JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 104, NO. B3, PAGES 4755–4768, MARCH 10, 1999

Foreshock occurrence before large earthquakes

Paul A. Reasenberg
U.S. Geological Survey, Menlo Park, California

Abstract. Rates of foreshock occurrence involving shallow $M \geq 6$ and $M \geq 7$ mainshocks and $M \geq 5$ foreshocks were measured in two worldwide catalogs over ~20-year intervals. The overall rates observed are similar to ones measured in previous worldwide and regional studies when they are normalized for the ranges of magnitude difference they each span. The observed worldwide rates were compared to a generic model of earthquake clustering based on patterns of small and moderate aftershocks in California. The aftershock model was extended to the case of moderate foreshocks preceding large mainshocks. Overall, the observed worldwide foreshock rates exceed the extended California generic model by a factor of ~2. Significant differences in foreshock rate were found among subsets of earthquakes defined by their focal mechanism and tectonic region, with the rate before thrust events higher and the rate before strike-slip events lower than the worldwide average. Among the thrust events, a large majority, composed of events located in shallow subduction zones, had a high foreshock rate, while a minority, located in continental thrust belts, had a low rate. These differences may explain why previous surveys have found low foreshock rates among thrust events in California (especially southern California), while the worldwide observations suggests the opposite: California, lacking an active subduction zone in most of its territory, and including a region of mountain-building thrusts in the south, reflects the low rate apparently typical for continental thrusts, while the worldwide observations, dominated by shallow subduction zone events, are foreshock-rich. If this is so, then the California generic model may significantly underestimate the conditional probability for a very large ($M \geq 8$) earthquake following a potential ($M \geq 7$) foreshock in Cascadia. The magnitude differences among the identified foreshock-mainshock pairs in the Harvard catalog are consistent with a uniform distribution over the range of observation.

1. Introduction

Short-term earthquake clustering, including the occurrence of foreshocks and aftershocks, is a widely observed phenomenon in shallow crustal seismicity. Nearly half of all earthquakes (most of them aftershocks) are included in short-term clusters [*Reasenberg*, 1985; *Davis and Frohlich*, 1991; *Ogata et al.*, 1995]. Moderate or strong earthquakes are occasionally followed by stronger shocks nearby within a few days. This clustering makes possible short-term probabilistic earthquake forecasts after any earthquake based on triggered stochastic models [*Kagan and Knopoff*, 1987] in which the occurrence times of the triggered events are Poissonian (with varying rate), while their rate, magnitude distribution, and geographic distribution are empirically determined from the clustering behavior observed in the historic seismicity. After a moderate earthquake in an urbanized area, the increased probability of an imminent, larger earthquake is of great concern. For example, after a moderate (M = 5–6) earthquake in either the San Francisco or Los Angeles area, where large earthquakes are considered likely in the long term, there follows a transient probability gain for a ($M \geq 7$) earthquake of 10^2–10^3, relative to the respective regional long-term probabilities. In absolute terms, the conditional probability of such an earthquake initially ranges from a few tenths of a percent to a few percent, and decreases to half its initial value after about 1 day.

Two approaches have been taken to develop triggered stochastic models of earthquake occurrence. The first is based directly on the observed frequency of foreshock-mainshock (fs-ms) pairs. Empirical estimates the of fs-ms pairing rate and the distribution of fs-ms magnitude differences have been made with numerous earthquake catalogs [*Jones and Molnar*, 1979; *Von Seggern*, 1981; *Jones*, 1984; *Bowman and Kisslinger*, 1984; *Lindh and Lim*, 1995; *Abercrombie and Mori*, 1996; *Michael and Jones*, 1998]. These measurements, which characterize the clusters' transient, pairwise interactions, are analogous to the a value and b value in the Gutenberg-Richter (time-independent, noninteracting events) seismicity model. Such foreshock studies provided the basis for conditional probability estimates for the next Parkfield, California, earthquake [*Michael and Jones*, 1998] and for characteristic earthquakes on selected fault segments in California [*Agnew and Jones*, 1991].

In another approach, *Reasenberg and Jones* [1989, 1994] introduced a stochastic model of aftershock occurrence based on the observed rate, magnitude distribution, and temporal decay of 62 aftershock sequences following $M \geq 5$ mainshocks in California. Their model represents the rate of aftershocks of magnitude M or larger as

$$\lambda(t, M) = 10^{a+b(M_m-M)}(t + c)^{-p}$$

where t is time after the mainshock, M_m is the mainshock magnitude, and a, b, p, and c are constant parameters. Their

Paper number 1998JB900089.

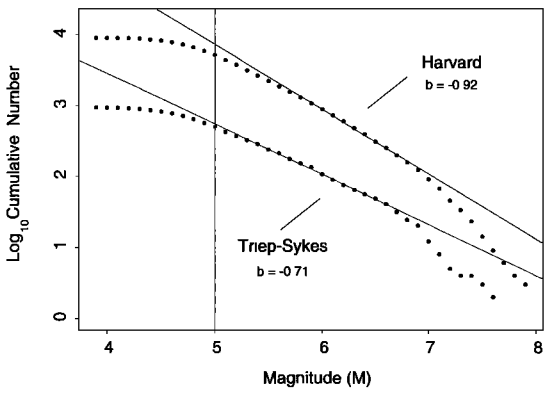

Figure 1. Magnitude distribution of earthquakes used, shown as $\log_{10} N(m \geq M)$. Data were drawn from Harvard catalog (1977–1996, 0–50 km depth) and Triep-Sykes compilation of shallow intracontinental earthquakes (1978–1994, 0–50 km depth). Gutenberg-Richter distributions (straight lines) represent least squares fits to the points shown over the range $5.8 \leq M \leq 6.8$ (Harvard) and $5.3 \leq M \leq 6.6$ (Triep-Sykes). Intersections with vertical line reveal cumulative deficits of earthquakes ($M \geq 5.0$) of approximately 28% and 10%, respectively, relative to models.

fit to the California data was achieved with parameter values $a = -1.67$, $b = 0.91$, $p = 1.08$, and $c = 0.05$ day. In applying this model, they allow the "aftershocks" to take on magnitudes larger than their "mainshock," in effect, applying an aftershock model to fs-ms pairing. This approach had the advantage that the model and its uncertainty were readily determined (even in a limited region such as California) owing to the large number of recorded mainshock-aftershock sequences. For estimating the expected numbers of $M \geq 3$ aftershocks following the 1994 Northridge, California, earthquake, the California generic model worked reasonably well [*Reasenberg and Jones*, 1994]. However, in the type of application for which the interest is greatest, namely, the case of potential $M \geq 5$ foreshocks prior to a larger earthquake in California, the model has not been directly tested through comparisons to observed earthquake occurrence because insufficient numbers of such pairs have been recorded. Recently, the question was raised as to whether the California generic model accurately calculates the average, short-term conditional probability of large earthquakes in California (California Earthquake Prediction Evaluation Council, minutes of meeting, February 28, 1997).

Here, I use two worldwide earthquake catalogs to estimate rates of fs-ms pairing and compare these rates to the corresponding probabilities calculated with the California generic model. Because the goal is to test a model derived from (and applied to) California seismicity, which is largely confined to the upper crust, I consider only shallow earthquakes in the global catalogs. The empirical foreshock rates are determined separately for mainshocks with $M \geq 6$ and $M \geq 7$; for mainshocks with thrust, normal, and strike-slip mechanisms; and for events located worldwide and events confined to intracontinental areas. Then the observed rates are compared to the California generic model probabilities and their ranges of uncertainty. Finally, I examine the magnitude differences among the fs-ms pairs involving $M \geq 7$ and $M \geq 6.5$ mainshocks in the Harvard centroid moment tensor (CMT) catalog and use them to constrain a model for their distribution.

2. Data

Data for this study were taken from the Harvard catalog of CMT solutions for large earthquakes [*Dziewonski et al.*, 1994] and from *Triep and Sykes'* [1997] compilation of shallow intracontinental earthquakes (also see http://www.ldeo.columbia.edu/seismology/triep/intra.expl.html, 1996). In both the Harvard CMT and the Triep and Sykes catalogs, events are listed with one or more magnitude estimates from among M_S and m_b (as originally reported by National Earthquake Information Service (NEIS) or International Seismological Centre (ISC), Edinburgh) and M_W (or M_0). For this study, I use M_S when it is given (90% of the shallow events) and m_b when M_S is not given (10% of the shallow events), and refer to this magnitude simply as M (Figure 1). The Harvard catalog is complete from 1977 to 1996 for $M_W \geq 5.6$–5.8 events [*Kagan*, 1997; *Molchan et al.*, 1997] and the Triep-Sykes catalog is complete from 1978 to 1994 for $M_W \geq 5.5$ events [*Triep and Sykes*, 1997]. Adjusting for apparent offset in the magnitude scales, these completeness levels correspond to $M \geq 5.6$–5.8 (Harvard catalog) and $M \geq 5.3$ (Triep-Sykes catalog).

For this study, I selected earthquakes from both catalogs with $M \geq 5.0$, below the completeness levels of the catalogs. This choice requires some justification. The motivation for it stems from the ultimate aim of the study, which is to test the assumption that an aftershock model [*Reasenberg and Jones*, 1989] may be used to forecast larger earthquakes (i.e., may be used as a fs-ms model). Because such an application is currently carried out by the U.S. Geological Survey (USGS) for all $M \geq 5$ earthquakes in California, I tailored the study to this magnitude range. The expected effect of choosing the foreshock magnitude cutoff below the completeness level is to underestimate the actual foreshock rate by an amount approximately equal to the fraction of events that are missing. The deficits were estimated graphically (see Figure 1) by comparing Gutenberg-Richter models (fit to each catalog above their respective completeness levels) to the actual cumulative counts ($M \geq 5.0$). The deficits were found to be approximately 28% (Harvard catalog) and approximately 10% (Triep-Sykes catalog). Hence the empirical foreshock rates determined below can be expected to be low by these amounts.

