DISCUSSION PAPER

April 2007 ■ RFF DP 07-21

Virtual Experiments and Environmental Policy

Stephen M. Fiore, Glenn W. Harrison,
Charles E. Hughes, and E. Elisabet Rutström

Hosted by Resources for the Future
1616 P Street, NW
Washington, DC 20036
202.328.5000
www.rff.org/frontiersconference

Virtual Experiments and Environmental Policy

by

Stephen M. Fiore, Glenn W. Harrison, Charles E. Hughes and E. Elisabet Rutström*

April 2007

Prepared for the RFF/EPA Conference
Frontiers of Environmental Economics, Washington, February 26-27, 2007

Abstract. We develop the concept of virtual experiments and consider their application to environmental policy. A virtual experiment combines insights from virtual reality in computer science, naturalistic decision-making from psychology, and field experiments from economics. The environmental policy applications of interest to us include traditional valuation tasks and less traditional normative decision-making. The methodological objective of virtual experiments is to bridge the gap between the artefactual controls of laboratory experiments and the naturalistic domain of field experiments or direct field studies. This should provide tools for policy analysis that combine the inferential power of replicable experimental treatments with the natural "look and feel" of a field domain.

* Cognitive Sciences Program, Department of Philosophy (Fiore); Department of Economics, College of Business Administration (Harrison and Rutström); and School of Electrical Engineering and Computer Science (Hughes), all at the University of Central Florida, Orlando, Florida. E-mail: sfiore@ist.ucf.edu, gharrison@research.bus.ucf.edu, ceh@cs.ucf.edu, and erutstrom@bus.ucf.edu.
We thank the U.S. National Science Foundation for research support under NSF/HSD 0527675; Miko Charbonneau, Mark Colbert, Jared Johnson, Jaakko Konttinen, Greg Tener and Shabori Sen for research assistance; and James Brenner, Don Carlton, Zach Prusak and Walt Thomson for helpful comments.

Table of Contents

1. Introduction . . . . . -1-

2. Experimental Economics in the Laboratory and the Field . . . . . -5-

3. Virtual Environments . . . . . -11-
   3.1 Virtual Reality . . . . . -11-
   3.2 Virtual Trees and Virtual Forests . . . . . -13-
   3.3 Extending Experiences to Mixed Reality . . . . . -15-

4. Naturalistic Decision-Making . . . . . -19-
   4.1 Learning from Experts . . . . . -19-
   4.2 Ecological Rationality . . . . . -21-
   4.3 Experts Behaving Badly . . . . . -23-

5. Virtual Choices . . . . . -26-
   5.1 The Value-Added of Virtual Experiments . . . . . -28-
   5.2 Scenario Rejection . . . . . -32-
   5.3 Voting in a Referendum . . . . . -33-

6. Virtual Lotteries . . . . . -39-
   6.1 Task Representation and Perception . . . . . -40-
   6.2 Foreground and Background Risk . . . . . -51-
   6.3 Policy Lotteries . . . . . -54-

7. A Case Study: Evaluating the Risks of Wildfires . . . . . -59-
   7.1 Representation as a Policy Lottery . . . . . -61-
   7.2 Generating the Probabilities . . . . . -65-
   7.3 The Virtual Experimental Task . . . . . -73-

8. Conclusions . . . . . -78-

References . . . . . -80-

1. Introduction

We develop the concept of virtual experiments and consider their application to environmental policy. A virtual experiment (VX) combines insights into virtual reality from computer science, naturalistic decision-making from psychology, and field experiments from economics. The environmental policy applications of interest to us include traditional valuation tasks and less traditional normative decision-making. The methodological objective of virtual experiments is to bridge the gap between the artefactual controls of laboratory experiments and the naturalistic domain of field experiments or direct field studies. This should provide tools for policy analysis that combine the inferential power of replicable experimental treatments with the natural "look and feel" of a field domain.

We begin with an exposition of the current state of experimental economics, and the relationship between lab and field experiments. Experimental methods have reminded economists about several simple facts of behavioral life that we sometimes neglect: that institutions matter, that cognitive constraints can affect the behavior of individuals, that the samples we observe in the field often select or sort themselves, and that people use field referents to help them make decisions. We also knew these things from game theory, from social psychology, from labor economics, and from common sense. They are not new insights, but the evolution of the application of experimental methods has forced a consistent perspective on these matters.

The result is now something of a surprising backlash against the use of lab methods, in

favor of natural experiments, social experiments and field experiments. Following Harrison and List [2004], we argue that this backlash is a mistake, and that these alternative types of experiments should be viewed as complementary. We develop that step of the argument in Section 1, since it generates a derived demand for virtual experiments.

If lab experiments *can* be complementary, what does one have to do to *make* them complementary? In some cases simple redesigns can do the trick, as we illustrate with some examples. This is important to recognize, so that we do not apply fancy technology when it is not really needed. But in other cases we argue that we have to re-think the technology used to undertake lab experiments, and this is where virtual reality (VR) technology comes in. What we want to retain from lab experiments, and that cannot be easily guaranteed in the field alternatives, is cost effective replicability and orthogonal controls. However, we want these features of an experiment without sacrificing the natural look and feel of the field. We explain in section 2 how VR can provide this.

Once we have the technology in place, we need to be much more systematic about what we mean by the "natural look and feel" of the experimental task. There are two reasons for this. First, we want to do better than the judge that claimed to be unable to define pornography but knew it when he saw it: if we are going to call on fancy VR technology to do something to make our analyses more rigorous, then we need to have some metrics for evaluating if we are achieving that goal.

The second reason has to do with the inherent complexity of many decision-making domains, common in environmental policy, and the role of experts. One often finds conflicting positions by experts and affected non-expert citizens, as well as between experts. Building on the field of naturalistic decision-making (NDM) in psychology, we hypothesize that much of this conflict is created because experts and non-experts perceive the problem differently and use different cognitive processes when evaluating and forming decisions. This theme is also a common one from the field experimental literature in economics, where "market specialists" generally behave differently than others.[1] We propose the use of VR simulation methods designed to facilitate a convergence in these cognitive processes, and hypothesize that this will lead to a reduction in the degree of conflict. In effect, the VR technology might allow us to cue the non-expert to factors on which experts automatically focus.

There are many potential applications of the idea of virtual experiments in environmental economics. Subjects can be given virtual choices that mimic those underlying the conceptual experiments used in the travel cost method, hedonic pricing tasks, and stated choice tasks. These VR versions can add naturalistic cues in a controlled manner, allowing researchers to see how that affects valuations and choices. In particular, VR typically requires that one construct some simulation of the natural physical process generating the environment, and we will show that this has an important and unexpected benefit, restricting the domain of choice to coherent

[1] For example, Bohm and Lind [1993], Harrison and List [2003] and List [2004].

environments.

Virtual experiments can also be used to illustrate to subjects the latent implications of choices made today. In agricultural applications, for instance, what is the effect of fertilizing decisions today on the crop outcomes or water quality of tomorrow? What are the effects of environmental policies today on health tomorrow? Visualizing these long-run effects correctly should improve short-run decisions with long-run consequences.

In summary, we review and illustrate the use of virtual experiments as a new tool in environmental economics, with an emphasis on the methodological issues involved. In section 2 we examine the relationship between laboratory experiments and field experiments, focusing on the features of each that make them complementary. In sections 3 and 4 we introduce the areas of VR and NDM to environmental economists, since each is a well-developed sub-discipline with lessons for the design of VX. We then survey potential applications to environmental economics in a general manner, focusing on valuation tasks (section 5) and the representation of uncertainty (section 6). Section 7 then presents the plan of a case study in which these techniques are applied to the assessment of the consequences of wildfire risks and fire management options in Florida. Section 8 draws some conclusions.

2. Experimental Economics in the Laboratory and the Field

It is tempting to think of field experiments as being akin to laboratory experiments, but with more relevance and less control. According to this view, lab experiments maximize internal validity, but at the cost of external validity. Greater external validity comes at the cost of internal validity, and that is just a tradeoff we have to make. We argue that this view is too simple, and does not do justice to the nature of the controls that are needed in experiments of *any kind* in order for them to be informative.

Perhaps the problem is just with the expression "external validity." The term normally refers to the validity of applying the inferences drawn from the lab to field behavior. Nevertheless, from a scientific viewpoint, the validity of experimental inferences must first and foremost depend on the theoretical framework that is being used to draw inferences from the observed behavior in the experiment. If we have a theory that (implicitly) says that hair color does not affect behavior, then any experiment that ignores hair color is valid from the perspective of that theory. On the other hand, if our experiment confirms casual field observations that hair color does matter for behavior, but we have no theory that logically predicts any causation, how can we claim that the experiment has external validity? In order to identify what factors make an experiment valid we have to possess some priors from a theoretical framework, which is crossing into the turf of "internal validity." Furthermore, the "theory" at issue here should include the assumptions required to undertake statistical inference with the experimental data, and not just treat them as an

unfortunate, residual appendage, as stressed by Ballinger and Wilcox [1997]. Thus, external and internal validity are not independent.

Harrison and List [2004] argue that lab experimenters may actually lose control of behavior when they use abstract instructions, tasks and commodities; use procedures which are unfamiliar to subjects; or impose information which subjects are not accustomed to processing. The same argument has been made by Ortmann and Gigerenzer [1997], and older traditions in psychology that they reference. If much of decision making depends on decision heuristics developed through years of experience in the field, and involves relying on field cues for selecting good options, it should come as no surprise that behavior may become seemingly irrational when no field cues are present, or when inappropriate ones are given. Rather than ensuring generality of the conclusions about behavior, such common "sterilizing" devices only serve to encourage subjects to import their own context and subjective field referents. Without knowledge of that context and set of referents, the experimenter has actually *lost* control of the behavior under study. The outcome of an experimental design process that intentionally sterilizes the task is to introduce potentially fatal sample selection biases similar to those that arise when using non-random subject recruitment processes (Hertwig and Ortmann [2004]). Parallels exist to continuing debates over the relative validity of *in vitro* and *in vivo* techniques in biology, particularly in the realm of enforced animal and voluntary human testing of new drugs. *In vitro* tests use glass beakers and culture dishes, and therefore occur outside of a living organism; *in vivo* tests occur within the living organism.

It may be useful to have a concrete example to illustrate. In some experimental designs subjects have very restricted access to cognitive tools and aids that they would naturally use in the field. Clark [1997] discusses the cognitive process as involving not just resources embedded in our bodies, but also other tools and resources that we naturally use in decision making, problem solving, and learning. Consider the *Tower of Hanoi* game which has been extensively studied by cognitive psychologists (e.g., Hayes and Simon [1974]) and more recently by economists (McDaniel and Rutström [2001]). The physical form of the game, as found in all serious Montessori classrooms, is shown in Figure 1.

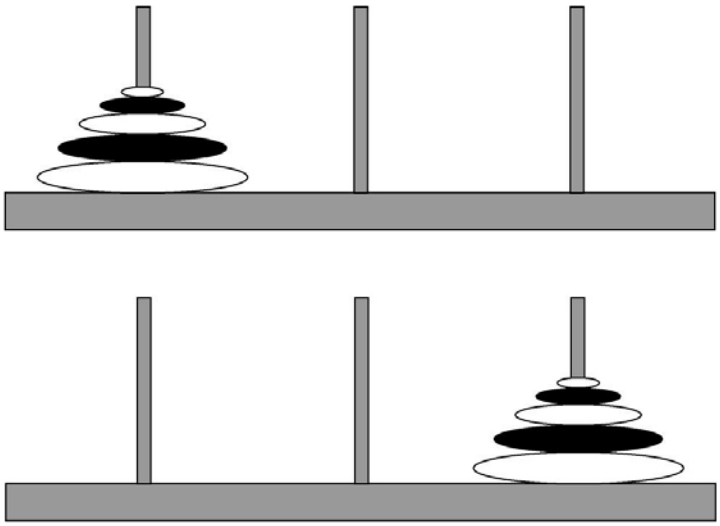


Figure 1: The *Tower of Hanoi* Puzzle

The top picture shows the initial state, in which n disks are on peg 1. The goal is to move all of the disks to peg 3, as shown in the goal state in the bottom picture. The constraints are that only one disk may be moved at a time, and a larger disk can never be placed on a smaller disk. The objective is to reach the goal state in the least number of moves. The "trick" to solving the *Tower of Hanoi* is to use backwards induction: visualize the final, goal state and use the constraints to figure out what the penultimate state must have looked like (viz., the tiny disk on the top of peg 3 in the goal state would have to be on peg 1 or peg 2 by itself). Then work back

from that penultimate state, again respecting the constraints (viz., the second smallest disk on peg 3 in the goal state would have to be on whichever of peg 1 or peg 2 the smallest disk is *not* on). One more step in reverse and the essential logic should be clear (viz., in order for the third largest disk on peg 3 to be off peg 3, one of peg 1 or peg 2 will have to be cleared, so the smallest disk should be on top of the second smallest disk).

Observation of students in Montessori classrooms makes it clear how they (eventually) solve the puzzle, when confronted with the initial state.[2] They shockingly violate the constraints and move all the disks to the goal state *en masse*, and then physically work backwards along the lines of the above thought experiment in backwards induction. The critical point here is that they temporarily violate the constraints of the problem in order to solve it "properly."

Contrast this behavior with the laboratory subjects in McDaniel and Rutström [2001]. They were given a computerized version of the game, and told to try to solve it. However, the computerized version did not allow them to violate the constraints. This was an appropriate constraint for the limited purposes of that research project, which were to understand how financial incentives affect the effort applied to the task and the efficacy of that effort in unfamiliar situations. Nevertheless, if the purpose is to more broadly understand the cognitive process that humans apply to problem solving then such a constraint would be inappropriate. The laboratory subjects were unable to use the classroom Montessori method, by which the student learns the

[2] This type of observation has not been ignored in the psychology and cognitive science: see Bull et al. [2004], Goldin-Meadow and Wagner [2005], and Senn et al. [2004], for example.

idea of backwards induction by exploring it with unconstrained physical referents. Without the physical analogue of the final goal state being allowed in the experiment, the subject was forced to visualize that state conceptually, and to likewise imagine conceptually the penultimate states. Although that may encourage more fundamental conceptual understanding of the idea of backwards induction, if attained, it is quite possible that it posed an insurmountable cognitive burden for some of the experimental subjects.

It might be tempting to think of this as just two separate tasks, instead of seeing it as a real object and its abstract analogue. But we believe that this example does identify an important characteristic of ideal field experiments: the fact that they select the subjects and the tasks jointly rather than separately. If cognition is not independent of the environment of the subject, and if the environment plays an integral role in cognition, then it would be impossible to make general statements about context-free cognition. Thus, joint selection of subjects and tasks imply a natural matching that should maximize cognitive efficacy.[3]

The example also illustrates that experiences of off-equilibrium states, in which one is not optimizing in terms of the original constrained optimization task, may indeed be critical to the attainment of the equilibrium state.[4] Such off-equilibrium experiences play an integral part in

---

[3] *Eventually* maximize cognitive efficiency. It is assumed that we are observing the end result of some lengthy Darwinian process, even if the process occurs at the social level with so-called "memes" instead of at the genetic level with genes. Blackmore [1999] and Boyd and Richerson [2005] point to a large literature on cultural evolution.

[4] This is quite distinct from the valid point made by Smith [1982; p.934, fn.17], that it is appropriate to design the experimental institution so as to make the task as simple and transparent as

learning theories (Erev and Roth [1998], Camerer and Ho [1999]), although they can raise formidable problems of statistical inference (Wilcox [2006]). Thus we should not ignore field devices that allow subjects to explore off-equilibrium states, even if those states are ruled out in our null hypotheses based on equilibrium predictions.

Finally, the example illustrates one of our main themes, that lab and field are complements. Recognizing the way in which students solve the puzzle in naturally occurring settings, one *might* simply re-design the computer representation of the task to allow subjects to try constraint-violating moves or to access natural field-relevant cues. The general point, which we propose as the rationale for Virtual Experiments, is that a science of undertaking such representations already exists, and we should pay attention to it. Just as the commercial game playing industry is benefitting greatly from advances in the science and technology of VR, so should social scientists.

There are many parallels to these concerns in the literature. The father of modern field experiments, the late Peter Bohm, was quite clear on this matter.[5] And it is well known from the pioneering work of Charles Plott and co-authors that the "simple" choice of parameters can add

---

possible providing one holds constant these design features as one compares experimental treatments. Such designs may make the results of less interest for those wanting to make field inferences, but that is a tradeoff that every theorist and experimenter faces to varying degrees.

[5] Bohm [1972][1979][1984a][1984b][1994] repeatedly stressed the importance of recruiting subjects who have some field experience with the task or who have an interest in the particular task. His experiments have generally involved imposing institutions on the subjects who are not familiar with the institution, since the objective of the early experiments was to study new ways of overcoming free-rider bias. But his choice of commodity has usually been driven by a desire to confront subjects with stakes and consequences that are natural to them. In other words, his experiments illustrate how one can seek out subject pools for whom certain stakes are meaningful.

significant field context to lab experiments.[6] In their applications Plott and co-authors selected parameters that were relevant to field applications and policy issues, and took these parameters into the lab. Thus the lab behavior they observed was in an interesting domain of elasticities, market power configurations, and institutional constraints. To extend these attempts to include a focus on providing field cues and cognitive scaffolds is quite natural.

3. Virtual Environments

3.1 Virtual Reality

In a broad sense, a Virtual Reality (VR) is any computer-mediated synthetic world that can be interactively experienced through sensory stimuli. The range of senses supported nearly always includes sight and very often hearing. Touch, smell and taste are less often supported, unless required by the VR system's application. For example, touch is often provided when using VR to train a person in mechanical skills.

Experiencing a VR requires that the user have some means to navigate through the virtual world. Although not strictly required, such experiences typically allow interaction, and not just observation, such that the actions of the user affect the state of the virtual environment. For all cases of allowed interaction, including navigation, the actions of the user must result in a response that occurs as quickly as in the physical world. Thus, navigation should be smooth and natural.

---

[6] See Grether, Isaac and Plott [1981][1989], Grether and Plott [1984], and Hong and Plott [1982]

Where interaction is allowed, the action of a user must result in a reaction in the virtual world that is appropriate, within the framework of the virtual world's model of truth (its physics, whether real or fanciful), and as responsive as interaction would be with the virtual object's real world counterpart, if such a counterpart exists.

An immersive VR refers to one that dominates the effected senses. As VR is typically visual, the primary sense of vision is often dominated by having the user wear a head-mounted display (HMD) or enter a CAVE (a "box" in which the surrounding visual context is projected onto flat surfaces). The user's position and orientation are then tracked so that the visual experience is controlled by normal user movement and gaze direction. This supports a natural means of navigation, as the real movements of the user are translated into movements within the virtual space.

The realistic navigation paradigm described above works well within constrained spaces and simple movements, but is not as appropriate for motions such as rising up for a god's eye view and then rapidly dropping to the ground for a more detailed view. There are HMD and CAVE-based solutions here, but such experiences are more often provided with either a dome screen, illustrated in Figure 2, or a flat panel used to display the visual scene. Since dome screens and flat panels are clearly less immersive than HMDs and CAVEs, experiences that are designed for these lower-technology, lower-cost solutions need to employ some means to overcome the loss of physical immersion. This is generally done through artistic conventions that emotionally draw the

user into the experience. In effect, the imagination of the user becomes one of the tools for immersion (Stapleton and Hughes [2006]). This is somewhat akin to the way a compelling book takes over our senses to the point that we "jump out of our skins" when a real world event occurs that seems to fit in with the imagined world in which we find ourselves.

Figure 2: A Dome Screen

3.2 Virtual Trees and Virtual Forests

Our case study considers wildfire management policies, and demands some consideration to the rendering of virtual trees and forests. Computer modeling of trees has been an active research area for a number of decades. Cohen [1967] and Honda [1971] pioneered computer models for tree generation. Prior to that, Ulam [1962] and then Lindenmayer [1968] studied purely mathematical models of plant growth. Lindenmayer proposed a set of rules for generating text strings as well as methods for interpreting these strings as branching structures. His method, now referred to as L-systems, was subsequently extended to allow random variations, and to account for the interaction of growing plants with their environment (Mech and Prusinkiewicz [1996]; Prusinkiewicz, James and Mech [1999]; Deussen et al. [1998]). A wide range of environmental factors, including exposure to sunlight, distribution of water in the soil, and competition for space

with other plants, were incorporated into the model, leading to biologically accurate ecosystems.

