occur in unexpected ways Moreover, as financial markets have become increasingly interconnected, they can be viewed as 'systems of systems.' This means that a failure in one or more constituent parts (market venues) could have widespread repercussions. It also implies that system-level failure is difficult to predict, not only because both humans and computers can adapt their behaviour over time (and can do so at high speed), but also because of the sheer number of possible interactions between humans and computers, both within and across venues. These complexities make it worth asking whether the concept of a protection or viability envelope would be helpful and at the same time these complexities add enormously to the challenge of designing and validating such an approach. We have identified the following specific questions to help articulate these issues:

1. What would the protection and control envelopes look like?
2. What would be the parameters that need to be measured and what would we infer from them? How are they related to existing controls such as price limits or circuit breakers?
3. What would the availability and reliability requirement be for such a system e.g. the probability of failure on demand, the frequency of spurious activation?
4. What is the balance between automation and operator recovery?
5. What additional understanding (and research) is needed given the complex adaptive systems nature of markets? How would the markets adapt to having such protection?

6. What additional analysis techniques and data are needed to assess the risks arising from correlated failures and to design risk controls to guard against their impact?

The questions raised could be useful in further exploring the challenge of developing viability envelopes and designing protection systems.

5.3. Computer assurance

Computer based systems, and ICT in general, are of course essential for high-frequency trading and for algorithmic trading more broadly. In all cases, systems will need to be sufficiently trusted: how do we describe that level of trust and how is this evaluated?

Our comparison with the nuclear sector leads to the following questions:

1. What would be the advantages/disadvantages of having an explicit assessment of the trust needed in computer-based systems and prospective protection and control measures?
2. What are the trade-offs between providing protection mechanisms at a venue level vs. those on individual users of algorithms?
3. What different levels of trust for individual, collective and cross-market behaviours are required?
4. What software engineering techniques would be appropriate to assure future algorithmic systems?

Our analysis in this paper underlines the importance of trust in computer-based systems. The questions outlined above may be helpful in exploring this topic in the context of financial markets and to assess whether it would be worthwhile to use some of the nuclear assurance strategies and techniques as a basis for innovative approaches in the financial sector.

5.4. Some final observations

In both industries, there are limitations to the amount of systemic risk that can be mitigated: both industries bring benefits and have, inherent in their activities, risks to society as a whole. As discussed in the paper, some of these limitations a may arise from difficulties in monitoring risks, difficulties in assessing the effectiveness of risk controls that are seldom used, or difficulties in anticipating the full range of systemic events. But other limitations arise endogenously from the decisions taken by human operators and the licensees.

First, technological innovation permits efficiency savings and improved risk management (thus pushing out the protection envelope), but it may also create new risks (Foresight (2001) DR3). Rapid technological innovation could also limit the ability of regulators, operators and licensees to monitor and understand new risks, and apply new risk controls.

Second, competition between key participants may create incentives to pursue profit-maximising activities that are not necessarily risk-reducing. In the finance industry, competition between market participants, and between trading venues, has created incentives to invest in high-speed technology. This has caused some to worry about an ‘arms race’ (Haldane (2012)). In the nuclear industry licensees are not incentivised to take risks, in fact the reverse, and there is of course strong regulatory oversight[21]. However misaligned economics incentives might

[21] Some commentators would disagree, see e.g. Perrow (2007) and the US nuclear industry

lead to lack of profitability with consequential reductions in investment with potential impact on safety culture and state of plant.

Third, over time, risk frameworks may become obsolete as market participants and computer algorithms adapt their behaviour, risk frameworks may become obsolete. But adaptive behaviour has positive features too: feedback and learning are essential parts of a safety culture. The nuclear industry has its share of accidents and incidents (from Windscale to Fukushima) that cause reflection and reanalysis of its risk frameworks.

Fourth, part of this adaptive behaviour may be a response to regulatory differences that at times exist between jurisdictions. In spite of efforts to create internationally-consistent risk frameworks, participants may have incentives to develop operations in jurisdictions where the regulatory burden is perceived to be lighter. This so-called regulatory arbitrage is constrained by fixed costs (e.g. developing new trading systems). Nonetheless, it is a concern in the finance industry and to a much lesser extent in the nuclear industry as moving plant between jurisdictions is not usually feasible.

And finally, although both industries are so different in terms of the culture, technology, regulation, incentives, geography, history, rate of evolution, and their fundamental purpose, the fact that they both focus on societal significant systemic risks has provided the authors with a stimulating perspective on how risks might be evaluated, controlled and communicated in the future.

