Accepted Manuscript

Agent-based computational modelling of social risk responses

J.S. Busby , B.S.S. Onggo , Y. Liu

Please cite this article as: J.S. Busby , B.S.S. Onggo , Y. Liu , Agent-based computational modelling of social risk responses, *European Journal of Operational Research* (2015), doi: 10.1016/j.ejor.2015.12.034

Highlights

- Managing major societal risks involves the need to understand public risk responses.

- The social amplification of risk framework has been our main theoretical approach.

- We explore how to model endogenised risk observation, behaviour and communication.

- Agent simulation shows characteristic outcomes like peaks and drift in risk beliefs.

- The model indicates the key areas where further empirical research is needed.

Agent-based computational modelling of social risk responses

J.S. Busby, B.S.S. Onggo and Y. Liu

Department of Management Science

Lancaster University

Lancaster

LA1 4YX

UK

Corresponding author: J.S.Busby; +44 1524 594447; j.s.busby@lancaster.ac.uk

Abstract

A characteristic aspect of risks in a complex, modern society is the nature and degree of the public response – sometimes significantly at variance with objective assessments of risk. A large part of the risk management task involves anticipating, explaining and reacting to this response. One of the main approaches we have for analysing the emergent public response, the social amplification of risk framework, has been the subject of little modelling. The purpose of this paper is to explore how social risk amplification can be represented and simulated. The importance of heterogeneity among risk perceivers, and the role of their social networks in shaping risk perceptions, makes it natural to take an agent-based approach. We look in particular at how to model some central aspects of many risk events: the way actors come to observe other actors more than external events in forming their risk perceptions; the way in which behaviour both follows risk perception and shapes it; and the way risk communications are fashioned in the light of responses to previous communications. We show how such aspects can be represented by availability cascades, but also how this creates further problems of how to represent the contrasting effects of informational and reputational elements, and the differentiation of private and public risk beliefs. Simulation of the resulting model shows how certain qualitative aspects of risk response time series found empirically – such as endogenously-produced peaks in risk concern – can be explained by this model.

Keywords

OR in societal problem analysis; multiagent systems; risk management

1 Introduction

Managing the major risks experienced by a complex society – the risks of epidemic disease, climate change, food and drug contamination, the catastrophic failure of hazardous installations and so on – almost invariably involves managing public anxiety or public complacency as well as containing physical threat (Leiss, 2001). Not only are public perceptions pivotal in shaping public behaviour, and therefore exposure to the threat, but characteristically produce further risks. A recent analysis of the Fukushima nuclear power accident argued that:

> 'There was a rushed evacuation response to the accident… this evacuation actually led to more premature deaths, by a factor of at least ten, than it gave protection from radiation… The reaction was driven to a large extent by the public's sense of the scale of the hazard, which was not close to the reality of the risk… The studies of many previous accidents have come to a similar conclusion. Even for an accident as significant as Chernobyl it can be shown that the vast majority of the public health impacts are caused by mental stress relating to the fear of the event, rather than the effects caused by the amount of ionising radiation released...' (Cahart, 2013).

Thus public risk perceptions have often mattered more than objective assessments of risk, as seen in such celebrated cases as Love Canal, Alar and TWA 800 in the US (Kuran and Sunstein, 1999) and the Sudan 1 and Hatfield scandals in the UK (Busby and Alcock, 2008). It has become essential for organizational decision making to be founded on an understanding of societal risk responses, and for decision makers to theorise, however loosely, about how such responses arise.

The formation of these responses has a number of defining features. Most if not all of the public, and many managers, have no first-hand technical knowledge of the risk and rely on other social actors – including the media – of whom they are often sceptical if not cynical (Petts and Niemeyer, 2004). These actors in turn generally have a clear appreciation of this cynicism and anticipate it in the way they act and communicate (Busby and Duckett, 2012). The responses of a wide range of actors, including risk managers and the general public, typically influence the character of the threat and the risk bearers' exposure to it (Busby and Onggo, 2013). Responses are shaped by the way in which such groups inter-communicate within their social networks (Scherer and Cho, 2003). And the responses become events in their own right, to which social actors further respond (Kasperson et al, 1988).

3

Credible models of social risk responses need to incorporate such features. They need to endogenise observation, representing the way in which actors, despite their heterogeneity, often base their own responses in part on how they see peers or neighbours responding, not on direct experience or knowledge of the risk. They need to endogenise behaviour, representing the way actors adapt their behaviour to changing observations of a risk, thus changing their exposure and the risk itself, and thereby also changing subsequent perceptions of this risk. And they need to endogenise risk communication, representing the way actors base their risk perceptions on the communications of others whose apparent biases they correct for, but who in turn can anticipate such corrections in formulating their communications. Yet, as Rahmandad and Sterman (2008) point out, we typically model disease outbreaks as though contact rates were fixed, ignoring the way people change their behaviour as prevalence grows. And, as Busby and Onggo (2013) argue, we have typically ignored the way in which actors communicating about risk anticipate each other's biases, and even anticipate the anticipation of each other's biases. The aims of this paper are to explore how an agent based model can incorporate these characteristics, to explore what we can say about model validity, and to explore what quantities need to be known in order to parameterise such a model.

Our main theoretical foundation for doing this is the 'social amplification of risk framework' (Kasperson et al, 1988). The development of this framework followed earlier lines of work on individual risk perception (broadly starting with Fischhoff et al, 1978), and on cultural risk selection (broadly starting with Douglas and Wildavsky, 1982). It has probably been the only mainstream attempt to try to synthesise this prior work, to deal with social emergence and to capture the importance of social communication in explaining risk behaviour (Renn, 1991). Much empirical work on risk responses in various domains, ranging from nuclear waste to terrorism, has been done under the heading of social risk amplification. But, as we attempt to show in the next section, the social amplification of risk remains largely un-modelled and under-specified. This means that various empirical findings that have emerged over the last 25 years remain ambiguous, and the implications for decision makers unclear.

There are several observations in the literature that motivate the use of agent models in particular. First, in the social risk amplification framework, risk has been seen as first and foremost a matter of social communication (Luhmann, 1993; Renn, 1991). The essence of the social amplification framework is that some risk event is experienced by a very small number of

4

social actors, and communication about the risk then spreads through a system of heterogeneous actors seen as 'amplification stations' (Kasperson et al, 1988). Second, empirical work – notably Scherer and Cho's (2003) article and more recently Muter et al's (2013) in the risk literature, but also work such as that of Kohler et al (2007) in the demography literature – has shown how important social interactions are in the development of risk perceptions. An individual's risk beliefs tend to be strongly correlated to those of others with close social connections, and individuals' reports tend to acknowledge how those others have influenced them. Third, non-linearities are central to how risk amplification arises. Some of the few prior attempts at modelling risk amplification (Burns and Slovic, 2007; Busby and Onggo, 2012) have shown how complex are the feedback loops between the perceptions, behaviours and communications among the different actors in a risk issue, making analytical modelling infeasible. Fourth, the actors respond heterogeneously. Much of the later work on individual risk perception (for example Marris et al 1997, Langford et al 1999) has stressed individual differences. And individual risk sensitivity (Sjoeberg, 2000) has been an important explanatory variable for differences in individual risk perception. This all strongly points to agent-based modelling as the appropriate medium for modelling social risk amplification.

In this article we first review the literature on social risk amplification in an attempt to draw out the main theoretical and empirical contributions that have arisen since it was first proposed. We then describe the construction of a model, justifying its content by reference to the literature. We base this model on Kuran and Sunstein's (1999) account of availability cascades, and show how we can model central aspects of such cascades – particularly the ideas that individuals have both espoused and expressed risk beliefs, that they have both informational and reputational reasons for responding to beliefs common in a social discourse, and that there are availability 'entrepreneurs' who knowingly exploit the possibility of such cascades. We present typical results of simulating the model, and we discuss issues of model validity by reference to empirical work on time series of risk perceptions and concerns in the literature. We conclude with a discussion of the study's implications and limitations.

2 Literature review

The social amplification of risk framework (SARF) was first proposed by Kasperson et al (1988) as a way explaining the often apparently mistaken responses of populations to risks in modern society. The original framework was intended to show 'that risk events interact with psychological, social, and cultural processes in ways that can heighten or attenuate public perceptions of risk and related risk behaviours'. And it stressed the 'ripple' effects through which risk perceptions led people to behave in ways that created secondary impacts beyond the harmful effects of the original risk. Generally, it has been applied to study excessively high rather than excessively low risk perceptions, although the need for symmetry has long been recognised (Rip, 1988).

It has been used in a wide variety of contexts, including wildfire risk (Brenkert-Smith et al, 2013), the siting of potentially hazardous installations (Binder et al, 2011), environmental risk from tunneling (Chung, 2011), disease outbreaks (Lewis and Tyshenko, 2009, Raude et al, 2004, Busby and Duckett, 2012), genetically modified foods (Frewer et al, 2002), the dismantling of hazardous installations (Bakir, 2005), chemical accidents (Porto and de Freitas, 1996), climate change (Renn, 2011), nuclear weapons facility accidents (Metz, 1996), inoculation risks (Petts and Niemeyer, 2004) and general levels of violence in society (Hill, 2001). In such situations, the framework has provided a way of describing how discrepancies between the risk beliefs of different groups, and between experts and lay communities especially, can arise.