The above approaches focus on biologically accurate generation of vegetation, not necessarily on the graphical rendering. A number of methods have been proposed for generating plants with the goal of visual quality without relying on botanical knowledge. Oppenheimer [1986] used fractals to design self-similar plant models. Bloomenthal [1985] assumed the availability of skeletal tree structures and concentrated on generating tree images using splines and highly detailed texture maps. Weber and Penn [1995] and Aono and Kunii [1984] proposed various procedural models. Chiba et al. [1997] utilized particle systems to generate images of forest scenes.

These and other approaches have found their way into commercial products over the last decade. In general, L-systems are used when each model must be unique and when a single user experience can involve noticeable tree growth, such as when simulated years pass by during a single visit to the virtual world. When the degree of reality that is needed can be provided without uniqueness, and where dynamic changes are limited, virtual forests are generally created using a limited number of trees, with instances of these trees scaled, rotated and otherwise manipulated to give the appearance of uniqueness.[7]

---

[7] For our case study, we have selected commercial library of trees and its associated renderer, *SpeedTree*, to provide a realistic-enough forest (see Section 7 for further details). However, we plan to revisit this issue as this project progresses, with a critical eye towards extending the experiences we develop to allow not only fire spread but also forest changes based on policies that can affect nutrients, pollution, tree diversity and other disruptions to the natural cycles. Our confidence in successfully using this seemingly more computationally intensive approach is based on the results we have obtained in a related forest walk-through project (Micikevicius and Hughes [2007]; Micikevicius et al. [2004]).

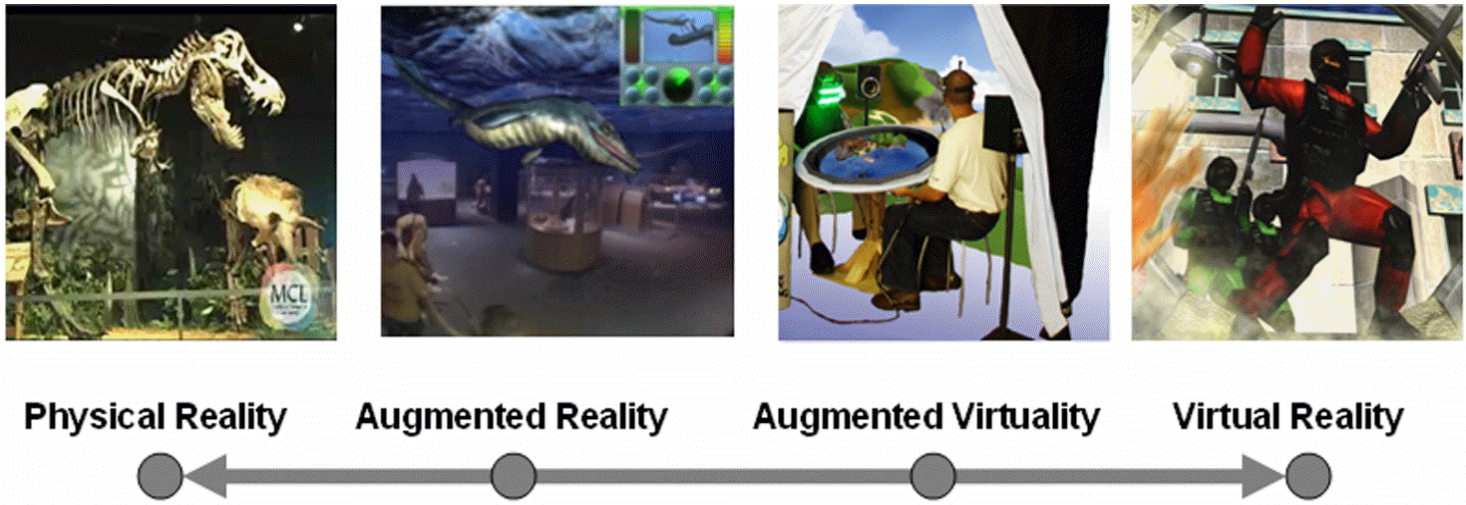


Figure 3: The Mixed Reality Continuum

3.3 Extending Experiences to Mixed Reality

While a VR experience is totally synthetic, a Mixed Reality (MR) experience involves elements from the real and virtual worlds (Hughes et al. [2005]). In MR we capture the real world and merge its sensory inputs with those of a virtual world. When this is done in such a way as to make the setting real, with virtual augmentation, the experience is called an Augmented Reality (AR). When an essentially synthetic world is augmented with objects extracted from the real world, the experience is called an Augmented Virtuality (AV). The continuum from purely real to purely virtual, illustrated in Figure 3, is the domain in which MR operates (Milgram and Kishino [1994]; Tamura et al. [2001]).

There are some experiences that just beg for a physical presence. Fidopiastis et al. [2006] provide a compelling example of this in an application to cognitive rehabilitation as depicted in Figure 4. The left monitor shows one view through the HMD; the middle monitor shows real-time

Figure 4: Mixed Reality Used in Cognitive Rehabilitation

tracking of the movements of the patient; and the right monitor shows the subject at home. In this experience, the goal is to help a person with traumatic brain injury return home and regain some measure of independence, for instance in the kitchen. Because a pure VR experience is disembodied, it cannot support activities that require subtle senses of touch, such as making coffee. In contrast, an MR experience can contain a real coffeemaker, grinder, cups, etc. The use of virtuality here is to make a generic kitchen look like the home environment of the subject. The hypothesis is that this makes transfer of learning easier since the practice is occurring in context.

Figure 5: The Technical Infrastructure of Mixed Reality

Moreover, with no technical effort, this MR experience can satisfy the olfactory and taste senses (you smell and taste the coffee).

The technical infrastructure of MR is illustrated in Figure 5, and provides a useful summary of the components involved. First we start with some video capture of reality, in this case acquired through cameras set into the front of an HMD. Second, we have to track the position and orientation of the subject in the environment. This tracking must be done very precisely in order to properly merge virtual and real content. Third, we have to generate the experience using one or more software "engines" to modify the environment. This is the stage in which the real environment is blended with virtual content. Underlying this can be 3D computer-generated

images, video of some physical environment, 2D computer-generated images, special effects of various kinds (e.g., real water vapor emanating from a virtual tree or actual warning lights going on and off), and additional audio enhancements (e.g., ambient sounds like trucks in the distance or spatial sounds like burning trees crackling nearby). Fourth, the visual aspects of the merged scene are delivered to the subject through some rendering device, in this case LCDs embedded on the inside of the HMD. Of course, not all of these components are needed in every application, but all these capabilities must be present in any system that purports to provide MR experiences.

MR is not strictly needed to create the experiences we are developing for the case study presented in section 7. But it can bring another dimension to the task: collaboration (Hughes and Stapleton [2005]). In MR we can place people in the rich context of a forest that surrounds them, a physical table on which they see a top-down view of the same forest, and a social setting in which they can discuss decisions and consequences with neighbors and family who are jointly setting policies and experiencing their consequences. The hypothesis is that this immersion and the social responsibility of discussing decisions with others might lead naturally to decisions with more consideration given to the preferences of others. There is abundant evidence for this hypothesis from laboratory and field experiments conducted with common pool resource extraction problems (e.g., Ostrom et al. [1992][1994]).

4. Naturalistic Decision-Making

4.1 Learning from Experts

In NDM researchers study expert decision makers in their natural task environment. Economists are a much more cynical lot, typically thinking about experts in some strategic, lobbying or legal sense. We are not using the term in that legalistic sense as an advocate for someone else's position, but rather in the sense that Keynes famously referred to the humble dentist who simply knows his craft.

Applications often involve time pressure, high stakes, inadequate information, ill-defined goals, and dynamic conditions. The focus is on studying the matching of cues to decisions using pattern recognition and template cognitive processes developed through years of training.[8] Contextual cues are critical to this pattern matching process in which the expert engages in a type of mental simulation. Theories emerging out of this literature suggest that experts employ forward thinking simulations in order to ensure that the best option is selected. We propose that lessons from this literature can be extended to human decision makers generally as long as they are in familiar task environments, even when they would not naturally be labeled "experts." We hypothesize that the VR environment can be used to generate cues that are sufficiently natural and familiar that decisions will be significantly more like those that would be generated in the field with sufficient expertise.[9]

---

[8] Klein [1998] and Salas and Klein [2001] are standard references in the area.

[9] Although NDM research suggests that option generation and evaluation is not used by experts in

The automatic nature of these cue and pattern matching heuristics implies that the decisions can be significantly sub-optimal when the decision maker is put into an unfamiliar context. One of the reasons that experts and non-experts have conflicting views on the appropriateness of different policies is the extent to which they differ in how their decision heuristics are matched to the context. The context of the expert may be based on specific goal functions, perhaps inappropriate to the policy decision, or the non-expert may be lacking in experience in professional and scientific judgment required for the context. The VR environment enables one to control and manipulate the field cues in ways that allows a blend of the perceptual and conceptual processes that decision makers use. For example, in the case study of forest fire management options in section 7, we propose comparing the decisions made by forestry and fire management experts to those of non-expert members of society.

To the extent that differences in knowledge and attitudes between experts and non-experts lead to different conclusions, implementing solutions to economic environmental problems may be difficult, leading to political conflict and other social costs. Developing mechanisms for understanding and resolving the sources of such differences is likely to have a high expected return

---

the classical sense of decision making, that research has explored the rapid decision processes engaged by experts in time-stressed situations. Although experts in such situations may rely more heavily on perceptual and pattern-recognition processes, the simulation process thought to be engaged by experts after the first option is generated, based upon pattern matching, can be construed of as a form of evaluation. Our point here is that NDM studies of experts suggests that there is a heavy reliance on perceptual information in the decision process. Our interest is in the extent to which a lack of this knowledge, for naïve or novice decision makers, can be overcome through the use of advanced technologies such as VR.

because they can lead to solutions that are both economically optimal and politically acceptable. Here we are proposing the use of VR simulation technologies for augmenting the scientific capabilities of non-experts, thereby reducing the conflicts that result from differing perspectives.

4.2 Ecological Rationality

Another perspective on what extensions of NDM into experimental research can generate is a consideration of the vast "cognitive illusions" literature popularized by behavioral economists. This literature, reviewed by Gigerenzer et al. [2007], has evolved primarily within cognitive psychology, and generated major debates that have been undeservedly ignored by most behavioral economists. To provide just one example, it has been shown that instances of the "base rate fallacy" of Bayesian reasoning, the neglect of the base rate when making inferences, is significantly reduced if one simply presents tasks to subjects using natural frequencies rather than probabilities. Natural frequencies involve referring to 550-in-1000 women, rather than the probability 0.550. The argument is that humans have developed ecologically valid heuristics for handling some operations with natural frequencies, but that those heuristics have not evolved to work with probabilities.[10] VR

[10] One should also add, particularly for an audience of environmental economists, that hypothetical bias is a major confound in behavioral economics, despite assertions to the contrary. For example, Conlisk [1989] conducted simple lottery choices testing the Allais Paradox in which he dropped the zeroes from the prizes used by Allais in his hypothetical tasks, and paid subjects in a between-subjects design (each sub-sample only received one of the binary choices). What happened was simple: when motivated, subjects apparently just cranked out the expected value of the lotteries and chose accordingly. Of course, in more complex environments, which might just be more than 2 or 3 prizes per lottery, this heuristic may be hard to apply. See Harrison [2006; §4.3] for further discussion of the manner in which behavioral economists

enhances the ability of behavioral researchers to provide decision makers with information using presentation frames with which they are more familiar.

Research in memory reminds us that all subjects will impose a natural context on a task even if it literally involves nonsense. Long traditions in psychology, no doubt painful to the subjects, involved detecting how many "nonsense syllables" a subject could recall. The logic behind the use of nonsense was that the researchers were not interested in the role of specific semantic or syntactic context as an aid to memory, and in fact saw those as nuisance variables to be controlled by the use of random syllables. Such experiments generated a backlash of sorts in memory research, with many studies focusing instead on memory within a natural context, in which cues and frames could be integrated with the specific information in the foreground of the task (e.g., Neisser and Hyman [2000]). A healthy counter-lashing was offered by Banaji and Crowder [1989], who concede that needlessly artefactual designs are not informative. But they conclude (p. 1193) that

> ...we students of memory are just as interested as anybody else in why we forget where we left the car in the morning or in who was sitting across the table at yesterday's meeting. Precisely for this reason we are driven to laboratory experimentation and away from naturalistic observation. If the former method has been disappointing to some after about 100 years, so should the latter approach be disappointing after about 2,000. Above all, the superficial glitter of everyday methods should not be allowed to replace the quest for generalizable principles.

It is time to build the bridge between the lab and the field. Extending NDM concepts into the

---

have characterized the effects of hypothetical bias.

controlled and replicable environment of VR-enhanced experimental methods is one way of building such a bridge.

4.3 Experts Behaving Badly

Experts do not always make better decisions in an unqualified sense, and blind watchmakers make lots of bad watches before they get it right. So we should not always assume that the expert's methods are the best. In the case of ecological rationality, we know that we have to wait until the Darwinian process has had some time to work. But there are more serious issues with the interpretation of expert or specialist behavior as a normative standard.

The obvious qualification is when experts do not respect the preferences of affected parties to risk, time delay, or equity. In large measure, this was the point of the controversial *World Development Report* issued by the World Bank [1992] to examine the environmental priorities of the planet prior to meetings in Rio de Janeiro on global warming (a precursor to the proposed Kyoto agreement). It argued that if one was to adopt a global priority ordering that reflected international population instead of GDP, and one recognized that proximity to the absolute poverty line tends to raise one's discount rate dramatically, global warming would not be the #1 priority. Instead, the effects from water pollution and scarcity, and air pollution, were more important. This prompted great debates, not the least over the validity of using a "discounting metaphor" for what is better viewed as a matter of inter-generational and cross-national equity

(Schelling [1995]). These important debates continue to this day with the *Stern Review of the Economics of Global Warming* (Stern [2006]).

An equally obvious qualification is that expertise in one domain, even within economics, does not obviously transfer to other domains. Undisputed experts in one area might well make a complete hash of their advice in another area, as some environmental economists claim to have occurred in the *NOAA Expert Panel on Contingent Valuation* (Arrow et al. [1993]). Similarly, in the legal arena someone that is an expert at researching and publishing in an area need not be able to communicate these insights to a jury, making them useless as an "expert witness." The remedy here is to just define expertise accurately for the purpose.

A more subtle problem occurs when there is "too much" expert consensus on some matter, perhaps because the cognitive processes that everyone is using are biased in some manner. An important policy illustration comes from computable general equilibrium modeling of multilateral trade policy.

After the Tokyo Round, it was thought desirable to get all the major modelers to examine some common policies and trace through the reasons for different conclusions. The result was a sprawling mess, in which differences in raw data as much as modeling assumptions quickly accounted for the different policy conclusions (Srinivasan and Whalley [1986]). This led to the development of common data bases, at the Global Trade Analysis Project (GTAP) (https://www.gtap.agecon.purdue.edu/) that virtually every modeler used for the assessments of the

Uruguay Round. This dramatically enhanced the ability to see what assumptions explained different conclusions: see Harrison et al. [1997] for an extended illustration. But one of the maintained, untested assumptions in this comparison was that the GTAP database was correct. There is no evidence *ex post* that the use of a common database led to systematic errors that could have been avoided at the time, but this was not something that a contemporary policy-maker would have any way of knowing.

The ease with which a consensus was reached in debates over the Uruguay Round has not been replicated in debates over the Doha Round, but that is primarily due to the strategic incentives for promoting contrarian modeling assumptions in the interests of advocacy. Similar issues confound modeling efforts over global warming, as illustrated by pre-Kyoto debates over the correct value of the "autonomous aggregate energy efficiency improvements" parameter.

Thus we do not claim that experts always get it right. The contexts that we have in mind in the present applications make it easy to think of experts that almost always do dramatically better: for example, compare an average pilot landing a plane to an average economist landing the same plane.

5. Virtual Choices

One of the mainstays of environmental valuation is the hypothetical or actual choice process. In a referendum setting the respondent is asked if they want to vote for policy A or policy B, and in some cases whether they want to vote at all. In a conjoint choice setting the respondent is asked if they want to purchase bundle A, bundle B, or neither, where bundles are exogenously defined in terms of characteristics of the environment. In a travel cost setting the respondent is observed to have chosen location A over other known locations B and C, given that they had to come from location Z.

In a VX setting one can represent the choice mechanism and the consequences virtually. The respondent can be placed in an environment in which the voting, purchasing or travel task is rendered in a naturalistic manner, and the subject experiences some of the affectual consequences of the choice. This might seem to be a trivial extension, but such devices can have a major impact on responses.

To provide an exemplar of the extreme VR environment that might be used for VX in the future, consider the online community known as *Second Life* (see http://secondlife.com/). The web site describes *Second Life* in this manner:

> *Second Life* is a 3-D virtual world entirely built and owned by its residents. Since opening to the public in 2003, it has grown explosively and today is inhabited by a total of 3,121,832 people from around the globe. From the moment you enter the World you'll discover a vast digital continent, teeming with people, entertainment, experiences and opportunity. Once you've explored a bit, perhaps you'll find a perfect parcel of land to build your house or business. You'll also be surrounded by

> the Creations of your fellow residents. Because residents retain the rights to their digital creations, they can buy, sell and trade with other residents. The Marketplace currently supports millions of US dollars in monthly transactions. This commerce is handled with the in-world currency, the Linden dollar, which can be converted to US dollars at several thriving online currency exchanges.

Many other such environments exist, although most are set in some mythical world in which the laws of physics and body proportionality do not apply. Castronova [2005] provides an entertaining review of the industry of these environments, with some eye to the economics of the synthetic environment.[11]

We discuss the value added from extending response frames that are currently adopted to VX. We discuss how VX enables researchers to match the information perspective to the preferences of the respondents, such as when allowing subjects to choose whether to view a geographic area from above or from the ground. VX also allows researchers to engage respondents in counter-factual dynamic scenarios as they evolve. Thought experiments that require subjects to imagine the future can be simulated and visualized, maintaining both the requirement of naturalistic cues and the requirement of scientific accuracy. We then review some important problems in all valuation exercises, and how they also impact the design of experiments based on VR.

---

[11] These online games are known as massively multiplayer online role-playing games (MMORPGs), since many humans can be online at the any time and assume certain roles in the game. There are some extreme statements about how MMORPGs differ from VR environments, such as Castronova [2005; p.286-294].