Acknowledgments

Robin Bloomfield would like to thank the high frequency trading experts that made their expertise available to brief and tutor him. We would both like to thank the anonymous reviewers as well as the experts who provided us with feedback on early drafts: P Bond, D Cliff, A S L Guerra, B Littlewood and L Strigini.

References

Areva (2011) Overview of the UK EPR GDA Submission, UKEPR-0013-001 Issue 01, Areva 2011

Areva (2011b) UK EPR PCSR – Sub-chapter 14.7 – Fault and Protection Schedule, Areva 2011

Barro, R. and Ursua, J. (2009), Stock Market Crashes and Depressions, NBER working paper 14760.

Basel Committee of Banking Supervision (BCBS) (BIS 2009), Revisions to the Basel II Market Risk Framework – Final Version, July.

Basel Committee of Banking Supervision (BCBS) (BIS 2011), Messages from the Academic Literature on Risk Management for the Trading Book, January.

Basel Committee of Banking Supervision (BCBS) (2012), Fundamental Review of the Trading Book, consultative document, May.

Billio, M., Getmansky, M., Lo, A. and Pelizzon, L. (2010), Econometric measures of systemic risk in the finance and insurance sectors, NBER working paper 16223.

Bishop P (2010) and R Bloomfield, “Safety and Assurance Cases: Past, Present and Possible Future”, Safety Critical Systems Symposium, Bristol, UK, 9-11 Feb 2010

Bloomfield R E (1998), P G Bishop, C C M Jones, P K D Froome, ASCAD—Adelard Safety Case Development Manual, Adelard 1998, ISBN 0-9533771-0-5.

Borio, C and Drehman, M (2009), Towards an operational framework for financial stability: “fuzzy” measurement and its consequences, BIS working paper no. 284, June.

Bordo, M. (2003), Stock market crashes, productivity booms, busts and recessions: some historical evidence, mimeo Rutgers University.

Brunnermeier, M and Pedersen, L. (2009), Market Liquidity and Funding Liquidity, Review of Financial Studies, vol 22, pp 2201-2238

Bsiais, D., Flood, M., Lo, A, and Valavanis, S. (2012), A survey of systemic risk analytics, Office of Financial Research, Working paper 1, January.

Campbell, J., Lo, A. and MacKinlay, C. (1996), The Econometrics of Financial Markets, Princeton University Press, Princeton.

Cliff, D., Brown, D. and Treleaven, Ph. (2011), Technology trends in the financial markets: A 2020 vision, Foresight, DR 3.

Cliff, D. and Northrop, L. (2011), The global financial markets: an ultra-large-scale systems perspective, Foresight, DR 4.

Committee on Payment and Settlement Systems (CPSS) (2003), A glossary of terms used in payments and settlement systems, March.

Commodity Futures Trading Commission (CFTC) and Securities and Exchange Commission (SEC) (2010), Findings Regarding the Market Events of May 6, 2010, Report to the Joint Advisory Committee on Emerging Regulatory Issues, September.

Cuttler, D., Poterba, J. and Summers, L. (1989), What moves stock prices, Journal of Portfolio Management, Summer, pp. 4-12.

Danielson, J., Shin, H. S. and Zigrand, J.P. (2009), Risk Appetite and Endogenous Risk, mimeo, London School of Economics.

Deffuant, Guillaume "Viability and Resilience of Complex Systems: Concepts, Methods and Case Studies from Ecology and Society (Understanding Complex Systems)", ISBN-13: 978-3642204227, Springer 2011.

Easley, D., Lopez de Prado, M. and O'Hara, M. (2011), The Microstructure of the "Flash Crash": Flow Toxicity, Liquidity Crashes, and the Probability of Informed Trading, Journal of Portfolio Management, Winter.

D. E. Eckhardt, A. K. Caglayan, J. C. Knight, L. D. Lee, D. F. McAllister, M. A. Vouk and J. P. J. Kelly, "An Experimental Evaluation of Software Redundancy as a Strategy for Improving Reliability", IEEE Transactions on Software Engineering, 17 (7), pp.692-702, 1991.

European Securities Markets Authority (ESMA) (2011), Guidelines on systems and controls in an automated trading environment for trading platforms, investment firms and competent authorities, Final report, December.

Financial Services Authority (FSA) (2010), The FSA's markets regulatory agenda, May.