For the 5695 events ($M \geq 5.0$; depth $\leq$50 km) taken from the Harvard CMT catalog, I used the hypocentral locations listed as having originated from either NEIS or ISC. From 1977 to 1992 I used only ISC hypocenters (G. Smith, CMT associated with ISC, http://tempo.harvard.edu/~smith/CMT-ISC.html, 1995). Between 1993 and 1996, I used the combination of ISC and NEIC hypocenters listed in the Harvard catalog, which consists primarily of NEIS locations. Earthquake locations for the 647 ($M \geq 5.0$; depth $\leq$45 km) events taken from the Triep-Sykes catalog are as given by *Triep and Sykes* [1997]. Because many of the depths listed in both catalogs are fixed or poorly determined, depth information was used only in the selection of events to include in the study; calculated distances between events are epicentral distances. The choice of a 50 km depth cutoff (45 km in the case of the Triep-Sykes catalog), while somewhat arbitrary, accommodates our desire to work with "shallow" events and includes the fixed depths of 10 and 33 km assigned to a large number of events in the Harvard catalog for which more precise depth determination was impossible.

3. Definition of Foreshocks

Foreshock-mainshock pairs used in this study were defined according to the following rules.

1. The magnitudes of the first and second events in each pair (M_f and M_m, respectively) exceed fixed thresholds: $M_f \geq M_f^{\min}$ and $M_m \geq M_m^{\min}$.
2. The second event is not smaller than the first: $M_f \leq M_m$.
3. Interevent epicentral distance in a pair does not exceed dX, and the interevent time does not exceed dT.
4. When two or more mainshocks are themselves clustered within dX and dT of each other, only the largest is considered a mainshock.

In the estimates of foreshock rates below, tied magnitudes are allowed in the definition above, in conformance with the definition by *Reasenberg and Jones* [1989, 1994], so that the present results should be directly comparable to their model. But in the subsequent analysis of the distribution of magnitude differences among fs-ms pairs, I require that $M_f < M_m$, in conformance with the definition used by *Agnew and Jones* [1991], *Lindh and Lim* [1995], and *Michael and Jones* [1998], to facilitate comparison with those studies. For each qualifying mainshock, all qualifying events within the time dT and distance dX of it are identified as potential foreshocks. If more than one foreshock is thus associated with a given mainshock, only the largest foreshock is counted.

Rule 4 resulted in the removal of seven out of a total of 156 potential ($M \geq 7$) mainshocks identified in the Harvard catalog between 1978 and 1996. Two of these were magnitude ties ($M = 7$, which were treated arbitrarily as fs-ms pairs), two were foreshocks ($M = 7.3$ followed by $M = 7.5$ and $M = 7$ followed by $M = 7.3$), and three were $M = 7$ or $M = 7.1$ aftershocks. Their removal reduced the number of mainshocks available for analysis by 4%. For the $M \geq 6$ set, 15% of the potential mainshocks were deleted under rule 4.

The use of rule 4 to handle swarms has some undesirable consequences. The treatment of a swarm in real-time may depend strongly on small differences among early estimates of magnitude or epicenter (which may initially be large). Such sensitivity introduces ambiguity in the identification of the "mainshock" and hence ambiguity in the starting time of the model. This problem is seen in the treatment of the 1980 swarm at Long Valley, California, where the first event is identified here as the mainshock (Table 3), while the Caltech and Berkeley catalogs list the second and fourth main events, respectively, in the swarm as the largest. Since the four main events ($6.0 \leq M \leq 6.2$) in this sequence occurred over a 46-hour period, Omori's law of decay is probably a poor model for the hazard in this case. Such problems are inherent in this type of investigation; it is impossible for closely clustered events to define the fs-ms attribution unambiguously. One way out of this difficulty is through the application of the theory of stochastic processes, which allows one to treat all interearthquake interactions with a consistent algorithm. Unfortunately, the results of such approaches are difficult to interpret (Y. Kagan, written communication, 1998).

Current application of short-term forecasts by the USGS is consistent with rule 4 (D. Oppenheimer, personal communication, 1998). If a large earthquake were recognized to be an aftershock (for example, the 1992 Landers and Big Bear, California, events), it would be treated as an aftershock, not as a new mainshock; revised short-term probabilities based on its magnitude, time, and location would not be estimated. However, if a large earthquake were to spawn a short-term forecast and then be followed by a nearby, larger event, the short-term probabilities would be recalculated, based on the later (and larger) earthquake, and the forecast would be revised. Thus the largest event in a sequence is always used as the mainshock. Rule 4 parallels this procedure.

3.1. Choice of Spatial and Temporal Windows

The spatial and temporal windows used to define the foreshock-mainshock pairs were selected so as to capture the strongest and most obvious foreshock clustering activity in the Harvard CMT catalog. Figure 2 shows the earthquake pairing corresponding to three choices of M_f and M_m, with interevent separations up to 500 km and 60 days. Figure 3 shows the corresponding cumulative distributions of interevent time and distance among the pairs. At interevent distances greater than ~75 km and interevent times greater than ~10 days, the density of pairs apparent in Figure 2 is approximately uniform with respect to both time and distance, and the cumulative curves in Figure 3 are approximately linear, consistent with noninteracting earthquakes occurring uniformly in time and epicentrally distributed uniformly along fault zones [e.g., *Kagan and Knopoff*, 1980]. At interevent distances and times less than 75 km and 10 days, the dense clustering of points in Figure 2 and steep assent of the cumulative distributions in Figure 3 reveal the foreshock clustering process. The apparent range of fs-ms distance appears to be independent of magnitude of the mainshock; the median and 90 percentile interevent distances are approximately 15 and 30 km, respectively, for pairs involving $M \geq 5$, $M \geq 6$, and $M \geq 7$ mainshocks (Figure 4). The consistency of these ranges suggests that they are essentially controlled by location errors and should be considered upper bounds. *Bowman and Kisslinger* [1984] found an apparent interevent distance range (which includes an unstated location uncertainty) of 20–40 km for events located near Adak Island. *Ogata et al.* [1995] observed a range of apparent interevent distances between $M \geq 4$ foreshocks and mainshocks in the catalog of the Japan Meteorological Agency of 10–20 km, but they warned that event location errors may be contributing to these distances. Among observations of well-located events, there is evidence that foreshocks often occur much closer to their mainshocks. Using standard catalog locations, *Jones* [1985] found that most foreshocks ($M \geq 3$) in southern California located within approximately 1 km of the epicenters of their respective mainshocks. Using precise relative event locations (uncertainties of 0.1–0.3 km) obtained with a waveform correlation technique, *Dodge et al.* [1996] found that foreshocks before mainshocks ($4.7 \leq M \leq 7.3$) in California located ~1 km or less from their mainshock's epicenter. Returning to the present study, the median epicentral error in the Harvard CMT catalog was estimated by *Smith and Ekström* [1997] to be 25 km or greater for $M_w = 6.5$ events. Relative locations between foreshocks and mainshocks are probably better determined than this owing to common unmodeled velocity structure and station site effects. While we have neither relocated nor estimated uncertainty in the relative locations of these event pairs, it is reasonable to assume that at least some of the apparent interevent range of 15–30 km seen here reflects location errors. Clearly, the 75 km window used here is larger than necessary and was adopted to ensure that badly mislocated events would not be excluded from the analysis; it does not imply that the actual event separations are this large. To