The methods used in such studies have been wide-ranging. Some are qualitative, analyzing rich verbal accounts among the public from interviews (for example Masuda and Garvin, 2006) and discussion groups (Petts and Niemeyer, 2004; Busby and Duckett, 2012), or analysing media content (Bakir, 2005). These have revealed how the worldview of individuals affects their tendency to amplify risk, and how particular actors use the media to convey their view of the risk and influence opinion. Quantitative studies have occasionally used economic measures of risk responses, such as property values and business activity (Metz, 1996), and there has been some content analysis of the news media (Lewis and Tyshenko, 2009). But most quantitative work has been based on public surveys (for example Brenkert-Smith et al, 2013; Binder et al, 2011, Frewer et al, 2002). These are generally directed at the public, but some involve surveying the specific groups involved in a particular risk issue, such as physicians

6

dealing with a potential disease outbreak (Raude et al, 2004). Surveys have generally been analysed by regressing risk perceptions, and sometimes amplified risk perceptions, against expected correlates. These include types of information source and social interaction (Brenkert-Smith et al, 2013), attitudes of support or hostility toward a technology (Binder et al, 2011), engagement in public meetings (McComas, 2003) and the volume of reporting (Frewer et al, 2002). Occasional studies have also looked at the covariation of behavioural changes, such as consumption of foods thought to carry disease, with risk perception (Raude et al, 2004).

Mostly these studies use data suggesting the presence of risk amplification, but Lewis and Tyshenko (2009) and Metz (1996) both found expected amplification not to be present, and so were primarily concerned with correlating the *absence* of risk amplification with specific circumstances. Some of the more ambitious studies of social risk amplification, attempting to develop a comprehensive view of the factors causing amplification, have had to use mixed methods. Burns et al (1993) in particular needed a combination of public survey, expert ratings, a Delphi panel and analysis of media attention to develop structural equations linking risk amplification to physical risk and social processes. Both they, and subsequently Freudenberg (2003), emphasized the importance of perceived managerial incompetence or misconduct in amplifying public risk perceptions.

However, a basic limitation of this empirical work is its concentration on the statistical correlates of amplification, rather than the mechanism that produces it. The core of social risk amplification in Kasperson et al's (1988) original formulation is a mechanism rather than a law linking structural variables. The kind of modelling that explicitly represents this mechanism has been very limited to date. There has been some systems dynamics work, based on the observation that social risk amplification involves complex feedback loops connecting risk responses with decisions and behaviours that in turn modify risk and the perceptions of risk. Burns and Slovic (2007) modelled public perception and amplification of terrorism risk, while Busby and Onggo (2013) recently modelled socially amplified public responses to zoonotic disease outbreaks. But these studies assumed essentially homogenous populations. As suggested earlier, social risk amplification was from the start conceived as a communications phenomenon taking place in a network of 'amplification stations' of different kinds. And more recent work on risk perception (Scherer and Cho, 2003; Muter et al, 2013) has shown empirically that individuals' risk perceptions are strongly determined by their social connections. It therefore

7

seems essential that modelling should involve heterogeneous agents interacting within a social network.

There have also been some basic criticisms of the social risk amplification framework itself. In particular, Rayner (1988) argued that the concept naively assumed risks to exist objectively. The criterion for testing whether risk was amplified at any time was whether social beliefs were greater than the objective level. Rayner argued that risks do not exist outside the social system, and those with the expertise to have the most complete and accurate understanding of a risk were still social actors, subject to social processes. A recent study (Busby and Duckett, 2012) made an attempt to deal with this. It involved an empirical analysis of how people form beliefs about distortions in the risk beliefs of others. Most people have to get their information about societal risks from other social actors, rather than direct experience. But, while relying on other actors, they also anticipate how their views might be biased. Some earlier studies pointed in the same direction. For example, Petts and Niemeyer (2004) referred to how people expect the news media to exaggerate risks knowingly. Frewer (2004) argued that people correct and even over-correct for the biases they expect in sources of information that promote their own vested interests (see also Frewer et al, 2003). And Renn and Levine (1991) claim that information receivers by default tend to assume that risk communicators are trying to deceive their audience. Modelling risk amplification should therefore represent not just actors communicating risk beliefs to each other, but actors mutually theorising about and compensating for each other's apparent biases.

We therefore argue that there are several important reasons for modelling social risk amplification generally, and for using agent based modelling specifically. The general justification for using agent based modelling is that it is, as Axelrod (1997: 3) has suggested, 'a third way of doing science' - not proving theorems but making inductions from what are effectively thought experiments. Gilbert and Troitzsch (1999: 5) suggest that building simple simulation models is valuable both as a process of formalizing social theory and discovering its consequences in an artificial society. Computational agent-based models have a specific relevance because 'the consequences of adaptive processes are often very hard to deduce when there are many interacting agents following rules that have non-linear effects...' (Axelrod,1997: 4). They are especially appropriate for studying processes that lack central coordination (Macy and Willer, 2002) and where the effects of diversity in individuals' attributes and behaviours are

8

relevant to the behaviour of a system as whole (Macal and North, 2010). Social risk amplification is a phenomenon in which heterogeneous agents are engaged in parallel processes without a well-defined order of action (Gilbert and Troitzsch, 1999: 6), and in which we cannot reasonably assume homogeneity and perfect mixing within compartments (Rahmandad and Sterman, 2008). Agent based modelling makes a commitment to methodological individualism, tracing all collective phenomena back to individuals but, as Epstein and Axtell (1996: 16) point out, it still allows emergent institutions to have feedback effects on individual agents. This, as we have suggested, is an important characteristic of social risk amplification. Although agent-based approaches appear not to have been used in the development and application of risk amplification, they have been reported in somewhat similar areas – for example the modelling of warning message dissemination through multiple channels (Nagarajan et al., 2012), the mutual influence of a society's members on their choice of transport modes (Sunitiyoso and Matsumoto, 2009), and the diffusion of new products (Amini et al, 2012)).

In the agent model development that follows, we first deal with the core model of risk perceivers' social interaction in Section 3. We then explore how to represent the role of behaviour in Section 4, and then add the representation of inter-group communication in Section 5.

3 The core process of availability and the endogeneity of risk observation

The first key element is the proposition that social actors find out about a societal risk almost entirely from other actors, not from direct physical observation of experience. The 'observations' that shape risk beliefs, come from within the social system. Most obviously actors' beliefs are shaped by those of their peers and a model therefore requires a commitment to a specific mechanism through which social actors interact. Our choice is to draw on Kuran and Sunstein's (1999) notion of availability cascades, which itself derived from Kahnemann and Tverky's (1972) work on the heuristics and biases that arise in human decision making under uncertainty. Kuran and Sunstein (1999) show that an expressed risk perception has increasing plausibility through rising availability in public discourse. They argue that, of all the cognitive heuristics, the availability heuristic is the most fundamental to the social response to risk events

9

in which people tend to form their risk judgments largely on the basis of information exchanged through a social process. Kuran and Sunstein's framework contains several basic elements that need to be modelled: 1) people have internal or espoused beliefs; 2) people have external or expressed beliefs, which may differ from their espoused beliefs; 3) espoused beliefs are influenced by the availability of a risk perception in public discourse for informational reasons (they tell people something about the risk); 4) expressed beliefs are influenced by the availability of a risk perception in public discourse because people are motivated for reputational reasons to conform; 5) there are actors, organizational or individual, who knowingly exploit the availability effect to raise or lower societal risk perceptions. Kuran and Sunstein provide a compelling argument for these elements. They describe how a focus on purely informational effects (found in some parts of the prior literature) is invariably under-socialised, neglecting the way in which people have to manage social reputations; and how a focus on purely reputational effects is over-socialised, neglecting the rationality of relying on social discourse for information about a risk

The availability of a risk perception exists in some social network in which N agents are connected according to some graph G by bi-directional links such that $G_{ij} = 1$ if i and j are social neighbours but 0 otherwise. We fix this network for the rest of the analysis, which seems reasonable in the context of an individual risk issue, but is clearly a simplification of a reality in which unacquainted individuals can be brought together in dealing with a common threat. Each agent i has an espoused (that is, internal) belief, or not, in some proposition in public discourse, $b_{esp,i} \in \{0, 1\}$, and an expressed belief, $b_{exp,i} \in \{0, 1\}$. For example, in the BSE crisis (Beck et al, 2005) the central risk proposition was that a disease crossed the species barrier from cattle to human. In the UK MMR triple vaccine crisis (Stroud, 2005) the risk proposition was that the vaccine caused autism. The defining aspect of the availability heuristic (Tversky and Kahnemann, 1973) is that the probability of a proposition is estimated by its availability, so at any given time t at which i is active its belief is a function of the proportion q_i of its social neighbours expressing that belief, $q_i = (\sum_{j \mid Gij = 1} b_{exp,j}) / |\{ j \mid G_{ij} = 1 \}|$.

Kuran and Sunstein (1999) link espoused beliefs to the informational role of risk communications and expressed beliefs to the reputational role. Internal beliefs about the truth of some hazard are based on the information contained in public discourse. But expressed beliefs are based on reputational motives: on the extent to which individuals feel they should agree with

those around them. At around the same time that Kuran and Sunstein's work emerged another study, from Gardner and Kleinman (2000), showed that informational and reputational motives led to different functional forms of response. Informational motives made the likelihood of influence a concave function of how widely an idea was accepted among a local reference group. Reputational motives made it a convex function. It therefore becomes natural to model the availability effect in this way. Given some positive parameter C we simply update i's beliefs after a discrete time interval to $b_{esp,i}(t+1) = 1$ with probability $q_i(t)^{1/C}$ and $b_{exp,i}(t+1) = 1$ with probability $q_i(t)^C$ and otherwise set them to zero.

