JOURNAL OF ENVIRONMENTAL ECONOMICS AND MANAGEMENT **18**, 247–262 (1990)

Should We Try to Predict the Next Great U.S. Earthquake?[1]

WILLIAM D. SCHULZE

University of Colorado, Boulder, Colorado 80309-0257

DAVID S. BROOKSHIRE

University of Wyoming, Laramie, Wyoming 82071

RENATTE K. HAGEMAN

San Diego State University, San Diego, CA 92182

AND

SHAUL BEN-DAVID

University of New Mexico, Albuquerque, New Mexico 87131

Received April 13, 1987; revised August 29, 1989

Short term earthquake prediction has become technically possible. However, the desirability of such predictions depends both on the probability of successfully predicting an event and on the odds of false predictions. This paper investigates the economic feasibility of earthquake prediction as a function of program performance for the Los Angeles area. Based on the now increasing probability of a great earthquake in the region, the paper concludes that, given current best estimates of program performance, such predictions may well provide expected benefits which exceed expected costs. © 1990 Academic Press, Inc.

Earthquake prediction in the 1980s has entered the realm of "real-time geology"—the scientific possibility of measuring geologic processes as they occur (Wesson and Wallace [20]). Special investigators for the National Security Council have made a long term prediction that the next most likely and most destructive geologic event in the United States will be a catastrophic earthquake on the San Andreas fault in California.[2] Yet, as the scientific technology is developed for making a more precise short term prediction, little research has weighed the potential economic benefits against the costs of hazard prediction.

An aggressive prediction program pursued in the People's Republic of China has saved many lives. Hamilton [5] reports that a strong earthquake of magnitude

[1]The research presented here was supported by the U.S. Geological Survey. Thanks go to Rich Bernknopf and John Filson of the U.S.G.S. for their help and encouragement as well as to two anonymous reviewers. We thank Glenn Russell and David Schenk for research assistance and Melinda Berg for manuscript preparation. All opinions and conclusions are, of course, the sole responsibility of the authors.

[2]Recent damaging earthquakes in central and southern California have not been on the San Andreas fault and have been much smaller in intensity than the event discussed in the following presentation.

0095-0696/90 $3.00

7.3, which occurred in February of 1975 in northeast China, was successfully predicted. His account of the process leading up to the prediction is as follows:

> The prediction was made by a gradual refining or homing in, on the place, time, and magnitude of the upcoming shock by using a variety of techniques. As early as 1970, the area of Liaoning Province in northeast China, where the shock took place, was identified as an area of possible risk, apparently on the basis of long-term variations in seismicity. This concern was reaffirmed in June 1974 when the State Seismological Bureau called for increased vigilance in the area. This warning was based on a combination of observations, including migrations of seismic activity, tilting of the ground surface, changes in the water level in wells, changes in electric currents in the ground, and strange animal behavior. These observations prompted the Chinese to move more seismographs and tiltmeters into the area. On December 20, 1974, the local government was warned to expect a large earthquake. Apparently this warning resulted in a false alarm on the part of local officials, and people slept outside in the snow for 2 days. In mid-January 1975, the State Seismological Bureau met again, concluded than an earthquake was imminent, and on January 28, the villages were warned to be prepared. Extra seismographs were set up. On February 1, anomalous earthquake activity begin, which was interpreted as foreshocks, and it increased markedly on February 3. At 2 p.m. on February 4, people were told to expect a major quake within 2 days. Shops were shut in the town of Yingkow, and general evacuation of buildings was ordered in Yingkow and Haicheng Counties. The quake came at 7:36 p.m. that evening.

Fundamental social and economic differences exist between China and the United States. Although it is clear that lives can be saved by a successful prediction, a prediction program may entail large costs as well. Research has shown that long range hazards predictions are generally ignored by the public, and Turner *et al.* [18] have found that incorrect long term predictions can also reduce public responsiveness if a later event occurs. Mileti *et al.* [11] found that the formulation of prediction scenarios can affect individual responses. However, neither of these studies explicitly addressed the type of prediction considered in this paper.

The U.S. Geological Survey has set forth the goal of issuing short term predictions covering a period of only a few days. As a first step in this effort, scholarly research in 1980 led to an indication that an earthquake would occur along a section of the Calaveras fault near San Jose, California, The earthquake struck on April 24, 1984 (Kerr [7]). Encouraged by this long term forecast, on April 5, 1985, federal and state panels "endorsed studies that indicate a moderate earthquake is likely to strike near Parkfield in central California within several years of 1988" (Kerr [8]). More recently, the focus has moved towards short term predictions. John Filson, Chief of the Office of Earthquakes, Volcanoes and Engineering, U.S. Geological Survey, summarized this goal:

> The USGS earthquake prediction research program includes an experiment in Parkfield, CA, as a test of concepts and instruments to be considered as the basis for a larger network in urban areas in S. California and elsewhere. The object of the Parkfield, CA test is to provide a warning *within a few days or hours* of a magnitude 6 event expected in the region within the next few years." [Emphasis added][3]

It is this goal which motivates this study.

In Section II, we develop a model of the expected benefits and costs of a short term earthquake prediction program which has been proposed by the U.S. Geological Survey for southern California. We focus on a short term "prediction window"

[3]J. Filson, Chief of the Office of Earthquakes, Volcanoes and Engineering, Geologic Division, U.S. Geological Survey, Reston, Virginia; personal communication, 1985.

designed by U.S.G.S. researchers who believe that this type of prediction is technically feasible. This short term prediction is based upon detection of immediate (2 days hence) geologic precursors to a catastrophic event.