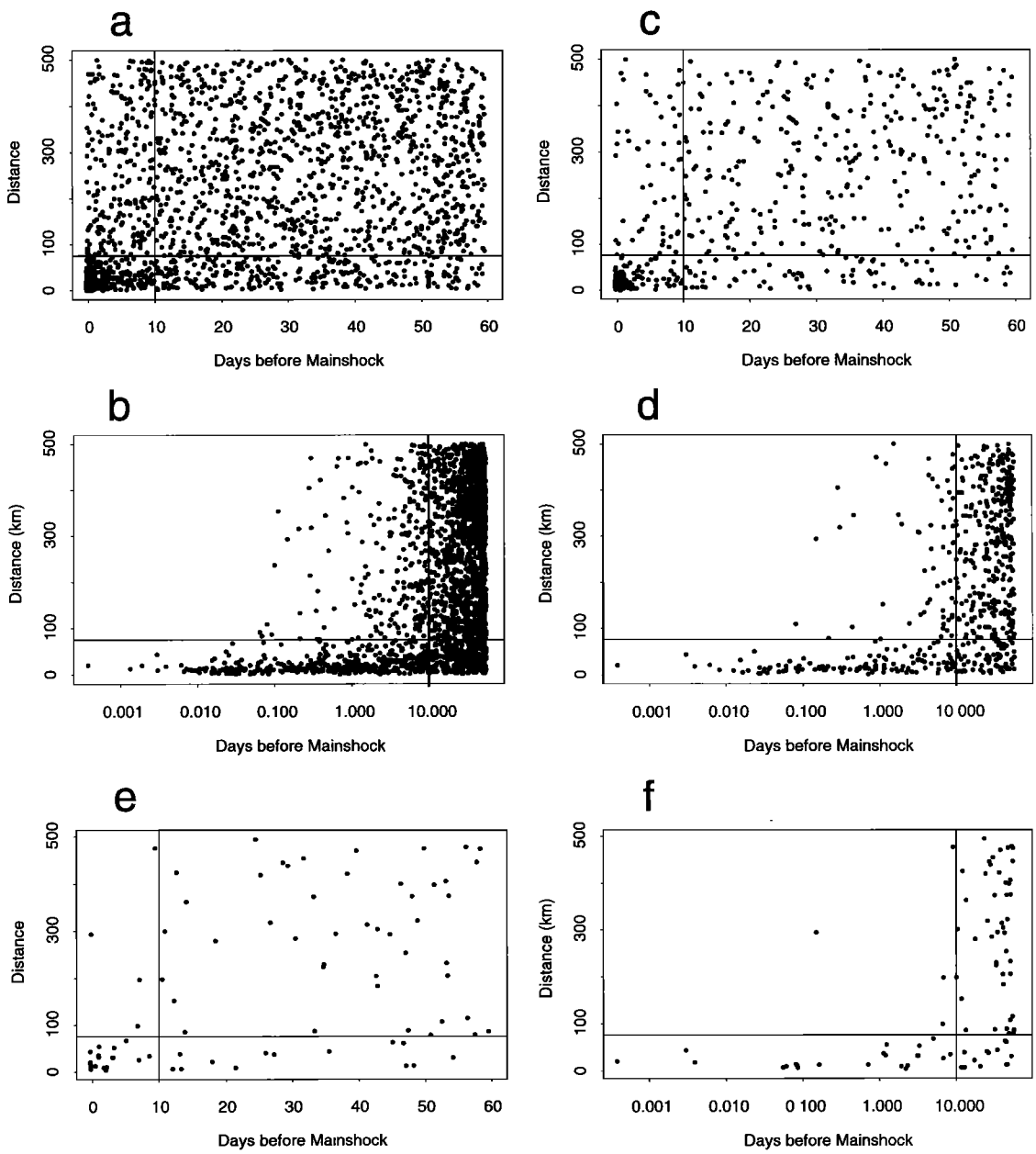

Figure 2. Foreshock-mainshock pairs in the Harvard CMT catalog (1977–1994) for various mainshock and foreshock magnitude thresholds, shown as a function of the pair's interevent distance and interevent time, plotted with (a, c, e) linear time axis and (b, d, f) log time axis. (a, b) Pairs with foreshock magnitude $M_f \geq 5$ and mainshock magnitude $M_m \geq 5$; $N = 1966$. (c, d) Pairs with $M_f \geq 5$ and $M_m \geq 6$; $N = 562$. (e, f) Pairs with $M_f \geq 5$ and $M_m \geq 7$; $N = 84$. The large ranges of interevent time and distance shown here reveal both the foreshock-related clustering and background or "chance" pairings. Pairs with interevent times <10 days and distances <75 km (the lower left quadrants formed by the solid lines) were used in the subsequent foreshock rate analysis.

correct the resulting assays for this oversized window, a "background" rate, corresponding to chance pairings, must be estimated and removed from the pair counts.

3.2. Correction for "Background" Events

A certain number of earthquakes are expected to fall within our interevent distance and time windows due to the "background" seismicity in the regions of the mainshocks. The effect of the background seismicity is seen in the linear trends in Figure 3 at large interevent times and distances. To correct for this effect, background rates were estimated from the seismicity by counting earthquakes located within 75 km of each mainshock and occurring during the time period 300–100 days before each mainshock. Based on these rate estimates, the numbers of background earthquakes expected to fall into the actual foreshock windows were subtracted from the counts of possible foreshocks. This method may slightly overestimate the background rate because $M \geq 5$ aftershocks, which were not removed from the catalog, could be included in the background count, while foreshocks were limited to one per main-

shock. However, inspection of the background events revealed none to be $M \geq 5$ aftershocks.

3.3. Focal Mechanisms

Both the Harvard CMT and Triep-Sykes catalogs provide detailed information about the focal mechanism of the earthquakes. For each catalog, I categorized the mainshocks as either thrust, strike-slip, or normal. For events in the Triep-Sykes catalog, I used the mechanism types given in the catalog. For the Harvard CMT catalog, I categorized events according to the plunge of the tension axis (p1) and plunge of the null axis (p2), as done by *Triep and Sykes* [1997]: thrust (p1 > 45°), normal (p1 < 45° and p2 < 45°), and strike-slip (p2 > 45°). These definitions include some oblique thrust events as "thrust" and some oblique normal events as "normal." The use of the plunges of principal axes for classifying the focal mechanism leads to some clear misclassifications. These arise because of fault-plane ambiguity in the solution. In cases where the slip plane is known, the error can be obvious. For example, in Table 3, the 1979 Imperial Valley earthquake is listed as strike-slip because it is known that the northwest striking plane is the slip plane; classification using the axis plunges would put it in the normal group. It is not known how many other events "crossed over" from one mechanism category to another because information about the identity of the slip plane was not used in the analysis. Presumably, such crossovers occur in both directions, however, so their net effect on our assays may be minor. Alternative methods for classifying and displaying earthquake focal mechanisms using catalogs of centroid moment tensor solutions have been proposed and applied to some worldwide data sets [*Frohlich*, 1992; *Frohlich and Apperson*, 1992; *Sipkin*, 1993; *Kaverina et al.*, 1996].

Table 1 shows the distribution of earthquake focal mechanisms among the mainshocks used in this study. To characterize the fs-ms pairs, I used the mainshock's focal mechanism rather than the foreshock's, assuming it to be more representative of the regional faulting style. Thus, in interpreting the results, this classification may be applied to the region in which a potential foreshock is located but may not be used to condition the probability of a mainshock based on the potential foreshock's focal mechanism.

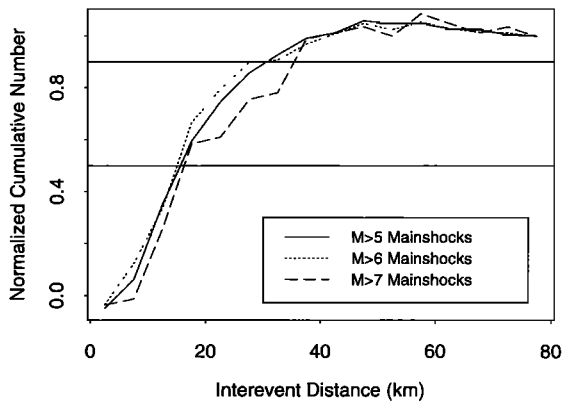

Figure 4. Normalized cumulative distributions of interevent distances among foreshock-mainshock pairs shown in Figure 3b, after correcting for "background rate." Background rate is defined as the average spatial density of pairs observed in each data set over the interevent distance range 200–500 km (and corresponds to the slopes of the flat portions of curves in Figure 3b). Removal of these constant background rates accounts for the decreases in "cumulative" number seen in the upper tails of the distributions. Horizontal lines indicate the median and 90 percentile levels. For all three data sets (mainshocks with $M \geq 5$, $M \geq 6$, and $M \geq 7$), the median and 90 percentile interevent distances are approximately 15 and 30 km, respectively. Numbers of pairs represented are 518, 161, and 26 for the $M \geq 5$, $M \geq 6$, and $M \geq 7$ mainshocks, respectively.

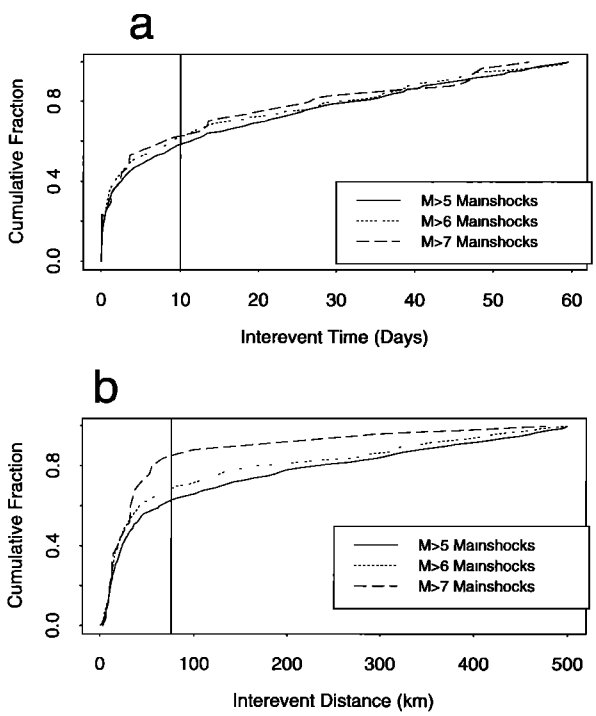

Figure 3. Interevent distances and times among foreshock-mainshock pairs found in the Harvard CMT catalog. (a) Cumulative distribution of interevent times among foreshock-mainshock pairs with interevent distances of 75 km or less. (b) Cumulative distribution of interevent distances among foreshock-mainshock pairs with interevent times of 10 days or less. Background seismicity, which contributes to the linear trends at large interevent times and distances, is not removed in this figure. Vertical lines indicate windows of 10 days and 75 km used in calculating foreshock rates.

Table 1. Earthquakes Used in the Study, Classified According to Magnitude and Focal Mechanism

Mainshock Category	$M \geq 5$		$M \geq 6$		$M \geq 7$	
	N	%	*N*	%	*N*	%
Harvard Catalog (1977–1996, depth ≤50 km)						
Thrust	3083	47.0	658	50.8	100	64.1
Normal	1362	20.7	208	16.1	15	9.6
Strike-slip	2120	32.3	429	33.1	41	26.3
All mechanisms	6565	100.0	1295	100.0	156	100.0
Triep-Sykes Catalog (1978–1994, depth ≤50 km)						
Thrust	253	39.2	60	39.5	13	54.2
Normal	154	23.8	34	22.4	2	8.3
Strike-slip	239	37.0	58	38.2	9	37.5
All mechanisms	646	100.0	152	100.0	24	100.0

Percentages shown are relative to total number of events in each magnitude class.