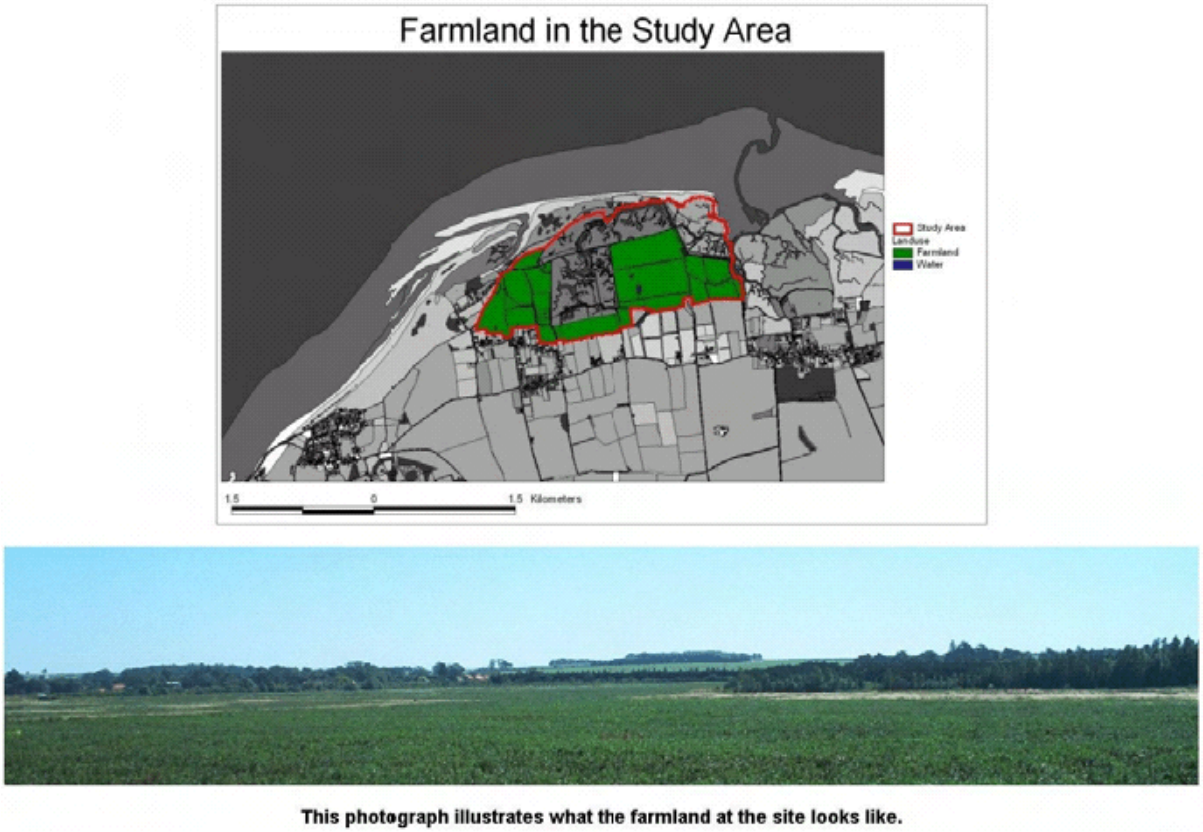


Figure 6: Introductory Displays from Bateman et al. [2006]

5.1 The Value-Added of Virtual Experiments

Many environmental valuation techniques provide respondents with some visualization of the choice options. VX provides a rich extension to such practices by adding dynamic rendering over time, properly constrained by scientific theories that predict the future paths, possibly in a stochastic manner.[12]

[12] There is also a valuable role for VR simulations of some latent process, even if one does not take the extra step of rendering the view. Bin et al. [2006] illustrate how one can use detailed GIS information on a set of parcels of land to simulate attributes of the view that is likely correlated with the land and a building of known height. In effect this is simulating a virtual person looking out of the property, in this case in the direction of an amenity, the ocean. They measure the number of degrees of ocean view associated with each property from the perspective of this virtual viewer, and use that information in a traditional hedonic model to identify the value of the amenity. This also provides better estimates of the value of disamenities, such as flood surge risk, correlated with that amenity. The fact that the virtual viewer takes into account virtual obstacles to the view makes these statistical controls significantly less correlated that if one did not use the VR simulation, enhancing statistical inferences.

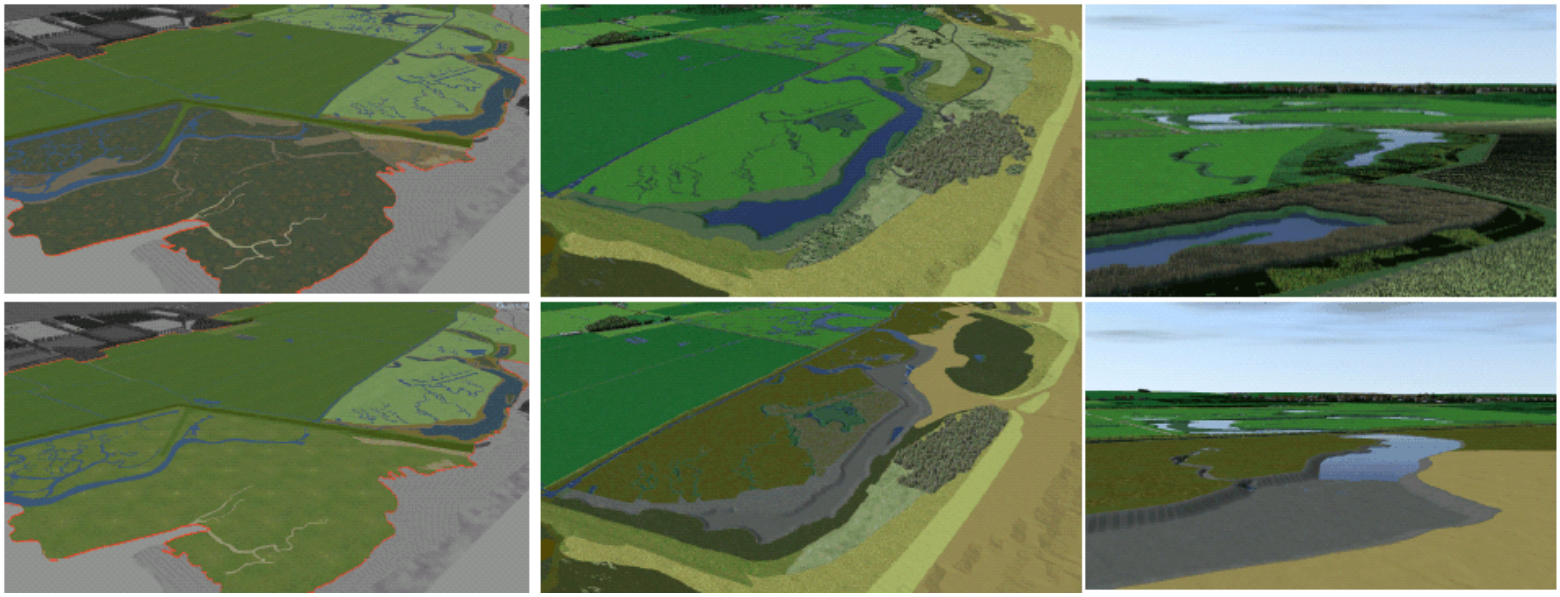


Figure 7: GIS rendering of alternative land uses, from Bateman et al. [2006]

Bateman et al [2006] use visualization methods for coastal areas to design hypothetical choice tasks for the subject, with the dynamic nature of the consequences implemented as "before and after" scenarios. The idea in this method is to import Geographic Information Systems (GIS) data for a naturally occurring area, render that area to the subject in some consistent manner, and then use a simulation model to generate a counter-factual land use and render that "side by side" to the subject.[13] Figure 6 shows some of the introductory stimuli presented to their subjects, to help locate the property on a map, and to have a photographic depiction of the property. Figure 7 shows the critical "before and after" comparisons of the property, using the VR rendering technology. The top panel in Figure 7 shows the property as it is now, from three perspectives. The bottom panel in Figure 7 shows the property in the alternative land use, rendered from the

[13] The technology used to render these images is discussed in more detail by Jude et al. [2006]. More detailed images can be accessed at http://www.uea.ac.uk/~e154/research/visualisations/, including some animated renderings.

same perspective. In this manner the subject can be given a range of counter-factual alternatives to consider.

Using VX extends this methodology by allowing subjects to manipulate the perspective of the visual information. Subjects can wander around the property and examine perspectives that interest them. Some people like the perspective of a helicopter ride, some like to walk on the ground, and most like to be in control of where they go.[14] It is desirable not to "force feed" the subject with pre-ordained perspectives so that they can search and find the cues that are most valuable to their decision making process. Of course, this puts a much greater burden on the underlying VR rendering technology to be able to accommodate such real-time updates, but that technology is available (as modern gaming illustrates). To an economist it also raises some interesting questions of statistical methodology, in which subjects self-select the attributes of the choice set to focus on. There is a rich literature on choice-based sampling, and such methods will be needed for environments that are endogenously explored (e.g., see Cameron and Trivedi [2005; §14.5]).

In fact, such extensions have in part been undertaken. Bishop [2001] and Bishop et al. [2001] allowed their subjects to visualize a recreational area in Scotland using 3D rendering, and walk virtually wherever they wanted within that area. They recorded the attributes of the locations

[14] In VR one refers to this distinction more formally in terms of egocentric and exocentric views. These distinctions are explored in order to understand their impact on general immersion in the environment, as well as the impact on performance in a variety of tasks (Darken and Cevik [1999]).

the subjects chose, and then modeled their choices statistically. The attributes of interest to them, such as how many pixels are in the foreground trees as compared to the background trees, are not of any interest to environmental economists, and we have our own way of modeling latent choices such as these, but the methodological contribution is clear. Moreover, these exercises were undertaken as one part of a broader conception of what we would call virtual experiments (see Bishop and Gimblett [2000]).

One elicitation problem that is common to many environmental issues is the *temporally latent* nature of the impacts of choices made today. Examples in terms of health impacts are clear, as we know all too well from tobacco, where the decision to start smoking is invariably made in the teens, and the consequences invariably felt after the age of 50. One role for VR is to be able to speed up such processes in a credible, naturalistic manner. It is one thing to tell a teen smoker that there will be consequences from a lifetime of smoking, but if one can show the teen these consequences then one would expect, or at least hope for, a more informed decision.

Perhaps a simpler example, at least in terms of the demands on rendering technology, would be to display the longer-run consequences for the environment of certain policies. Many researchers are providing numerical projections of the effects of failing to mitigate the risk of global warming, but it has been a standard complaint from advocates of the need for action that people find it hard to comprehend the nature or scope of the possible consequences. If we have access to VR systems that can display forests, for example, and we have projections from climate

change models of the effects of policies on forests, then one can use the VR setting to "fast forward" in a manner that may be more compelling than numbers (even monetary numbers).[15] Of course, one positive side-benefit of undertaking this activity properly would be a reflection of the uncertainty over those future predictions.

The potential value-added of including VR in the information given to subjects is clear. Nevertheless, many problems that are present in traditional valuation exercises are not resolved by the use of VR, and in some cases responses may become even more sensitive to the presence of these problems. We discuss some of these problems next.

5.2 Scenario Rejection

For some valuation scenarios the counter-factuals may produce perceptually very similar consequences. In order to generate greater statistical power researchers may then simulate physical or economic changes that are not consistent with scientific predictions. That is, the counterfactual scenarios may not make sense from a physical and economic perspective. Violating such constraints, however, can lead subjects to "scenario rejection."[16] This problem is bound to be

[15] An equally interesting possibility might be to turn the ecological clock backwards, allow subjects to change past policies, and then run the clock forward conditional on those policies instead of the ones that were actually adopted. Then the subject can see the world as it is now and compare it to what it might have been, perhaps with greater salience than some notional future world (however beautifully rendered).

[16] This term is an unfortunate short-hand for situations in which the subject views the scenario presented in an instrument as incredible for some reason, and hence rejects the instrument. There is an important semantic problem, though. If "incredible" just means "with 0 probability," then that is something that deserves study and can still lead to rational responses. There is nothing about a zero probability that

especially severe in situations with which the subject is very familiar, such as when expert decision makers are put into VR simulations with naturalistic rendering. Underlying the rendering of the visual cues there must therefore be a scientifically consistent simulation model.

Another reason for "scenario rejection" may be an inappropriate, or a vaguely stated choice process. Why am I being asked to choose between these two options? Is this a referendum vote? If so, am I the pivotal voter? Or am I the King (it is good to be King), who gets to say what the land use will be no matter what others think? Every one of these questions needs some answer for the responses to be interpretable. Of course, this is the same comment one would have for any hypothetical survey, and is independent of the application of VR technology. But VR technology allows one to fill in that choice structure in a natural way. We consider next a survey designed to value the damage caused by the *Exxon Valdez* oil spill, in which respondents were given a voting box and the option of not voting. One could readily imagine such referenda occurring in *Second Life*.

5.3 Voting in a Referendum

A voting referendum is a popular way to elicit responses regarding valuations of public

---

causes expected utility theory to be invalid at a formal level. But from an experimental or survey perspective, the term refers to the instrument itself being rejected, so that one does not know if the subject has processed the information in it or not. This semantic issue is important for our methodology, since we are focussed precisely on mitigating degrees of scenario rejection in artefactual instruments. Hence we do not want to view it as an all or nothing response, as it is in some literature in environmental valuation using surveys.

projects because they are natural and familiar ways for people to make such choices. In the United States there are often local bond issues for specific projects, and some fraction of the population exercises the right to vote.[17] The use of referendum processes is therefore consistent with NDM and provides a useful starting point for considering virtual choice. Even with such a familiar process, however, responses can be very sensitive to the details of the framing.

Carson et al. [1998] and Krosnick et al. [2002] report on a replication and extension of the original *Exxon Valdez* oil spill survey that Carson et al. [1992][2003] conducted. The replication was close, in the sense that essentially the same survey instrument was used from the original study and the same in-person survey method used. Three treatments were added that make the 1993 replication particularly interesting for present purposes, since they arguably served to make the voting task more realistic and familiar to the subject.

One treatment was the use of a "ballot box" in which the respondent could privately deposit their response, and did not have to reveal it to the interviewer. Another treatment was the addition of a "no vote" option, a choice that could be made prior to the voting stage. A third treatment was a cross of the last two: the "no vote" option was also implemented with a ballot

[17] Political scientists discuss at length the extent to which the sample of the population that votes is statistically representative of the population, even if it is legally representative for the purposes of the vote. Rarely does one see a referendum survey correct for the sample one would likely see in the field in the voting booth. It is not at all obvious that they should undertake that correction, but the issue is ignored. This is separate from how one treats those that respond "do not know if I would support this hypothetical referendum" or "I choose not to vote," which we consider later. It concerns the statistical re-weighting of those who do actually record a vote.

box. The overall sample was 1,182, with roughly 25% allocated at random to each version.

Figure 8 displays the raw responses in each of the four versions. To simplify, the light grey lines with a circle marker indicate responses "for" the proposed policy to avoid another oil spill over the next 10 years, and the black lines with square markers indicate responses "against" the policy. Dark grey lines with triangle markers indicate "no vote" responses when asked for. These responses have not been adjusted for any population weights, or complex survey sample design features; nothing of consequence here is affected by those corrections.

The top left panel of Figure 8 shows that respondents tend to favor the policy on balance when the tax-price is lowest, and that this net vote for the policy diminishes as the tax-price increases. These results would suggest that the largest tax-price that would just receive a majority vote would be around $60. The exact determination of the point at which the "for" responses equal 50% would be determined by applying one or other statistical estimators to the raw responses. Although there is some debate over the best estimators to apply, for now we will simply use the piecewise linear estimates connecting the raw responses.[18]

[18] One could use a simple probit estimator, in which the tax-price was an independent variable and the "for" responses the dependent variable. In this case the estimated tax-price can be solved from the estimated equation, which could control for individual characteristics in the sample to ensure comparability across versions. Using such a model, one can calculate damages for versions 1 through 4 of $1.390 billion, $2.587 billion, $1.641 billion and $3.659 billion, respectively. These calculations employ synthetic tax prices between $1 and $500 in increments of $1. There are short-hand expressions for such inferences, but this long-hand approach has other advantages.

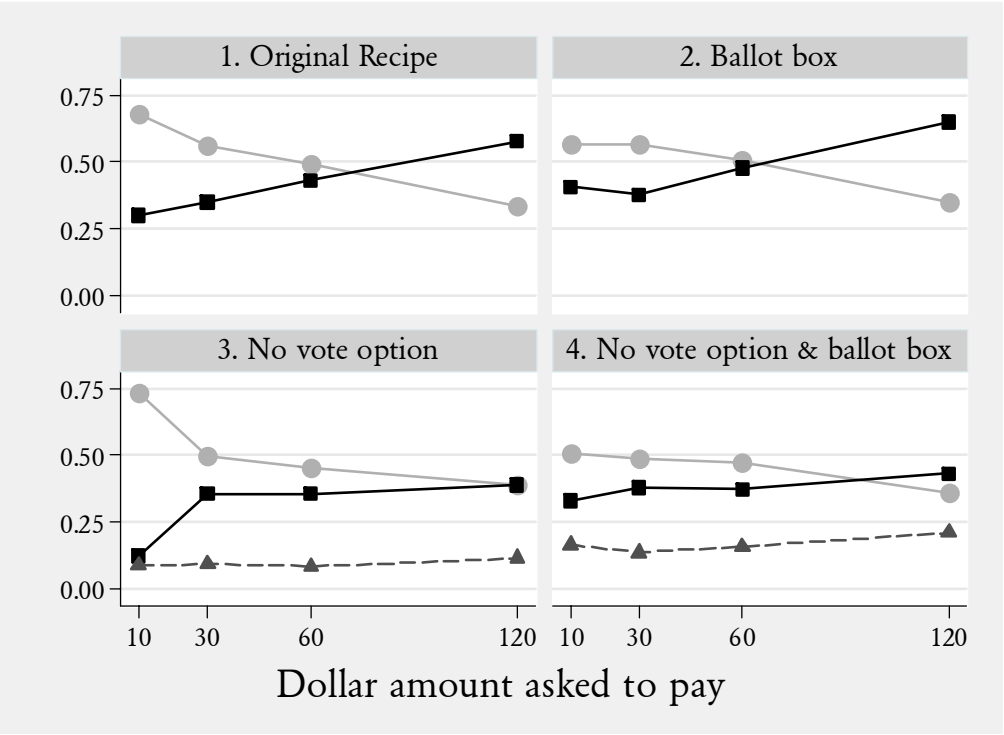


Figure 8: Raw Responses from 1993 *Exxon Valdez* CVM Survey

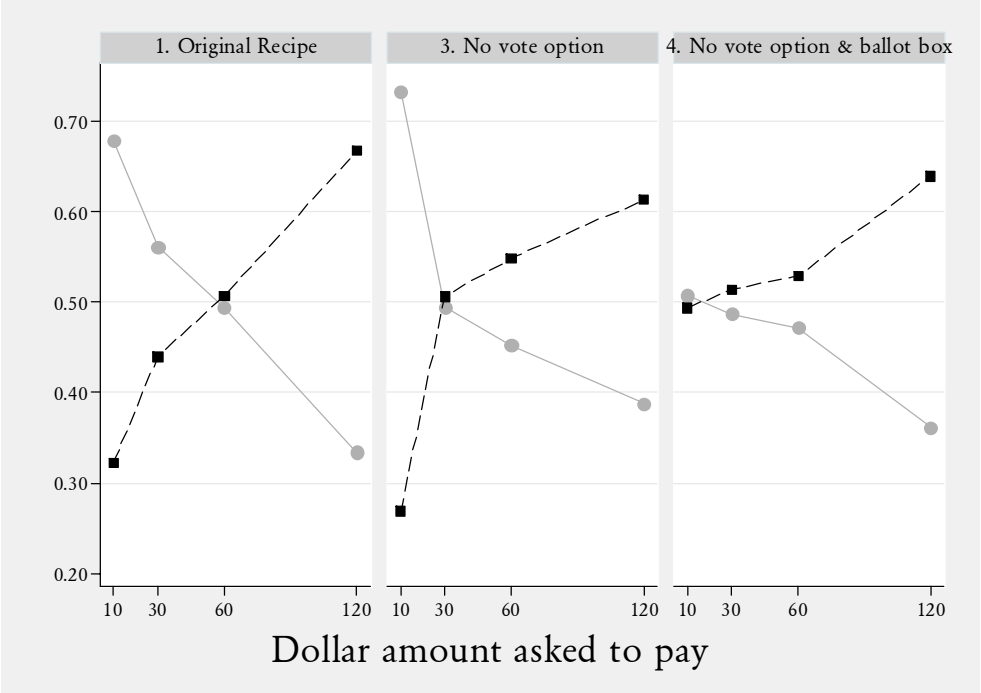


Figure 9: Responses to *Exxon Valdez* CVM Survey if "No Vote" or "Unsure" Means "Against"

The use of the anonymous ballot box, in the top right panel of Figure 8, reduces the positive and increases the negative responses at the lowest tax-price of $10, but otherwise has relatively little effect in relation to the "original recipe" used in version #1. As a consequence there is much less *net* support for the proposal at that $10 level, which might become an issue if one recognized sampling errors around these responses. But the main effect of the ballot box is to dampen enthusiasm for or against the proposal at the lowest tax-price, suggesting that those responses may have been affected by the presence of an observer of the vote.

The raw responses with the explicit "no vote" option are shown in the bottom left panel of Figure 8. Compared to version #1, above it, there is a very slight increase in the "for" responses at $10, and a marked reduction in the "against" responses. The other large change is the reduction in the "against" responses at the $120 tax-price. Although the fraction of "no vote" responses increases slightly with the tax-price, from 9% at $10 to 11% at $120, it is relatively constant with respect to tax-price. Taken together, these changes suggests that this option may be picking up respondents who originally voted against the proposal, perhaps because they reject one or more aspects of the scenario presented in the survey referendum.