The primary benefit of short term prediction is that public officials can take immediate action; in our case, it is assumed the population is required to remain in the relative safety of their homes for a 48-hour period. Although extensive emergency preparations probably cannot be made due to the short period between geologic indicators and an imminent event, many lives can be saved by keeping people away from dangerous structures. Even so, losses in economic activity associated with false predictions might well overwhelm the value of possible lives saved if a prediction program performs poorly. Though great value is placed on saving lives in the United States, we pose the question, does such value justify the costs of a short term hazard prediction program?

The answer for the Los Angeles area may well be yes. In the following sections, we use our model to develop an empirical example of a benefit/cost analysis based upon U.S.G.S. estimates for program performance. At the heart of any benefit/cost analysis of hazards prediction lies the likelihood of an event's occurrence. We focus attention on this aspect in Section III, where we describe our estimation of the probability of the next catastrophic earthquake on the San Andreas fault. Inspection of the historical record (beginning in the year 261 A.D.) of large earthquakes on the San Andreas led us to apply statistical failure theory (a Weibull distribution) to this estimation problem, and the results of this statistical analysis show that the probability of a large earthquake's occurring on the San Andreas is increasing annually.

In Section IV, we describe the assumptions made and data employed for estimating the benefits/costs of a short term prediction program for the Los Angeles area. Estimates of potential economic benefits, as measured by reduced risk to lives, are based upon U.S.G.S. estimates of potential loss of life in a catastrophic earthquake on the San Andreas. A Monte Carlo simulation model of earthquake prediction is presented in Section V. This model incorporates U.S.G.S. estimates of the prediction program's performance reliability. However, any hazards prediction program, whether it addresses adverse weather conditions or geological phenomena, will be characterized by some uncertainty and risk for decisionmakers responsible for making a prediction. Thus, in this section we develop an iso-benefit mapping for program evaluation where benefits and costs are compared over a range of feasible success rates for the program's prediction capabilities. Within this policy context, conclusions and caveats are presented in Section VI.

I. A THEORETICAL BASIS FOR EVALUATING EARTHQUAKE PREDICTION

Benefit–cost analysis of an earthquake prediction program for the Los Angeles region depends in great part on the probabilities of success of the prediction program resulting from the proposed seismic monitoring network along the Southern San Andreas fault. Given an earthquake prediction program, four possible states may occur. These states are denoted $j = 1, 2, 3, 4$, where the respective probabilities for each state, P_j, which sum to one, are defined as

P_1 = probability that an earthquake prediction will be made and an earthquake will occur;
P_2 = probability that an earthquake prediction will be made but no earthquake will occur;
P_3 = probability that no prediction will be made but an earthquake will occur;

and

P_4 = probability that no prediction is made and no earthquake occurs.

In theory, these probabilities could be directly related to the scale, expense, and quality of an earthquake prediction program—the better the program, the more closely P_2 and P_3 should approach zero. These state probabilities may be related to three additional probabilities that define both earthquake risk and prediction program capabilities (Collins [3]). These probabilities are, where a historical record for prediction over many periods is available,

P_E = probability of a specified earthquake in a given year;[4]

P_S = probability that the specified event will be successfully predicted $= \dfrac{\text{number of successful predictions}}{\text{number of earthquake events}}$;

and

P_F = probability that a prediction will be false $= \dfrac{\text{number of false predictions}}{\text{number of predictions}}$.

We assume an event is successfully predicted if a prediction is made and an event occurs. A false prediction occurs if a forecast is made and no event occurs. Only an event of fixed size or intensity is considered. Given these assumptions, using Bayes' Theorem,[5] probabilities of occurrence for the four states of the world are now defined as

$$P_1 = P_S P_E \tag{1}$$

$$P_2 = P_F P_S P_E/(1 - P_F) \tag{2}$$

$$P_3 = (1 - P_S) P_E \tag{3}$$

and

$$P_4 = 1 - P_E - P_F P_S P_E/(1 - P_F). \tag{4}$$

A "good" earthquake prediction program plausibly has $P_F \to 0$ and $P_S \to 1$; so, from the relationships shown above, we would wish to have $P_1 \to P_E$, $P_2 \to 0$, $P_3 \to 0$, and $P_4 \to (1 - P_E)$. Thus, in a world with perfect prediction capabilities

[4]Note that $P_1 + P_3 = P_E$ and so $(1 - P_E) = P_2 + P_4$ since $\Sigma_{j=1}^{4} Pj = 1$.

[5]Bayes' Theorem is required to derive P_2. Bayes' Theorem implies that $1 - P_F = P_1/(P_1 + P_2)$ which yields Eq. (2), given Eq. (1).

only two states are relevant; state 1, where a predicted event occurs, and state 4, where no prediction is made and no event occurs. However, this is an implausible situation, so we must be concerned with state 2, wherein a false prediction is made. Thus, benefits and costs of a prediction program depend on P_1, P_2, P_3, and P_4, which are defined by (1) to (4) above. These in turn depend on P_E, the odds of an event, which is obtained for Los Angeles from the analysis of Section III, and on P_S and P_F, which are the two parameters describing the success of a prediction program. We now turn to development of a economic model incorporating these probabilities to structure our benefit–cost analysis of the prediction program.