Financial Stability Board (FSB) (2009), Guidance to assess the systemic importance of financial institutions, markets and instruments: initial considerations, Report to the G-20 Finance Ministers and Central Bank Governors, October.

Flannery, M., Glasserman, P., Mordecai, D. and Rossi, C. (OFR) (2012b), Forging Best Practices in Risk Management, Office of Financial Research, Working Paper 2, March.

Foresight (2011) The Future of Computer Trading in Financial Markets, Working paper 11-1276, Foresight, Government Office for Science 2011.

Foresight (2011b) Marco De Luca, Charlotte Szostek, John Cartlidge, & Dave Cliff, Studies of Interactions Between Human Traders and Algorithmic Trading Systems, The Future of Computer Trading in Financial Markets - Foresight Driver Review – DR 13, Government Office for Science 2011.

Guerra, S (2010) , P Bishop, R Bloomfield and D Sheridan, "Assessment and Qualification of Smart Sensors", 7th American Nuclear Society International Topical Meeting on Nuclear Plant Instrumentation, Control and Human-Machine Interface Technologies NPIC&HMIT 2010, Las Vegas, Nevada, November 7-11, 2010.

Haldane, A., Saporta, V., Hall, S. and Tanaka, M. (2004), Financial stability and macroeconomic models, Financial Stability Review, June, pp. 80-88.

Haldane, A. (2012), Financial Arms Races, speech, Bank of England website.

HSE (1992) The tolerability of risk from nuclear power stations, The Stationery Office 1992 ISBN 0 11 886368 1, web version: www.hse.gov.uk/nuclear/tolerability.pdf

HSE (2001) Health and Safety Executive “Reducing risk, protecting people: HSE’s decision making process”, ISBN 0 7176 2151 0, HMSO 2001

HSE (2005) Health and Safety Executive, “Technical Assessment Guide: Guidance on the Purpose, Scope and content of Nuclear Safety Cases,” T/AST/051, Issue 001, 13 May, 2005

HSE (2006) Health and Safety Executive, “Nuclear Safety Assessment principles”, http://www.hse.gov.uk/nuclear/SAPs/SAPs2006.pdf

HSE (2010) Health and Safety Executive, “ALARP at a glance”, http://www.hse.gov.uk/risk/theory/alarpglance.htm

HSE (2011) New Reactor Build EDF/AREVA EPR Step 2 PSA Assessment from www.hse.gov.uk/newreactors/reports/eprpsa.pdf

IAEA (2000), IAEA NS-R-1 Safety of nuclear power plants: design: safety requirements, September 2000.

IMF (2010), Global Financial Stability Report, April.

IMF (2009), Global Financial Stability Report, April.

IAEA (2012) INES Factsheet http://www.iaea.org/Publications/Factsheets/English/ines.pdf

IOSCO (2010), Objectives and Principles of Securities Regulation, June.

IOSCO (2011a), Regulatory Issues Raised by the Impact of Technological Change on Market Integrity and Efficiency, final report, December.

IOSCO (2011b), Mitigating Systemic Risk. A Role for Securities Regulators, Discussion Paper, February 2011.

Johanssen, A. and Sornette, D. (2001), Large Stock Market Price Drawdowns are Outliers, Journal of Risk, 4, pp. 69-110.

Johnson, N., Zhao, G., Hunsader, E., Meng, J. Ravindar, A., Carran, S. and Tivnan, B. (2012), Financial black swans driven by ultrafast machine ecology, mimeo, University of Miami.

Kannan, P., Pabanal, P. and Scott, A. (2009), Macroeconomic Patterns and Monetary Policy in the Run-up to Asset Price Bubbles, IMF Working Paper 09/52, November.

Kay, J. (2012), The Kay Review of UK Equity Markets and Long-Term Decision Making, Interim Report, February.

Kelly, T P (2004) and R A Weaver, “The Goal Structuring Notation - A Safety Argument Notation”, Proceedings of the Dependable Systems and Networks 2004 Workshop on Assurance Cases, July 2004.

Khandany, A.E. and Lo, A. W. (2011) What Happened To The Quants In August 2007?, Journal of Financial Markets, 14, pp. 1-46.

Kyle, A. (1985), Continuous Auctions and Insider Trading, Econometrica, vol. 53, pp. 1315-35.

Knight, J.C and N.G. Leveson, “An Empirical Study of the Failure Probabilities in Multi-Version Software”, Proc. FTCS 16, Vienna, July 1986.