When the no-vote and ballot box treatments are crossed, in version #4 shown in the bottom right panel of Figure 8, the fraction of "no vote" responses goes up to around 15% for the lowest tax-prices and then 21% for the highest tax price. The direct "for" and "against" responses at the lowest tax-price are akin to those observed for version #2, in which the ballot box was the sole

treatment. However, both are lower due to the fraction siphoned off into a "no vote" response. In fact, *the policy scenario barely garners a majority vote even at the lowest tax-price level of $10.*

One can go further and make the simple assumption that the "no vote" and "unsure" responses should be interpreted as "against." This is a natural interpretation, and adopted by Carson et al. [1998]. The results are dramatic, and displayed in Figure 9. Clearly the manner in which the choice is framed matters for the outcome, and in a non-trivial manner.[19] Consider however the effect of moving from version #1 to version #4: the referendum fails, or comes close to failing, at the lowest tax price of $10. It definitely fails at any higher tax price. No statistical tests are needed in this case to see that the interpretation of the CVM, plus the use of different versions, makes a huge difference to the implied damages.[20]

---

[19] Carson et al. [1998] conclude that the two left-hand panels in Figure 9 are not statistically different. Since they look different to the naked eye, it is useful to go through the statistical tests they use to draw their conclusion. The first test is a simple contingency table test of independence, where one treatment is the "for" or "against" response and the other treatment is the provision of a "no vote" option or not. These tests are then undertaken separately for each tax-price. Thus, for the $10 tax-price for example, the 68% and 32% responses in the left panel of Figure 9 (which are 59 individuals and 28 individuals) are compared to the 73% and 27% responses in the middle panel of Figure 9 (or 60 individuals and 22 individuals), using standard tests. The χ^2 statistic is 0.58 in this case, with a *p*-value of 0.45, so one cannot reject the null that these treatments are independent. Similarly for the other tax-prices. The problem is that these tests have extremely low power at these sample sizes. Using a *p*-value of 5% for a two-tailed test, one would need samples of over 1500 per cell in order to reject the above null hypothesis with a power of 0.90 or greater. Alternatively, with the actual samples of only 87 and 82, the power for tests of this null hypothesis is only 0.086. One implication is that, since the sample sizes cannot be increased, we must be prepared to make some parametric assumptions about responses in order to detect differences due to these treatments. This is what Carson et al. [1998] do, by considering parametric statistical models of responses.

[20] Uncertainty in valuation does not imply that the valuation is always uninformative. Bohm [1984], in particular, stressed the use of multiple measuring instruments that might bound any (strategic) bias to misrepresent value. If all of those methods yield values that completely exceed or completely fall below the cost of a policy, crisp cost-benefit inferences can still be made.

Referenda can be easily implemented using VR. As illustrated by this review, however, it is important to understand where attention to detail is needed in order to model a naturalistic setting in which the cues to the subject are contextually familiar. The use of VR does not provide a solution to this problem, but instead generates an even greater need to pay attention to such details.

6. Virtual Lotteries

Many economic policy assessments involve a great deal of uncertainty, either because the impact is stochastic in nature or because of uncertainty in the estimation of impacts. VX can facilitate the representation of uncertain prospects and can do so in a naturalistic way that allows respondents to use cues that are familiar to them.

In this section we review some of the representation problems of present methodologies for eliciting risk perceptions, risk attitudes, and valuations of risky objects. We will focus on the commonly studied "humble lottery" of expected utility theory. We then argue that VX is an attractive alternative, or complement, to these methods. An important aspect of VX is the ability to incorporate background risk into scenarios which focus on elicitation of preferences over foreground risk options. We also discuss the use of VX for policy choices where the estimates of the policy effects, rather than the effects themselves, are uncertain.

6.1 Task Representation and Perception

Our primary argument is that cognition is not independent of the environment, and that the success of the heuristics used depends crucially on how familiar the context appears to the decision maker. It is obvious, therefore, that the representation of the task should have a significant impact on decisions and inferred valuations. VX is proposed as a flexible and powerful method for designing tasks and environments that are familiar to subjects, particularly when there is some uncertainty about outcomes.

The essential problem is illustrated by the cartoon in Figure 10, taken from Altmann [1997; p.82]. The poor dog just keeps perceiving the object in front of him in terms of those objects he is familiar with. Assuming that they avoid deception, the challenge experimenters face is to get the subject to see what the object really is, and to be sure that the subject in games of interaction has some sense that other subjects see the same object.[21]

Figure 10: Dog and Bone

When subjects are presented with a lottery in an experiment, the experimenter has made some choice as to how to represent it. The prospects and probabilities could be written out in

[21] We focus here on settings in which there is no interaction between subjects, since that interaction raises a host of strategic issues to do with common knowledge and the perceived benefits of joint mis-representation of the game. A recent trend in experimental economics is to develop computer interfaces for tasks that rely on a minimum amount of formal "instruction" (e.g., Feiler and Camerer [2006]). Just as one often considers a software product well-designed if the user does not need to consult a manual to use it, such developments will likely make for more replicable experiments.

words, they could be written out in numbers, they could be displayed in pie charts, they could be displayed in bar charts, and so on. It would be wonderful if these alternatives had no effect on behavior, but unfortunately there is evidence that they do. Moreover, even when one derives valuation inferences that might seem *a priori* plausible, they might entail implausible changes in the underlying probabilities or implausible insensitivity to the underlying risks (e.g., Hammitt and Graham [1999]). Thus one has to "unpack" the constituent components of the humble lottery to know where things might be going wrong.

From the perspective of traditional EUT, the lottery consists of two objects. The first object is a probability, and the second object is a prize. In some cognitively degenerate cases, of course, some of these objects are not formally needed (e.g., when the certain prize $10 is represented as the lottery, "$10 with probability 1").[22] But we will not presume that subjects perceive the lottery the way we might manipulate them mathematically, since that begs the question of interest here.

6.1.1 Representation and Perception of the Prize

The prize is typically represented as a monetary prize, and we will focus on that representation to simplify matters. One central issue is whether the subject perceives the prize as a gain or a loss, since many models of choice under uncertainty are sign-dependent. Experimenters are

[22] This example is not chosen casually: there is considerable evidence that subjects behave differently when making choices in the interior of the Marschak-Machina triangle compared to making choices on the borders. The implications of this tendency are subtle (Humphrey [2001]).

often restricted in the ability to represent losses by human ethics considerations, not being able to impose actual negative net earnings on participants.[23] Thus, most losses are modeled in relation to some initial endowment of money, under the assumption that such endowments may generate natural reference points to subjects from which they then evaluate changes as gains or losses. Or one can have the subjects perform some task, pay them, and assert that they have integrated the earnings into their lifetime wealth before confronting them with lotteries involving losses from those initial earnings. That assertion may be true, but it is just one more auxiliary hypothesis.[24]

It is likely, though, that in many situations subjects bring their "homegrown" reference points to the experiments. For example, in our laboratory subjects expect around $15 on top of their show-up fee, and there is considerable subjective heterogeneity in those expectations.[25] That expectation should arguably be factored into their assessments as to whether prizes are gains or losses.

VX may affect the reference points of subjects if they are contextual, which is likely. Since

---

[23] We know of only one experiment where actual losses were credibly imposed on subjects and where the experimenters had human subject approval for these methods (Bullock and Rutström [2007]).

[24] Some instances may be more credible than others. For example, Mason et al. [2005] provided their subjects with an endowment of $100, as a gift from the experimenters. The size of this gift must have caused most subjects to pay attention to the fact that they now had such a large amount, in this experimental context, to risk. In the end average earnings were around $72.50, which is still a healthy amount of money. Thus the auxiliary hypothesis in this instance is likely to be more plausible *a priori* than if the endowments were much smaller. And there is considerable evidence from experiments that subjects behave differently if they are asked to make choices over stakes they have earned from some non-trivial activity, as compared to when they are asked to make choices over money they are randomly endowed with. In particular, Rutström and Williams [2000] and Cherry et al. [2002] showed that such subjects did not exhibit "social preferences" towards others.

[25] See Andersen et al. [2006] for this evidence.

the design provides a familiar naturalistic setting, reference points should be more salient to subjects than in a lab setting using stylized lottery tasks. Identification of reference points can therefore be facilitated by using VX.

6.1.2 Representation and Perception of Probabilities

There are two representational issues with probabilities. The first is that subjects may base their decisions on concepts of *subjective* probabilities such that we should expect them to deviate in some ways from objective probabilities. The second is that *perceptions* of probabilities may not correspond to the actual probabilities. Only with a theory that explains both the perception of probabilities and the relationship between subjective and objective probabilities would we be able to identify both of these deviations.[26] Nevertheless, careful experimental design can be helpful in generating some robustness in subjective and perceived probabilities, and a convergence in both of these on the underlying objective ones when that is normatively desirable. There is a large literature on the experimental methodology for achieving these goals, which we briefly review here. We argue that the naturalistic setting of VX can further facilitate achieving both of these goals.

---

[26] These issues are separable from the use that subjects put their perceived, subjective probabilities when weighing alternatives. Under EUT they use them directly, to linearly weight the utility of outcomes. Under many alternatives to EUT these perceived, subject probabilities are used in some non-linear, rank-dependent and/or sign-dependent manner to define decision weights used to linearly weight the utility of outcomes. We do not believe that it is productive to view EUT and non-EUT models as competitive alternatives, in which any one model should be expected to explain all behavior, but as complements in the sense that either could explain some behavior: see Harrison and Rutström [2005].

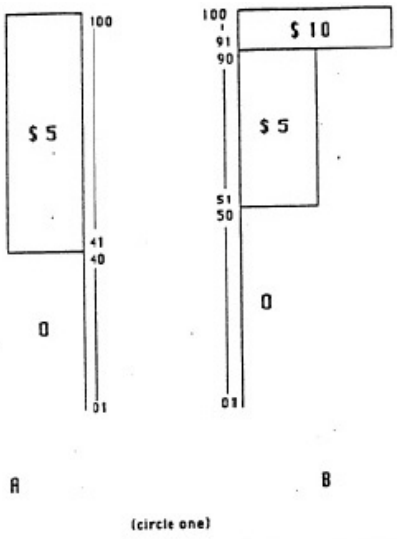


Figure 11: Lottery Display Used by Camerer [1989]

Camerer [1989] used a stacked box display to represent his lotteries to subjects. The length of the box provided information on the probabilities of each prize, and the width of the box provided information on the relative size of the prizes. The example in Figure 11 was used in his written instructions to subjects, to explain how to read the lottery. Those instructions were as follows:

> The outcomes of the lotteries will be determined by a random number between 01 and 100. Each number between (and including) 01 and 100 is equally likely to occur. In the example above, the left lottery, labeled "A", pays nothing (0) if the random number is between 01 and 40. Lottery A pays five dollars ($5) if the random number is between 41 and 100. Notice that the picture is drawn so that the height of the line between 01 and 40 is 40% of the distance from 01 to 100. The rectangle around "$5" is 60% of the distance from 01 to 100.
>
> In the example above the lottery on the right, labeled "B", pays nothing (0) if the random number is between 01 and 50, five dollars ($5) if the random number is between 51 and 90, and ten dollars ($10) if the random number is between 91 and 100. As with lottery A, the heights of the lines in lottery B represent the fraction of the possible numbers which yield each payoff.. For example, the height of the $10 rectangle is 10% of the way from 01 to 100. The widths of the rectangles are proportional to the size of their payoffs. In lottery B, for example, the $10 rectangle is twice as wide as the $5 rectangle.

This display is ingenious in the sense that it compactly displays the "numbers" as well as visual referents for the probabilities and relative prizes. The subject has to judge the probabilities for each prize from the visual referent, and is not directly provided that information numerically. There is a valuable literature on the ability of subjects to accurately assess quantitative magnitudes from visual referents of this kind, and it points to the need for individual-specific calibration in experts and non-experts (Cleveland et al. [1982][1983][1984]).

Battalio, Kagel and Jiranyakul [1990] and Kagel, MacDonald and Battalio [1990] employed purely numerical displays of their lotteries. For example, one such lottery was presented to subjects as follows:

A:	Winning $11 if 1-20	(20%)
	Winning $5 if 21-200	(80%)
B:	Winning $25 if 1-6	(6%)
	Winning $5 if 7-100	(94%)

Answer: (1) I prefer A. (2) I prefer B. (3) Indifferent.

This display presents all values numerically, with no visual referents. The numerical display does show the probability for each prize, rather than require the subject to infer that from the cumulative probabilities.

Hey and Orme [1994] used pie charts to display the probabilities of the lotteries they presented to subjects. A sample display from their computer display to subjects is shown in Figure 12. There is no numerical referent for the probabilities, which must be judged from the pie chart. As a check, what fraction would you guess that each slice is on the left-hand lottery? In fact, this lottery offers £10 with probability 0.625, and £30 with probability 0.385. The right-hand lottery offers the same probabilities, as it happens, but with prizes of £10 and £20 respectively.

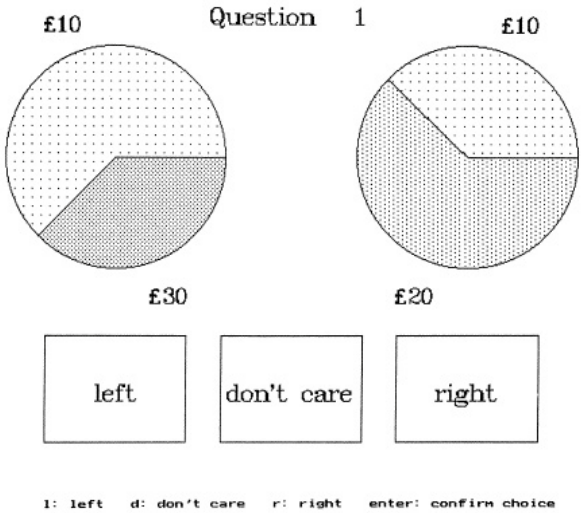


Figure 12: Lottery Display Used by Hey and Orme [1994]

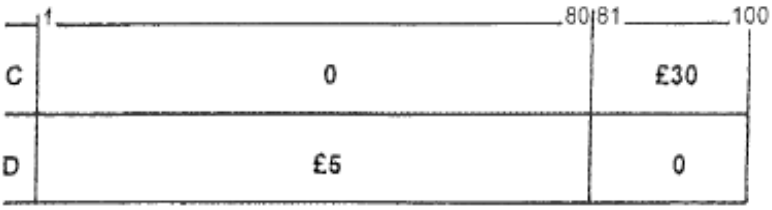


Figure 13: Lottery Display Used by Beattie and Loomes [1997]

Beattie and Loomes [1997] used displays that were similar to those employed by Camerer [1989], although the probabilities were individually comparable since they were vertically aligned with a common base. Figure 13 illustrates how they presented the lotteries to subjects. In addition, they provided text explaining how to read the display.

An example of the representation of probability using a verbal analogical scale is provided by Calman and Royston [1997; Table 4], using a distance analogue. For risks of 1 in 1, 1 in 10, 1 in 100, 1 in 1000, for example, the distance containing one "risk stick" 1 foot in length is 1 foot, 10 feet, 100 feet, and 1,000 feet, respectively. An older tradition seeks to "calibrate" words that are found in the natural English language with precise probability ranges. This idea stems from a concern that Kent [1964] had with the ambiguity in the use of colloquial expressions of uncertainty by intelligence operatives. He proposed that certain words be assigned specific numerical probability ranges. A study reported by von Winterfeldt and Edwards [1986; p.98ff.] used these expressions and asked a number of NATO officials to state the probabilities that they would attach to the use of those words in sentences. The dots in Figure 14 show the elicited probability judgements, and the shaded bars show the ranges suggested by Kent [1964]. The fact that there is a poor correspondence with untrained elicitors does not mean, however, that one could not undertake such a "semantic coordination game" using salient rewards, and try to encourage common usage of critical words.

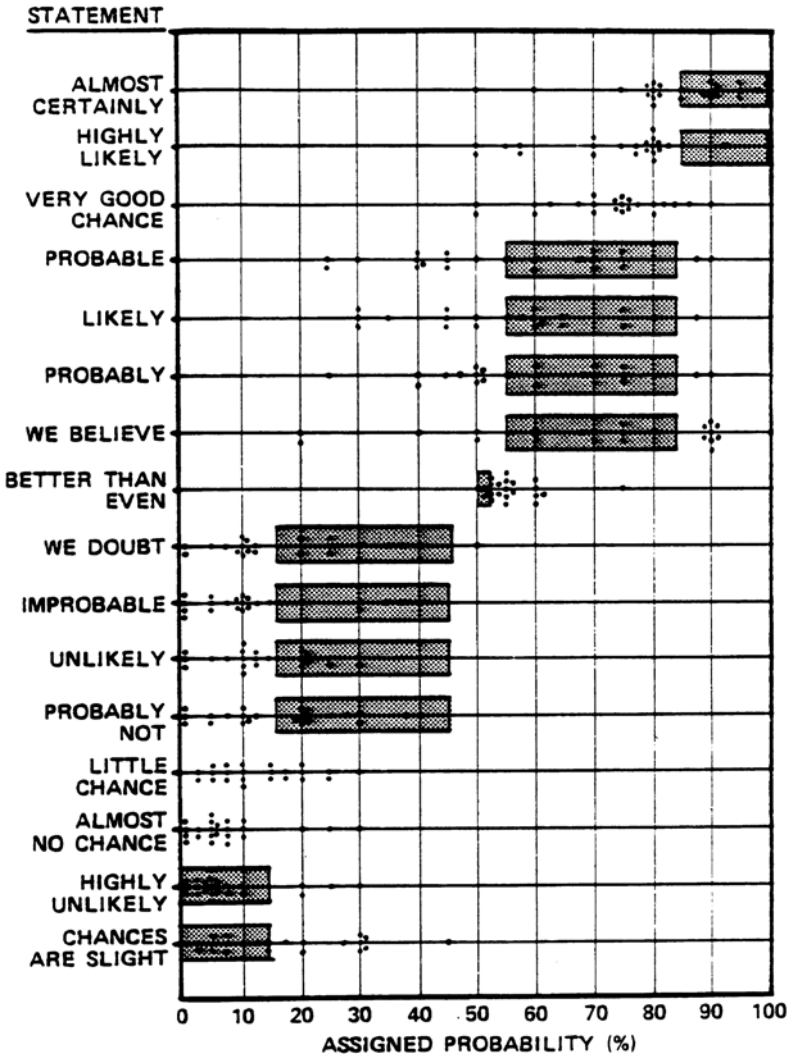


Figure 14: Representing Risk with a Verbal Analogical Scale

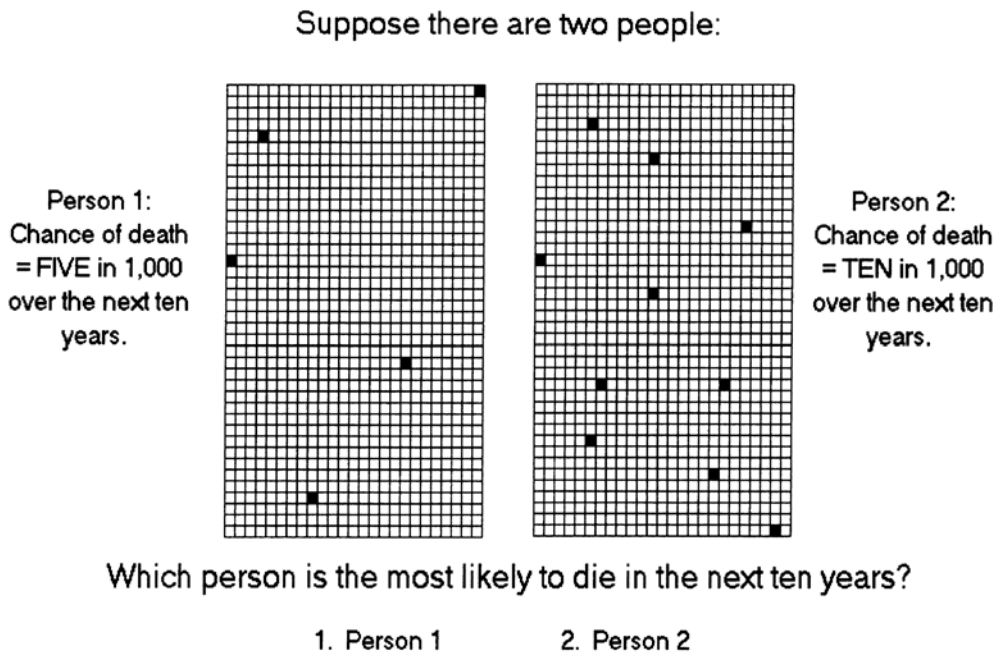


Figure 15: Representing Risk With Dots

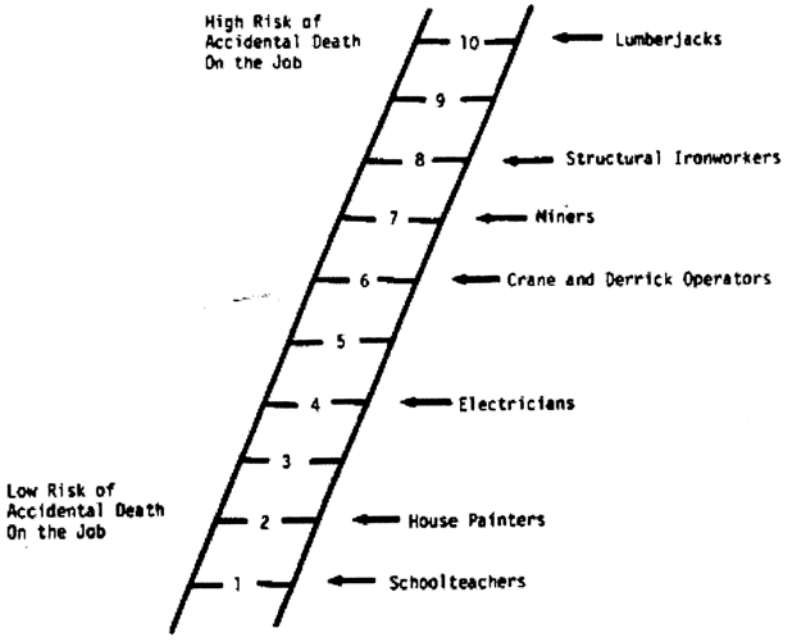


Figure 16: Representing Risk with a 2D Ladder

The visual dots method is employed by Krupnick, Alberini, Cropper and Simon [1999], and provides a graphic image to complement the direct fractional, numerical representation of probability. An example of their visualization method is shown in Figure 15.