Benefit–cost analysis traditionally has employed expected utility theory to justify the weighting of benefits and costs in different states of the world by probabilities of those states as a first approximation of a social welfare function. Although questions have been raised concerning the behavioral predictions of the expected utility model (Schoemaker [13] and Kunreuther *et al.* [9]), positive evidence does exist that real estate values have capitalized the expected value of potential earthquake damage (Brookshire *et al.* [2]). Further, in recent work by McClelland, Schulze, and Coursey [10], involving purchase of insurance in laboratory experiments, probabilities above 0.1 do not induce behavior divergent from the expected utility model. This work is consistent with Kahneman and Tversky's [6] prospect theory which argues that very low probabilities can be drastically overweighted in decisionmaking but that intermediate probabilities receive about the right "weight." With regard to the problem at hand, individual behavior enters only in the response to a prediction which, as we show below, requires a probability of success for the 2 day window above 0.1 for economic feasibility. It is also worth noting that the character of the problem is analogous to hurricane warnings which have been quite successful. It is possible that successful predictions work because they are structured so as to communicate relatively high probabilities, where people behave more rationally, as opposed to low probabilities, where much research shows that people have severe cognitive problems. Further, the question is not one of observed rationality so much as one of enforced rationality. That is, often police organizations and the national guard enforce a warning through evacuation procedures or limits placed on travel. The similarities between hurricane warnings and proposed short term earthquake warnings are in many ways quite compelling.

Thus, the framework for benefit–cost analysis we propose assumes that an earthquake prediction would shut down economic activity in the Los Angeles area while forcing people to remain at home, where studies show risk is greatly reduced. We use the following additional notation:

VLS = value of lives saved by a successful prediction;

C_S = incremental economic cost due to a successful prediction, i.e., the expected value of production lost between the time the warning occurs (so economic activity stops) and the time at which the earthquake occurs (when the economic shutdown becomes attributable to the earthquake event, not the prediction);

C_F = incremental economic cost of a false prediction, i.e., the value of production lost from the time the warning is announced until the warning is called off plus any further losses incurred during start up of the economy (e.g., from a reluctance of people to return to work after a false prediction);

$C(P_S, P_F)$ = discounted costs of the prediction program itself, an increasing function of P_S and a decreasing function of P_F;

r = discount rate;

n = rate of regional population growth;

t = time ($0 \leq t \leq T$).

Discounted expected benefits of the prediction program can then be defined as

$$B = \int_0^T e^{-(r-n)}[P_1(\text{VLS} - C_S) - P_2 C_F]\, dt. \tag{5}$$

In this expression the value of lives saved *net* of the expected cost of a successful prediction is weighted by the probability of a successful prediction, P_1. Thus, we adjust the value of lives saved for the expected economic losses occasioned by shutting down the local economy while waiting for the predicted event to occur. Since, in this state of the world, the event does occur, the benefit of saving lives is realized. This residual benefit must, however, be further adjusted to account for the possibility of a false prediction where the event does not occur. Thus, economic costs of a false prediction, weighted by P_2, the probability of a false prediction, are subtracted in (5). It should be noted that VLS, C_S and C_F are all specified at the level of the initial program start up ($t = 0$). They are assumed to grow over time at an exponential rate equal to the regional rate of population growth, n. Since we discount benefits over time at a discount rate r, we simply subtract n from r in the discounting process to account for the effect of growth in local population and, we assume, values.

The total net benefits of the prediction program can then be written as

$$B - C(P_S, P_F), \tag{6}$$

where B is taken from (5). Thus, an optimal prediction program would choose P_S and P_F, the design probability that an event will be predicted and the design probability that a prediction will be false, such that net benefits will be maximized.

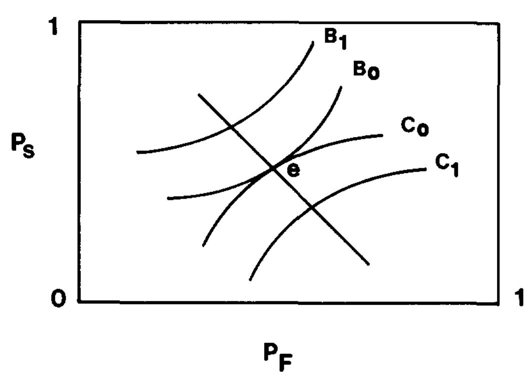


FIG. 1. Iso-benefit and iso-cost curves for earthquake prediction.

Substituting P_S and P_F into (5) and then (6) gives net benefits as

$$\text{(7)} \quad \left[\int_0^T e^{-(r-n)t} P_E(t)\, dt\right] P_S \left[\text{VLS} - C_S - \left(\frac{P_F}{1 - P_F}\right) C_F\right] - C(P_S, P_F).$$

Figure 1 shows the tradeoff between P_S and P_F both on the benefit side which yields convex iso-benefit lines (B_0, B_1) and on the cost side which yields (we assume) concave iso-cost lines (C_0, C_1). Clearly, efficient points like "e" are defined by the tangency between iso-cost and iso-benefit lines like C_0 and B_0. One point along the program expansion path which moves from the lower right to the upper left made up of efficient choices of P_F and P_S, such as the one marked e in the figure, will maximize net benefits as defined in (7). Unfortunately, little is known about the cost function, $C(P_F, P_S)$. Rather, we have a point estimate of P_F, P_S, and program costs from the U.S.G.S. Thus, in what follows, we attempt to develop the necessary information to approximate benefits for any point in Fig. 1 Thus, with estimates of $P_E(t)$, the annual odds of an event which we develop in Section III, and of VLS, C_S, and C_F, which we develop in Section IV, we can determine if benefits exceed costs for the U.S.G.S. performance estimates and provide benefit estimates which may be useful if better information is developed on the costs and tradeoffs associated with earthquake prediction programs.