Table 2. Summary of Retrospective Foreshock Rates

Mainshock Category	Number of Mainshocks[a,b]	Number of Possible Foreshocks[b,c]	Expected Number of Background Events	Fraction of Mainshocks Preceded by a Foreshock,[d] %	95% Confidence Range for Foreshock Frequency,[e] %	Foreshock Rate per Unit Magnitude Difference, %
			Harvard CMT Catalog ($M \geq 6$) 1978–1996			
All	1108	161	14.4	13.2	11.3–15.4	13.2
Thrust	533	103	9.7	17.5[f]	14.4–21.0	17.5
Normal	172	23	2.1	12.2	7.7–18.0	12.2
Strike-slip	397	34	2.4	8.0[g]	5.5–11.1	8.0
			Harvard CMT Catalog ($M \geq 7$) 1978–1996			
All	149	26	1.5	16.5	10.9–23.4	8.3
Thrust	95	19	1.2	18.7	11.5–28.0	9.4
Normal	13	3	0.1	22.3	···	11.2
Strike-slip	41	4	0.3	9.0	···	4.5
			Triep-Sykes Compilation of intracontinental events ($M \geq 6$) 1979–1994			
All	116	17	0.3	14.4	8.6–22.1	14.4
Thrust	41	7	0.1	16.8	7.0–31.8	16.8
Normal	27	5	0.05	18.3	···	18.3
Strike-slip	47	5	0.15	10.3	···	10.3

[a]After removing "clustered" mainshocks according to rule 4 (see text).
[b]Event counts in thrust, normal, and strike-slip categories may not sum to count in "All" due to exclusive definitions of categories used (see text).
[c]A spatial window of 75 km and a temporal window of 10 days were used.
[d]Corrected for estimated background rate.
[e]Confidence range is for an observed frequency in a binomially distributed population.
[f]Foreshock rate among thrusts is greater than foreshock rate among all $M \geq 6$ earthquakes in the Harvard CMT at the 95% confidence level.
[g]Foreshock rate among strike-slip events is lower than foreshock rate among all $M \geq 6$ earthquakes in the Harvard CMT catalog at the 99% confidence level.

4. Retrospective Foreshock Frequencies

The retrospective foreshock frequency associated with a set of mainshocks is defined as the fraction of the mainshocks that are preceded by a foreshock. To estimate this rate, I used time and distance windows of 0–10 days and 0–75 km, set the foreshock magnitude threshold $M_f^{\min} = 5$, and considered mainshock magnitude thresholds of $M_m^{\min} = 5$, 6, and 7. The resulting sets of fs-ms pairs have magnitude differences distributed between zero and an upper limit approximately given by $\Delta M = M_m^{\min} - M_f^{\min}$. I call ΔM the magnitude difference aperture. The larger the aperture, the greater is the number of foreshocks that will be counted. This is true for any assumed distribution of magnitude difference. Therefore, to quantitatively compare different foreshock assays, it is necessary to normalize for the magnitude apertures used. Other differences among studies, including the choice of time and spatial windows used to define the foreshocks, are of less importance because fs-ms pairs are tightly clustered in space and time, and thus self-defining. Oversized windows include too many background events, but this error can be approximately corrected for. However, the case for clustering of magnitude differences into a distribution with a central tendency is not supported by most previous studies or by the current one (see discussion below); consequently, variations in the observable range of magnitude difference (aperture) can be expected to have a first-order effect on the resulting assay.

To compare the present results to other studies, I assume that the magnitude differences among the fs-ms pairs have a uniform distribution over their range of observation (or aperture) and calculate a "unit foreshock rate" by dividing the apparent foreshock rate by the magnitude difference aperture. In section 6 I show that the actual distribution of magnitude differences among the fs-ms pairs in the Harvard CMT catalog is consistent with this assumption.

The $M \geq 6$ mainshocks in the Harvard CMT catalog between 1978 and 1996 ($N = 1108$) have a foreshock rate ($M \geq 5$ foreshocks, corrected for estimated background seismicity) of 13.2% (Table 2). The thrust earthquakes among these ($N = 533$) have significantly higher rate (17.5%), and the strike-slip earthquakes ($N = 397$) have a significantly lower rate (8.0%). Among the $M \geq 7$ mainshocks in the Harvard CMT catalog ($N = 149$), both the overall frequency of foreshocks (16.5%) and the departures from this rate observed among subsets defined according to focal mechanism are similar to the rates in the $M \geq 6$ set. However, because of the smaller numbers of fs-ms pairs, the differences among the $M \geq 7$ subsets are not statistically significant.

The overall foreshock rate of 13.2% found for $M \geq 6$ mainshocks preceded by $M \geq 5$ foreshocks (13.2% per magnitude unit) is comparable to results obtained in other studies. *Jones* [1984] found that seven out of 20 ($M \geq 5$) earthquakes occurring in California between 1966 and 1980 were preceded by ($M \geq 2$) foreshocks (a foreshock rate density of 12% per magnitude unit). *Jones and Molnar*'s [1979] study of worldwide ($M \geq 5$) foreshock activity before 161 $M \geq 7$ mainshocks (1914–1973) detected a rate of 24.8%, or approximately 12% per magnitude unit. More recently, *Abercrombie and Mori* [1996] examined 59 ($M \geq 5$) mainshocks in California and Nevada and found that 26 of them were preceded by ($M \geq 2$) foreshocks (15% per magnitude unit). *Michael and Jones* [1998] looked at 33 $M \geq 5$ strike-slip earthquakes along the San Andreas fault physiographic province and found that 17 were preceded by ($M \geq 2$) foreshocks, corresponding to a rate density of 17% per magnitude unit. A similar result was also

reported by *Lindh and Lim* [1995] for the same region, time period, and magnitude ranges. *Bowman and Kisslinger* [1984] determined that 14% of $m_b \geq 5.0$ earthquakes in the Adak thrust zone were preceded by foreshocks. While they estimated a completeness threshold of $M_f^{\min} = 2.5$ (m_b), inspection of their Figure 2 suggests an alternate threshold of, perhaps, $m_b = 3.8$, which would correspond to a foreshock rate density of about 12% per magnitude unit. Variations among these studies involving the region studied and the choice of spatial and temporal windows make it difficult to compare these results. Given these differences, the range of results is surprisingly narrow, from 12% to 17% per magnitude unit. This range coincides closely with the 95% confidence range estimated here for foreshock rate among $M \geq 6$ mainshocks in the Harvard catalog) and may be considered a robust, worldwide meta-result.

The present results for $M_m^{\min} = 7$ do not follow this trend, however, and instead indicate a foreshock rate density of between 5.5% and 11.7% per magnitude unit (95% confidence range) (Table 2). However, because of the small number of $M \geq 7$ mainshocks ($N = 149$), these rates are not significantly different from the corresponding rates determined with the more numerous, $M \geq 6$ events. Similarly, the foreshock rates observed among the $M \geq 6$ mainshocks in the Triep and Sykes catalog ($N = 116$) are indistinguishable from the Harvard catalog $M \geq 6$ rates.

Comparison of the present results to those of *Von Seggern* [1981] is done with the recognition that Von Seggern's analysis did not adequately correct for background seismicity. While Von Seggern's assay was known to include background events, he did not attempt to quantify their presence. Instead, he interpreted the relative rates but cautioned against assigning much significance to the absolute rates. He concluded that "the rate ... for shallow foreshocks is significant but still small," and that "the true rate of foreshock occurrence was less than 20%." Von Seggern's Table 4 reports a foreshock rate of 18% (nine out of 51) among $M_S \geq 7$ mainshocks of all kinds (mostly 0–100 km depth). This result may be compared most directly to the apparent rate of foreshocks before $M \geq 7$ mainshocks (19 out of 95, or 20%) observed in the present study before the estimated rate of background activity is removed (Table 2). The spatial and temporal distribution of foreshocks relative to their mainshocks observed by Von Seggern is similar to those seen in this study. Most of Von Seggern's foreshocks occurred within 10 days of their mainshocks, and most fell within 0.5° of the mainshocks.

4.1. Dependence on Focal Mechanism

Foreshock rate varied significantly among subsets of mainshocks defined by their focal mechanisms. Among the $M \geq 6$ thrust earthquakes in the Harvard CMT catalog, 17.5% were preceded by a $M \geq 5$ foreshock, this rate being significantly higher (95% confidence) than the corresponding rate of 13.2% observed before all $M \geq 6$ earthquakes. Also, strike-slip mainshocks were preceded by foreshocks only 8.0% of the time, a rate significantly lower than that for all ($M \geq 6$) events at the 99% confidence level. Thus foreshocks occurred before shallow $M \geq 6$ thrust events at approximately twice the rate as they did before shallow $M \geq 6$ strike-slip events.

The relatively high foreshock rates observed before thrust earthquakes appears to contradict a result of *Abercrombie and Mori* [1996, Figure 2b], who reported a lower frequency of foreshocks before thrust mainshocks than before strike-slip mainshocks ($M \geq 5$) in California and western Nevada. They also showed that the thrust events tended to be deeper than strike slip events in California and that deeper events tend to have fewer foreshocks than shallower ones. Hence their result might be interpreted as reflecting a dependence of foreshock rate on depth rather than on focal mechanism. In the Harvard catalog, as in California, thrust events tend to be deeper (median depth 33 km) than strike-slip events (median depth 15 km) in this study. But the foreshock rate is higher among thrust events, unlike California.

The present result appears to contradict *Jones* [1984], who found no foreshocks preceding four ($M_L \geq 5$) thrust mainshocks in southern California and who tentatively suggested that the foreshock rate for thrust earthquakes in California might be lower than the rate for strike-slip events (15% per magnitude unit, based on 16 mainshocks), although they cautioned that their data set was too small to definitively distinguish between these rates.

Also, my result appears at first to contradict *Jones and Molnar* [1979], who concluded that there was a higher foreshock rate before large earthquakes in nonsubduction zones than in subduction zones. They reported a foreshock rate of 20% (10% per magnitude unit) among 120 $M \geq 7$ earthquakes in subduction zones, and a rate of 29.3% (15% per magnitude unit) among 41 $M \geq 7$ events in nonsubduction zones. (Here I assumed a magnitude aperture of 2 in Jones and Molnar's study.) However, there are too few data in this study to infer a significant rate difference with confidence; the 90% confidence ranges are 7–14% per magnitude unit (subduction zone events) and 9–22% per magnitude unit (nonsubduction zone events).