This heuristic process of belief updating reflects the fact that agents are characteristically adaptive rather than rational, given their cognitive limitations (Doran et al, 1994), and follow simple procedures when giving and receiving influence (Axelrod, 1997: 153). The basic notion that individuals are willing to set aside what beliefs they formerly had in favour of those received from social interaction underlies thinking about information cascades that has gone on for some time: Kuran and Sunstein cite a number of earlier studies, including the commonly-cited work of Banerjee (1992) and Bikhchandani et al (1998). In practice, the availability effect is not determinate. It will vary between issues and individuals, and quite probably within individuals from one time to another. And individuals may sometimes misperceive the expressed beliefs of others, given the ambiguity of natural language and limited bandwidth of human communications – especially when conducted via social media. It is therefore necessary to incorporate noise in the belief updating process, for example by switching from belief to non-belief or vice versa at random with some relatively small probability v.

Kuran and Sunstein also make an important point about the existence of 'availability entrepreneurs'. Certain agents have incentives to amplify or attenuate risk in public discourse. Providers of protective products and services, for instance, may benefit from a widespread belief in some hazard, whereas providers of products and services that put people at risk may benefit from an attenuated hazard perception. In a detailed case study, Lofstedt (2008) clearly portrays an organizational actor (a cancer research foundation) as an availability entrepreneur, deliberately raising the availability of a specific hazard (the carcinogenic effects of aspartame) in the public discourse. The modelling of this requires that certain agents can spontaneously develop expressed beliefs, rather than simply express beliefs determined by availability. It also requires that these agents have a strong social capacity to promote their beliefs. The simplest,

11

general representation of this is – for every agent – to set some low probability Q with which it spontaneously has an expressed belief value, which is 1 or 0 with equal probability. The entrepreneur's deliberately engineered social influence can then be represented in various ways. Perhaps most simply, availability in all interactions is biased by the entrepreneur's spontaneous expressed belief value by taking a mean of the social neighbourhood availability q_i with the entrepreneur's expressed belief in the S subsequent social interactions, where S is a model constant. Much more complex representations of availability entrepreneurship can readily be envisaged – for example, allowing entrepreneurs to search in the social network for highly connected hubs and influence their expressed beliefs. But in this simple form Figure 1 illustrates the agent's decision rules, at this point, using pseudocode.

Figure 1 here

The outcome of this representation, before a more meaningful level of complexity is introduced, is indicated in Figure 2 by two sample traces from a simulation of 1 000 agents of the proportion believing in a risk proposition over 20 000 model periods, in each of which one agent is selected at random with equal probability for activation. The relevant population will vary from case to case, depending on the specificity and localisation of the risk event in question. Responses to even relatively global crises such as the *BSE* outbreaks were sometimes studied as socially local phenomena in small networks (Lehmkuhl, 2008). Using relatively small numbers also helps avoid 'epistemic opacity' (Miller, 2015). And there is some evidence that important emergent behaviours appear insensitive to network size (Santos and Pacheco, 2005). But modelling a *particular* situation will require a specific judgment of the appropriate N. The black line shows the time series of mean espoused beliefs, and the grey line that of expressed beliefs. Two successive runs of the model with identical parameters are shown in the upper and lower parts of the Figure, indicating the extent to which the belief trajectories are path dependent. For the agent network structure we use a scale-free network with a power law distribution of link numbers k in which the number of nodes with k links is proportional to $k^{-\gamma}$. This seems to apply to many actual networks of social contacts, in which γ generally lies in the range of 2 to 3 (Santos and Pacheco, 2005; Barabasi, 2009). At $t = 0$, agents are randomly assigned an espoused belief of 1 with some probability I_{esp}, otherwise 0; and an expressed belief of 1 with some low probability I_{exp},

12

otherwise 0.

Figure 2 here

The availability mechanism specified so far produces growth in risk perception up to some quite high, but not saturated level with very little distinction between expressed and espoused beliefs. The process shows some drift with fairly well-defined turning points. Parameter values are $C = 2$, $v = 0.01$, $Q = 0.001$, $S = 200$, $\gamma = 2.5$. Sensitivity to these and subsequent parameters is discussed later.

4 Ripple effects and the endogeneity of behaviour

A central element of Kasperson et al's (1988) social risk amplification framework is that social actors act on their perceptions of risk. These actions are themselves observed by other social actors, and responded to. Behaviour is the product of amplified risk perceptions and it simultaneously shapes those perceptions. For example, in relation to risks such as pharmaceutical and food contamination (Wolnik et al, 1984; Ingelfinger, 2008), heightened risk perception typically reduces consumption, and so reduces exposure, the reported prevalence of harm, and therefore subsequent perception. Liu et al's (1998) study of a milk contamination crisis showed an obvious connection between heightened risk perception and reduced consumption. If the risk is to the availability of a commodity, rather than its contamination, heightened risk perception might increase or bring forward consumption, and ultimately magnify perception. But in the example that follows we concentrate on the first case.

Although we refer to the risk bearer's exposure path as being 'consumption' it is really any discretionary aspect of the risk bearer's behaviour that shapes the risk. Then what we call consumption is most simply another binary variable $c_i \in \{0, 1\}$, equated with the negation of internal belief, $c_i = 1 - b_{esp,i}$, such that if i believes in some hazard it does not consume, but otherwise does consume. Consumption in a population is, to varying degrees, observable by the same population. In the case of major risk events, global consumption levels are widely reported with reasonable credibility – for example, travel and tourism levels and their changes during the

13

SARS crisis (Brahmbhatt and Dutta, 2008). In the case of more minor events, individuals may only observe consumption behaviours in their direct social network. When general reports are available, individuals will be able to estimate the mean public consumption $\sum_{i \in [1, N]} c_i / N$, and thereby the mean espoused belief. Probably the most parsimonious model of how this affects perception is then to revise an individual's espoused beliefs (and its state of consumption) if the discrepancy between the mean public consumption and its current espoused belief (0 or 1) exceeds some threshold T, a model constant. Thus:

$$b_{esp,i} (t + 1) = 0 \text{ if } b_{esp,i} (t) - (\sum_{j \in [1, N]} (1 - c_{esp,j}(t))) / N > T;$$

$$b_{esp,i} (t + 1) = 1 \text{ if } (\sum_{j \in [1, N]} (1 - c_{esp,j} (t))) / N - b_{esp,i} (t) > T;$$

$$b_{esp,i} (t + 1) = b_{esp,i} (t) \text{ otherwise.}$$

Figure 3 illustrates the additional pseudocode.

Figure 3 here

Example traces are shown in Figure 4 in which $T = 0.5$ and the other parameters set at previous values. The main effect is to separate the espoused and expressed beliefs more clearly in value.

Figure 4 here

This way of representing the effect of observed behaviour on perception is simplistic in various aspects. For example, we have taken no account of the way in which people might avoid inconsistency between expressed beliefs and consumption behaviour. It therefore illustrates only one way of making the commitment to incorporating the perception-behaviour link in a risk amplification model.

5 Risk principals, news and the endogeneity of communication

Probably in all risk cases there will be one or more risk principals. These are the social actors, generally organizations, responsible for a risk. Often they include producers of good and services who are causally responsible for a risk – for example through contamination events in food

14

processing installations, disease outbreaks in healthcare facilities and so on. They may well also include state agencies that regulate such producers and carry some moral if not causal responsibility. The members of a society are clearly influenced in their risk beliefs not only by their peers but also by such risk principals.

Risk principals are, generally, either sources or users of expert, objective risk assessment – typically estimating the probability of harm given consumption, $\sigma \in [0, 1]$. This might be the probability of fatality per unit of consumption (for example 1 kg of contaminated foodstuff, or period of exposure to an air-borne disease, or a single inoculation). In most cases there will be some non-zero background level of hazard, σ_{normal}, but during a crisis (for example a food contamination event, or disease outbreak) lasting for some period $t_{initiation}$ to $t_{resumption}$ some heightened level, σ_{raised}. In complex risk issues, the group of experts may well be ill-defined, consisting of independent academics, individuals employed by state regulatory agencies and those employed by commercial interests. There is evidence that expert risk beliefs converge with increasing inter-communication (Muter et al, 2013) but clearly in any particular case there may be no consensus. In what follows we assume a single, consensual expert assessment. In some contexts, however, a better approach might well be a more socialised model of the expert process.

The public are not exposed to expert risk assessment continually, and perhaps the simplest representation of the timing of risk communication is that it is triggered when there is a discrepancy greater than some threshold D between the harm probability σ and the probability a member of the public has a positive expressed belief $\sum_{i \in [1, N]} b_{exp,i} / N$, for example as indicated by public surveys. Again, it is possible to represent this in quite different ways, but some commitment needs to be made to how risk communications from risk principals are triggered. However, it is unlikely that general members of the public receive expert estimates in an unmoderated form. We made the point earlier that the way in which actor A communicating about a risk to actor B does so with some expectation of whether B is under- or over-sensitive to communications from A. We described how Rayner's (1988) early critique of social risk amplification, and a later revision of the framework (Busby and Duckett, 2012), requires that we think of social actors as deciding how to interpret risk messages based on their expectations about bias and distortion in those messages. We also cited other studies showing individuals

15

correct for others' expected distortions (Petts and Niemeyer, 2004; Frewer, 2004; Renn and Levine, 1991). This involves actors engaged in 'theory of mind' (for example De Weerd et al, 2013): holding representations of what is in the mind of other actors when interpreting what they hear from them. A commitment therefore has to be made in any modelling process to what these theories of mind are, and how they influence agents' communicative actions.