III. THE PROBABILITY OF A LARGE EVENT ON THE SAN ANDREAS NEAR LOS ANGELES

This section uses statistical failure theory and evidence on the history of the San Andreas fault to estimate the odds over time of a large earthquake in the Los Angeles area. In terms of expected levels of ground shaking, a large event on the San Andreas in southern California contributes the largest fraction of total seismic risk for the Los Angeles area. Sieh [14] has estimated that large events occur on the San Andreas fault in southern California with an average recurrence interval of about 145 years based on excavations of later Holocene marsh deposits at Pollett Creek.

Table I presents approximate dates of past large events on the San Andreas taken from Sieh. Figure 2 shows the distribution of the 11 intervals between the 12 events that Sieh has identified. This distribution strongly suggests the use of statistical failure theory, which is typically applied to aircraft wings, automotive tires, and manufactured parts. This statistical approach to mechanical or structural failures from strain and wearing out uses the Weibull distribution which is a cumulative distribution over time $(0, \infty)$ of the form

$$\text{(8)} \qquad F(t) = 1 - e^{-\alpha(t-t_0)\beta} \qquad \text{for } t \geq t_0.$$

F is the cumulative fraction in a given sample which has failed up to time t, from time zero. The rate at which failure occurs, $f(t)$, is given by the probability density function which is the time derivative of (8):

$$\text{(9)} \qquad f(t) = \frac{dF}{dt} = \alpha\beta(t - t_0)^{\beta-1} e^{-\alpha(t-t_0)\beta} \qquad \text{for } t \geq t_0.$$

TABLE I
Approximate Dates for the Last Twelve Major Events on the Southern San Andreas (All Dates A.D.)

Date	Time Since The Last Event	Error
261	—	±91
349	88	±78
588	239	±57
733	145	±62
843	110	±74
935	96	±86
1013	78	±99
1080	67	±65
1350	270	±50
1550	200	±70
1720	170	±50
1857	137	0

Note $F(t) = f(t) \equiv 0$ for $t \leq t_0$. Also if $\beta > 1$ then the cumulative distribution given in (8) is S shaped and asymptotically approaches one. The probability density function is bell-shaped and asymptotically approaches zero.

Viewing catastrophic earthquakes on the San Andreas fault in southern California as stress-related failures, we can take the interval between large earthquakes to be the length of time, t, until failure occurs. These cumulative data from Sieh are plotted in 25 year intervals in Fig. 3. For the 11 recorded intervals between failures

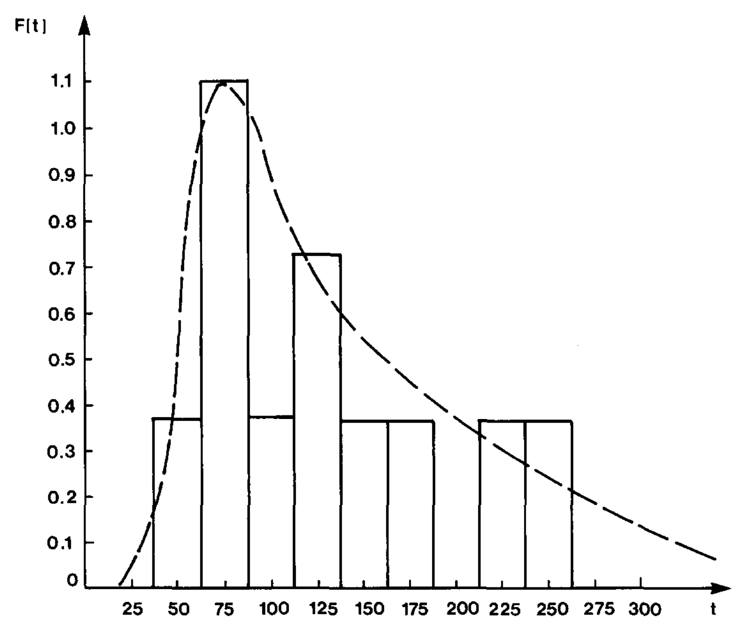


FIG. 2. Probability distribution (in percentage), 25 year intervals.

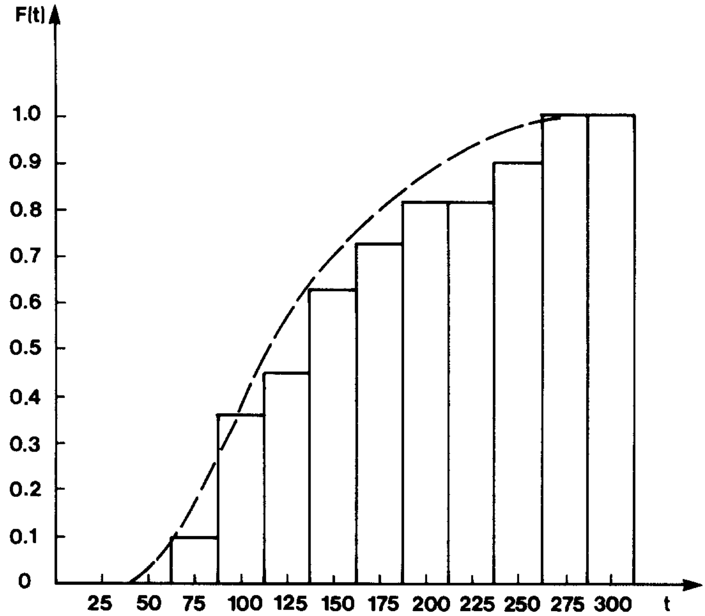


FIG. 3. Cumulative density function, 25 year intervals.

(earthquakes), zero out of 11 occurred up to 50 years after the last event, one out of eleven (1/11) occurred prior to 87.5 years, and so on. Twenty-five year intervals were chosen to reflect some of the uncertainty over the precise date of historic earthquake events since 1 year intervals would exaggerate the precision of Sieh's dating techniques. Similarly, 50 year intervals would be too wide, reducing the number of observations for analysis below the number of intervals.