In spite of the difficulties in comparing these studies, the apparent discrepancies between them and the present worldwide result suggest that California foreshocks may be a special case, owing to the particular regional tectonics there and are not representative of worldwide foreshock rates. This possibility is explored below.

4.2. Subduction Zones and Intracontinental Regions

It is well known that worldwide catalogs are dominated by subduction zone events, and that these events extend in depth well below the crust. Even when only shallow earthquakes are considered, as in this study, shallow subduction zone events are a significant constituency. If foreshocks occur at significantly different rates in subduction zones and continental regions, a worldwide assay will be unrepresentative of the continental areas. While most earthquakes in the world occur in subduction zones, most earthquakes in California (with the exception of rare, large events in southern Cascadia, which were not included in the studies cited herein) occur in continental crust. To investigate the dependence of foreshock rates on location in shallow subduction zones and continental crustal regions, I used *Triep and Sykes'* [1997] worldwide catalog of shallow (depth ≤45 km) intracontinental earthquakes.

Of the 152 $M \geq 6$ earthquakes in the Triep-Sykes catalog (1978–1994), 116 satisfied all of the requirements necessary for their inclusion as mainshocks in the foreshock assay. Approximately equal numbers of these events have thrust and strike-slip focal mechanisms (Table 1). For the working set of 116 mainshocks in the Triep-Sykes catalog of intracontinental earthquakes, a foreshock rate of 14.7% was obtained, a rate not significantly different from the rate in the Harvard $M \geq 6$ data set. A breakdown of the foreshock rate according to the

focal mechanism of the Triep-Sykes mainshocks yielded small numbers of foreshocks in each category and statistically insignificant departures from the overall rate, from 10.3% for strike slip events to 18.3% for normal events (Table 2). Again, however, strike-slip events displayed the lowest rate of foreshock occurrence. The foreshock rates in the Triep and Sykes catalog do not significantly differ from their counterparts in the Harvard CMT catalog.

4.3. Subduction Zones and Continental Thrusts

In the Harvard $M \geq 6$ assay, 17.5% of the thrust-type mainshocks were preceded by foreshocks, and most of these are located in the shallow portions of the major circum-Pacific subduction zones (Figure 5a). These events involve the subduction of water-saturated oceanic crust in the uppermost portion of descending slabs and have a median depth of 30 km. In contrast, only three foreshocks were found among the 35 thrust earthquakes in the Himalayan collision belt involving the Indo-Australia and Eurasian plates, and none were found before the seven thrust (and oblique thrust) earthquakes in California. These events are mountain-building earthquakes involving the transpression of continental, crystalline rock and have a median depth of 15 km. The contrast suggests that the foreshock rate may be higher in shallow subduction zones than in continental thrust belts, although the number of events in these regional subsets are generally too small to prove it. This could explain both the high foreshock rate for thrust earthquakes observed in this study and the low rates observed among California thrust earthquakes [*Jones*, 1984; *Abercrombie and Mori*, 1996].

Foreshocks preceding normal mechanism earthquakes occurred along the East African rift system, on the spreading centers in the Indian Ocean, and along the western Pacific subduction zones from New Zealand to Japan (Figure 5b). Neither of the $M \geq 6$ normal mechanism mainshocks in California had a foreshock (Table 3).

Most of the foreshocks that preceded ($M \geq 6$) strike-slip mainshocks were along the convergent boundary between the Pacific and Indo-Australian plates, from the New Guinea Thrust to the New Hebrides Thrust (Figure 5c). Three were located in California (Table 3). In other regions, few strike-slip earthquakes were preceded by foreshocks.

5. Prospective Foreshock Frequencies

From the perspective of real-time, short-term earthquake forecasting, it is the probability of earthquakes following earthquakes, not preceding them, that is of interest. The prospective foreshock frequency $F_p(m_1, m_2, T, R)$ is defined as

$$F_p = \frac{N(m_1, m_2, T, R)}{N(m_1)}$$

where $N(m_1, m_2, T, R)$ is the number of earthquake pairs (each pair consisting of an event with magnitude $M \geq m_1$ followed by a larger event with $M \geq m_2$, with interevent times up to T and interevent distances up to R) and $N(m_1)$ is the number of individual events in the catalog with $M \geq m_1$. These frequences were calculated for $m_1 = 5$; $m_2 = 5, 6, 7$; $R = 75$ km; and $0.25 \leq T \leq 10$ days and compared to the corresponding probabilities calculated with a generic model of aftershock activity in California [*Reasenberg and Jones*, 1989, 1994].

The observed frequency of larger ($M \geq 5$) earthquakes following an $M \geq 5$ earthquake is 7.5% in a week, approximately 25% lower than the California generic model (Figure 6a). In contrast, the observed frequency of $M \geq 6$ earthquakes following an $M \geq 5$ earthquake is 2.3% in a week, approximately twice the generic model and outside its 95% confidence range (Figure 6b), and the observed frequency of $M \geq 7$ earthquakes following a $M \geq 5$ earthquake is approximately 0.4%, again approximately twice the generic model and outside its 95% confidence range (Figure 6c). Why are the observed global foreshock rates double the California model rates for $M \geq 6$ and $M \geq 7$ mainshocks but not for $M \geq 5$ mainshocks? The reason may be due, in part, to catalog incompleteness in the range $5.0 \leq M \leq 5.7$, which directly enters into the expression for F_p when $m_1 = m_2 = 5$ but essentially cancels out when $m_2 = 6$ or $m_2 = 7$. In Figure 6a the cross indicates the probability usually stated in USGS forecasts in California after an $M \geq 5$ mainshock ("10% probability of an equal or larger event in the next 7 days"). While the corresponding frequency observed in the Harvard catalog is 7.5%, extrapolation from the $M \geq 6$ and $M \geq 7$ results suggests that the actual worldwide rate may be as high as 15%.

The observed frequencies of $M \geq 6$ and $M \geq 7$ earthquakes in the Harvard catalog were used to foward fit a new model (which I call the "world generic model") having the same form as that of *Reasenberg and Jones* [1989] but defined by new parameter values ($a = -1.5$, $b = 0.8$, $p = 1.0$, and $c = 0.05$). The new parameter values were obtained by informal (trial and error) fitting of a, b, and p, based on visual comparisons with data points in Figures 6b and 6c. Numerically, the biggest difference between the California and world models is in the a value. But there are important other differences as well. The world generic model may be better determined than the California generic model because it is based on 171 fs-ms pairs, compared to only 62 aftershock sequences in the *Reasenberg and Jones* [1989] compilation. In addition, the world generic model directly reflects the behavior of large earthquakes involved in fs-ms sequences, while the California generic model is an extrapolation from smaller aftershock activity.

6. Foreshock-Mainshock Magnitude Differences

I investigate here the distribution of magnitude differences among fs-ms pairs in the Harvard catalog (Figure 7). While all observed pairs for which $M_m \geq 6$ and $M_f \geq 5$ are included in this figure, only those within the range of complete observation ($M \geq 5.6$–5.8) are suitable for this analysis. I selected two subsets for analysis, indicated by the parallelogram-shaped boxes in Figure 7: pairs with $M_m \geq 7$ and $M_f \geq M_m - 1.5$, and pairs with $M_m \geq 6.5$ and $M_f \geq M_m - 1.0$. The first set provides a wider magnitude difference aperture but includes fewer pairs ($N = 19$), while the second provides a smaller aperture but includes more data ($N = 46$). In order to make the magnitude apertures as wide as possible, these subsets were permitted to include some foreshocks with magnitude slightly below the completeness level. For the $M_m \geq 7$ set (Figure 8a) the distribution of magnitude differences appears rather uniform; a Kolmogoroff-Smirnoff test fails to reject the uniform distribution with 80% confidence ($N = 19$, $D = 0.15$). Among the more numerous $M_m \geq 6.5$ set of pairs (Figure 8b) there is a slight apparent deficit of pairs with small-magnitude differences, but a uniform distribution still cannot be rejected with 80% confidence ($N = 46$, $D = 0.12$).

To put these results in perspective, I compare them to others obtained for data having different mainshock magnitude