Visual ladders have been used in previous VSL research by Gerking, de Haan and Schulze [1988] and Gegax, Gerking and Schulze [1991]. One such ladder, from their survey instrument, is shown in Figure 16. An alternative ladder visualization is offered by Calman and Royston [1997; Figure 1], and is shown in Figure 17.

One hypothesis to emerge from this review of the representation of lotteries in laboratory settings is that there is no single task representation for lotteries that is perfect for all subjects. It follows that some of the evidence for framing effects in the representation of risk may be due to the implicit assumption that one form of representation works best for everyone: the "magic bullet" assumption. Rather, we should perhaps expect different people to perform better with different representations.

To date no systematic comparison of these different methods have been performed and there is no consensus as to what constitutes a state of the art representation. VR provides a salient alternative to these methods because it gives the subject an opportunity to experience a range of possible outcomes generated by simulation from an underlying objective probability distribution. This can be seen either as a substitute for present methods, feasible in circumstances where each simulation would be fairly short and the range of possibilities narrow, or as a complement.

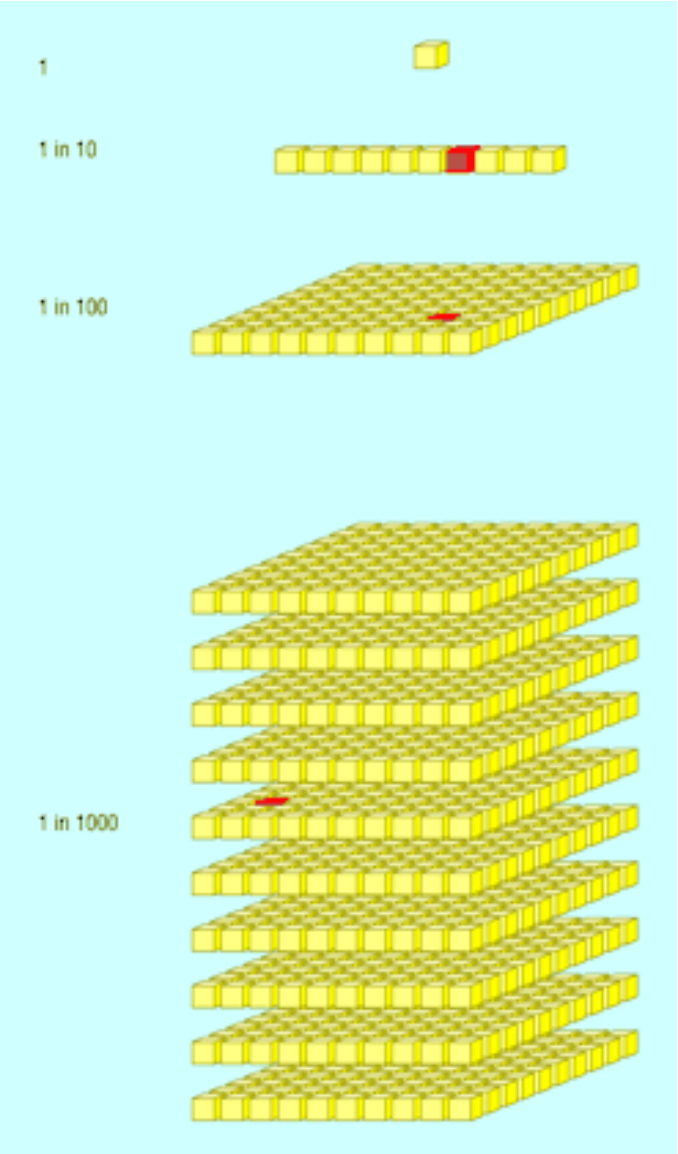


Figure 17: Representing Risk with a 3D Ladder

6.2 Foreground and Background Risk

One of the most severe problems in the communication and assessment of risk is the existence of "background risk" that might be correlated with the "foreground risk" of interest. To take a canonical case in environmental policy, consider the elicitation of the value that a person places on safety, a critical input in the cost-benefit assessment of environmental policy such as the Clean Air Act (U.S. Environmental Protection Agency [1997]). Conventional procedures to measure such preferences focus on monetary values to avoid mortality risk, by asking subjects to value scenarios in which they face different risks of death. The traditional interpretation of responses to such questions ignores the fact that it is hard to imagine a physical risk that could kill you with some probability but that would have no effect whatsoever on your health if you survived.[27] Of course, such risks exist, but most of the environmental risks of concern for policy do not fall into such a category. In general, then, responses to the foreground risk question should allow for the fact that the subject likely perceived some background risk.

In the traditional paradigm, risk aversion is viewed in terms of diminishing marginal utility of the final prize in some abstract lottery. The concept of a lottery here is just a metaphor for a real lottery, of course, although in practice the metaphor has been used as the primary vehicle for laboratory and field elicitation of risk attitudes. In general there is some commodity x and various

[27] Many VSL studies now jointly consider mortality and morbidity risk. The need to do so was an early theme from Gerking et al. [1988], who elicited information on morbidity and mortality risk in their surveys.

levels i of x, x_i, that depend on some state of nature which occurs with a probability p_i that is known to the individual whose preferences are being elicited. Thus the lottery is defined by $\{x_i ; p_i\}$. Traditional measures of risk aversion are then defined in terms of the curvature of the utility function with respect to x.

In many field settings, however, it is not possible to artificially identify attitudes towards one risk source without worrying about how the subjects view that risk as being correlated with other risks. As noted above, for example, mortality risks from alternative occupations tend to be highly correlated with morbidity risks. It is implausible to ask subjects their attitude toward one risk without some coherent explanation as to why a higher or lower level of that risk would not be associated with a higher or lower risk of the other. Theoretically, the effect of a background risk can be either an increase in the apparent risk aversion (so called risk vulnerability, see Gollier and Pratt[1996]) or a decrease (Quiggin [2003]).

The field experiment of Harrison, List and Towe [2007] used numismatists at a coin show as experimental subjects, and looked at the response of decision-makers to the addition of independent background risk. The foreground risk was a lottery with the prizes denominated in collectors' coins rather than cash currency, with the background risk represented by uncertainty about the quality of the coins. They show that the use of artificial monetary prizes provides a reliable measure of risk attitudes when the natural counterpart outcome has minimal uncertainty, but that it can provide an unreliable measure in a natural context when the natural counterpart

outcome has background risk. These results are consistent with the available theory from conventional expected utility theory for the effects of background risk on attitudes to risk.[28]

Another important inference gained from these results concerns the applicability of laboratory results to the field. The good news for the relevance of findings from previous laboratory studies is that there is little evidence to suggest that risk attitudes vary significantly due to the nature of the commodity. Yet, tempering this optimism is the fact that the observed risk posture depends critically on the level of background risk. This empirical insight calls into question the generalizability of the extant literature purporting to measure risk attitudes in the lab *and* the field, if they fail to measure or control for the potential confound of background risk. It is not the case that field estimates of risk attitudes are correct and lab estimates are inherently flawed: both are conditional on assumptions made about background risk and other theoretically relevant variables. However, it is the case that eliciting risk attitudes in a natural field setting requires one to consider the nature and degree of background risk, since it is inappropriate to ignore.

A further virtue of extending lab procedures to the field and, as proposed here, to the virtual field, is to encourage richer lab designs by forcing the analyst to account for realistic features of the natural environment that have been placed aside. In virtually any market with asymmetric information, whether it is a coins market, an open air market, or a stock exchange, a central issue is the quality of the object being traded. This issue, and attendant uncertainty, arises naturally. In

[28] For example, see Pratt and Zeckhauser [1987], Pratt [1988], Kimball [1993], Gollier and Pratt [1996] and Eeckhoudt et al. [1996].

many markets, the grade of the object, or professional certification of the seller, is one of the critical variables determining price. Thus one could scarcely design a test of foreground risk in these markets without attending to the background risk. In a VX one can hope to gain some control over this characteristic of risk in the field. VX introduces a natural way to simulate the presence of background risk in ways that are salient.

6.3 Policy Lotteries

Our final example of the use of VR for representing decision tasks under uncertainty is to consider the evaluation of policy uncertainty. To do so we first introduce the intermediate notion of a "policy lottery." This is a representation of the simulated effects of a policy in which the uncertainty of the simulated impact is explicitly presented to the policy maker. Thus when the policy maker decides that one policy option is better than another, the uncertainty in the estimate of the impact has been taken into account. Note that this is uncertainty in the *estimate of the impact*, and not necessarily uncertainty in the *impact itself*. But we submit that in the limited information world of practical policy-making such uncertainties are rife.[29] We illustrate this idea using simulations from a computable general equilibrium model developed for the Danish government to evaluate precisely this type of policy. In the next section we expand this idea to

[29] For example, see Desvousges et al. [1999]. The limitation on information can derive from the inherent difficulty of modeling behavioral or physical relationships, from the short time-frame over which the model has to be developed or applied, or both.

show how it forms a central place in a VX based on a policy scenario.

The specifics of the model are documented in Harrison, Jensen, Lau and Rutherford [2002]. It is basically a static model of the Danish economy calibrated to data from 1992. The version we use has 27 production sectors, each employing intermediate inputs and primary factors to produce output for domestic and overseas consumption. A government agent raises taxes and pays subsidies, and the focus of our policy simulation is on the level of indirect taxes levied by the Danish government. Apart from a representative government household, which consumes goods reflecting public expenditure patterns in 1992, the model differentiates several private household types. The model is calibrated to a wide array of empirical and *a priori* estimates of elasticities of substitution using nested constant elasticity of substitution specifications for production and utility functions. More elaborate versions of the model exist in which inter-temporal and inter-generational behavior are modeled (e.g., Lau [2000]), but this static version is ideal for our illustrative purposes.

The model represents several different private households, based on the breakdown provided by *Statistics Denmark* from the national household expenditure survey. For our purposes, these households are differentiated by family type into 7 households: singles younger than 45 without children, singles older than 45 without children, households younger than 45 without children, households older than 45 without children, singles with children, households with children and where the oldest child is 6 or under, and households with children and where the oldest child is between 7 and 17. The model generates the welfare impact on each of these households measured

in terms of the equivalent variation in annual income for that household. That is, it calculates the amount of income the household would deem to be equivalent to the policy change, which entails changes in factor prices, commodity prices and expenditure patterns. Thus the policy impact is measured as some number of Danish kroner, which represents the welfare gain to the household in income terms.

This welfare gain can be viewed directly as the "prize" in a policy lottery. Since there is some uncertainty about the many parameters used to calibrate realistic simulation models of this kind, there is some uncertainty about the calculation of the welfare impact. If we perturb one or more of the elasticities, for example, the welfare gain might well be above or below the baseline computation. Using randomized factorial designs for such sensitivity analyses, we can undertake a large number of these perturbations and assign a probability weight to each one (Harrison and Vinod [1992]). Each simulation involves a random draw for each elasticity, but where the value drawn reflects estimates of the empirical distribution of the elasticity.[30] For the purposes of this illustration, we undertake 1,000 simulations with randomly generated elasticity perturbations, so it is as if the household faces a policy lottery consisting of 1,000 distinct prizes that occur with equal probability 0.001. The prizes, again, are the welfare gains that the model solves for in each such simulation.

---

[30] For example, if the empirical distribution of the elasticity of substitution is specified to be normal with mean 1.3 and standard deviation 0.4, 95% of the random draws will be within $\pm 1.96 \times 0.4$ of the mean. Thus one would rarely see this elasticity take on values greater than 3 or 4 in the course of these random draws.

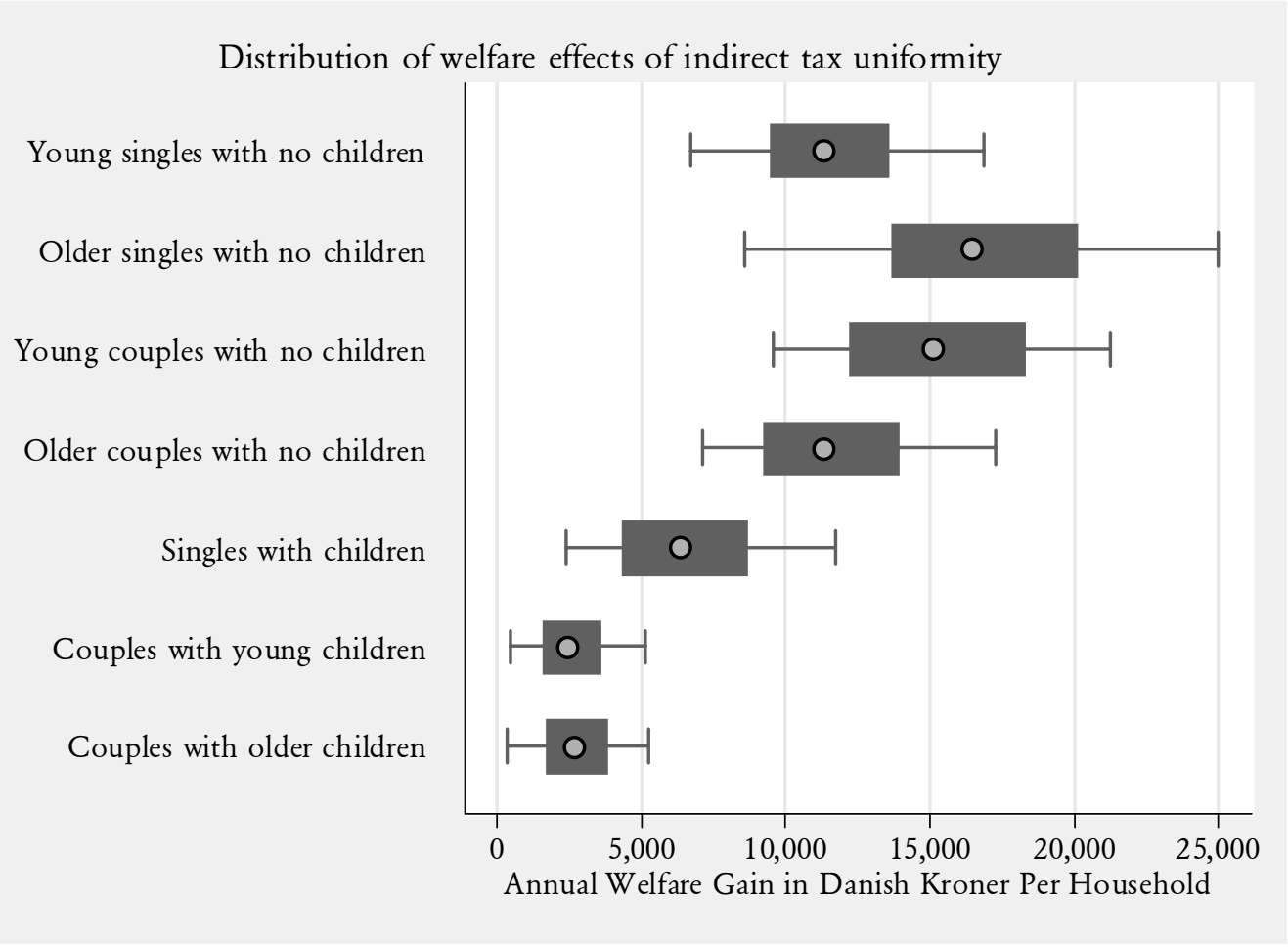


Figure 18: An Illustrative Policy Lottery

Figure 18 illustrates the type of policy lottery that can arise. In this case we consider a policy of making all indirect taxes in Denmark uniform, and at a uniform value that just maintains the real value of government expenditure. Thus we solve for a revenue-neutral reform in which the indirect tax distortions arising from inter-sectoral variation in those taxes are reduced to zero. Each box in Figure 18 represents 1,000 welfare evaluations of the model for each household type. The large dot is the median welfare impact, the rectangle is the interquartile range,[31] and the whiskers represent the range of observed values. Thus we see that the policy represents a lottery for each

[31] Defined by the 25th and 75th percentiles, this range represents 50% of the observations around the median.

household, with some uncertainty about the impacts.

The general point is that we have a tool to use in the design of a wide range of VX applications. If one can come up with a simulation model that represents the policy effects, or for that matter a set of such models, then it is a relatively easy matter to generate distributions of policy impacts such as our Figure 18. This just recognizes that our policy models are never certain, and that they contain standard errors, and in fact lots of standard errors. But that uncertainty should not be ignored when the policy maker uses the model to decide on good policies.

This idea is a simple one, and well known in the older simulation literature in computable general equilibrium modeling.[32] The methods developed to address it amounted to Monte Carlo analysis on repeated simulations in which each uncertain parameter was perturbed around its point estimate. By constraining these perturbations to within some empirical or *a priori* confidence region, one implicitly constrained the simulated policy outcome to that region. The same idea plays a central role in the *Stern Review on the Economics of Climate Change* (Stern [2006]). It stresses (p.153) the need to have a simulation model of the economic effects of climate change that can show stochastic impacts. In fact, any of the standard simulation models can easily be set up to do that, by simply undertaking a systematic sensitivity analysis of their results. The *Review* then proposes an "expected utility analysis" of the costs of climate change (p.161) which is effectively the same as viewing climate change impacts as a lottery. When one then considers alternative policies to

[32] See Harrison et al. [1992][1993].

mitigate the risk of climate change, the "expected utility analysis" is the same as our policy lottery concept.

To the best of his knowledge, choosing one policy over another is the same thing as if the policy maker is choosing one distribution of impacts over another. We next show how this idea can be used in a specific application, but it is quite general.

7. A Case Study: Evaluating the Risks of Wildfires

The policy context we consider, to illustrate the application of virtual experiments, is the assessment of how ordinary households and trained experts evaluate the risks and consequences of forest fire. Fires can occur naturally, such as from lightning, or they can occur in a controlled burn. The policy of the U.S. Forest Service has been to undertake controlled burns as a way of reducing the "fuel" that allows uncontrolled fires to become dangerous and difficult to contain. Additionally, many private citizens and local communities undertake such controlled burns for the same purpose. The benefit of a controlled burn is the significant reduction in the risk of a catastrophic fire; the costs are the annoyance that smoke causes to the local population, the scarring of the immediate forest environment for several years, and the low risk of the burn becoming uncontrolled. So the policy decision to proceed with a controlled burn is one that involves balancing uncertain benefits against uncertain costs. These decisions are often made by individuals and groups with varying levels of expertise and input, often pitting political, scientific, aesthetic and moral views against

each other.