A Weibull distribution can be fitted to the 11 observations ranging from 67 to 270 years shown in Fig. 2 using the maximum likelihood technique. The shift parameter t_0 in the eleven variate density functions of (9) shifts the Weibull distribution away from the origin. Maximum likelihood estimates are obtained by iterating through values of t_0 and then finding the values of α and β which maximize the likelihood that our observations come from the hypothesized function. These values are found by taking the first derivatives of the natural log of the likelihood function with respect to α and β, then setting them equal to zero and solving simultaneously for the estimated α and β. Maximum likelihood estimation yields the following parameters: $\alpha = .0119$, $\beta = 1.113$, and $t_0 = 64.33$.

Given these estimates of α, β, and t_0, the resulting estimated probability density function[6] is the solid line plotted in Fig. 4. The interpretation of the probability density function can be taken from this figure. Starting in the year 1857, just after the Fort Tejon earthquake, the probability of an event occurring 125 years later in 1982 is about .6%. Since a large earthquake has not occurred since 1857, the probability of an event in 1982 was actually greater than .6 percent. This statement is the logical consequence of Bayes' theorem, which states that the probability of

[6]Note that since the value of β is very close to unity, the Weibull distribution estimated here is close to an exponential distribution wherein convergence occurs at $\beta = 1$.

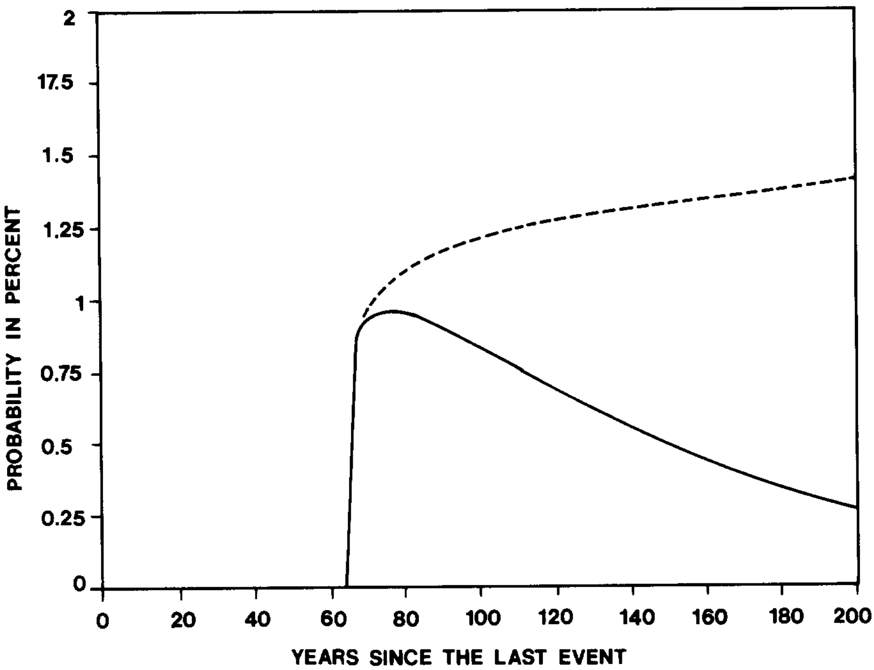


FIG. 4. Weibull density and associated failure rate function; parameter estimates are maximum likelihood. — Probability at T; --- Conditional probability at T.

an event in the year T^*, given that no event has occurred prior to T^*, is

$$p(T^*) = \frac{f(T^*)}{\int_{T^*}^{\infty} f(t)\, dt}. \tag{10}$$

This relationship is shown as the dotted line in Fig. 4 and implies that in 1982, 125 years after the last event, the annual odds of an earthquake on the San Andreas in southern California were about 1.2 percent. If no new event occurs by the year 2000, 143 years since the last event, the annual probability will have risen slightly to 1.3 percent.

IV. BENEFITS AND COSTS OF AN EARTHQUAKE PREDICTION ON THE SOUTHERN SAN ANDREAS FAULT

The prediction used for the benefit–cost analysis takes the following form: "A great earthquake is predicted to occur within the next 48 hours on the southern San Andreas fault near Los Angeles." This short-term prediction is aimed at saving lives, not at reducing property losses, and has the advantage of minimizing pre-event disruption. The appropriate response to the prediction described above is not a panic evacuation; rather, residents are simply told to stay home, since most homes in the Los Angeles area are wood framed structures with a very low risk of death in a major earthquake. It is assumed that public officials (e.g., the National Guard) require the populace to remain inside residences for the 48-hour prediction period. We discuss this particular prediction scenario because U.S.G.S.

researchers have estimated that while a catastrophic earthquake could cause 3,000 deaths in the middle of the night with most people at home, 10,000 additional deaths could result if the event occurred midday without warning. Thus, our response strategy is to create a lower risk situation for the 48-hour prediction window which is similar to that of the late night geographic distribution of population.