The simplest possible approach is to assume that public actors believe that a risk principal, in the context of a specific type of belief, has a fixed level of bias in its communications, $\beta_{principal}$, where $\beta_{principal} = 0.5$ indicates neutrality, 1 indicates maximal exaggeration, and 0 maximal understatement of risk. Similarly a risk principal attributes a fixed bias β_{public} to the public as a group. Putting the $\beta_{principal}$ constant assumes that actors learn their theories of mind from one crisis to another, rather than within the course of a specific crisis. For the principal a basic decision rule about what to communicate is to say that if ($\sum_{i \in [1, N]} b_{exp,i} (t)/ N - \sigma$) is positive (the public as a group are over-stating the risk) and $\beta_{public} > 0.5$ (they are normally expected to do so), the communication m after some delay A should be $m(t + A) = 0$, telling the public there is nothing to worry about. If ($\sum_{i \in [1, N]} b_{exp,i} (t) / N - \sigma$) is negative (the public are under-stating the risk) and $\beta_{public} < 0.5$ (they normally do so) the communication $m(t + A) = 1$, telling them there *is* something to worry about. But if ($\sum_{i \in [1, N]} b_{exp,i} / N - \sigma$) is positive and $\beta_{public} < 0.5$ there should be some non-zero probability that the communication $m = 1$, rather than 0, because it is generally believed the public under-estimate even though, on this occasion, it appears to be exaggerating. The closer the value of β_{public} to 0 the more it under-estimates, so probability($m(t + A) = 1$) = $0.5 - \beta_{public}$ in this case. Similarly, if ($\sum_{i \in [1, N]} b_{exp,i} (t) / N - \sigma$) < 0 and $\beta_{public} > 0.5$, probability($m(t + A) = 0$) = $\beta_{public} - 0.5$.

The risk communication m is broadcast to the public agents as 'news'. An agent has a finite probability that it will be influenced by the news. This probability is a function of individual susceptibility, $\lambda_i \in [0, 1]$, a quantity that is fixed for a given agent but randomly endowed. News is persistent, and there can be some difference in its age τ_{news} when attended to by different agents, and it is reasonable to discount it by some universal factor δ, so that the probability of an agent i updating its belief to be consistent with the news is $\lambda_i (1 + \delta)^{-\tau_{news}}$. Only

16

the most recent item of news is considered by a public agent.

Given this probability that a public agent pays attention to the news, it needs to be given some rule for interpreting communications from the risk principal. For instance, if $b_{esp,i}(t) = 0 = m(t)$, or if $b_{esp,i}(t) = 1 = m(t)$, the agent's current belief and the news are consistent so it should retain its current belief. But if $b_{esp,i} = 0$ and $m = 1$ then, if also $\beta_{principal} > 0.5$, the principal tends to amplify so the communication that the belief is true is simply a manifestation of this bias, indicating that the agent's posterior belief should definitely stay at zero. But if $\beta_{principal} < 0.5$, the principal tends to attenuate and this gives strength to the news that the belief is true. So with some finite probability $0.5 - \beta_{principal}$ the agent should update its belief to 1. If $b_{esp,i} = 1$ and $m = 0$ then, if also $\beta_{principal} < 0.5$, the principal tends to attenuate and the communication is simply a manifestation of this bias, indicating that the agent's posterior belief should definitely stay at 1. But if $\beta_{principal} > 0.5$, the principal tends to amplify so it seems reasonable to follow its communication stating that the belief is false with some probability $\beta_{principal} - 0.5$.

Figure 5 shows the pseudo-code for the risk communications processes. It seems likely that in practice these theories of mind will be highly context specific, varying between actors and situations, because they are a product of individual histories. The main point is that some kind of commitment has to be made to a mechanism describing how people adjust for their preconceptions about other people when acting on and forming their communications.

Figure 5 here

Figure 6 shows a typical trace with $D = 0.1$, $A = 100$, and the operative risk levels σ_{raised} and σ_{normal} set at 0.8 and 0.01 with $t_{initiation} = 10\ 000$ and $t_{resumption} = 13\ 000$. The attributed biases β_{public} and $\beta_{principal}$ are 0.9 and 0.1. The figure shows the public expressed value on the black trace, the amplification of the actual risk level in mid-grey (showing mostly attenuation in this case), and the density of changes in belief in light grey. The change density is the proportion of the population changing either or both of their expressed and espoused beliefs within each 10 cycle period of the model. As expected, belief changes occurred much more frequently around the change of the objective risk change, but also occurred from endogenous activity – at somewhat lower levels – at two distinct points in the trace before and after the objective risk change.

17

Figure 6 here

----

6 Model validity and sensitivity

6.1 Microvalidity

A general view (Moss and Edmonds, 2005; Midgley et al, 2007) is that the validation of agent based models is essentially of two kinds: micro, the validation of the component assumptions and decision rules, and macro, the validation of the outcomes that emerge in the model's behaviour. But as Midgley et al (2007) argue such validation is inherently difficult, given the heterogeneity of agents and the shaping of behaviour at the macro level by interaction at the micro level. They suggest that 'even the "simple" step of establishing face validity may itself be a significant challenge'. In Table 1 we summarise the elements on which Sections 3, 4 and 5 were based. The table does not list all the work we draw in section 3, and it obscures the fact that some of the key contributions – such as Kuran and Sunstein's (1999) – themselves draw on other research, most of it empirical. In some ways this is beneficial, as it helps synthesise and interpret prior work. Its limitation is that it creates more distance between original empirical work and the modelling that is ultimately based on it. In all cases the evidence lacks 1) support for specific functional forms, and 2) indication of specific values for model parameters. A large part of the problem is context specificity: as the table indicates, the various studies arise in various, quite different contexts.

Table 1 here

However, modelling only where calibrating data exist would unduly restrict the extent to which modelling could contribute to theory development (Miller, 2015). Part of its value is to identify the empirical research needed on social constants, and to show what decision makers need to know about specific, contextual parameters in order to understand the potential for social risk amplification in specific cases. The effort devoted to measuring such parameters can be guided by some kind of sensitivity estimate, such as that shown in Table 2. Our outcome measure $\hat{Y}$ is

18

the mean discrepancy over time between objective risk level and the mean public expressed belief during the period indicated on the traces in previous figures (in which there is a single episode of an increased objective risk level). This discrepancy is expressed instantaneously as a probability, with values greater than 0.5 indicating amplification and less than 0.5 attenuation,

$Y(t) = 0.5 + 0.5 \left(\left(\sum_{i \in [1, N]} b_{exp,i} (t) \right) / N - \sigma \right)$. Ford and Flynn (2005) suggest using the simple product moment correlation between model outcome and model constants to measure the relevance of each constant's role, so we randomly sample the constants 100 times from their plausible ranges, assuming uniform distributions for each, and run the model 10 times for each sample. The question of how to establish plausible ranges for model constants is problematic, and ultimately subjective. The table suggests that in the current setup the outcomes are insensitive to the network parameter (the exponent defining the link degree distribution). They are moderately sensitive to specific parameters defining the social context, notably the 'noise' or probability of random changes in belief, and to the bias attributed to the public by a communicating risk principal.

Table 2 here

6.2 Macrovalidity

The basis for macro-validity needs to be empirical time series of risk perceptions, risk responses or indications of risk beliefs in social groups, in response to identifiable risk events of societal significance. What we know from the literature is that such time series can exhibit very different forms, with few qualitative features in common. Deploying a model in practice, to support decisions in the course of specific risk episodes, requires a comparison of model output with observations in episodes that are obviously similar in some relevant way.

Loewenstein and Mather's (1990) empirical work records time series of public concern in a variety of different issues, all of which they regard as societally significant risks. Table 4 summarises the measures used and the qualitative features of these series, in our words. All the issues show measures of concern that fluctuate substantially more than the underlying, objective conditions – sometimes to extreme degrees. But the qualitative features vary substantially across contexts, illustrating how important it is to avoid over-generalisation. Loewenstein and Mather's

19

work is now dated, given changes in social communication and interaction. The measures of concern are, in many cases, only indirectly connected with the extent of some risk belief in a society. The variety in patterns of response for different issues is striking, however.

Table 3 here

It is notable that there are no flat peaks in Loewenstein and Mather's (1990) data, nor in those of Chung (2011) – nor in those of our simple simulation model. The exogenous threat takes the form of a finite impulse, and so has a rectangular profile, but the dynamics of the social response mean that risk beliefs grow toward a peak, sometimes repeated, before immediately declining. Similarly, in the absence of threat the processes show a continual drift, with turning points. Klimek et al (2011) also found evidence for public attention more commonly arising for endogenous reasons than in response to a clear exogenous event. And Midden and Verplanken (1990) found that risk perceptions around nuclear power showed considerable instability over time.

A more recent empirical study of the dynamics of risk perceptions around a pandemic influenza outbreak (Ibuka et al, 2010) produces a similar indication. Over the survey period, the objective incidence of cases increased monotonically. But two measures of aggregate risk perception show substantial instability. The former broadly looks like 'noise' – an aperiodic, low amplitude, apparently random movement; the latter looks more like 'instability' – movements of large amplitude with an approximate periodicity. Generally, though, longitudinal data on risk perception is hard to come by. And, because typically it only covers a period in which some crisis is established, it cannot show the dynamics prior to an exogenous change in risk. A good example is Lau et al's (2003) study during the SARS outbreak, which was used as an empirical comparison for a system dynamics model of social risk amplification (Busby and Onggo, 2012). It shows a clear peak, but the sparsity of observations and the short window of data collection preclude the possibility of showing the kind of endogenous fluctuation in risk perception that our model indicates.