Layfield (1987), Sir Frank, “Sizewell B public enquiry report”, London H.M.S.O, 1987.

Littlewood, B (1993) and L Strigini (1993), “Assessment of ultra-high dependability for software-based systems”, Communications of the ACM, 36 (11), pp.69-80.

Littlewood, B (1998), “The Use of Computers in Safety-Critical Applications, Final Report of the Study Group on the Safety of Operational Computer Systems constituted by the Advisory Committee on the Safety of Nuclear Installations”, HSE Books London 1998 also available at www.hse.gov.uk/nuclear/computers.pdf.

Littlewood, B (2010)., Bishop, P., Bloomfield, R., Povyakalo, A., & Wright, D. Towards a formalism for conservative claims about the dependability of software-based systems. IEEE Transactions on Software Engineering. doi:10.1109/TSE.2010.67

Longin, F (2000), From VaR to Stress Testing: the Extreme Value Approach, Journal of Banking and Finance, vol. 24, pp. 1097-1130.

Meulen (2008) Meine J.P. van der Meulen, and Miguel A. Revilla The Effectiveness of Software Diversity in a Large Population of Programs, IEEE Transactions on Software Engineering Vol. 34, No. 6, November/December 2008.

McNeil, A, Frey, R. and Embrechts, P. (2005), Quantitative Risk Management: Concepts, Techniques and Tools, Princeton Series in Finance, Princeton, N.J.

Oberkampf W L (2004) and J.C. Helton, “Alternative Representations of Epistemic Uncertainty,” Reliability Eng. and System Safety, vol. 85, special issue, 2004.

ORNL (2000) Reported at http://www.ornl.gov/info/reporter/no19/scram.htm

Perrow, C (2007), The Next Catastrophe: Reducing Our Vulnerabilities to Natural, Industrial, and Terrorist Disasters, Princeton University Press, ISBN: 9781400838516

Schwaab, B., Koopman, S. and Lucas, A. (2011), Systemic risk diagnostics. Coincident indicators and early warning signals.

Securities and Exchange Commission (SEC) (2011), Recommendations Regarding Regulatory Responses to the Market Events of May 6, 2010, Summary Report, February.

Shiller, R.J. (2000), Irrational exuberance, Princeton University Press, Princeton, N.J..

Sornette, D. (2003), Why Stock Markets Crash: Critical Events in Complex Financial Systems, Princeton University Press, Princeton, N.J.

Sornette, D. And Johanssen, A. (2001), Significance of log-periodic precursors to financial crashes, Quantitative Finance,1 (4).

Toulmin, S E (1958,) The Uses of Argument, Cambridge University Press, 1958.

Tucker, P (2011), Macroprudential policy: building financial stability institutions, Bank of England, April.

Glossary

Term/Abbreviation	Explanation
ACSNI	Advisory Committee on the Safety of Nuclear Installations
BE **BSL** **BSO**	British Energy Basic Safety Level Basic Safety Objective
CAE	Claims Arguments Evidence
FSB	Financial Stability Board
GSN	Goal-Structuring-Notation
HSE	Health and Safety Executive
HSW	Health and Safety at Work Act, 1974
IAEA	The International Atomic Energy Agency
MoD	Ministry of Defence
NIA	Nuclear Installations Act 1965
NII	Nuclear Installations Inspectorate
PFD	Probability of failure on demand
SAPs	Safety Assessment Principles
TAGs	Technical Assessment Guides

The following table is extracted from Health and Safety Executive, “Nuclear Safety Assessment principles” (HSE (2006)) unless otherwise indicated.

Term	Explanation
Accident	Any unintended event, including operator errors, equipment failures or other mishaps, the consequences or potential consequences of which are not negligible from the point of view of protection or safety (IAEA Safety Glossary). In this document, and when used generally, the term ‘accident’ includes any undesired circumstances which give rise to ill health or injury; damage to property, plant, products or the environment; production losses or increased liabilities. When referring to nuclear safety, ‘accident’ refers to a fault sequence resulting in a dose greater than 0.1 mSv to a worker, or greater than 0.01 mSv to a person outside the site, or in a substantial unintended relocation of radioactive substances within the facility.
Accident management	The strategies which are developed to reduce the risks arising from accidents, and bring the facility to a safe, controlled state.
Alarm	An automatic visual or audible indication to personnel of when a specific plant variable or condition has reached a pre-set limit or state.
Availability	The fraction of time for which a system is capable of fulfilling its intended purpose (IAEA Safety Glossary).
Barrier	A means to: • prevent or inhibit the movement of people or radioactive substances, or some other phenomenon (eg fire); • provide shielding against radiation; • protect against some other potentially hazardous event.
Best estimate	When used to describe analysis, this refers to an analysis expected to provide the most accurate description of the fault and its consequences that could be achieved within the limitations of the analytical model employed without any deliberate bias being introduced. When used to describe the