The benefits of a prediction program can be estimated in terms of the value of reducing the risk to life in Orange and Los Angeles Counties if a large event on the San Andreas fault is predicted successfully. Following the methodology established in the value of safety literature (see Mishan [12]), the benefits of a successful prediction are derived by estimating the number of lives saved by an announced prediction when a large earthquake occurs. The value of safety benefit is taken as the product of the estimated number of lives saved and the marginal value of safety, assumed to be one million dollars. The one million dollar value is an inflation adjusted central value based on studies by Thaler and Rosen [17], Viscusi [19], Smith [15], and Blomquist [1].[7]

Steinbrugge *et al.* [16] have estimated the expected number of deaths in Orange and Los Angeles Counties if an earthquake of magnitude 8.3 occurs on the southern San Andreas fault. Updated to reflect 1980 Census data, the following numbers of lives are lost for different times of occurrence: (1) 2:30 a.m., 3,080 lives; (2) 2:00 p.m., 11,906 lives; (3) 4:30 p.m., 13,007 lives. Given no prediction, it is assumed that on average the population spends 4 hours on freeways and entering/leaving work daily (the 4:30 p.m. risk), 8 hours at work (the 2:00 p.m. risk), and 12 hours at home (the 2:30 a.m. risk). Thus, the time of occurrence weighted average number of deaths which would occur as the result of a large earthquake, given no warning, is approximately 7,800.

If a prediction is made that a large event will occur within 48 hours and the population is required to return and remain home in response to the prediction, the 4:30 p.m. risk is assumed to apply for 1.5 hours during the initial response period, while the at-home (2:30 a.m.) risk applies for the remainder of the 48-hour period. The weighted average number of deaths that would occur, given a successful prediction warning, is 3,400 in this case.[8]

The difference between the two cases (7,800 − 3,400) implies that approximately 4,400 deaths would be saved in Los Angeles and Orange Counties by an earthquake warning. Assuming a marginal value of safety of $1 million, the safety benefits from a successful earthquake prediction in the Los Angeles area are about $4.4 billion.

There are several reasons why this value may well underestimate the possible savings in lives and resulting level of benefits. First, we do not account for possible evacuation of high risk residences. Second, some mobilization of public emergency programs and personnel might be accomplished. However, because the short term prediction is based upon immediate geological precursors to the event, this type of

[7]These studies found the marginal value of safety to be from .43 to 2.5 million dollars (stated in 1980 terms). Though we also conducted empirical analyses using a $2.5 million figure, we choose to report results for the $1 million value in order to remain conservative in this exercise.

[8]The number of deaths shown here is based exclusively upon an extensive analysis by NOAA of potential earthquake losses in the Los Angeles area (Steinbrugge *et al.* [16]). The behavioral response by the public to the warning itself cannot be completely projected, though we have made the simplifying assumption that authorities could avoid panic by requiring the public to remain within their residences.

prediction would not provide enough time to undertake major emergency preparations such as lowering the water levels of regional dams. (Although these types of benefits may be large when the scientific technology is developed for making more precise, longer term earthquake predictions.) Last, only the benefits associated with the prediction program for a catastrophic earthquake are estimated. The program might provide additional benefits by also predicting more frequent, smaller earthquakes.

An approximate cost associated with the two-county response to an earthquake prediction can be estimated as the loss in local output due to the forced return of the populace to their residences. In 1980 dollars, per capita output per day is $36.26 and $36.65 in Orange and Los Angeles Counties respectively. Using 1980 census populations of 1.9 and 7.5 million, respectively, the two-county total daily output is $.34 billion. In our analysis, one day's loss in output is used to approximate the cost of a successful prediction. The rationale is that with a 48-hour prediction window, the event will take place on average after 24 hours if the prediction is correct. On the other hand, if a prediction proves to be false, we assume that the local disruption could amount to up to one week's loss in output. Thus, as an order of magnitude approximation, the cost of a false earthquake prediction in Orange and Los Angeles Counties may be seven times as great as in a successful prediction, or $2.4 billion.

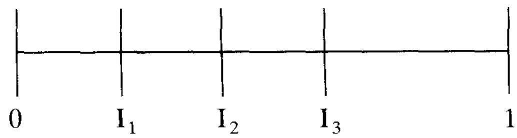
These cost estimates associated with community response to a prediction are probably exaggerated. First, no allowance is made for the partial offset from leisure benefits which accrue when the working population receives unexpected work days off. Second, evidence exists which indicates that the community disruption costs of a false prediction may be quite low. In late 1980, two U.S. scientists predicted that a devastating earthquake, the most powerful in the world this century, would occur in Peru and northern Chile on one of three specific days in 1981. After the forecasted event did not take place, Echevarria *et al.* [4] surveyed a sample population in Lima and concluded that although 99% of the populace knew of the forecast before the prediction dates, only 2–3% of those surveyed indicated they had either left the area or stayed away from work temporarily. On the other hand, a significant number of individuals did seek information on emergency procedures or stocked up on foodstuffs and medicines. These responses suggest that only limited economic and psychic costs may have been imposed on the population from the false prediction.

However, individuals possibly would be willing to pay some amount to avoid needless fear and worry associated with a false earthquake prediction. No measures of such psychic costs are available. Even so, in light of the Peruvian experience, the assumption that an entire week's worth of local GNP would be lost due to the disruption caused by a false forecast may be an overestimate of the response cost.

V. A MONTE CARLO SIMULATION OF THE NET BENEFITS OF AN EARTHQUAKE PREDICTION PROGRAM IN LOS ANGELES

The discounted present value of net benefits of the earthquake prediction program are determined by using a Monte Carlo simulation of the four states of the world which can occur over the specified time horizon. This approach was taken to allow non-constant rates of discount or population growth to be used in future studies for which better data on benefits and costs may become available.

For a time horizon of 50 years, a random number process was used to simulate which of the four states of the world would occur in each year. Since the four probabilities sum to one, the cumulative probabilities can be placed on a 0–1 scale with four intervals



where the intervals are defined as $I_1 = P_1$, $I_2 = P_1 + P_2$, and $I_3 = P_1 + P_2 + P_3$.