Finally, Moss and Edmonds (2005) argue that it is such turning points in time series, unpredictably clustered volatility, and consequent leptokurtosis, that appears beyond the reach of econometric models. Yet such a feature is common in the behaviour of aggregated social

20

phenomena, and arises because agents are socially embedded and have 'meta-stable' decision rules that only respond when thresholds are crossed. Figure 7 shows the distribution of proportional period-by-period changes in the mean public espoused risk belief from a single simulation with modal parameter values, for windows of 10 model 'ticks', $(b(t + 10) – b(t))/b(t + 10)$. It clearly shows strong leptokurtosis. The value of excess kurtosis (defined in terms of fourth and second central moments, μ_4 and μ_2, as $\mu_4/\mu_2^2 – 3$) is 815, which is very high.

Figure 7 here

Clearly, the possibilities for validation of any specific model of social risk response are limited. But this reflects the state of our knowledge generally about the subject in question. The model produces certain features that simply have not been examined empirically, despite their potential theoretical interest, such as the drift in a population's mean belief that occurs for internal, systemic reasons rather than as a result of external signals or events.

7 Discussion and conclusion

7.1 The main implications

The outcomes of the modelling and simulation point to the need to update our theorising about the social amplification of risk in two main ways. First, the simple finding that dynamics are interesting matters to theory. Amplification is not a simple case of constant multiplication of some risk level over time, even when that risk level is constant. Past work (with the exception of the systems dynamics studies of Burns and Slovic (2007), and Busby and Onggo (2013)) has said little about dynamics. The original framework (Kasperson et al, 1988) points to 'ripple effects' that by definition follow some initiating risk event, but pays little attention to how much amplification fluctuates over a crisis, to what extent it pre-figures it, and to what degree there is an aftermath. Our simulation shows one possible pattern, and the simple, underlying structure of agents interacting through availability-based rules provides a mechanism that can be assimilated in the theory of social risk amplification.

The internally-produced change evident in the simulation of social risk response also matters. Endogenous peaks in the simulation are not as large as peaks that follow crises in which objective risk suddenly increases (for example in a disease outbreak) – but they look significant

21

and would create a material consequence for a risk manager, whether economic or political. The social amplification of risk literature does not deal with this, perhaps because the original framework (Kasperson et al, 1988) stresses the way in which social risk amplification is initiated by some 'risk event'. Our theory of social risk amplification needs updating to allow for the possibility of a public anxiety that has no basis in some specific, objective occurrence. Such a proposition is almost impossible to confirm in practice, because in a complex society there is always some event occurring near the start of some growth of public concern that, in hindsight, can be claimed as its trigger. But neither is there any theoretical justification in assuming the opposite: that an identifiable event is necessarily required for an acute growth in social concern about some risk.

The way to revise theory to take account of such outcomes is to incorporate the commitments that have had to be made in the basic model proposed here. These commitments may need to be made differently in different contexts, reflecting the diversity of outcomes found, for example, by Loewenstein and Mather (1990). Such commitments include specific mechanisms of interaction, the relationship between espoused and expressed beliefs, and the role of informational and reputational motivations during interactions (Kuran and Sunstein, 1999). They include processes of deliberate distortion (Lofstedt, 2008; Busby and Duckett, 2012). And they include some representation of theory of mind, in which observers of risk responses base their own responses not only on the responses they observe but the models they have of those producing such responses.

It is also worth commenting on the connection between the basic notion of social risk amplification and the dissemination of warnings – as also analysed in the literature with agent based models (Nagarajan et al, 2012). Warning dissemination through peer networks as well as broadcast channels is an important mechanism not just for communicating information but also exerting influence, given limited compliance with official warnings. Zechman's (2011) model of water contamination events provides an example. The notion of risk amplification is that dissemination does not occur without distortion of some initial risk signal. But, equally, distortion does not occur without dissemination. So although social amplification can be seen as undesirable and problematic for risk managers it also achieves a dissemination of risk understanding that is often desirable if not vital in certain risk events, particularly if large-scale public action such as evacuation is needed. Both kinds of problem – anticipating distortions of

public risk beliefs, and achieving active public response to risks – are faced by risk managers as aggregate effects across a population. But they are produced by individual agents interacting with one another and exchanging information about something that is exogenous to each of them yet endogenous to the population of which they are members. Agent-based modelling is logically central to our understanding of both facets of social responses to danger: productive dissemination and unproductive distortion.

7.2 The intended contribution

The intended contribution is to make the understanding of social risk amplification more precise and therefore more operational – both as a basis for more directed empirical work and as a basis for theorising about what we mean by social risk amplification. As Miller (2015) has recently argued, a critical realist view calls for explanations in terms of underlying mechanisms, and agent-based modelling fundamentally represents the interaction of social actors, rather than variables, providing a process perspective rather than a variance perspective. An agent-based model such as ours explains social risk amplification as mechanism, and the simulation of that mechanism allows us then to identify interesting consequences that have so far received little attention. The model is not intended to be definitive, but to show how some key elements of social risk responses can be represented:

1. The endogeneity of risk observation. Most risks of societal significance are not experienced or observed physically by most people: most people rely on other people, and the beliefs they express.

2. The endogeneity of risk behaviour. Risk beliefs shape risk behaviours (such as withdrawing from consumption of activities or products that expose people to the risk in question), and behaviours are observed, so shape risk beliefs.

3. The endogeneity of risk communication. Actors receive communications that have been adjusted or moderated to reflect the communicator's experiences and expectations about those communicated to, and communications received influence communications subsequently made.

We showed simple ways in which these elements can be represented, drawing particularly on Kuran and Sunstein's (1999) well-developed notion of availability cascades, itself built on more fundamental observations of the role of the availability bias in human responses to uncertainty.

23

This gives insight into the mechanisms that risk managers dealing with public responses have to deal with, and indicate the range of outcomes they should be prepared for.

7.3 Some limitations

There are some obvious limitations of the approach taken here. First, there are important types of actor and related social processes that have been omitted from the model. The most obvious is the role of the media, quite often referred to in the literature (for example Petts and Niemeyer, 2004), and often exploited by availability entrepreneurs – as demonstrated by Bakir (2005). Second, the model does not deal with homophily, the affinity of individuals within groups, and related issues such as cultural polarisation (for example Flache and Macy, 2011). The literature on cultural risk selection (for example Douglas and Wildavsky, 1982) suggests that this is an important next step in developing the model. And the model has concentrated on the representation of single issues. Any real society has to come to terms with many risk issues that come into prominence at broadly the same time, making it important to understand how public anxiety can be attenuated simply through distraction. There were also limitations in verification and validation. As we discussed, these strongly limit any claims that can be made about the outcomes from the simulation, and clearly indicate that making models of this kind operational to support managerial practice or policy making has to involve investigation of what happens in specific contexts. Relying on generalised models with tentative parameters could be highly misleading.

Finally, there are of problems in the use of agent based computational modelling generally. Macy and Willer (2002) argue that 'Computer simulation is more tractable (but less generalizable) than mathematical modelling and more rigorous (but less nuanced) than natural language'. In terms of representational richness and fidelity, on the one hand, and transparency, on the other, it is a half-way house and does not avoid the need for more formal treatments of social risk amplification. But it nonetheless represents a reasonable direction for representing amplification more precisely, and working out the consequences of such a representation. It broadly avoids falling into the 'trap of verisimilitude' (Doran and Gilbert, 1994) – of trying to say too much about the world – or the 'trap of tractability' – saying too little in order to be analytical. And it helps to enhance our explanatory spaces. If we were asked to explain a peak in risk perceptions, or a drift one way or the other, we would probably to try to find a news item, or

some event in the world, that could cause such an outcome. But an equally plausible alternative is that peaks and drift come from the structures and mechanisms of social observation and reaction alone.

List of notation

G	Non-directed graph s.t. $G_{ij} = 1$ if agent i and agent j are neighbours, else 0
N	Number of public agents
$b_{esp,i}(t)$, $b_{exp,i}(t)$	Agent i's espoused (private) and expressed (public) beliefs at time t
q_i	Fraction of agent i's neighbours expressing a belief
ν	Probability of random belief switching when an agent active
C	Exponent determining probability of influence from belief availability
Q	Spontaneous adoption of belief as an availability entrepreneur
S	Number of continuing interactions of availability entrepreneur's influence
γ	Exponent defining distribution of link degrees in network
$c_i(t)$	Consumption of agent i at t
T	Margin at which consumption observation revises a belief
$\sigma(t)$	Objective risk level at t
D	Margin at which risk principal communicates
$\beta_{principal}$, β_{public}	Bias in expressed risk attributed to risk principal and public
A	Lag before communication from risk principal
λ_i	Agent i's susceptibility to influence by risk principal's communication
δ	Age discount factor for risk principal's communication

Acknowledgements

25

References

Amini, M., Wakolbinger, T., Racer, M. and Nejad, M.G. (2012). Alternative supply chain production-sales policies for new product diffusion: an agent-based modelling and simulation approach. *European Journal of Operational Research*, *216*, 301-311.

Bakir, V. (2005). Greenpeace v. Shell: media exploitation and the Social Amplification of Risk Framework (SARF). *Journal of Risk Research*, *8*(7-8), 679-691.

Banerjee, A.V. (1992). A simple model of herd behavior. *Quarterly Journal of Economics*, *107*(3), 797-817.

Barabasi, A.-L. (2009). Scale-free networks: a decade and beyond. *Science*, 325, 412-413.

BBC. (2013). Fonterra products free of botulism, says New Zealand. *BBC Business News, 28th August. Last accessed at* http://www.bbc.co.uk/news/business-23860452 on 23rd April 2014.

Beck M, Asenova D and Dickson G (2005). Public administration, science and risk assessment: A case study of the UK bovine spongiform encaphalopathy crisis. *Public Administration Review*, *65*(4), 396–408.