Decisions about controlled burns are not, however, typically made directly by residents. Instead, there are teams of experts from various disciplines and government agencies involved in these decisions. To be sure, they are often charged with listening to residents and taking their concerns into account. Ultimately, however, residents exert significant influence on the policy outcome through the political process. The use of controlled burns as a forest management tool is extremely controversial politically, and for quite understandable reasons. Many homeowners oppose controlled burns if they are too close to their residences, since it involves some risk of unplanned damage and ruins the aesthetics of the area for several years. Conversely, other residents want such burns to occur because of the longer-term benefits of increased safety from catastrophic loss, which they are willing to put against the small chance of localized loss if the burn becomes uncontrolled. And there are choices to be made about precisely where the burns should go, which can have obvious consequences for residents and wildlife. Thus, the tradeoffs underlying household decisions on whether to support controlled burns involve their risk preferences, the way they trade-off short-term costs with long-term gains, their perception of the risks and aesthetic consequences of the burns, and their view of the moral consequences of animal life lost in human-controlled burns versus the potentially larger losses in uncontrolled fires.

So the policy context of this case study has most of the characteristics that one finds in other environmental policy settings. The standard economic model of decision making, EUT,

specifies that decisions are determined by the subjective valuation of each possible outcome, perceptions of the value of outcomes over time as well as their subjective discounted values, perceptions of the extent and distribution of the risk, and subjective risk preferences. A key question is how the experts' decisions differ from those that residents make and what influences these differences. Is it due to differences in their risk preferences, their willingness to trade-off short-term costs (to others) and longer-term benefits (to others), or their perception of the risks and aesthetic consequences of the burns? Or is it due to the extensive experience that the experts may have in forest management, making it perhaps easier for them to conceive the range of consequences and the risks involved in various actions? Most likely it is some combination of these factors. It may also be influenced by how they "pool" all of the information: even if they have the same information and beliefs as residents, they might combine these differently to arrive at a different decision. If we can identify the individual components of decisions in this context, as well as observe the final decision, we can infer how much of the final decision is due to better use of data and how much is due to better knowledge about combining the data correctly.

7.1 Representation as a Policy Lottery

Forest fire management options can be viewed as choices with uncertain outcomes. In the language of economists these options are "lotteries" or "prospects" that represent a finite range of final outcomes, each with some probability, and hence can be viewed as a policy lottery in the

sense described earlier. A final outcome in this instance might be "my house did not burn down in the 5 years after I made my choice" or "my house burned down in the 5 years after I made my choice." Many more outcomes are possible, and we just pick two extreme ones to illustrate concepts. The probability that each outcome has can be subjective. Even if the probability is objective, in the sense that the process generating the outcome is known (e.g., a roll of a fair die), it is possible that the individual uses "decision weights" that transform the probability in some manner. The use of decision weights has a long history in cognitive psychology and economics.

In general we therefore expect risk preferences to affect the fire management option preferred. Someone may be neutral to risk, or want to avoid risk. Some might even be risk loving. It is not possible to say what is a good choice or a bad choice unless one knows the risk preferences of the decision-maker. What might look like an irrational choice to a risk neutral person could be a rational choice to a risk averse person, and *vice versa*. Moreover, risk preferences depend on the way in which the individual perceives and uses all of the information about the lottery: understanding of the consequences of the final outcome as well as the way in which the probability of each outcome is incorporated into the evaluation of lotteries.

The canonical choice task that economists use to identify risk preferences in a setting like this is to present individuals with simple binary choices. One choice option might be a "safe" lottery that offers outcomes that are close to each other but different: for example, $16 and $20. Another choice option would then be a "risky" lottery that offers outcomes that are more extreme:

for example, $1 or $38.50. If each outcome in each lottery has a ½ chance of occurring, and the decision-maker is to be paid off from a single realization of the lottery chosen, then one cannot say which is the best choice without knowing the risk preferences of the decision maker. In this example a risk neutral individual would pick the risky lottery, since the expected value is $19.75, compared to the expected value of $18 for the safe lottery. So someone observed picking the safe lottery might just be averse to risk, such that the expected increment in value of $1.75 from picking the risky lottery is not enough to compensate for the chance of having such a wide range of possible outcomes. A risk averse person is not averse to having $38.50 over $1, but just to the uncertain prospect of having one or the other prior to the uncertainty being realized.

Armed with a sufficient number of these canonical choices tasks, where the final outcomes are monetary payments and the probabilities are objective, economists can estimate the parameters of alternative latent choice models. These models generally rely on parametric functional forms for key components, but there is now a wide range of those functional forms and a fair understanding of how they affect inferences.[33] Thus one can observe choices and determine if the individual is behaving as if risk neutral, risk averse, or risk loving in this situation. And one can further identify what component of the decision making process is driving these risk preferences: the manner in which final outcomes are valued by the individual, and/or the manner in which probabilities are transformed into decision weights. There remains some controversy in this area, but the alternative

[33] For example, see Holt and Laury [2002] and Harrison et al. [2005][2007].

approaches are now well established.

Our research may be viewed as considering the manner in which the lottery is presented to subjects, and hence whether that has any effect on the way in which they, in turn, perceive and evaluate it. Outcomes in naturally occurring environments are rarely as crisp as "you get $20 today." And the probability of any one outcome is rarely as simple and discrete as ½, ⅓, ¼ or ⅛, but usually take the form of a probability density function which the individual has to represent in some manner. Following Harrison and List [2004] and Harrison, List and Towe [2007], we may think of the traditional lottery choice task as artefactual, in the sense that it uses artificial but precise counterparts to the outcomes and probabilities in a naturally occurring environment.

One difficulty with jumping to more natural counterparts to prizes or probabilities is that one loses control over the underlying stimuli. This makes it hard, or impossible, to identify and estimate latent choice models. The advantage of VX is that one can smoothly go where no lab experiment has gone before: the same virtual stimuli can be set to be artefactual and crisp, and then steadily one can make the stimuli more and more natural in affect and comprehension. For example, if the probability distribution is actually a smooth density defined over a continuous set of possible outcomes, one can "discretize" it and present that to the subjects. By varying the coarseness of the discrete representation one can see how that affects behavior.

Figure 19: *SpeedTree* rendering from *The Elder Scrolls IV: Oblivion*

7.2 Generating the Probabilities

One of the things that is immediately apparent when working with VR, but unexpected to outsiders, is the "R" part. Mimicking nature is not a matter left solely to the imagination. When a subject is viewing some scene, it is quite easy to generate images in a manner that appears unusual and artefactual. That can occasionally work to one's advantage, as any artist knows, but it can also mean that one has to work hard to ground the simulation in certain ways. At one affectual level there might be a concern with "photo-realism," the extent to which the images look as if they were generated by photographs of the naturally occurring environment, or whether there is rustling of trees in the wind if one is rendering a forest.

We utilize software components that undertake some of these tasks for us. In the case of trees and forests, a critical part of our case study, we presently use *SpeedTree* to undertake most of

the rendering (http://www.speedtree.com/). This package requires that one specify the tree species (including miscellaneous weeds), the dimensions of the tree, the density of the trees in a forest, the contour structure of the forest terrain, the prevailing wind velocity, the season, and a number of other characteristics of the forest environment. We discuss how that is done in a moment. This package is widely used in many games, including *Tiger Woods PGA Tour*, *Call of Duty 3*, and of course numerous gothic games. The forest scene in Figure 19 illustrates the capabilities we needed for a walk-through, and is from the game *The Elder Scrolls IV: Oblivion*. The image that is rendered allows for real-time updating as the subject moves through the forest, and handles light and shade particularly well from this walking perspective. As the subject gets closer to a particular tree, it renders that in greater detail as needed. Figure 19 is a professional product, exploiting many of the detailed options of the rendering software, but illustrates the type of images that are possible.

One can also take a helicopter tour of a forest. Figures 20 and 21 were each rendered by *SpeedTree* for the same forest, first from a distance, and then close-up.[34] This is "just the trees" and nothing else, to illustrate the pure effect of rendering the forest. Figure 20 employs a 2-mile visibility horizon and displays roughly 45,000 trees, but is based on a virtual forest covering 230 square miles and roughly 4 million trees. So the helicopter could travel some distance without falling off the end of this virtual earth. Figure 21 is a close-up of the same forest, again *sans* soil and weeds. The point is that this is the same forest, just seen from two different perspectives, and the "unseen forest"

[34] This is not the same forest as in Figure 19.

would be rendered in a consistent manner if the user chose to go there.

This last point is not trivial. There is one underlying computer representation of the entire forest, and then the "window" that the subject looks at is rendered as needed. In the case of the Figures 20 and 21, the rest of the forest is rather boring and repetitive unless you are a squirrel. But one can add numerous variations in vegetation, topography, landmarks, and so forth. To do this one links a separate simulation software package to the graphical rendering software that displays the evolution of a landscape. Since we are interested in forest fires, and how they spread, we can use one of several simulation programs used by professional fire managers. There are actually many that are available, for different fire management purposes. We wanted to be able to track the path of a wildfire in real time, given GIS inputs on vegetation cover, topography, weather conditions, points of ignition, and so on. Thus, the model has to be able to keep track of the factors causing fires to

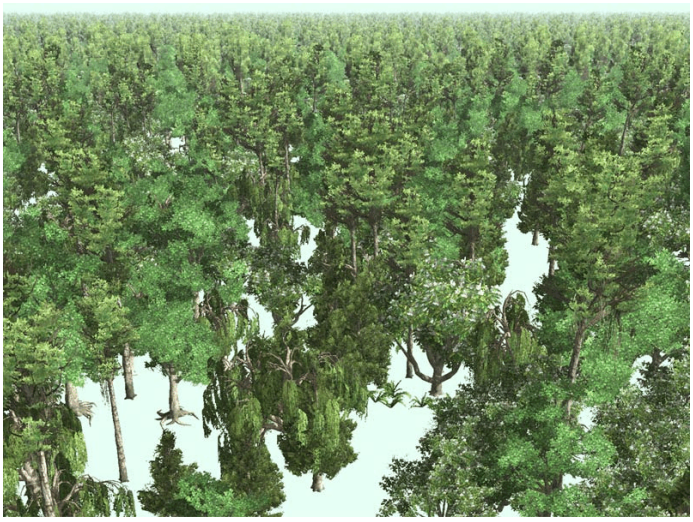


Figure 20: *SpeedTree* rendering of a forest from a distance

Figure 21: *SpeedTree* rendering of the same forest up close.

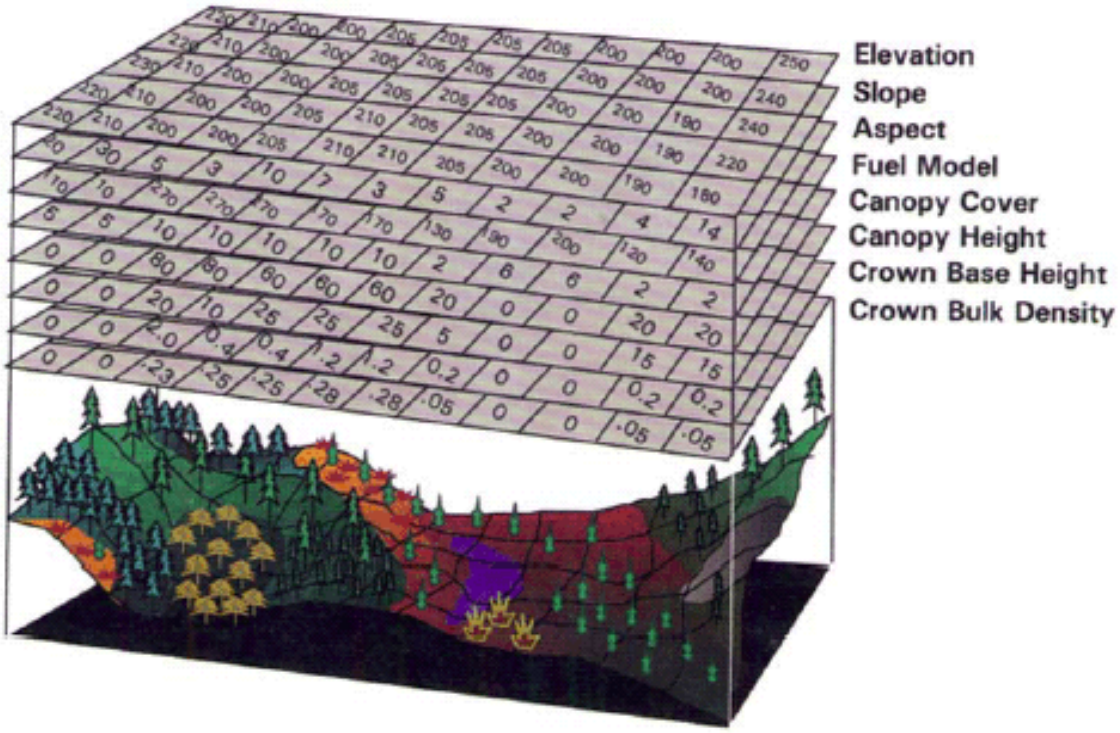


Figure 22: GIS Input Layers for *FARSITE* Simulation of Path of a Fire

accelerate, such as variations in slope, wind or vegetation, as well as differences between surface fires, crown fires, and the effects of likely "spotting" (when small blobs of fire jump discretely).

For our purposes the *FARSITE* software due to Finney [1998] is ideal. It imports GIS information for a given area and then tracks the real-time spread of a fire ignited at some coordinate, say by lightning or arson. Figure 22 shows the GIS layers needed to define a *FARSITE* simulation. Some of these, such as elevation and slope, are readily obtained for most areas of the United States. The "fuel model" and vegetation information (the last four layers) are the most difficult to specify. Vegetation information can be obtained for some sites, although it should

ideally be calibrated for changes since the original survey.[35] The fuel model defines more precisely the type of vegetation that is present at the location. The most recent tabulation of 40 fuel models is documented by Scott and Burgan [2005]. One formal type of vegetation is water, for example, which does not burn. Two of the other fuel models are shown in Figures 23 and 24: some photographs help one identify what is described, and then the fuel model implies certain characteristics about the speed with which a fire burns.

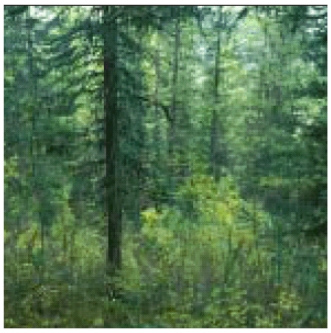
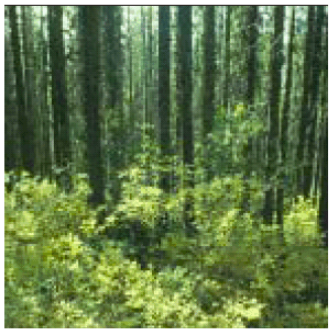
The output from a *FARSITE* simulation is illustrated in Figure 25. The fire is assumed to be ignited at the point indicated, and the path of the fire as of a certain time period is shown. The output from this simulation consists of a series of snapshots at points in time defining what the location and intensity of the fire is. By modifying *FARSITE* to output its state at a sufficiently fine granularity, we can take that output and use it for detailed visual rendering of the fire as it burns.

We can then take that output and render images of the forest as it burns. Using *SpeedTree* as the basic rendering module, we have adapted its software for our needs. This has included the incorporation of other landscape components, smoke, fire and the effects that fire has on the appearance of the forest, both in altering illumination and causing damage. The images in Figure 26

[35] This is an important feature of the data underlying the *Florida Fire Risk Assessment System* of the Division of Forestry of the Florida Department of Agriculture & Consumer Affairs, documented at http://www.fl-dof.com/wildfire/wf_fras.html.

TU5 (165)

Very High Load, Dry Climate Timber-Shrub

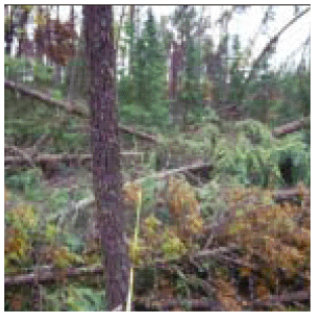



Description: The primary carrier of fire in TU5 is heavy forest litter with a shrub or small tree understory. Spread rate is moderate; flame length moderate.

Figure 23: Illustrative fuel load model from Scott and Burgan [2005]

SB4 (204)

High Load Blowdown



Description: The primary carrier of fire in SB4 is heavy blowdown fuel. Blowdown is total, fuelbed not compacted, most foliage and fine fuel still attached to blowdown. Spread rate very high; flame length very high.

Figure 24: Illustrative fuel load model from Scott and Burgan [2005]

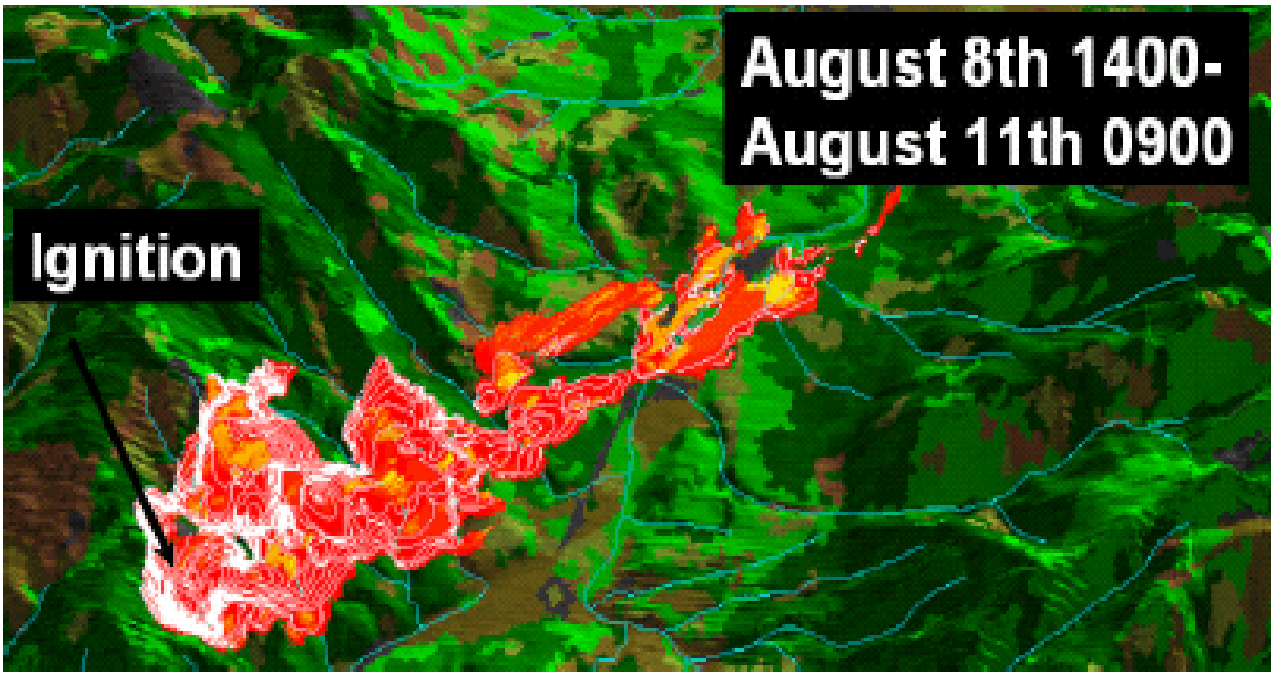


Figure 25: Illustrative *FARSITE* Simulation

display the evolution of a wildfire, initially from a distance, then close up, and finally after the front of the fire has passed. These images have been generated using *SpeedTree* and our own programming, based on the simulated path of the fire from *FARSITE*, conditional on the landscape. Thus we have simulated a GIS-consistent virtual forest, with a fire spreading through it in a manner that is consistent with one of the best models of fire dynamics that is widely used by fire management professionals, and then rendered it in a naturalistic manner. The subject is free to view the fire from any perspective, and we can track that; or we can force feed a series of perspectives chosen beforehand, to study the effect of allowing endogenous information accumulation.