Term	Explanation
	data, it refers to the most accurate value of the data item derived from experiment, operating experience, judgement etc as appropriate. Where there is inadequate evidence, and no credible best estimate is possible, then bounding or conservative values should be used.
Bounding case	The case that represents the extreme consequences of a class of accidents.
Class of accident	A group of fault sequences that follow paths that are sufficiently similar to justify analysis of the sequences together as a class.
Common cause failure (CCF)	Failure of two or more structures, systems or components due to a single specific event or cause (IAEA Safety Glossary).
Common mode failure (CMF)	Failure of two or more structures, systems or components in the same manner or mode due to a single event or cause (IAEA Safety Glossary).
Conservative estimate	The use of models, data and assumptions which would be expected to lead to a result that bounds the best estimate (where known) on the safe side. The degree of conservatism should be proportionate to the level of uncertainty, and the overall significance of the estimate to the safety case.
Containment	Methods or physical structures designed to prevent the dispersion of radioactive substances (IAEA Safety Glossary).
Countermeasures	An action aimed at alleviating the radiological consequences of an accident (IAEA Safety Glossary).
Defence in depth	An approach to designing and operating nuclear facilities that prevents and mitigates accidents that release radiation or hazardous materials. The key is creating multiple independent and redundant layers of defense to compensate for potential human and mechanical failures so that no single layer, no matter how robust, is exclusively relied upon. Defense-in-depth includes the use of

Term	Explanation
	access controls, physical barriers, redundant and diverse key safety functions, and emergency response measures. (NRC glossary)
Design basis	The range of conditions and events that should be explicitly taken into account in the design of the facility, according to established criteria, such that the facility can withstand them without exceeding authorised limits by the planned operation of safety systems (IAEA Safety Glossary).
Design basis fault	A fault (sequence) which the plant is designed to take or can be shown to withstand without unacceptable consequence, by virtue of the facility's inherent characteristics or the safety systems.
Diversity	The presence of two or more systems or components to perform an identified function, where the systems or components have different attributes so as to reduce the possibility of common cause failure, including common mode failure (IAEA Safety Glossary).
Dutyholder	A person or corporate body who has a duty in law.
Essential service	Essential services are all those resources necessary to maintain the safety systems in an operational state at all times and may include electricity, gas, water, compressed air, fuel and lubricants.
External hazard	External hazards are those natural or man-made hazards to a site and facilities that originate externally to both the site and the process, i.e. the dutyholder may have very little or no control over the initiating event.
Failure	A failure has occurred when a structure, system or component fails to meet its safety function, or functions spuriously
Failure modes	The manner or state in which a structure, system or component fails (IAEA Safety Glossary).
Fault	Any unplanned departure from the specified mode

Term	Explanation
	of operation of a structure, system or component due to a malfunction or defect within the structure, system or component or due to external influences or human error.
Fault condition	When used without qualification, this means design basis fault conditions and includes, where appropriate and as far as reasonably practicable, beyond design basis conditions. A combination of an initiating fault and any additional failures.
Fault sequence	A combination of an initiating fault and any additional failures, faults and internal or external hazards which have the potential to lead to accidents.
Hazard	The potential for harm arising from an intrinsic property or disposition of something to cause detriment (R2P21). See also internal and external hazards.
Hazard potential	The propensity for the harm from a hazard to be realised.
Incident	An undesired circumstance or ‘near miss’ that has the potential to cause an accident.
Inherent safety	Preventing a specific harm occurring by using an approach, design or arrangement that ensures that the harm cannot happen, for example a criticality safe vessel. This is not the same as passive safety.
Initiating fault	The starting event of a fault sequence. This may be an internal failure or a fault caused by an internal or external hazard or by human action. This does not include pre-existing latent failures that may be revealed when safety equipment is called upon to function during a fault sequence.
Intelligent customer	An intelligent customer is the capability of an organisation to have a clear understanding and knowledge of the product or service being supplied.
Internal hazard	Internal hazards are those hazards to plant and