Bikhchandani, S., Hirshleifer, D. and Welch, I. (1998). Learning from the behavior of others: conformity, fads, and informational cascades. *Journal of Economic Perspectives, 12*(3), 151-170.

Binder, A.R., Scheufele, D.A., Brossard, D. & Gunther, A.C. (2011). Interpersonal amplification of risk? Citizen discussions and their impact on perceptions of risks and benefits of a biological research facility. *Risk Analysis*, *31*(2), 324–334.

Brahmbhatt, M. & Dutta, A. (2008). On SARS type economic effects during infectious disease outbreaks. World Bank Policy Research Working Paper 4466, The World Bank East Asia and Pacific Region Chief Economist's Office, January.

Brenkert-Smith, H., Dickinson, K.L., Champ, P.A. & Flores, N. (2013). Social amplification of wildfire risk: the role of social interactions and information sources. *Risk Analysis*, *33*(5), 800-817.

Burns, W.J., Slovic, P., Kasperson, R.E., Kasperson, J.X., Renn, O. & Emani, S. (1993). Incorporating structural models into research on the social amplification of risk: implications for theory construction and decision making. *Risk Analysis*, *13*(6), 611-623.

Burns, W.J. & Slovic, P. (2007). The diffusion of fear: modelling community response to a

terrorist strike. *Journal of Defense Modeling and Simulation*, *4*(4), 298–317.

Busby, J.S. & Alcock, R.E. (2008). Risk and organizational networks: making sense of failure in the division of labour. *Risk Management*, *10*(4), 235-256.

Busby, J. & Duckett, D. (2012). Social risk amplification as an attribution: the case of zoonotic disease outbreaks. *Journal of Risk Research*, *15*(9), 1049-1074.

Busby, J.S. & Onggo, S. (2012). Managing the social amplification of risk: a simulation of interacting actors. *Journal of the Operational Research Society, 64*, 638–653

Cahart, N. (2013). Risk communication – the future. *The Hazards Forum Newsletter*, Issue 78, Spring, 2-10.

Chung, I.J. (2011). Social Amplification of Risk in the Internet environment. *Risk Analysis*, *31*(12), 1883–1896.

De Weerd, H., Verbrugge, R. & Verheij, B. (2013). How much does it help to know what she knows you know? An agent-based simulation study. *Artificial Intelligence*, *199-200*, 67-92.

Doran, J. & Gilbert, N. (1994). Introduction. In Gilbert, N. and Doran, J. (eds.) *Simulating Societies: The Computer Simulation of Social Phenomena*, UCL Press (London), 1-18.

Doran, J., Palmer, M., Gilbert, N. & Mellars, P. (1994). The EOS project: modelling Upper Paleolithic social change. In Gilbert, N. and Doran, J. (eds.), *Simulating Societies: The Computer Simulation of Social Phenomena*, UCL Press (London), pp.195-221.

Douglas, M. & Wildavsky, A. (1982). *Risk and Culture: An Essay on the Selection of Technological and Environmental Dangers.* University of California Press (Berkeley, CA).

Epstein, J.M. & Axtell, R. (1996). Growing Artificial Societies: Social Science from the Bottom Up. (Washington DC: Brookings Institution Press).

Fischhoff, B., Slovic, P., Lichtenstein, S., Read, S. & Combs, B. (1978). How safe is safe enough? A psychometric study of attitudes towards technological risks and benefits. *Policy Sciences, 9*, 127-152.

Flache, A. & Macy, M.W. (2011). Small worlds and cultural polarization. *Journal of Mathematical Sociology*, *35*, 146-176.

Ford, A. & Flynn, H. (2005). Statistical screening of system dynamics models. *System Dynamics Review, 21*, 273–303.

Freudenberg, W.R. (2003). Institutional failure and the organizational amplification of risk: the need for a closer look. In Pidgeon N, Kasperson RE and Slovic P (eds). *The Social Amplification*

27

of Risk. Cambridge University Press (Cambridge, UK), pp. 102–120.

Frewer, L.J., Miles, S. & Marsh, R. (2002). The media and genetically modified foods: evidence in support of social amplification of risk. *Risk Analysis*, *22*(4), 701-711.

Frewer, L.J., Scholderer, J. & Bredahl, L. (2003). Communicating about the risks and benefits of genetically modified foods: effects of different information strategies. *Risk Analysis, 23*(6), 1117–113.

Frewer, L. (2004). The public and effective risk communication. *Toxicology Letters*, *149*, 391-397.

Gardner, H., Kleinman, N.L. & Butler, R.J. (2000). Workers' compensation and family and medical leave act claim contagion. *Journal of Risk and Uncertainty*, *20*(1), 89–112.

Gilbert N. & Troitzsch, K.G. (1999). *Simulation for the Social Scientist*. Open University Press (Buckingham, UK).

Ibuka, Y., Chapman, G.B., Meyers, L.A., Li, M. and Galvani, A.P. (2010). The dynamics of risk perceptions and precautionary behavior in response to 2009 (H1N1) pandemic influenza. *BMC Infectious Diseases*, **10**, 296-307.

Ingelfinger, J.R. (2008). Melamine and the global implications of food contamination. *New England Journal of Medicine*, *359*, 2745-2748.

Hill, A. (2001). Media risks: the social amplification of risk and the media violence debate. *Journal of Risk Research*, *4*(3), 209-225.

Kasperson, R.E., Renn, O., Slovic, P., Brown, H.S., Emel, J., Goble, R., Kasperson, J.X. & Ratick, S. (1988). The social amplification of risk: a conceptual framework. *Risk Analysis*, *8*(2), 177–187.

Klimek, P., Bayer, W. & Thurner, S. (2011). The blogosphere as an excitable social medium: Richter's and Omori's Law in media coverage. *Physica A: Statistical Mechanics and its Applications*, *390*(21-22), 3870–3875.

Kohler, H.-P., Behrman, J.R. & Watkins, S.C. (2007). Social networks and HIV/AIDS risk perceptions. *Demography*, *44*(1), 1-33.

Kuran, T. & Sunstein, C.R. (1999). Availability cascades and risk regulation. *Stanford Law Review*, *51*(4): 683-768.

Langford, I.H., Marris, C., McDonald, A., Goldstein, H.. Rasbash, J. & O'Riordan, T. (1999). Simultaneous analysis of individual and aggregate responses in psychometric data using

multilevel modeling. *Risk Analysis*, *19*(4), 675-693.

Lau, J.T.F., Yang, X., Tsui, H. & Kim, J.H. (2003). Monitoring community responses to the SARS epidemic in Hong Kong: From day 10 to day 62. *Journal of Epidemiology and Community Health*, *57*(11), 864–870.

Lehmkuhl, M.J. (2008). Congruency within rural social networks as an indicator of interpersonal influence on risk judgments: the great stir caused by BSE in a village in northern Germany. *Public Understanding of Science*, **17**, 484-502.

Lewis, R.E. & Tyshenko, M.G. (2009). The impact of social amplification and attenuation of risk and the public reaction to mad cow disease in Canada. *Risk Analysis*, *29*(5), 714–728.

Liu, S., Huang, J.-C., & Brown, G.L. (1998). Information and risk perception: a dynamic adjustment process. *Risk Analysis*, *18*(6), 689-699.

Lofstedt, R.E. (2008). Risk communication, media amplification and the aspartame scare. *Risk Management*, **10**, 257-284.

Loewenstein, G., & Mather, J. (1990). Dynamic processes in risk perception. *Journal of Risk and Uncertainty*, *3*(2), 155-175.

Luhmann, N. (1993). *Risk: A Sociological Theory*. Walter de Gruyter (Berlin).

Macal, C.M. & North, M.J. (2010). Tutorial on agent-based modelling and simulation. *Journal of Simulation*, *4*, 151-162.

Macy, M.W. & Willer, R. (2002). From factors to actors: computational sociology and agent-based modelling. *Annual Review of Sociology*, *28*, 143-166.

Marris, C., Langford, I., Saunderson, T. & O'Riordan, T. (1997). Exploring the 'psychometric paradigm': comparisons between aggregate and individual analyses. *Risk Analysis*, *17*(3), 303-312.

Masuda, J.R. & Garvin, T. (2006). Place, culture, and the social amplification of risk. *Risk Analysis*, *26*(2), 437-454.

McComas, K.A. (2003). Public meetings and risk amplification: a longitudinal study. *Risk Analysis*, *23*(6), 1257–1270.

Metz, W.C. (1996). Historical application of a social amplification of risk model: Economic impacts of risk events at nuclear weapons facilities. *Risk Analysis*, *16*(2), 185-193.

Midden, C.J.H. & Verplanken, B. (1990). The stability of nuclear attitudes after Chernobyl. *Journal of Environmental Psychology*, *10*(2), 111-119.

Midgley, D., Marks, R. and Kunchamwar, D. (2007). Building and assurance of agent-based models: An example and challenge to the field. Journal of Business Research, *60*, 884–893.

Miller, K.D. (2015). Agent-based modelling and organization studies: a critical realist perspective. *Organization Studies*, *36*(2), 175-196.

Moss, S. & Edmonds, B. (2005). Sociology and simulation: Statistical and qualitative cross-validation. *American Journal of Sociology*, *110*, 1095-1131.

Muter, B.A., Gore, M.L., & Riley, S.J. (2013). Social contagion of risk perceptions in environmental management networks. *Risk Analysis*, *33*(8), 1489-1499.

Nagarajan, M., Shaw, D. and Albores, P. (2012). Disseminating a warning message to evacuate: a simulation study of the behaviour of neighbours. *European Journal of Operational Research*, *220*, 810-819.