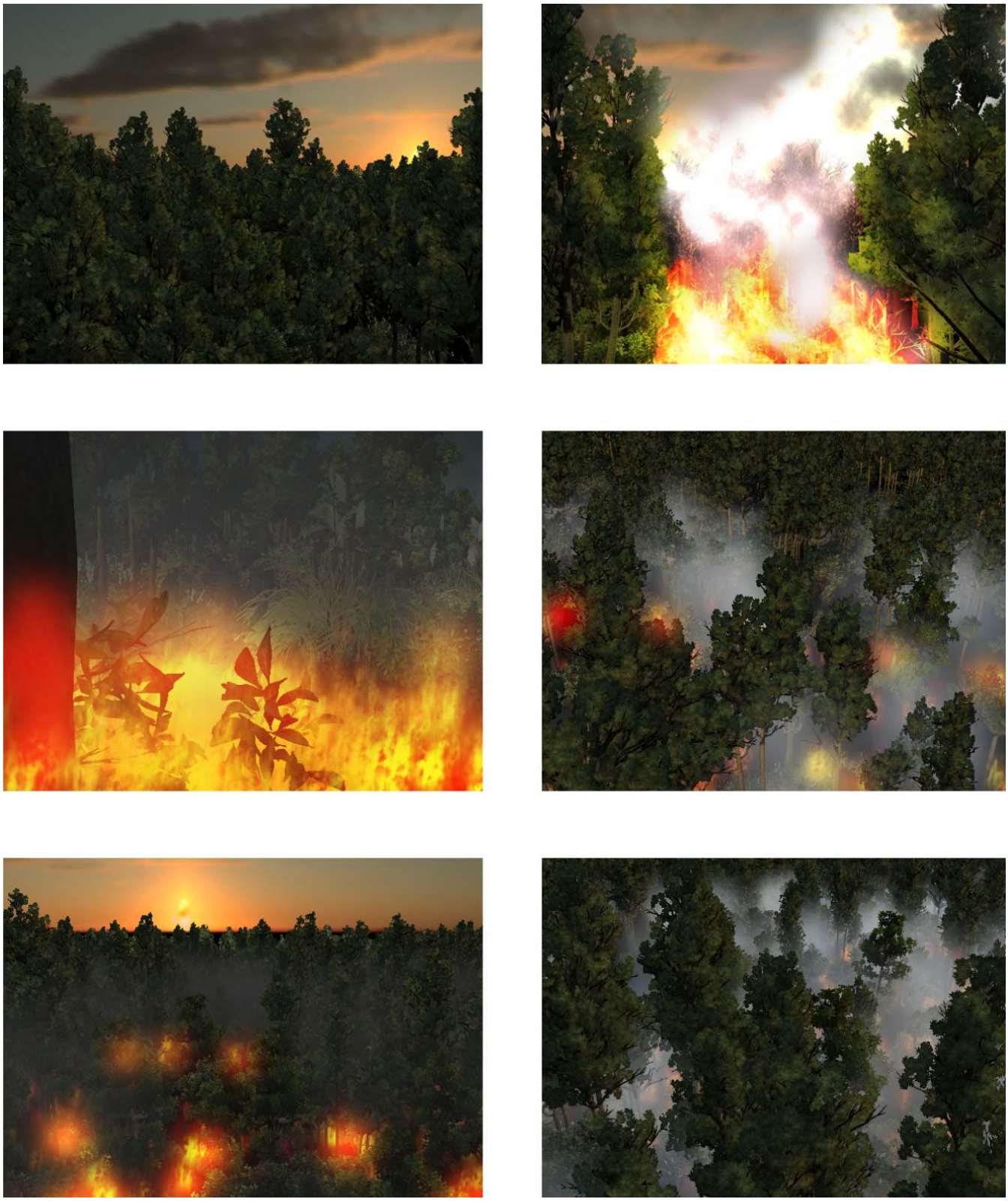


Figure 26: Authors' Rendering of Simulated Forest Fire

The static images in Figure 26 are actually snapshots of an animation which plays out in real time for the subject. This dynamic feature of the evolution of the fire is one reason that a model such as *FARSITE* is needed, since it computes the path and intensity of the fire, given the conditions of the environment. The dynamic feature also adds to the experience that the subject has, in ways that can again be studied orthogonally in a VX design. In other words, it is a relatively simple matter to study the effects of presenting static images versus a dynamic animation on behavioral responses.

7.3 The Virtual Experimental Task

We now have all of the building blocks to put together the task that will be given to subjects in our VX. The logic of the instrument we have developed can be reviewed to see how the building blocks fit together. We focus on our value added from a methodological perspective, building on existing contingent valuation surveys of closely related matters due to Loomis et al. [2002][2005].

The initial text of the instrument is standard, in the sense of explaining the environmental and policy context, the design of forest fire management in Florida:

> Florida typically has about 5000 wild fires each year, with the most active period between March and June. During the fires season of 2006 The Division of Forestry of Florida Department of Agriculture and Consumer Services reports over 4,000 wild fires covering over 200,000 acres at the end of July.
>
> Managing and preventing these wild fires is important. Significant economic costs, in terms of damage to property, as well as serious health and safety concerns from fire and

smoke, make this an issue of interest to everyone in the state.

Your views on this topic are very important to Florida fire managers and policy planners as they decide how best to proceed.

The scope of the forest fire management problem is then explained, and the likely range of impacts discussed. Screen-shot images from the VR simulations, such as shown in Figure 26, are included to illustrate fuel conditions and fire intensities. Subjects are also presented with some statistics on the economic damages from the well-known 1998 fires in Florida:

To give you a perspective on the economic damages caused by wildfire, here are some estimates from the 1998 fires in Florida. 340 homes, 33 businesses, and several cars and boats were destroyed, totaling $10-$20 million, with homes accounting for the majority of these costs. The average cost per house implied is $59,000. Over 500,000 acres of timber forest burned in Florida that year. Net timber damages are estimated at $300 - $500 million, implying a per acre loss of $600-$1,000. These are the two most significant sources of economic cost. In addition, there was increased health expenditures associated with increases in asthma incidences, particularly among children. The estimated total health costs for asthma was $210,000 - $650,000. Another significant loss in Florida was the lost tourism income estimated at $138 million.

The total economic cost estimated for the 1998 wild fires was $448 million for the 500,000 acres burnt. An average fire year, that burns about 200,000 acres, therefore costs about $180 million.

The idea of VR computer simulation is introduced, with explanations that the predicted path depends on things such as topography, weather, vegetation, and the presence of natural and man-made fire breaks. Participants are made familiar with the idea that fires and fire damages are stochastic and can be described through frequency distributions. The distributions that are presented to them are generated through Monte Carlo simulations using the *FARSITE* model, as described earlier. The subjects then experience two dynamic VR simulations, rendered from trhe

information supplied by a *FARSITE* simulation over a 30 year period. We selected these simulations to represent the upper and lower deciles of the distributions of fire damage, and the subjects are told this.

After the VR experience of this wild fire simulation, we discuss prescribed burning as a fire management tool that can reduce the frequency and severity of fires. The policy which they are asked to evaluate is an expansion of prescribed burns from 4% to 6% of the total forest area in Central Florida. It is described in general terms as follows:

Recently, prescribed burning has been used to prevent wild fires from becoming intensive and spreading quickly. Prescribed burning involves having fire professionals periodically set fires to clear the forest floor of the excess brush, dead branches and pine needles. These prescribed fires are easier to manage than wild fires since prescribed fires do not burn as intensely and they can be directed away from structures.

While prescribed fires do result in an increase in air pollution, they generally produce far less air pollution than would be expected from a wild fire on the same acreage. By timing prescribed fires with favorable weather and wind conditions, smoke can be directed away from the majority of the population.

Prescribed burns occur during winter, and therefore precede the wild fire season, which starts in March each year. Because the prescribed burns take place in the winter, when the vegetation is moist, the risk of an uncontrolled spread is extremely low.

The specific policy is described as:

Prescribed burn cost roughly $25 per acre. An average of almost a half million acres in Florida are already treated with prescribed burn each year whereof 13,000 are in Volusia county. This corresponds to about 4% of the total forest area. (This costs approximately $12.5 million statewide or $325,000 in Volusia County)

In this study we are investigating a proposed expansion of the prescribed burn program that increases the treated acreage from 4 to 6% of the total forest area, from half a million to three quarter million acres. This is expected to cost an additional $6.25 million per year.

To put the cost in perspective: The Florida state budget during 2005-2006 was $64.7 billion. The additional cost is about $0.35 per capita in Florida, or $1.40 for a typical four person household per year. For the 2006-7 budget year $80 million has been allocated to beach restoration after increased hurricane damages.

The subjects then get to experience two further dynamic VR simulations, rendered from the output of a *FARSITE* simulation over a 30 year period. These two simulations represent the upper and lower deciles of fire damages under the same conditions as the baseline simulations, but with the expanded burn program in place. Subjects are again shown the full distributions of damages from Monte Carlo simulations of this scenario.

Participants are then introduced to the choice problem they face. In order to make the choice salient, we pay them according to damages to one particular area in the region being simulated. They are assigned this region as "their property":

In the simulations we will designate one property as yours. This property has one house on it, valued at $59,000, and it also has 10 acres of timber forest, valued at $10,000.

If there is a fire within a 20 mile radius of your property you will also incur respiratory health costs of $1,900 for each member of the family. We will assume that you have a family of four, so the total health related cost is $7,600. Thus, the most you stand to lose in a fire is $76,600. These are very rough numbers, of course, but they are on the order of magnitudes that people actually lose during wild fires.

To make it interesting and salient to you, we are going to pay you for your participation in the study. Your earnings will depend on the outcome of one simulation that we will run at the end of the session.

At the start of this simulation you are credited with $X. Then you are asked to make a choice of fire management policy. You may choose one of these policies:

1. Present prescribed burn program
2. Expanded prescribed burn program

We are going to simulate 30 years of your chosen program on the computer that we have brought along. During this time period wild fires may get ignited in a realistic manner in any or all of the fire seasons. Whether or not your designated property is damaged depend on numerous factors such as a location of the ignition, wind factors, and moisture levels. We will simply deduct all of the damages to your designated property from your initial credit line of $X, and you will be paid the residual. The less wild fire damage there is to your property the more you will get paid.

If you choose not to expand the program there will be no additional costs to you. We will only deduct the wild fire damages from your initial credit. It is expected, though, that there will be a higher risk for more extensive damages.

The expanded prescribed burning program involves upfront additional cost to you. If you choose this program you have to make the payment out of your initial credit. It will decrease the risk of property damage or loss in timber or any health cost you incur in the event of a wild fire.

If one or more wild fires ignite during the 30 year simulation, the cost of the fires will depend on the intensity of the fire as well as spread.

Before the subject makes a choice they are also shown distributions of the damages to their designated property during the two Monte Carlo simulations discussed above, one for the case where no additional fire management has been introduced, and one where it has. They are then asked to make their choice, where X denotes values to be supplied parametrically:

Recall that you will be paid according to the damages to your designated property only. You are given an initial credit of $X, and we will pay you amount of this remaining after deducting wild fire damages and any costs for the expanded burn program. Each time your house burns down we will deduct $X. Each time part or all of your forest burns down we deduct $X per acre burnt. Each time smoke spreads into your property from nearby fires we deduct $X for health expenses.

You now get to make a choice between paying $X towards the expansion of prescribed burning program or maintaining the present fire prevention program at no additional cost to yourself.

Which do you choose? Please circle the number corresponding to your choice.

1. Present prescribed burn program at no upfront cost

2. Expanded prescribed burn program at an upfront cost of $X

The purpose of our experiment is to compare the responses in the VX to those participants make to those they make in more traditional survey instruments, such as used in a Contingent Valuation (CV) analysis. We therefore include a control treatment that is a standard CV instrument, where the consequences of their choices are purely hypothetical and simulation models play no role at all. In fact, this CV instrument is close to the one actually used by Loomis et al. [2002][2005] in Florida to study the same policy issue. We then separately test for the effect of learning about the fire damages through training simulations and from introducing monetary consequences by generating a VR simulation after their choice is made. We recruit participants from residential communities and from fire management experts in Central Florida and compare their responses to test whether the use of VX facilitates a convergence of responses.

8. Conclusions

There is one major assumption in all of the work estimating frontier models in econometrics: that one or more firms is actually *on* the frontier. In effect those methods calculate the production possibility frontier as an envelope of the observed input-output data from firms that actually exist, recognizing the special nature of truncated error terms with respect to individual firm efficiency (viz., that no firm, by definition, can be more than 100% efficient). So if one is truly to define the frontier, given existing inputs and technology, one has to consider counter-factual firms

and ask if they are feasible. We have done this with respect to the concept of virtual experiments, by speculating on the feasibility of a range of applications which do not exist so far. In this manner we hope to identify the likely frontier.

We believe that we have restricted attention to counter-factual applications that are feasible, and some have been completed. But time will tell in terms of the importance of virtual experiments for environmental policy. We believe that the concept of virtual experiments is on the frontier because it brings together core ideas from three different areas of research, does so in a manner that makes the whole more than just the sum of the parts, and can be used to address some major environmental policy issues. Our perspective is that of the environmental economist who wants the control of the laboratory with the realism and naturalistic cues of the field. There is something of a tension in the area now, with many of the apparent lessons from the laboratory being questioned because they are viewed as being generated in an abstract, artefactual setting. This tension leads to a derived demand for field experiments, but that methodology has its own problems of replicability and control that limit the range of issues that can be effectively examined. This tension resonates with CV practitioners that need to specify contextually rich scenarios to address the questions of interest to them, but who lack the salience of real rewards and consequences we find in laboratory and field experiments. We need some way to bridge these three islands of intellectual activity in environmental economics.

References

Altmann, Gerry T.M., *The Ascent Of Babel. An Exploration of Language, Mind, and Understanding*, (New York: Oxford University Press, 1997).

Andersen, Steffen; Harrison, Glenn W., and Rutström, E. Elisabet, "Choice Behavior, Asset Integration and Natural Reference Points," *Working Paper 06-04*, Department of Economics, College of Business Administration, University of Central Florida, 2006.

Aono, Masaki, and Kunii, Tosiyasu L., "Botanical tree image generation," *IEEE Computer Graphics and Applications*, 4(5), 1984, 10-34.

Arrow, Kenneth; Solow, Robert; Portney, Paul; Leamer, Edward E.; Radner, Roy; and Schuman, Howard, "Report of the NOAA Panel on Contingent Valuation," *Federal Register*, 58(10), January 15, 1993, 4602-4614.

Banaji, Mahzarin R., and Crowder, Robert G., "The Bankruptcy of Everyday Memory," *American Pyschologist*, 1989, 44, 1185-1193.

Bateman, Ian J.; Jones, Andrew P.; Jude, Simon, and Day, Brett H., "Reducing Gain/Loss Asymmetry: A Virtual Reality Choice Experiment (VRCE) Valuing Land Use Change," *Unpublished Manuscript*, School of Environmental Sciemces, University of East Anglia, December 2006.

Battalio, Raymond C.; Kagel, John C., and Jiranyakul, K., "Testing Between Alternative Models of Choice Under Uncertainty: Some Initial Results," *Journal of Risk and Uncertainty*, 3, 1990, 25-50.

Beattie, Jane, and Loomes, Graham, "The Impact of Incentives Upon Risk Choice Experiments," *Journal of Risk and Uncertainty*, 14, 1997, 155-168.

Becker, Gordon M.; DeGroot, Morris H., and Marschak, Jacob., "Measuring Utility By A Single-Response Sequential Method," *Behavioral Science*, 9, July 1964, 226-232.

Benjamin, Daniel K.; Dougan, William R.; and Buschena, David, "Individuals' Estimates of the Risks of Death: Part II - New Evidence," *Journal of Risk and Uncertainty*, 19, 2001, 35-57.

Benjamin, Daniel K., and Dougan, William R., "Individuals' Estimates of the Risks of Death: Part I - A Reassessment of the Previous Evidence," *Journal of Risk and Uncertainty*, 15, 1997, 115-133.

Bin, Okmyung; Crawford, Thomas W.; Kruse, Jamie B, and Landry, Craig E., "Viewscapes and Flood Hazard: Coastal Housing Market Responses to Amenities and Risk," *Working Paper,* Center for Natural Hazards Research, East Carolina University, November 2006.

Binswanger, Hans P., "Attitudes Toward Risk: Theoretical Implications of an Experiment in Rural India," *Economic Journal*, 91, December 1981, 867-890.

Binswanger, Hans P., "Attitudes Toward Risk: Experimental Measurement in Rural India," *American Journal of Agricultural Economics*, 62, August 1980, 395-407.

Bishop, Ian D., "Predicting Movement Choices in Virtual Environments," *Landscape and Urban Planning,* 56, 2001, 97-106.

Bishop, Ian D.; Wherrett JoAnna R., and Miller, David R., "Assessment of Path Choices on a Country Walk using a Virtual Environment," *Landscape and Urban Planning,* 52, 2001, 225-237.

Bishop, Ian D., and Gimblett, H.R., "Management of Recreational Areas: Geographic Information Systems, Autonomous Agents and Virtual Reality," *Environment and Planning B*, 27, 2000, 423-435.

Blackburn, McKinley; Harrison, Glenn W., and Rutström, E. Elisabet, "Statistical Bias Functions and Informative Hypothetical Surveys," *American Journal of Agricultural Economics*, 76(5), December 1994, 1084-1088.

Blackmore, Susan, *The Meme Machine* (New York: Oxford University Press, 1999).

Bloomenthal, Jules, "Modeling the mighty maple," *Proceedings of SIGGRAPH* 85, 1985, 305-311.

Bohm, Peter, "Behavior Under Uncertainty Without Preference Reversal: A Field Experiment," in J. Hey (ed.), *Experimental Economics,* (Heidelberg: Physica-Verlag, 1994).

Bohm, Peter, and Lind, Hans, "Preference Reversal, Real-World Lotteries, and Lottery-Interested Subjects," *Journal of Economic Behavior and Organization*, 22, 1993, 327-348.

Bohm, Peter, "Revealing Demand for an Actual Public Good," *Journal of Public Economics*, 24, 1984a, 135-151.

Bohm, Peter, "Are There Practicable Demand-Revealing Mechanisms?" in H. Hanusch (ed.), *Public Finance and the Quest for Efficiency* (Detroit: Wayne State University Press, 1984b).

Bohm, Peter, "Estimating Willingness to Pay: Why and How?," *Scandinavian Journal of Economics*, LXXXI, 1979, 142-153.

Bohm, Peter, "Estimating the Demand for Public Goods: An Experiment," *European Economic Review*, 3, June 1972, 111-130.

Boyd, Robert, and Richerson, Peter J., *The Origin and Evolution of Cultures* (New York: Oxford University Press, 2005).

Bull, Rebecca; Espy, Kimberly Andrews, and Senn, Theresa E., "A comparison of performance on the Towers of London and Hanoi in young children," *Journal of Child Psychology and Psychiatry*, 45, 2004, 743-754.

Bullock, David S., and Rutström, E. Elisabet, "Policy Making and Rent-Dissipation: An Experimental Test," *Experimental Economics,* 10, 2007 forthcoming.

Calman, Kenneth C., and Royston, Geoff, "Risk Language and Dialects," *British Medical Journal*, 315, 1997, 939-942.

Camerer, Colin, and Ho, Teck, "Experience weighted attraction learning in normal-form games," *Econometrica,* 67, 1999, 827-874.

Camerer, Colin F., "An Experimental Test of Several Generalized Utility Theories," *Journal of Risk and Uncertainty*, 2, 1989, 61-104.

Cameron, A. Colin, and Trivedi, Pravin K., *Microeconometrics: Methods and Applications* (New York: Cambridge University Press, 2005).

Carson, Richard T.; Mitchell, Robert C.; Hanemann, W. Michael; Kopp, Raymond J.; Presser, Stanley; and Ruud, Paul A., "Contingent Valuation and Lost Passive Use: Damages from the *Exxon Valdez*," *Environmental and Resource Economics*, 25, 2003, 257-286.

Carson, Richard T; Hanemann, W. Michael; Krosnick, Jon A.; Mitchell, Robert C.; Presser, Stanley; Ruud, Paul A., Smith, V. Kerry, "Referendum Design and Contingent Valuation: The NOAA Panel's No-Vote Recommendation," *Review of Economics and Statistics*, 80(2), May 1998, 335-338; reprinted with typographical corrections in *Review of Economics and Statistics*, 80(3), August 1998.