For each year a number was randomly generated from a uniform distribution between 0 and 1. In a given year, if the random value fell within the first interval, then state 1 was selected in the simulation. A value falling within the second interval $(I_2 - I_1)$ implied state 2, and so forth. Since all values between 0 and 1 had an equal chance of being selected, the state associated with the largest interval (i.e., the largest probability) had the greatest likelihood of occurring in the simulation. Similarly, the state with the smallest interval (i.e., the smallest probability) had the least likelihood of selection. Therefore, over many simulations the frequencies at which various states were selected reflect the relative probabilities that the states will occur.

For example, the proposed earthquake prediction program is estimated by the U.S.G.S. to produce a 40 percent rate of false prediction while successfully predicting 20–50 percent of large earthquakes. Based upon our maximum likelihood estimation of the Weibull distribution, the probability of having a large earthquake in 1980 is 1.27%, increasing slightly each year thereafter. Using these estimates where we assume P_S is 35% (the central estimate) in Eqs. (1)–(4) yields probabilities of $P_1 = .004$, $P_2 = .003$, $P_3 = .008$, and $P_4 = .984$ in the first year of the simulation. It can be seen that a disproportionately large number of random values in the simulation must fall in the last interval, .015 through 1. Thus, state 4 (no prediction, no earthquake) was selected the most often in the simulations.

During each simulation, annual benefits and costs were tabulated for the state selected in each of the 50 years. We assumed a 0% rate of population growth to be conservative. Note however, that the impact of constant rates of population growth can be incorporated in our results shown below by adjusting the discount rate for the rate of population growth (i.e., by interpreting the discount rate used as $r - n$). All values are stated in 1980 dollars. Annual values were then discounted and summed to obtain an estimate of the net present value of benefits from the prediction program for Los Angeles and Orange Counties. The 50-year simulation model was re-run continuously while keeping track of the running average of the net present value of benefits over all simulations. Due to the randomness of the procedure, the running average net benefit is itself a random number. As such, it has a distribution described by a mean and standard deviation. By the central limit theorem, as the number of simulations approaches infinity, the standard deviation of this running average approaches zero.

After 100,000 simulation runs the standard deviation of the running average net benefit value became statistically insignificant for additional runs. As an additional check, the frequencies of each state were calculated for each 50-year run. Since

TABLE II
Prediction Program Benefits and Costs (in 1980 dollars, millions; 50 year time horizon)

r	Present value of program benefits (net of community response cost)	Present value of direct program cost
0	416	185.0
.05	184	108.0
.10	111	85.0

Note. $P_F = .4$, $P_S = .35$.

these values also varied from run to run due to the randomness of the procedure, the average frequency for each state was calculated after each run. These averages can be interpreted as expected values. It was the case that given values of P_S and P_F, the expected value of the frequency of each state over the time horizon almost exactly equalled the true probability of each state. This implies that the process produced stable results after 100,000 computer runs. Therefore, results reported in the following section were derived from sets of simulation runs, where each set of 100,000 runs used a different discount rate or varied the assumptions about the prediction program's success/failure rate.

The U.S. Geological Survey has estimated the cost of the initial capital outlay to be about $60 million for the seismic monitoring network, with an estimated $2.5 million per year required in operating costs (measured in 1980 dollars). In Table II, the direct cost of establishing and operating the proposed program is shown for comparison with results obtained from the net benefit simulations.

Based on the benefit/cost criteria we have discussed, the results shown in Table II indicate the proposed program's potential for producing benefits to society in excess of program costs. However, an important caveat is in order. Because the psychic cost to society from false predictions has not been estimated in the analysis, there exists an element of political risk to the policy makers responsible for formulating and announcing earthquake predictions.

With regards to both political and economic feasibility, another useful result from the simulation model pertains to policy decisions about the performance of the prediction program. The program success and failure rates, P_S and P_F, can be treated as decision variables since program personnel can determine how often predictions are made. As an extreme example, P_S could be forced to equal one by merely predicting an earthquake every day. In this manner, every time an earthquake occurs there will have been a corresponding successful prediction. However, if such an overzealous policy were to be undertaken, observe what would happen to P_F. A prediction is made every day but since there will rarely be an earthquake, P_F is forced to approach unity. Thus, an important aspect of earthquake prediction policy is the weighing of the political/economic advantages and disadvantages of varying P_S and P_F.

Figure 5 shows three iso-benefit curves mapped in P_S, P_F space, each corresponding to a different discount rate. Each "breakeven" curve maps different combinations of P_S and P_F where total net benefits equal zero. In other words, given a discount rate and a value for P_F, they show the value which P_S must take

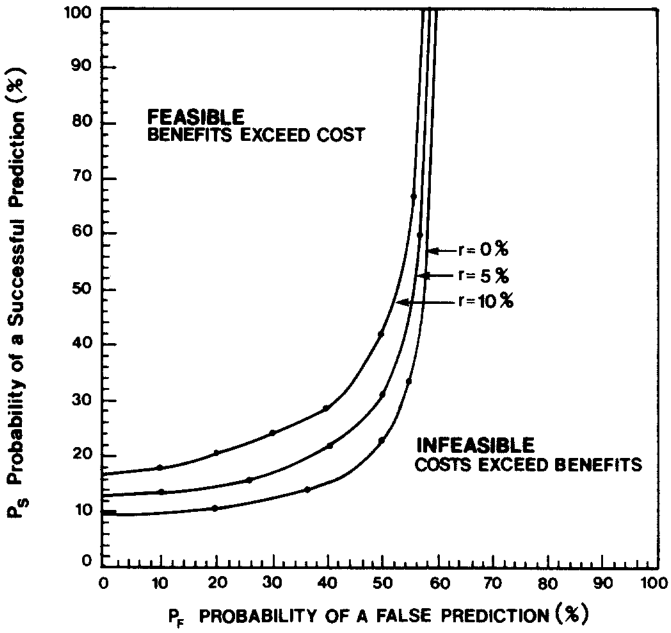


FIG. 5. Breakeven iso-benefit curves for an earthquake prediction program in Los Angeles. Time horizon = 50 years. Cost and benefits are in billions of 1980 dollars. r = real rate of discount. Costs include capital costs (.06 billion) and present discounted value of operating costs (which vary depending on r): for $r = 0\%$ total costs = .185 billion; for $r = 5\%$ total costs = .108 billion; for $r = 10\%$ total costs = .085 billion.