Petts, J. & Niemeyer, S. (2004). Health risk communication and amplification: learning from the MMR vaccination controversy. *Health, Risk & Society*, *6*(1), 7-23.

Porto, M.F. de S., de Freitas, C.M. (1996). Major chemical accidents in industrializing countries: the socio-political amplification of risk. *Risk Analysis*, *16*(1), 19-29.

Rahmandad, H. & Sterman, J. (2008). Heterogeneity and network structure in the dynamics of diffusion: comparing agent-based and differential equation models. *Management Science*, *54*, 998-1014.

Raude, J., Fischler, C., Lukasiewicz, E., Setbon, M. & Flahault, A. (2004). GPs and the social amplification of BSE-related risk: an empirical study. *Health, Risk & Society*, *6*(2), 173-185.

Rayner, S. (1988). Muddling through metaphors to maturity: A commentary on Kasperson et al. 'The social amplification of risk'. *Risk Analysis*, *8*(2), 201–204.

Renn, O. (1991). Risk communication and the social amplification of risk. In: Kasperson, R.E. and Stallen, P.J.M. (eds.), *Communicating Risks to the Public*. Kluwer (Dordrecht, NL), pp. 287–324.

Renn, O. & Levine, D. (1991) Credibility and trust in risk communication. In: Kasperson, R.E. and Stallen, P.J.M. (eds.), *Communicating Risks to the Public*. Kluwer (Dordrecht, NL), pp. 175-218.

Renn, O. (2011). The social amplification/attenuation of risk framework: application to climate change. *Wiley Interdisciplinary Reviews: Climate Change*, *2*(2), 154-169.

Rip, A. (1988). Should social amplification of risk be counteracted? *Risk Analysis*, *8*(2), 193-197.

Santos, F.C. & Pacheco, J.M. (2005). Scale-free networks provide a unifying framework for the emergence of cooperation. *Physical Review Letters*, **95**, 098104.

Scherer, C.W. & Cho, H. (2003). A social network contagion theory of risk perception. *Risk Analysis 23*(2), 261–267.

Stroud, L. (2005). MMR – public policy in crisis: whose tragedy? *Journal of Health Organization and Management*, *19*(3), 252-260.

Sunitiyoso, Y. and Matsumoto, S. (2009). Modelling a social dilemma of mode choice based on commuters' expectations and social learning. *European Journal of Operational Research*, *193*, 904-914.

Tversky, A. & Kahnemann, D. (1973). Availability: a heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2). 207-232.

Wolnik, K.A., Fricke, F.L., Bonnin, E., Gaston, C.M. & Satzger, R.D. (1984). The Tylenol tampering incident – tracing the source. *Analytical Chemistry*, *56*, 466A-474A.

Zechman, E.M. (2011). Agent-based modeling to simulate contamination events and evaluate threat management strategies in water distribution systems. *Risk Analysis*, *31*, 758-772.

Table 1: Micro-validating use of evidence for model elements

	Decision rule or model assumption	*Validating evidence or proposition*	*Evidence type*	*Context*
Model-wide	Social interaction shapes risk perception	Scherer & Cho (2003)	Empirical	Hazardous waste
		Muter et al (2013)	Empirical	Wildlife risk
		Kohler et al (2007)	Empirical	HIV/AIDS
	Non-linearity in social amplification processes	Burns & Slovic (2007)	Model	Terrorism
		Busby & Onggo (2012)	Empirical, model	Zoonotic disease
	Risk perceiver heterogeneity	Marris et al (1997)	Empirical	Multiple (locale specific)
		Langford et al (1999)	Empirical	Multiple
		Sjoeberg (2000)	Empirical	Radiation
	Social networks have small world properties	Barabasi (2009) etc.	General synthesis	Multiple
Endogen-ising observation	Social availability drives risk perception	Kuran & Sunstein (1999)	General synthesis	Multiple
	Learning/legitimation effects show diminishing/increasing returns to peer group size	Gardner & Kleinman (2000)	Empirical	Health insurance
Endogen-ising behaviour	Heightened perception reduces consumption and exposure	Liu et al (1998)	Empirical	Food contamination
		Brahmbhatt & Dutta (2008)	Empirical	SARS
		Rahmandad & Sterman (2008)	General synthesis	All disease
Endogen-ising commun-ication	Perceiver expects distortion in risk communications	Petts & Niemeyer (2004)	Empirical	Vaccination
		Busby & Duckett (2012)	Empirical	Zoonotic disease
		Frewer (2004)	General	All
		Frewer et al (2003)	Empirical	GM foods
		Renn & Levine (1991)	General synthesis	All

Table 2: Product moment correlations between mean discrepancy and model constants

Constant		Plausible range	r(discrepancy, constant)
Social context	Convexity factor, *C*	1.2 – 4.0	0.28**
	Noise level, *n*	0.001 – 0.1	0.48**
	Entrepreneur probability, *Q*	0.0001 – 0.001	0.30**
	Entrepreneur legacy period, *S*	100 – 700	0.14*
	Consumption revision gap, *T*	0.2 – 0.7	-0.08
	Communication triggering margin, *D*	0.05 – 0.2	-0.03
	Attributed principal's bias, $\beta_{principal}$	0.01 – 0.9	-0.11
	Attributed public bias, β_{public}	0.01 – 0.9	-0.46**
	News discount factor, δ	0.01 – 0.3	-0.04
	News latency period, *A*	20 – 300	0.03
Social network	Link exponent, γ	2 - 3	-0.01

* significant at 0.05; ** significant at 0.01

Table 3: Risk response patterns found by Loewenstein & Mather (1990)

Risk issue	Proxy measure of concern	Qualitative pattern in time series
AIDS	Frequency of national news articles	High frequency periodicity ≈ 3 months Double peak on smoothed trend Objective incidence monotonically rising and initially lagging the public response
Crime	General social survey % respondents afraid to walk at night	Single peak with non-periodic non-monotonic movements Objective incidence closely similar but smoothed and leading by ≈ 1 year
Drink driving	Difference between number of drink driving groups founded and disbanded	Single peak, width ≈ 7 years Objective incidence closely similar but smoothed and leading by ≈ 1 year
Herpes	Frequency of national news articles	Extreme-amplitude single peak with small-amplitude non-periodic non-monotonic movements Objective incidence closely similar without the extreme peak and lagging by ≈ 1 year
Price inflation	Opinion poll % respondents citing inflation as the primary problem	Double peak with non-periodic non-monotonic movements Objective incidence closely similar in most periods
Un-employment	Opinion poll % respondents citing unemployment as the primary problem	Multiple non-periodic peaks Objective incidence closely similar
Polio	Frequency of news articles in a specific outlet	Low frequency periodicity ≈ 4 years Objective incidence of similar periodicity with both close similarity and extreme difference at different times
Teenage suicide	Frequency of national news articles	Extreme-amplitude single peak with small-amplitude non-periodic non-monotonic movements Objective incidence similar with small monotonic trend
Teenage illegitimacy	Frequency of news articles in a specific outlet	Low frequency periodicity ≈ 6 years Objective incidence on shallow monotonically increasing curve