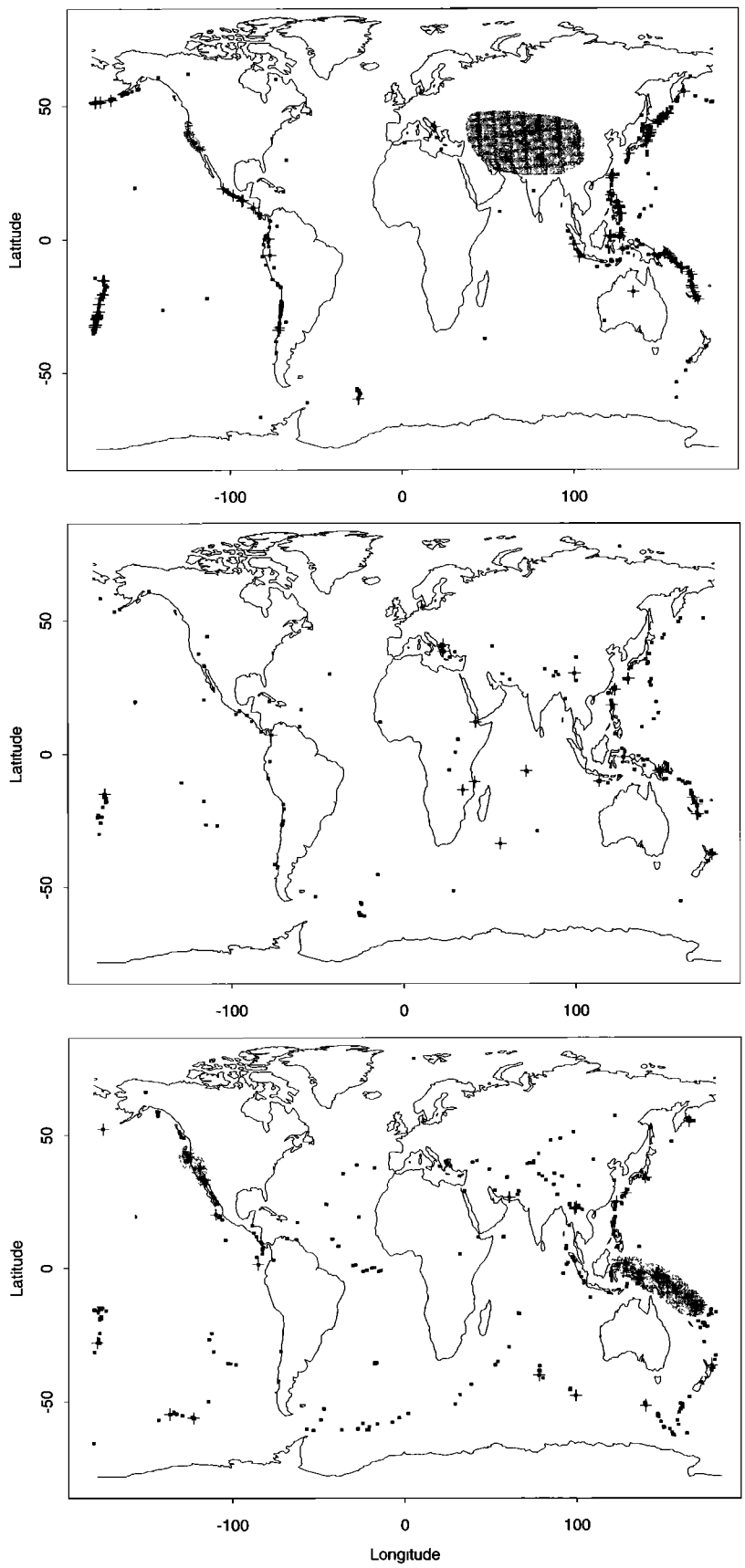

Figure 5. Locations of Harvard catalog $M \geq 6$ mainshocks (1978–1996) and their ($M \geq 5$) foreshocks, separated according to the focal mechanism of the mainshock. Solid squares indicate mainshocks; plus signs indicate foreshocks. (a) Thrust events. Shaded areas, which mark continental thrust zones, have a lower foreshock rate (three out of 35 mainshocks) than is found along the shallow portions of the major circum-Pacific subduction zones. (b) Normal events. (c) Strike-slip events. Shaded areas along the New Hebrides and New Guinea thrusts (presumably including strike-slip events on transform faults associated with the subduction zone) and the San Andreas fault zone have average or higher foreshock rates, compared to that for strike-slip events in the rest of the world.

Table 3. Harvard Catalog ($M \geq 6$) Earthquakes Located in California, 1978–1996

Region	Year	Month	Day	Hour	Minute	Magnitude	Depth	Latitude	Longitude	p1	p2	p3	rake1	rake2	Foreshock
Events With Thrust or Oblique Thrust Mechanism															
Coalinga	1983	5	2	23	42	6.6	7	36.24	−120.27	75	1	15	87	91	...
N. Palm Springs	1986	7	8	9	20	6.1	8	34.01	−116.73	47	34	23	156	55	...
Whittier Narrows	1987	10	1	14	42	6.0	17	34.05	−118.16	76	4	14	98	85	...
Loma Prieta	1989	10	18	0	4	7.1	8	37.06	−121.79	49	36	18	29	128	...
Honeydew	1991	8	17	19	29	6.1	12	40.27	−124.13	60	14	27	51	104	...
Cape Mendocino	1992	4	25	18	6	7.1	15	40.36	−124.05	53	3	37	68	93	...
Northridge	1994	1	17	12	30	6.8	18	34.21	−118.54	73	16	6	65	111	...
Events With Normal or Oblique Normal Mechanism															
Long Valley	1980	5	25	16	33	6.1	0	37.60	−118.80	24	5	66	−103	−85	...
Westmorland	1981	4	26	12	9	6.0	6	33.10	−115.66	26	44	35	−8	−135	...
Events With Strike-Slip Mechanism															
Imperial Valley[a]	1979	10	15	23	16	6.9	30	32.86	−115.46	33	39	33	−180	−51	...
Baja Calif.	1980	6	9	3	28	6.5	24	32.32	−114.92	0	90	0	180	0	...
W. of Eureka	1980	11	8	10	27	7.3	5	41.15	−124.30	20	62	18	178	27	...
S. Barbara Channel	1981	9	4	15	50	6.0	5	33.89	−119.04	0	90	0	180	0	...
Morgan Hill	1984	4	24	21	15	6.3	8	37.26	−121.72	10	76	9	179	0	...
Mendocino Fracture Zone	1984	9	10	3	14	6.7	5	40.39	−126.80	18	66	15	178	24	...
Chalfant Valley	1986	7	21	14	42	6.3	4	37.53	−118.44	10	56	32	−163	−31	M = 5.7, dT = 1.0 day, dX = 3.5 km
Superstition Hills	1987	11	24	13	15	6.6	2	33.12	−116.02	10	78	7	178	12	M = 6.3, dT = 0.5 day, dX = 36 km
Crescent City	1991	8	17	22	17	7.0	14	41.85	−125.41	22	62	16	176	28	M = 6.2, dT = 1.0 day, dX = 15 km
Joshua Tree	1992	4	23	4	50	6.3	12	33.91	−116.48	1	87	3	−1	−177	...
Landers	1992	6	28	11	57	7.6	1	34.25	−116.48	9	68	20	−172	−20	...
Mendocino Fracture Zone	1994	9	1	15	15	7.0	10	40.41	−125.65	21	64	15	176	25	...
Mendocino Fracture Zone	1995	2	19	4	3	6.8	10	40.56	−125.53	10	72	14	−177	−18	...

[a]The 1979 Imperial Valley earthquake was initially identified here as an oblique normal mechanism based on the plunges of the solution axes given in the Harvard catalog. The Harvard solution incorrectly resolves a dip on the northwest striking plane (with rake = −180), which causes a shallowing of p2. Local network observations [*Johnson and Hutton*, 1982] favor strike-slip motion on a vertical northwest striking plane, prompting its listing here among the strike-slip events.

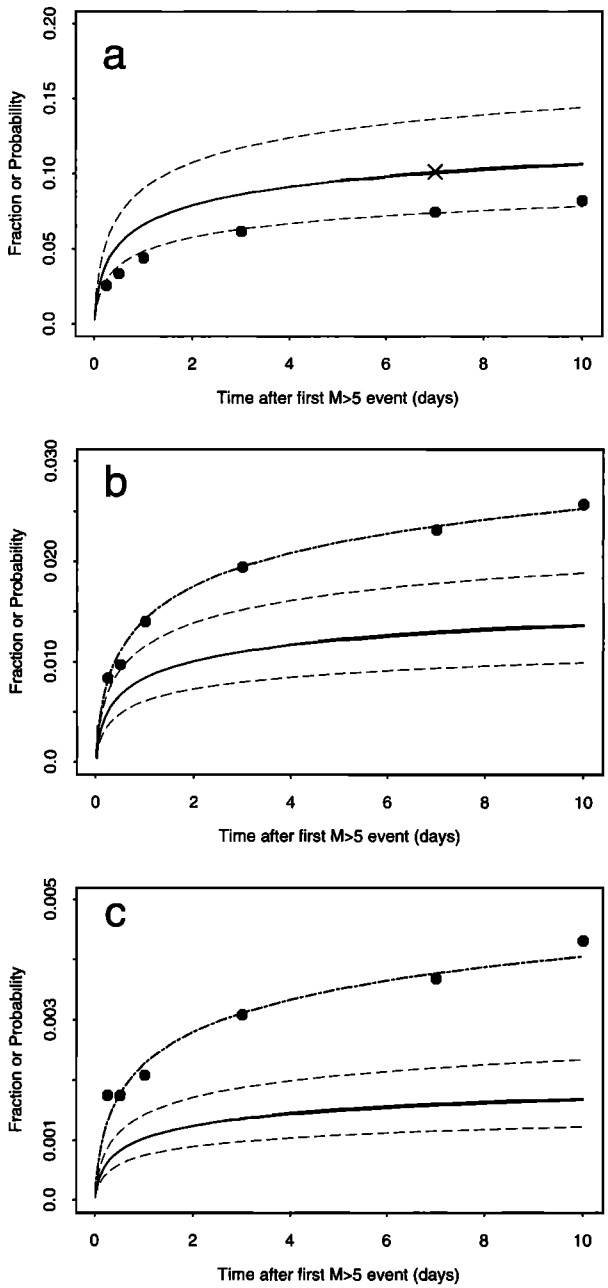

Figure 6. Observed cumulative frequencies (solid circles) and predicted probabilities based on California generic clustering model (solid lines) of larger earthquakes occurring after an $M \geq 5$ earthquake, as a function of the time after the $M \geq 5$ earthquake. Dashed lines indicate the 95% confidence range of the California generic model. Observed frequencies are calculated at times 0.25, 0.5, 1, 3, 7, and 10 days after the first ($M \geq 5$) event. (a) Case of any larger earthquake following an $M \geq 5$ event. (b) Case of an $M \geq 6$ earthquake following an $M \geq 5$ event. (c) Case of an $M \geq 7$ earthquake following an $M \geq 5$ event. Cross in Figure 6a marks a generic forecast often announced by the USGS after significant California earthquakes of "a 10% chance of a larger event in the next week." In Figures 6b and 6c, dash-dot lines indicate the world generic model (defined by the model parameters $a = -1.5$, $b = 0.8$, $p = 1.0$, and $c = 0.05$), which was foward-fitted to the data points shown.