in order for discounted benefits (net of community response costs) to just cover capital costs plus discounted operating costs. Thus, each curve represents the demarcation between the feasible region where benefits exceed costs and the infeasible region where costs exceed benefits.

VI. CONCLUSIONS AND CAVEATS

Two conclusions are apparent from examination of Fig. 5. First, choice of discount rate has little impact on the feasibility of earthquake prediction in the Los Angeles area. This surprising result is due to the fact that P_E, the odds of a great earthquake, are rising over time. Since the higher the P_E, the greater the likelihood of a predicted great earthquake, rising P_E implies a rising potential for prediction benefits over time which offsets the effect of discounting.

Second, the net benefits of the earthquake prediction program proposed by the U.S. Geological Survey fall, in the main, within the area of economic feasibility. Except when the success rate is extremely low in combination with the highest choice of discount rate, the U.S. Geological Survey estimates of $P_S \cong 20$–50% and $P_F \cong 40\%$ appear to yield positive total net benefits.

These results suggest that implementation of an earthquake prediction program for the southern San Andreas fault cannot be rejected on economic grounds and substantial net benefits might result to society. However, an important topic for future research is the need for more information on the public response to predictions of catastrophic events. Clearly, the costs and benefits of a prediction program will depend on its reputation for accuracy and the credibility of the government decisionmakers who announce such predictions—both of which should be subject to learning effects over time.

REFERENCES

1. G Blomquist, The value of human life: An empirical perspective, *Econom. Inquiry* **19** (1981).
2. D. S. Brookshire, M. Thayer, J. Tschirhart, and W. Schulze, A test of the expected utility model: Evidence from earthquake risks, *J. Polit. Economy* **93**, 369–389 (1985).
3. J. D. Collins, The probabilistic significance of earthquake prediction, *Bull. Seismol. Soc. Amer.* **67**(1) (1977).
4. J. A. Echevarria, K. Norton, and R. Norton, Community response to an earthquake forecast: A case study in Peru, *in* "Methods Development for Valuing Hazards Information" (D. S. Brookshire et al., Eds.), report prepared for the U.S. Geological Survey, Contract #14-08-0001-20458 (1983).
5. R. Hamilton, The Status of Earthquake Prediction, *in* "Earthquake Prediction—Opportunity to Avert Disaster," Geological Survey Circular 729, (1976).
6. D. Kahneman and A. Tversky, Prospect theory: An analysis of decision under risk, *Econometrica* **47**(2) 263–291, March (1979).
7. R. A. Kerr, An encouraging long term quake forecast, *Science* (July 20, 1984).
8. R. A. Kerr, Earthquake forecast endorsed, *Science*, **311** (April 19, 1985).
9. H. Kunreuther, R. Ginsberg, L. Miller, P. Sagi, P. Slovic, B. Borkan, and N. Katz, "Disaster Insurance Protection: Public Policy Lessons," Wiley-Interscience, New York (1978).
10. G. H. McClelland, W. D. Schulze, and D. L. Coursey, Valuing risk: A comparison of expected utility with models from cognitive psychology, manuscript, University of Colorado, Boulder (September, 1986).
11. D. S. Mileti, J. R. Hutton, and J. H. Sorensen, "Earthquake Prediction Response and Options for Public Policy," Program on Technology, Environment and Man, Monograph #31, Institute of Behavioral Science, University of Colorado (1981).
12. E. J. Mishan, Evaluation of life and limb: A theoretical approach, *J. Polit. Econom.* **79** (1971).
13. P. J. H. Schoemaker, The expected utility model: Its variants, purposes, evidence and limitations, *J. Econom. Lit.* **20**, 529–563 (June, 1982).
14. K. E. Sieh, Prehistoric large earthquakes produced by slip on the San Andreas fault at Pollett Creek, California," *J. Geophys. Res.* **83**(88) (1978).
15. R. S. Smith, Compensating wage differentials and public policy: A review, *Ind. Lab. Relat. Rev.* **32**, 338–352 (April, 1979).
16. K. V. Steinbrugge *et al.*, "Metropolitan San Francisco and Los Angeles Earthquake Loss Studies: 1980 Assessment," pp. 81–133, Department of the Interior, U.S. Geological Survey (1980).
17. R. Thaler and S. Rosen, The value of saving a life: Evidence from the labor market, *in* "Household Production and Consumption" (N. E. Terleckyj, Ed.), pp. 265–297, Columbia Univ. Press for NBER, New York (1975).
18. R. Turner, J. Nigg, and D. Paz, "Waiting for Disaster: Earthquake Watch in California," Univ. California Press, Los Angeles (1986).
19. W. F. Viscusi, Labor market valuations of life and limb: Empirical evidence and policy implications, *Public Policy* **26** (1978).
20. R. L. Wesson and R. E. Wallace, Predicting the next great earthquake in California, *Sci. Amer.* **252**(2), 35–44 (1985).