Carson, Richard T.; Mitchell, Robert C.; Hanemann, W. Michael; Kopp, Raymond J.; Presser, Stanley; and Ruud, Paul A., *A Contingent Valuation Study of Lost Passive Use Values Resulting From the Exxon Valdez Oil Spill* (Anchorage: Attorney General of the State of Alaska, November 1992).

Castronova, Edward, *Synthetic Worlds: The Business and Culture of Online Games* (Chicago: University of Chicago Press, 2005).

Castronova, Edward, "The Price of Bodies: A Hedonic Pricing Model of Avatar Attributes in a Synthetic World," *Kyklos*, 57(2), 2004, 173-196.

Cherry, Todd L.; Frykblom, Peter, and Shogren, Jason F., "Hardnose the Dictator," *American Economic Review*, 92(4), September 2002, 1218-1221.

Chiba, Norishige; Muraoka, Kazunobu; Doi, Akio, and Hosokawa, Junya, "Rendering of forest scenery using 3D textures," *The Journal of Visualization and Computer Animation* 8, 1997, 191-199.

Clark, Andy, *Being There: Putting Brain, Body, and World Together Again* (Cambridge, MA: MIT Press, 1997).

Cleveland, William S., and McGill, Robert, "Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods," *Journal of the American Statistical Association*, 79(387), September 1984, 531-554.

Cleveland, William S.; Harris, Charles S., and McGill, Robert, "Experiments on Quantitative Judgements of Graphs and Maps," *Bell System Technical Journal*, 62(6), July-August 1983, 1659-1674.

Cleveland, William S.; Harris, Charles S., and McGill, Robert, "Judgements of Circle Sizes on Statistical Maps," *Journal of the American Statistical Association*, 77(379), September 1982, 541-547.

Cohen, Dan, "Computer simulation of biological pattern generation processes," *Nature*, 216, 1967, 246-248.

Conlisk, John, "Three Variants on the Allais Example," *American Economic Review*, 79(3), June 1989, 392-407.

Darken, Rudolph P., and Cevik, Helsin, "Map Usage in Virtual Environments: Orientation Issues," *Proceedings of IEEE Virtual Reality,* 1999, 133-140.

Desvousges, William H.; Johnson, F. Reed, and Banzhaf, H. Spencer, *Environmental Policy Analysis With Limited Information: Principles and Applications of the Transfer Method* (New York: Elgar, 1999).

Deussen, Oliver; Hanrahan, Patrick; Lintermann, Bernd; Mech, Radomir; Pharr, Matt, and Prusinkiewicz,.Przemyslaw, "Realistic modeling and rendering of plant ecosystems," *Proceedings of SIGGRAPH* 98, 1998, 275-286.

Eeckhoudt, Louis; Gollier, Christian, and Schlesinger, Harris, "Changes in Background Risk and Risk-Taking Behavior," *Econometrica*, 64(3), May 1996, 683-689.

Erev, Ido, and Roth, Alvin E., "Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria," *American Economic Review,* 88, 1998, 848-881.

Feiler, Lauren, and Camerer, Colin F., "Code Creation in Endogenous Merger Experiments," *Unpublished Manuscript*, Division of Humanities and Social Sciences, California Institute of Technology, February 2006.

Fidopiastis, Cali M.; Stapleton, Christopher B.; Whiteside, Janet D.; Hughes, Charles E.; Fiore, Stephen M.; Martin, Glenn A.; Rolland, Jannick P., and Smith, Eileen M., "Human Experience Modeler: Context Driven Cognitive Retraining to Facilitate Transfer of Training," *CyberPsychology and Behavior*, 9(2), Mary Ann Liebert Publishers, New Rochelle, NY, 183-187.

Finney, Mark A., *FARSITE: Fire Area Simulator - Model Development and Evaluation*, Research Paper RMRS-RP-4, Rocky Mountain Research Station, Forest Service, United States Department of Agriculture, March 1998.

Gerking, Shelby; de Haan, Menno, and Schulze, William, "The Marginal Value of Job Safety: A Contingent Value Study," *Journal of Risk & Uncertainty*, 1, 1988, 185-199.

Gigerenzer, Gerd; Hertwig, Ralph; Hoffrage, Ulrich, and Sedlmeier, Peter, "Cognitive Illusions Reconsidered," in C. Plott and V. Smith (eds.), *Handbook of Experimental Results* (New York, NY: Elsevier, 2007).

Goldin-Meadow, Susan, and Wagner, Susan M., "How our hands help us learn," *Trends in Cognitive Science*, 9, 2005, 234-241.

Gollier, Christian, and Pratt, John W., "Risk Vulnerability and the Tempering Effect of Background Risk," *Econometrica*, 64(5), September 1996, 1109-1123.

Grether, David M.; Isaac, R. Mark, and Plott, Charles R., *The Allocation of Scarce Resources: Experimental Economics and the Problem of Allocating Airport Slots* (Boulder: Westview Press, 1989).

Grether David M., and Plott, Charles R., "The Effects of Market Practices in Oligopolistic Markets: An Experimental Examination of the Ethyl Case," *Economic Inquiry*, 22, October 1984, 479-507.

Grether, David M.; Isaac, R. Mark, and Plott, Charles R., "The Allocation of Landing Rights by Unanimity among Competitors," *American Economic Review (Papers & Proceedings)*, 71, May 1981, 166-171.

Hammitt, James K., and Graham, John D., "Willingness to Pay for Health Protection: Inadequate Sensitivity to Probability?" *Journal of Risk and Uncertainty*, 18(1), 1999, 33-62.

Harrison, Glenn W., "Hypothetical Bias Over Uncertain Outcomes," in J.A. List (ed)., *Using Experimental Methods in Environmental and Resource Economics* (Northampton, MA: Elgar, 2006).

Harrison, Glenn W.; Jensen, Jesper; Lau, Morten Igel, and Rutherford, Thomas F., "Policy Reform Without Tears," in A. Fossati and W. Weigard (eds.), *Policy Evaluation With Computable General Equilibrium Models* (New York: Routledge, 2002).

Harrison, Glenn W.; Jones, Richard; Kimbell, Larry J., and Wigle, Randall, "How Robust Is Applied General Equilibrium Analysis?" *Journal of Policy Modeling*, 15(1), 1993, 99-115.

Harrison, Glenn W.; Lau, Morten I., and Rutström, E. Elisabet, "Estimating Risk Attitudes in Denmark: A Field Experiment," *Scandinavian Journal of Economics,* forthcoming 2007.

Harrison, Glenn W.; Lau, Morten Igel; Rutström, E. Elisabet, and Sullivan, Melonie B., "Eliciting Risk and Time Preferences Using Field Experiments: Some Methodological Issues," in J. Carpenter, G.W. Harrison and J.A. List (eds.), *Field Experiments in Economics* (Greenwich, CT: JAI Press, Research in Experimental Economics, Volume 10, 2005).

Harrison, Glenn W., and List, John A., "Naturally Occurring Markets and Exogenous Laboratory Experiments: A Case Study of the Winner's Curse," *Working Paper 3-14*, Department of Economics, College of Business Administration, University of Central Florida, 2003.

Harrison, Glenn W., and List, John A., "Field Experiments," *Journal of Economic Literature*, 42(4), December 2004, 1013-1059.

Harrison, Glenn W.; List, John A., and Towe, Charles, "Naturally Occurring Preferences and Exogenous Laboratory Experiments: A Case Study of Risk Aversion,"*Econometrica*, forthcoming 2007.

Harrison, Glenn W.; Rutherford, Thomas F., and Tarr, David, "Quantifying the Uruguay Round," *Economic Journal*, 107, September 1997, 1405-1430.

Harrison, Glenn W., and Rutström, E. Elisabet, "Expected Utility Theory *And* Prospect Theory: One Wedding and A Decent Funeral," *Working Paper 05-18*, Department of Economics, College of Business Administration, University of Central Florida, 2005.

Harrison, Glenn W., and Vinod, H.D., "The Sensitivity Analysis of Applied General Equilibrium Models: Completely Randomized Factorial Sampling Designs," *Review of Economics and Statistics,* 74, May 1992, 357-362.

Hayes, J.R., and Simon, H.A., "Understanding Written Problem Instructions," in L.W. Gregg (ed.), *Knowledge and Cognition* (Hillsdale, NJ: Erlbaum, 1974).

Hertwig, Ralph, and Ortmann, Andreas, "The Cognitive Illusions Controversy: A Methodological Debate in Disguise That Matters To Economists," in R. Zwick and A. Rapoport (eds.), *Experimental Business Research* (Boston, MA: Kluwer, 2004).

Hey, John D., and Orme, Chris, "Investigating Generalizations of Expected Utility Theory Using Experimental Data," *Econometrica*, 62(6), November 1994, 1291-1326.

Holt, Charles A., and Laury, Susan K., "Risk Aversion and Incentive Effects," *American Economic Review*, 92(5), December 2002, 1644-1655.

Honda, Hisao,."Description of the form of trees by the parameters of the tree-like body: effects of the branching angle and the branch length on the shape of the tree-like body," *Journal of Theoretical Biology*, 31, 1971, 331-338.

Hong, James T., and Plott, Charles R., "Rate Filing Policies for Inland Water Transportation: An Experimental Approach," *Bell Journal of Economics*, 13, Spring 1982, 1-19.

Hughes, Charles E., and Stapleton, Christopher B., "The Shared Imagination: Creative Collaboration in Augmented Virtuality," *HCI International 2005*, Las Vegas, NV, July 22-27, 2005.

Hughes, Charles E.; Stapleton, Christopher B.; Hughes, Darin E., and Smith, Eileen M., "Mixed Reality in Education, Entertainment and Training: An Interdisciplinary Approach," *IEEE Computer Graphics and Applications*, 26(6), November/December 2005, 24-30.

Humphrey, Steven J., "Are Event-Splitting Effects Actually Boundary Effects?" *Journal of Risk & Uncertainty*, 22(1), 2001, 79-93.

Humphrey, Steven J., "More Mixed Results on Boundary Effects," *Economics Letters*, 61, 1998, 79-84.

Jude, Simon; Jones, Andrew P.; Andrews, J.E., and Bateman, Ian J., "Visualisation for Participatory Coastal Zone Management: A Case Study of the Norfolk Coast, England," *Journal of Coastal Research*, 22(6), 2006, 1527-1538.

Kagel, John H.; MacDonald, Don N., and Battalio, Raymond C., "Tests of 'Fanning Out' of Indifference Curves: Results from Animal and Human Experiments," *American Economic Review*, 80(4), September 1990, 912-921.

Kent, Sherman, "Words of Estimated Probability," *Studies in Intelligence*, 8, 1964, 49-65.

Kimball, Miles S., "Standard Risk Aversion," *Econometrica*, 61(3), May 1993, 589-611.

Klein, Gary, *Sources of Power: How People Make Decisions* (Cambridge, MA: MIT Press, 1998).

Krosnick, Jon A.; Holbrook, Allyson L.; Berent, Matthew K.; Carson, Richard T.; Hanemann, W. Michael; Kopp, Raymond J.; Mitchell, Robert C.; Presser, Stanley; Ruud, Paul A.; Smith, V. Kerry; Moody, Wendy R.; Green, Melanie C., and Conaway, Michael, "The Impact of 'No Opinion' Response

Options on Data Quality: Non-Attitude Reduction or an Invitation to Satisfice?" *Public Opinion Quarterly*, 66, 2002, 371-403.

Krupnick, Alan; Alberini, Anna; Cropper, Maureen, and Simon, Nathalie, "Mortality Risk Valuation for Environmental Policy," *Discussion Paper 99-47*, Resources For the Future, 1999; available as HTTP://WWW.RFF.ORG/CFDOCS/DISC_PAPERS/PDF_FILES/9947.PDF.

Lau, Morten Igel, "Assessing Tax Reforms When Human Capital is Endogenous," in G. W. Harrison, S. E. H. Jensen, L. H. Pedersen and T. F. Rutherford (eds.), *Using Dynamic General Equilibrium Models for Policy Analysis* (Amsterdam: North Holland, Contributions to Economic Analysis 248, 2000).

Lindenmayer, Astrid, "Mathematical models for cellular interactions in development, I & II," *Journal of Theoretical Biology*, 18, 1968, 280-315.

List, John A., "Neoclassical Theory Versus Prospect Theory: Evidence from the Marketplace," *Econometrica*, 72(2), 2004, 615-625.

Loomis, John B.; Bair, Lucas S., and Gonzáles-Cabán, Armando, "Language-Related Differences in a Contingent Valuation Study: English Versus Spanish," *American Journal of Agricultural Economics*, 84(4), November 2002, 1091-1102.

Loomis, John B., and du Vair, Pierre, "Evaluating the Effect of Alternative Risk Communication Devices on Willingness to Pay: Results from a Dichotomous Choice Contingent Valuation Experiment," *Land Economics*, 69, 1993, 287-298.

Loomis, John B.; Le, Hung Trong, and Gonzáles-Cabán, Armando, "Testing Transferability of Willingness to Pay for Forest Fire Prevention Among Three States of California, Florida and Montana," *Journal of Forest Economics,* 11, 2005, 125-140.

Lusk, Jayson L., and Coble, Keith H., "Risk Aversion in the Presence of Background Risk: Evidence from the Lab," in J.C. Cox and G.W. Harrison (eds.), *Risk Aversion in Experiments* (Greenwich, CT: JAI Press, Research in Experimental Economics, Volume 12, 2007 forthcoming).

Mason, Charles F.; Shogren, Jason F.; Settle, Chad, and List, John A., "Investigating Risky Choices Over Losses Using Experimental Data," *Journal of Risk & Uncertainty*, 31(2), 2005, 187-215.

McDaniel, Tanga M., and Rutström, E. Elisabet, "Decision Making Costs and Problem Solving Performance," *Experimental Economics*, 4(2), October 2001, 145-161.

McKee, Michael; Berrens, Robert P.; Jones, Michael; Helton, Ryan, and Talberth, John, "Using Experimental Economics to Examine Wildfire Insurance and Averting Decisions on the Wildland-Urban Interface, *Society and Natural Resources,* 17, 2004, 491-507.

Mech Radomir, and Prusinkiewicz, Przemyslaw, "Visual models of plants interacting with their environment," *Proceedings of SIGGRAPH* 96, 1996, 397-410.

Micikevicius, Paulius, and Hughes, Charles E., "Visibility-based Forest Walk-through Using Inertial Level of Detail Model," *Journal of Defense Modeling and Simulation*, 4, 2007, in press.

Micikevicius, Paulius; Hughes, Charles E.; Moshell, J. Michael; Sims, Valerie, and Smith, Hannah, "Perceptual Evaluation of an Interactive Forest Walk-through," *VR Usability Workshop: Designing and Evaluating VR Systems*, Nottingham, England, January 22-23, 2004.

Milgram, Paul, and Kishino, A. Fumio, "Taxonomy of Mixed Reality Visual Displays," *IEICE Transactions on Information and Systems*, E77-D (12), 1994, 1321-1329.

Neilsen, William S., "Some Mixed Results on Boundary Effects," *Economics Letters*, 39, 1992, 275-278.

Neisser, Ulric, and Hyman, Jr., Ira E. (eds.), *Memory Observed: Remembering in Natural Contexts* (New York: Worth Publishers, 2000, Second Edition).

Oppenheimer, Peter E., "Real time design and animation of fractal plants and trees," *Proceedings of SIGGRAPH* 86, 1986, 55-64.

Ortmann, Andreas, & Gigerenzer, Gerd, "Reasoning in economics and psychology: Why social context matters," *Journal of Institutional and Theoretical Economics,* 153, 1997, 700–710.

Ostrom, Elinor; Gardner, Roy, and Walker, James, *Rules, Games, & Common-Pool Resources* (Ann Arbor, MI: Michigan University Press, 1994).

Ostrom, Elinor; Walker, James, and Gardner, Roy, "Covenants With and Without a Sword: Self-Governance Is Possible," *American Journal of Political Science*, 86(2), June 1992, 404-417.

Pearl, Judea, *Heuristics: Intelligent Search Strategies for Computer Problem Solving* (Reading, MA: Addison-Wesley, 1984).

Pratt, John W., "Aversion to One Risk in the Presence of Others," *Journal of Risk and Uncertainty*, 1(4), December 1988, 395-413.

Pratt, John W., and Zeckhauser, Richard J., "Proper Risk Aversion," *Econometrica*, 55(1), January 1987, 143-154.

Prusinkiewicz, Przemyslaw; James, Mark, and Mech, Radomir, "Synthetic topiary," *Proceedings of SIGGRAPH* 94, 1999, 351-358.

Quiggin, John, "Background Risk in Generalized Expected Utility Theory," *Economic Theory*, 22(3), October 2003, 607-611.

Rabin, Matthew, "Psychology and Economics," *Journal of Economic Literature*, 36, 1998, 11–46.

Rutström, E. Elisabet, and Williams, Melonie B., "Entitlements and Fairness: An Experimental Study of Distributive Preferences," *Journal of Economic Behavior & Organization*, 43, 2000, 75-80.

Safra, Zvi, and Segal, Uzi, "Constant Risk Aversion," *Journal of Economic Theory*, 83(1), November 1998, 19-42.

Salas, E. & Klein, G. (Eds.), *Linking expertise and naturalistic decision making.* (Mahwah, NJ: Erlbaum, 2001).

Schelling, Thomas C., "Intergenerational Discounting," *Energy Policy*, 23(4/5), 1995, 395-401.

Senn, Theresa E.; Espy, Kimberly Andrews, and Kaufmann, Paul M., "Using path analysis to understand executive function organization in preschool children," *Developmental Neuropsychology,* 26, 2004, 445-464.

Scott, Joe H., and Burgan, Robert E., *Standard Fire Behavior Fuel Models: A Comprehensive Set for Use with Rothermel's Surface Fire Spread Model*, Research Paper RMRS-GTR-153, Rocky Mountain Research Station, Forest Service, United States Department of Agriculture, June 2005.

Smith, Vernon L., "Microeconomic Systems As An Experimental Science," *American Economic Review*, 72(5), December 1982, 923-955.

Srinivasan, T.N., and Whalley, John (eds.), *General Equilibrium Trade Policy Modeling* (Cambridge: MIT Press, 1986)

Stapleton Christopher B., and Hughes, Charles E., "Believing is Seeing," *IEEE Computer Graphics and Applications* 27(1), January/February 2006, 80-85.

Starbuck, C. Meghan; Berrens, Robert P., and McKee, Michael, "Simulating changes in the forest recreation demand and associated economic impacts due to fire and fuel management activities," *Forest Policy and Economics,* 8, 2006, 52-66.

Stern, Nicholas, *The Stern Review of the Economics of Global Warming,* available from web site at http://www.hm-treasury.gov.uk/independent_reviews/stern_review_economics_climate_change/sternreview_index.cfm, 2006.

Talberth, John; Berrens, Robert P.; McKee, Michael, and Jones, Michael, "Averting and Insurance Decisions in the Wildland-Urban Interface: Implications of Survey and Experimental Data for Wildfire Risk Reduction Policy," *Contemporary Economic Policy,* 24(2), April 2006, 203-223.

Tamura, Hideyui; Yamamoto, Hiroyuki, and Katayama, Akihiro, "Mixed Reality: Future Dreams Seen at the Border between Real and Virtual Worlds," *IEEE Computer Graphics and Applications,* 21 (6), 2001, 64-70.

Thaler, Richard H., *Quasi-Rational Economics* (New York: Sage, 1991)

U.S. Environmental Protection Agency, *The Benefits and Costs of the Clean Air Act, 1970 to 1990.* United States Environmental Protection Agency, Office of Policy Analysis and Review, Document # A.97.20, October 1997.

Ulam, Stanislaw, "On some mathematical properties connected with patterns of growth of figures," *Proceedings of Symposia on Applied Mathematics*, 14, 1962, 215-224.

von Winterfeldt, Detlof, and Edwards, Ward, *Decision Analysis and Behavioral Research* (New York: Cambridge University Press, 1986).

Weber, Jason, and Penn, Joseph, "Creation and rendering of realistic trees," *Proceedings of SIGGRAPH* 95, 1995, 119-128.

Wilcox, Nathaniel T., "Lottery Choice: Incentives, Complexity, and Decision Time," *Economic Journal*, 103, 1993, 1397-1417.

Wilcox, Nathaniel T., "Theories of Learning in Games and Heterogeneity Bias," *Econometrica,* 74, 2006, 1271-1292.

World Bank, *World Development Report 1992: Development and the Environment* (Washington, DC: Oxford University Press, 1992).