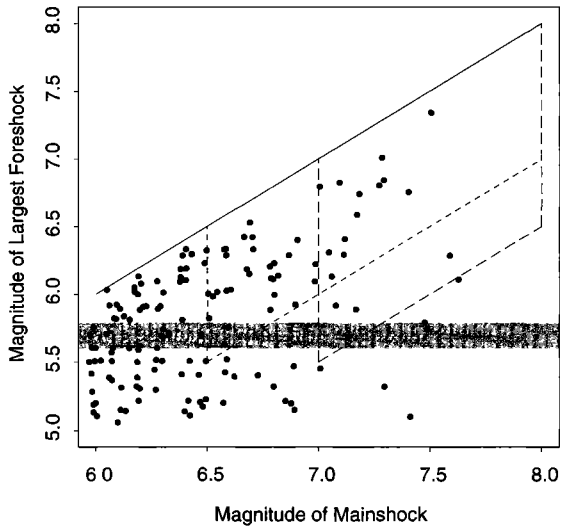

Figure 7. Magnitudes of foreshocks are shown in relation to the magnitudes of their respective mainshocks. Gray bar marks completeness level of Harvard catalog. Boxes drawn with dashed lines define subsets of pairs used to estimate distribution of magnitude differences (see Figure 8). Compare with Figure 9 of *Jones and Molnar* [1979].

ranges. The present results agree with *Agnew and Jones* [1991], who inferred from a set of several hundred fs-ms pairs ($M \geq 3$ mainshocks) in southern California that all foreshock magnitudes are equally likely for a given mainshock magnitude. The present results also agree with *Michael and Jones* [1998], who considered 38 fs-ms pairs in California (predominantly strike-slip, $M \geq 5$ mainshocks). They found that for the 29 pairs in their data set with magnitude differences up to 2.0 (the range for which complete observations of foreshocks were available), a uniform distribution for the magnitude differences could not be rejected with even 80% confidence. However, they also pointed out that, due to the small number of observations, their tests had little power and that their results depended on arbitrary choices of the binning interval. *Michael and Jones* [1998] also examined a larger set consisting of 60 $M \geq 4$ mainshocks with foreshocks up to 1 magnitude unit smaller than the mainshocks. An abundance of small-magnitude differences in this set allowed them to reject with 99.9% confidence a lognormal distribution for the magnitude differences. The present result disagrees with one reached by *Lindh and Lim* [1995], who considered 15 fs-ms pairs in the San Andreas fault zone between 1970 and 1995, involving predominantly strike-slip ($M \geq 5$) mainshocks. *Lindh and Lim* [1995] argued that the magnitude differences among these 15 pairs have a strong central tendency (located somewhere between 1.0 and 2.0, and that a lognormal distribution was preferable to the uniform distribution. The *Lindh and Lim* [1995] data set is the smallest among the studies cited here, and while it is qualitatively better fit with a lognormal than a uniform distribution, it is too small to confidently differentiate between these models.

7. Summary and Discussion

The overall foreshock rate of 13.2% per magnitude unit found here for all $M \geq 6$ earthquakes in the Harvard CMT catalog (1978–1996) lies within the range of estimates of foreshock rate density (12–17% per magnitude unit) obtained by

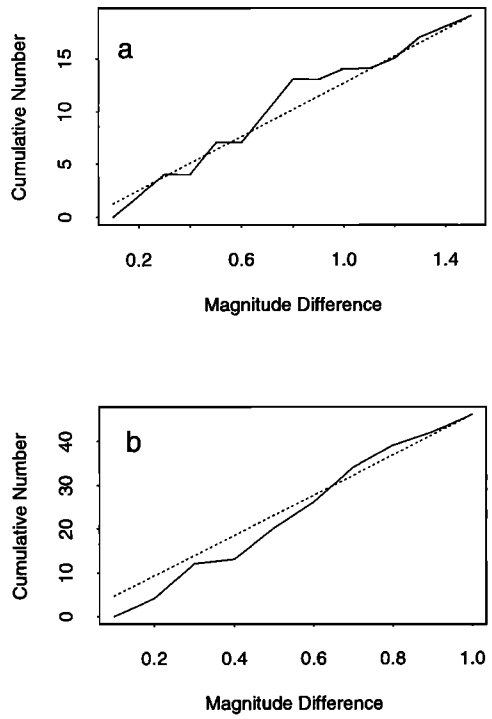

Figure 8. Cumulative number of foreshock-mainshock pairs, as a function of magnitude difference ($M_m - M_f$), observed among selected shallow earthquakes in the Harvard catalog, 1978–1996, is shown with solid line. Corresponding uniform distribution is shown with dashed line. (a) Pairs with $M_m \geq 7$ and $M_f \geq M_m - 1.5$ ($N = 19$). (b) Pairs with $M_m \geq 6.5$ and $M_f \geq M_m - 1.0$ ($N = 46$). For both sets of pairs (Figures 8a and 8b) the uniform distribution cannot be rejected by a Kolmogorov-Smirnof test at the 80% confidence level.

other studies using regional and worldwide catalogs. The higher rate estimates within this range (for example, those reported by *Michael and Jones* [1998] and *Lindh and Lim* [1995]) were obtained with relatively small numbers of $M \geq 5$ mainshocks (33 and 30, respectively) and were limited to strike-slip earthquakes in the San Andreas fault zone. Larger and broader studies [*Jones and Molnar*, 1979; *Jones*, 1984; *Agnew and Jones*, 1991; this study] find slightly lower overall rates in the range 12–15%.

This study found with high statistical confidence that the foreshock rate before $M \geq 6$ thrust mainshocks in the Harvard catalog is about twice the corresponding rate before strike-slip mainshocks. A lower foreshock rate among strike-slip mainshocks is also suggested by the $M \geq 7$ earthquakes in the Harvard catalog and by the $M \geq 6$ earthquakes in the Triep and Sykes catalog, although these data sets are too small to confidently resolve a dependence on focal mechanism. The apparent contradiction between the resolvable trend in the worldwide data of higher rates among thrusts than strike-slips and the opposite trend inferred in studies of California $M \geq 5$ earthquakes [*Jones*, 1984; *Abercrombie and Mori*, 1996] may be resolved by hypothesizing that most California thrusts (with the exception of offshore events associated with the subduction of Juan de Fuca plate) are typical of continental thrust belts involving unsaturated, crystalline continental rock (such as the Himalayan collision belt) but not typical of subduction zone thrust events occurring in young, water-saturated ocean sediments. This would imply that the generic California model may be representative of most earthquakes in California but may significantly underestimate the conditional probabilities following potential foreshocks in Cascadia, a region whose clusters may be more typical of shallow subduction zones.

The use by *Agnew and Jones* [1991] of a magnitude difference density function in their formulation for the conditional probability of a characteristic earthquake on a particular fault segment has stimulated interest about what form this distribution takes, or can be assumed to take. *Agnew and Jones* [1991] inferred from an analysis of several hundred fs-ms pairs observed in California (1932–1987) involving $M \geq 3$ mainshocks that, at least for the larger mainshock magnitudes, all foreshock magnitudes (within the available magnitude difference aperature) were equally likely, and assumed a uniform distribution in their calculations. Similarly, *Michael and Jones* [1998] concluded that a uniform distribution adequately described the magnitude differences among 30 fs-ms pairs in the San Andreas physiographic province (1933–1994) involving $M \geq 5$ mainshocks and $M \geq 2$ foreshocks, and they based a revised assessment of foreshock-related conditional short-term probabilities for the next Parkfield earthquake on such a model. *Lindh and Lim* [1995] considered a lognormal distribution function for the magnitude differences, with a peak between 1 and 2 magnitude units below the mainshock, based on their analysis of the same California data set. The resulting probabilities obtained with the lognormal model are higher for potential foreshocks in the range $M = 4.0$–5.0 (and lower outside this range) than the corresponding probabilities obtained with the uniform model [*Michael and Jones*, 1998].

The present study, which attempts to build on the previous ones by extending the observations to higher magnitude mainshocks, finds that the magnitude differences between 19 $M \geq 7$ mainshocks and their largest ($M \geq 5.5$) foreshocks in the Harvard catalog cannot reject a uniform distribution. The corresponding result for a larger set of $M \geq 6.5$ mainshocks ($N = 46$) also fails to reject a uniform distribution.

The worldwide occurrence of $M \geq 6$ and $M \geq 7$ earthquakes in the 10-day periods following (and 75-km distance range from) $M \geq 5$ earthquakes in the Harvard catalog exceeds by a factor of about 2 the rate predicted by the California generic model (Figures 5b and 5c). This difference may be understood, in part, by the dominance of shallow subduction thrusts in the Harvard data set (which were found to have a relatively high foreshock rate), the presence of an active strike-slip plate boundary in California (associated with a low foreshock rate worldwide), and the absence of an active subduction zone in most of California (an important exception being the Cascadia megathrust). In this sense, the California generic model can be considered "correct" for most of California, while the world generic model would better represent worldwide foreshock-related conditional probabilities. In certain other regions where sufficient seismological data are available, regionally determined generic clustering models [e.g., *Reasenberg et al.*, 1990; *Jones et al.*, 1995] may provide a better basis than the worldwide generic model for estimating short-term earthquake probabilities.

In the same way that the worldwide generic model was foward-fit to the complete set of worldwide observations in Figures 6b and 6c, a "generic subduction zone model" can be fit to the subset of observations corresponding to shallow subduction zones. Because it includes foreshock-rich, shallow subduction zone thrusts, such a model might be more suited than the California generic model for representing the conditional

probability of very large ($M \geq 8$) earthquakes along the Cascadia subduction zone (Juan de Fuca and Gorda Ridges) after the occurrence of potential ($M \geq 7$) foreshocks.

Short-term earthquake clustering is the strongest nonrandom feature observed in earthquake catalogs; it allows one to obtain reliable, short-term probabilistic forecasts of future earthquake activity. No other phenomenon currently provides such an ability. While the clustering apparent in Figure 2 clearly reveals the foreshock process, the resulting short-term probabilities for larger earthquakes amount to no more than a few percent; most large earthquakes are not preceded by observable foreshocks. Even if all large earthquakes had observable foreshocks, the probability of a large earthquake following a potential foreshock would still be small in areas having even moderate background seismicity [*Agnew and Jones*, 1991]. Unless potential foreshocks can somehow be recognized as being more likely to be foreshocks than background earthquakes, based on their characteristics or those of the smaller earthquakes with which which they may be clustered [e.g., *Ogata et al.*, 1995, 1996], their power to predict mainshocks is limited by the competing background seismicity: limited, but not insignificant.