Figure 1: Pseudo-code for the basic availability effect

```
publicAgent.applyAvailability()

//find mean of neighbour beliefs
beliefAvailability = meanOverNeighbours(expressedBelief)
//determine if this agent is to be an availability entrepreneur
with (EntrepreneurshipProbability)
        with (SpontaneousTrueProbability) entrepreneurBias = 1 else entrepreneurBias = 0
        entrepreneurBiasTransactions = 0
//determine if availability entrepreneurship is still active
if (entrepreneurBiasTransactions < entrepreneurBiasTransactionLimit)
        probabilityOfExpressedBelief = meanOf(beliefAvailability^EffectConvexity, entrepreneurBias)
        probabilityOfEspousedBelief = meanOf(beliefAvailability^(1/EffectConvexity), entrepreneurBias)
        entrepreneurBiasTransactions++
else
        probabilityOfExpressedBelief = beliefAvailability^EffectConvexity
        probabilityOfEspousedBelief = beliefAvailability^(1/EffectConvexity)
with (probabilityOfExpressedBelief) expressedBelief = 1 else expressedBelief = 0
with (probabilityOfEspousedBelief) espousedBelief = 1 else espousedBelief = 0

with (NoiseProbability)
        expressedBelief = ~expressedBelief
        espousedBelief = ~espousedBelief
```
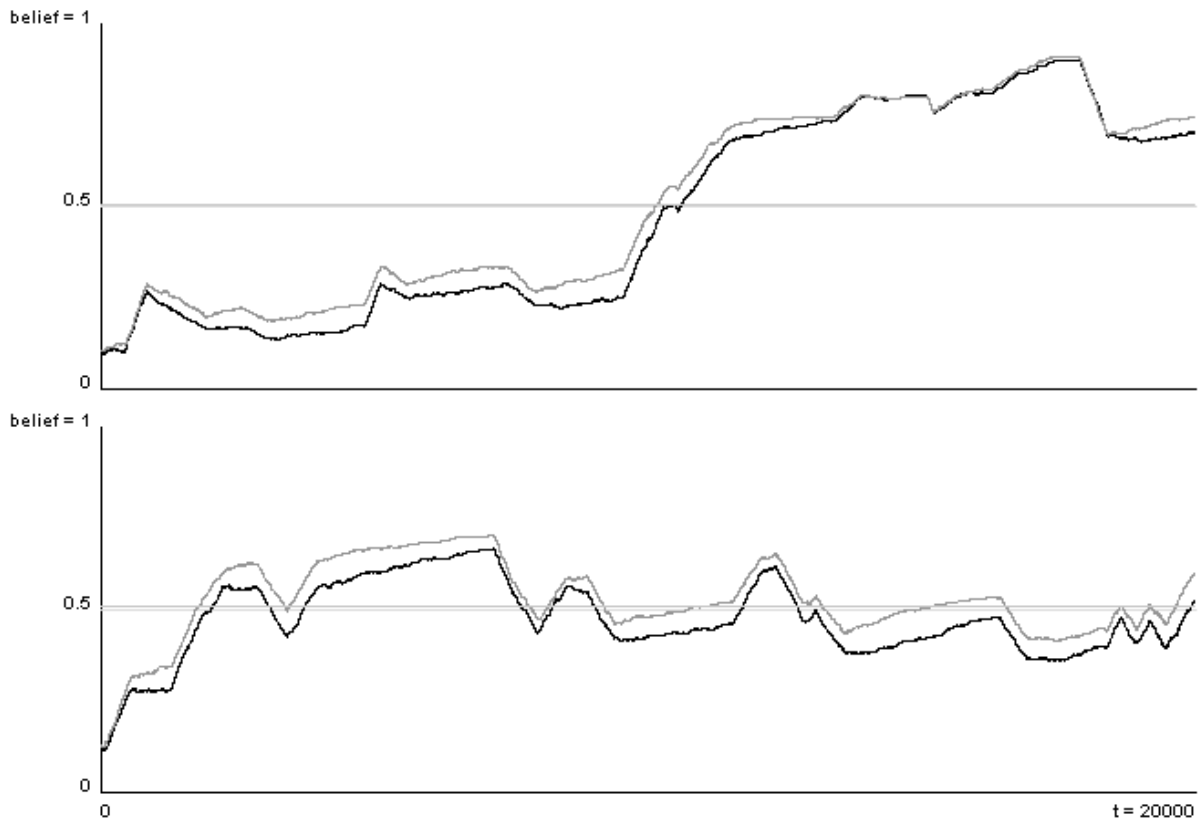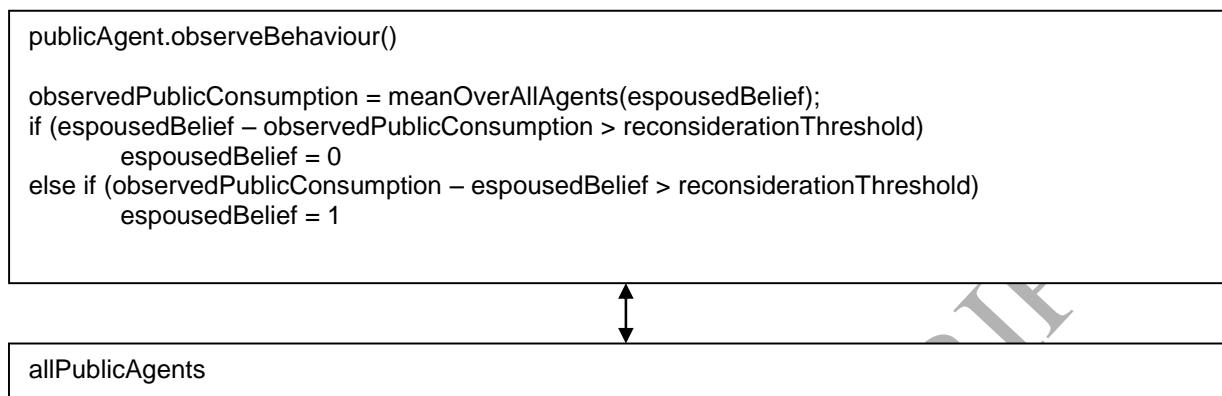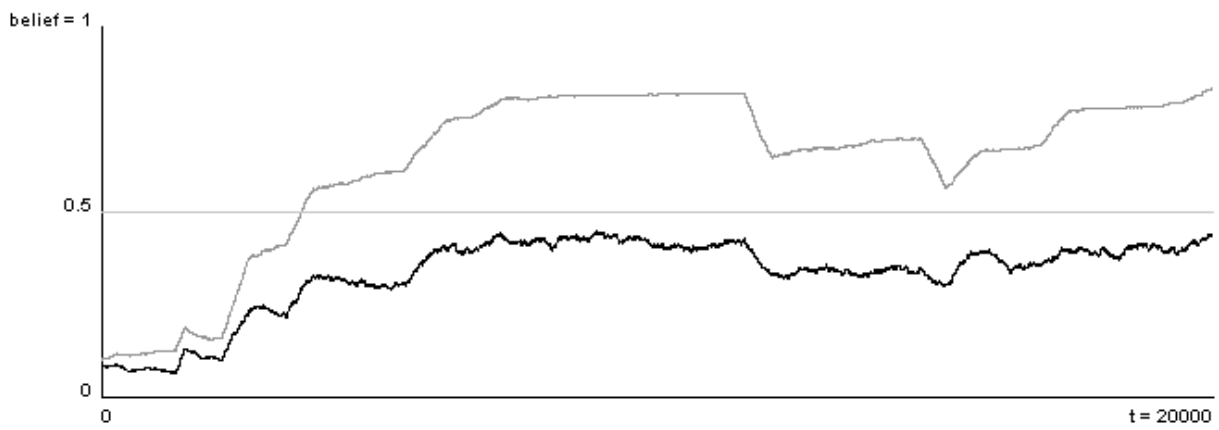
```
neighboursOf(public_agent)
```

Figure 2: Two sample traces of proportion of population having espoused (black line) and expressed (grey line) beliefs in some proposition

Figure 3: Pseudo-code for effect of behavioural observation

```
publicAgent.observeBehaviour()

observedPublicConsumption = meanOverAllAgents(espousedBelief);
if (espousedBelief – observedPublicConsumption > reconsiderationThreshold)
        espousedBelief = 0
else if (observedPublicConsumption – espousedBelief > reconsiderationThreshold)
        espousedBelief = 1
```

allPublicAgents

Figure 4: Sample traces with behavioural observation

Figure 5: Pseudo-code for risk principals and risk communications

```
riskPrincipal.communicateRisk()

publicDiscrepancy = meanOverAgents(expressedBelief) - objectiveRisk
if (|publicDiscrepancy| > communicationThreshold)
        if (publicDiscrepancy > 0)
                //definitely communicate belief to be false if public is exaggerating risk as usual
                if (amplificationOfPublic > 0.5) then probabilityOfTrue = 0
                else probabilityOfTrue = amplificationOfPublic – 0.5
        else
                //definitely communicate belief to be true if public is under-estimating risk as usual
                if (amplificationOfPublic < 0.5) then probabilityOfTrue = 1
                else probabilityOfTrue = 0.5 – amplificationOfPublic
        with (probabilityOfTrue) communicationContent = 1.0
        else communicationContent = 0.0
        communicateToPublic(communicationContent, communicationLag)
```

```
publicAgent.receiveCommunication()

newsAttentionProbability = agentSusceptibility*(1 + newsDiscountRate)^newsAge
with (newsAttentionProbability)
        if (existsCommunication(communicationContent) && espousedBelief ≠ communicationContent)
                //retain belief to be false if principal simply confirming a bias to amplify
                if (espousedBelief == 0 and communicationContent == 1)
                        if (amplificationOfPrincipal > 0.5) then probabilityOfTrue = 0.0;
                        else probabilityOfTrue = 0.5 - amplificationOfPrincipal;
                //retain belief to be true if principal confirming a bias to attenuate
                else if (espousedBelief == 1 and communicationContent == 0)
                        if (amplificationOfPrincipal < 0.5) then probabilityOfTrue = 1.0;
                        else probabilityOfTrue = 1.5 - amplificationOfPrincipal;
        with (probabilityOfTrue) espousedBelief = True else espousedBelief = False
```

39

Figure 6: Sample trace of public beliefs influenced by risk principal with change density and amplification
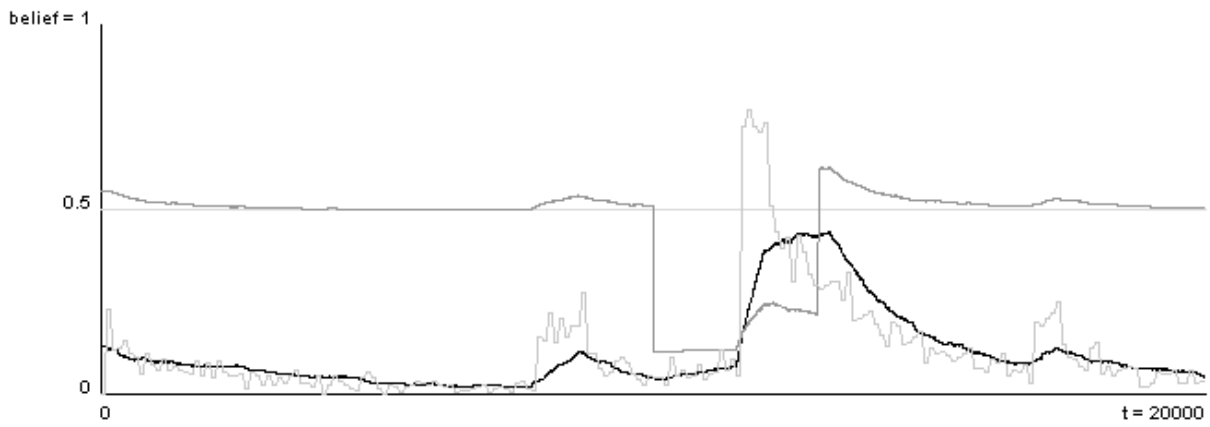
Figure 7: Distribution of fractional period-by-period changes in mean espoused risk beliefs