Term	Explanation
	structures that originate within the site boundary and over which the dutyholder has control over the initiating event in some form.
Licensee	The body corporate that has been granted a Nuclear Site Licence under the Nuclear Installations Act 1965 (as amended), which permits it to carry out a defined scope of activities on a delineated site (NIA).
Passive safety	Providing and maintaining a safety function without the need for systems to be actively initiated or for operator intervention, or other safety system support features. In the context of decommissioning and the storage of nuclear matter, providing and maintaining a safety function by minimising the need for active safety systems, monitoring or prompt human intervention. A passive safety system is not necessarily inherently safe.
Protection system	A system that monitors the operation of a facility and which, on sensing an abnormal condition, automatically initiates actions to prevent an unsafe or potentially unsafe condition (based on IAEA Safety Glossary).
Qualification	The process of demonstrating that a structure, system or component is fit for its intended purpose.
Redundancy	Provision of alternative (identical or diverse) structures, systems or components, so that any one can perform the required function regardless of the state of operation or failure of any other (IAEA Safety Glossary).
Risk	Risk is the chance that someone or something is adversely affected in a particular manner by a hazard (R2P21).
Safety actuation system	The collection of equipment required to accomplish the necessary safety actions when initiated by the protection system (IAEA Safety Glossary).

Term	Explanation
Safety case	In this document, 'safety case' refers to the totality of a licensee's (or dutyholder's) documentation to demonstrate safety, and any sub-set of this documentation that is submitted to NII.
Safety function	The safety function of a structure, system or component is the specific function required to maintain the facility within the safe operating limits and conditions determined by the fault analysis.
Safety measure	A safety system, or a combination of procedures, operator actions and safety systems that prevents or mitigates a radiological consequence, or a specific feature of plant designed to prevent or mitigate a radiological consequence by passive means.
Safety-related system	An item important to safety that is not part of a safety system (IAEA Safety Glossary).
Safety system	A system that acts in response to a fault to prevent or mitigate a radiological consequence.
Safety schedule	A schedule or other suitable means that identifies the minimum safety system requirements for each of the initiating faults, including internal and external hazards, identified within the design basis. A safety schedule may also be called a safety system or engineering or protection or fault and protection schedule. Other suitable means may include the use of configuration diagrams.
Segregation	The physical separation of components and systems, by distance or by some form of barrier that reduces the likelihood of common cause failures.
Severe accident	A fault sequence which leads either to consequences exceeding the highest radiological doses given in the BSLs of Target 4, or to a substantial unintended relocation of radioactive material within the facility which places a demand on the integrity of the remaining physical barriers.
Societal concerns	Societal concerns are the risks or threats from hazards which impact on society and which, if

Term	Explanation
	realised, could have adverse repercussions for the institutions responsible for putting in place the provisions and arrangements for protecting people.
Societal effects	A term used to describe those societal concerns that are capable of quantitative prediction such as numbers of deaths or injuries, numbers of people evacuated, area of land contaminated and general economic loss.
Societal risk	The risk of an accident causing the death of a specified number of people in a single event from a single major industrial activity, ie an activity from which risk is assessed as a whole and is under the control of one company in one location, or within a site boundary.
Validation	Dependent on context: 1. The process of determining whether a product or service is adequate to perform its intended function satisfactorily. Computer system validation: The process of testing and evaluation of the integrated computer system (hardware and software) to ensure compliance with the functional, performance and interface requirements. Model validation: The process of determining whether a model is an adequate representation of the real system being modelled, by comparing the predictions of the model with observations of the real system. System code validation: Assessment of the accuracy of values predicted by the system code against the relevant experimental data for the important phenomena expected to occur. 2. Confirmation by means of objective evidence that the requirements for a specific intended purpose and use or application have been fulfilled (IAEA Safety Glossary).
Verification	Dependent on context: 1. The process of determining whether the quality or performance of a product or service is as stated, as intended or as required.

Term	Explanation
	Computer system verification: The process of ensuring that a phase in the system life-cycle meets the requirements imposed on it by the previous phase. Model verification: The process of determining whether a computational model correctly implements the intended conceptual model or mathematical model. System code verification: Review of source coding in relation to its description in the system code documentation. 2. Confirmation by means of objective evidence that specified requirements have been fulfilled (IAEA Safety Glossary).
Veto	Inhibition of a safety system.

Foresight
1 Victoria Street
London SW1H 0ET
www.foresight.gov.uk

URN: 12/1059