Earthquake forecasts based on foreshock-mainshock are routinely communicated by the USGS to the public after $M \geq 5$ earthquakes in California. The societal response to such forecasts is molded by official policy (for example, the State of California Office of Emergency Services may issue an earthquake advisory, which triggers an alert status in the affected counties) and by individuals, who may make certain preparations in home and community, or alter industrial and utility decisions according to the forecasts. The usefulness of such actions depends on how one interprets and applies the probabilistic forecasts [*Jones*, 1996; *Mileti and O'Brien*, 1991; *Mileti et al.*, 1990].

Possible benefits of short-term probabilistic earthquake forecasts vary from case to case and may involve a wide variety of different responses, including the postponement of certain industrial, chemical, or nuclear industrial processes in which strong shaking could be hazardous or costly; the postponement of certain elective surgeries and nonessential travel in the affected area; the temporary reduction of the speed of trains; and the avoidance or postponement of certain unstable construction situations, such as erecting tilt-up buildings, carrying out high-steel work, scaffolding work, to name just a few. In home and community, securing breakable items and developing or reviewing family, school, and business emergency or contingency plans can mitigate risk with little cost.

The possible costs and benefits associated with a given response (e.g., postponing an industrial process for a few days) may be relatively easy to estimate in each case. Once the short-term probability of a potentially damaging earthquake is estimated, informed decisions about risk exposure and the possible benefits and costs of a particular response can be weighed and contingency plans can be formulated following conventional insurance analysis methods.

Possible costs associated with false alarms also should be considered. In terms of direct expenses, the cost is probably nil. The California Office of Emergency Services estimates that there is minimal cost to the state for issuing a short-term earthquake advisory and to the county for response teams that respond to its earthquake advisories, since only existing human and infrastructure resources are involved (R. Eisner, personal communication, 1997). A possible indirect cost of false alarms may be born by those who fail to appreciate the low-probability nature of the forecasts and who initially hold unrealistically high expectations for their fulfillment. Some may become disillusioned and switch from overassessing the forecasts' significance to completely ignoring them. Presumably, both of these alternatives would lead to bad decision making. In the combined San Francisco Bay area and Los Angeles conurbations, an $M \geq 5$ earthquake occurs every 2 years, on average. Over a period of 20 years, it would be expected that 10 short-term forecasts for a larger earthquake might be made and that one of them would be fulfilled while the other nine would be false alarms. Embracing the ability to forecast means living with false alarms. Ironically, some of the benefits associated with successful short-term earthquake forecasts, such as the stimulation of contingency planning and the raising of the general level of awareness of long-term regional earthquake hazards, may also accrue for false alarms. In this sense, earthquake forecast false alarms might be thought of not as failures, but as much-needed and naturally triggered practice drills.

Acknowledgments. I am indebted to many of my colleagues for valuable discussions and suggestions, including Rachel Abercrombie, Bill Bakun, Jim Davis, Rich Eisner, Bill Ellsworth, Lucy Jones, Allan Lindh, Mark Matthews, Andy Michael, Jim Mori, and David Oppenheimer. Thanks to Rachel Abercrombie, Lucy Jones, Andy Michael, and two anonymous referees for constructive, critical reviews of the manuscript.

References

Abercrombie, R. E., and J. Mori, Occurrence patterns of foreshocks to large earthquakes in the western United States, *Nature*, *381*, 303–307, 1996.

Agnew, D. C., and L. M. Jones, Prediction probabilities from foreshocks, *J. Geophys. Res.*, *96*, 11,959–11,971, 1991.

Bowman, J. R., and C. Kisslinger, A test of foreshock occurrence in the central Aleutian Island arc, *Bull. Seismol. Soc. Am.*, *74*, 181–197, 1984.

Davis, S. D., and C. Frohlich, Single-link cluster analysis of earthquake aftershocks: Decay laws and regional variations, *J. Geophys. Res.*, *96*, 6335–6350, 1991.

Dodge, D. A., G. C. Beroza, and W. L. Ellsworth, Detailed observations of California foreshock sequences: Implications for the earthquake initiation process, *J. Geophys. Res.*, *101*, 22,371–22,392, 1996.

Dziewonski, A. M., G. Ekstrom, and M. P. Salganik, Centroid-moment tensor solutions for January–March 1994, *Phys. Earth Planet. Inter.*, *86*, 253–261, 1994.

Frohlich, C., Triangle diagrams: Ternary graphs to display similarity and diversity of earthquake focal mechanisms, *Phys. Earth Planet. Inter.*, *75*, 193–198, 1992.

Frohlich, C., and K. D. Apperson, Earthquake focal mechanisms, moment tensors, and the consistency of seismic activity near plate boundaries, *Tectonics*, *11*, 279–296, 1992.

Johnson, C. E., and L. K. Hutton, Aftershocks and preearthquake seismicity, The Imperial Valley, California, Earthquake of October 15, 1979, *U.S. Geol. Surv. Prof. Pap.*, *1254*, 59–76, 1982.

Jones, L. M., Foreshocks (1966–1980) in the San Andreas system, California, *Bull. Seismol. Soc. Am.*, *74*, 1361–1380, 1984.

Jones, L. M., Foreshocks and time-dependent earthquake hazard assessment in southern California, *Bull. Seismol. Soc. Am.*, *75*, 1669–1679, 1985.

Jones, L. M., and P. Molnar, Some characteristics of foreshocks and their possible relationship to earthquake prediction and premonitory slip on faults, *J. Geophys. Res.*, *84*, 3596–3608, 1979.

Jones, L. M., R. Console, F. Di Luccio, and M. Murra, Are foreshocks mainshocks whose aftershocks happen to be big? Evidence from California and Italy, (abstract), *Eos Trans. AGU*, *76*(46), Fall Meet. Suppl., F388, 1995.

Kagan, Y. Y., Seismic moment-frequency relation for shallow earthquakes: Regional comparison, *J. Geophys. Res.*, *102*, 2835–2852, 1997.

Kagan, Y. Y., and L. Knopoff, Spatial distribution of earthquakes: The two-point correlation function, *Geophys. J. R. Astron. Soc.*, *62*, 303–320, 1980.

Kagan, Y. Y., and L. Knopoff, Statistical short-term earthquake prediction, *Science*, *236*, 1563–1567, 1987.

Kaverina, A. N., A. V. Lander, and A. G. Prozorov, Global creepex distribution and its relation to earthquake-source geometry and tectonic origin, *Geophys. J. Int.*, *125*, 249–265, 1996.

Lindh, A. G., and M. R. Lim, A clarification, correction and updating of Parkfield, California, earthquake prediction scenarios and response plans, *U.S. Geol. Surv. Open File Rep.*, *95-695*, 1995.

Michael, A. J., and L. M. Jones, Seismicity alert probabilities at Parkfield, California, revisited, *Bull. Seismol. Soc. Am.*, *88*, 117–130, 1998.

Mileti, D. S., and P. W. O'Brien, Public response to the Loma Prieta earthquake emergency and aftershock warnings, Hazard Assess. Lab., Colo. State Univ., Fort Collins, 1991.

Mileti, D. S., C. Fitzpatrick, and B. C. Farhar, Risk communication and public response to the Parkfield earthquake prediction experiment, Hazard Assess. Lab., Colo. State Univ., Fort Collins, 1990.

Molchan, G., T. Kronrod, and G. F. Panza, Multi-scale seismicity model for seismic risk, *Bull. Seismol. Soc. Am.*, *87*, 1220–1229, 1997.

Ogata, Y., T. Utsu, and K. Katsura, Statistical features of foreshocks in comparison with other earthquake clusters, *Geophys. J. Int.*, *121*, 233–254, 1995.

Ogata, Y., T. Utsu, and K. Katsura, Statistical discrimination of foreshocks from other earthquake clusters, *Geophys. J. Int.*, *127*, 17–30, 1996.

Reasenberg, P. A., Second-order moment of central California seismicity, *J. Geophys. Res.*, *90*, 5479–5495, 1985.

Reasenberg, P. A., and L. M. Jones, Earthquake hazard after a mainshock in California, *Science*, *243*, 1173–1176, 1989.

Reasenberg, P. A., and L. M. Jones, Earthquake aftershocks: Update, *Science*, *265*, 1251–1252, 1994.

Reasenberg, P. A., Y. Okada, and F. Yamamizu, Earthquake hazard after a mainshock in the Kanto-Tokai districts, Japan (abstract), *Eos Trans. AGU*, *71*, 908, 1990.

Sipkin, S. A., Display and assessment of earthquake focal mechanisms by vector representation, *Bull. Seismol. Soc. Am.*, *83*, 1871–1880, 1993.

Smith, G. P., and G. Ekström, Interpretation of earthquake epicenter and CMT centroid locations, in terms of rupture length and direction, *Phys. Earth Planet. Inter.*, *102*, 123–132, 1997.

Triep, E. G., and L. R. Sykes, Frequency of occurrence of moderate to great earthquakes in intracontinental regions: Implications for changes in stress, earthquake prediction, and hazards assessments, *J. Geophys. Res.*, *102*, 9923–9948, 1997.

Von Seggern, D., Seismicity parameters preceding moderate to major earthquakes, *J. Geophys. Res.*, *86*, 9325–9351, 1981.

P. A. Reasenberg, U.S. Geological Survey, 345 Middlefield Road, Menlo Park, CA 94025. (reasen@hayward.wr.usgs.gov)

(Received April 22, 1998; revised September 24, 1998;
accepted November 11, 1998.)