Space-time clustering of seismicity in California and the distance dependence of earthquake triggering

P. M. Shearer¹

Received 21 May 2012; revised 23 July 2012; accepted 11 September 2012; published 13 October 2012.

[1] Using two recent high-resolution earthquake catalogs, I examine clustering in California seismicity by plotting the average rate of earthquakes as a function of both space and time from target events of $M \ 2$ to 5. Comparisons between pre- and post-target event activity can be used to resolve earthquake-to-earthquake triggering associated with target events of different magnitudes. The results are more complicated than predicted by computer simulations of earthquake triggering that begin with background events occurring at random times. In particular, at least some of the temporal clustering of seismicity at short scales (0.1 to 5 km) does not appear to be caused by local earthquake triggering, but instead reflects an underlying physical process that temporarily increases the seismicity rate, such as is often hypothesized to drive earthquake swarms. Earthquake triggering for M < 4.5 earthquakes is only resolvable in average seismicity rates at times less than about one day and to distances of less than about 10 km, and its linear density decreases as $r^{-1.5}$ to $r^{-2.5}$, significantly steeper than some previous studies have found.

Citation: Shearer, P. M. (2012), Space-time clustering of seismicity in California and the distance dependence of earthquake triggering, *J. Geophys. Res.*, *117*, B10306, doi:10.1029/2012JB009471.

1. Introduction

[2] Earthquakes are observed to cluster strongly in time and space. Much of this clustering is caused by the aftershock sequences of major earthquakes. However, smaller earthquakes (M < 5) also generate aftershocks and at least some of the space-time clustering of earthquakes at short times and distances is likely caused by triggering among small earthquakes. Felzer and Brodsky [2006] attempted to quantify the distance dependence of triggering from earthquakes in southern California by examining the average seismicity rate in short time windows following target events within different magnitude ranges and found evidence for a decay in linear event density (the average number of events, in any direction, per distance interval Δr from the target) of about $r^{-1.35}$ over a wide magnitude and distance range, which they argued supported dynamic rather than static triggering of aftershocks. However, Richards-Dinger et al. [2010] criticized many aspects of this analysis and conclusion, noting, for example, that the Felzer and Brodsky study is flawed by the fact that they neglected to compare the posttarget event seismicity with the pre-target event seismicity within analogous time windows. For smaller events, the preand post-target rates are nearly the same, calling into question how much of the observed spatial clustering is triggering associated with the target event and how much has some other cause.

[3] My goal here is to use the target event approach to examine the space-time clustering of seismicity in California over a very broad range of times and distances in an effort to understand possible origins for the clustering and its implications for earthquake triggering. This paper builds on the results of Shearer [2012], which examined the distribution of apparent foreshock/aftershock magnitudes in southern California in the context of Båth's law and self-similar earthquake triggering models but did not explicitly consider range dependence. My analysis will use two recent relocated earthquake catalogs that apply waveform cross-correlation to provide very accurate locations among closely spaced events. The basic approach is similar to that used by *Felzer* and Brodsky [2006] and Shearer and Lin [2009], in which the average density of events is computed within distancetime windows around target events of varying magnitude. Unlike these prior studies, however, I will explicitly compare pre- and post-target event seismicity rates, which will help in resolving any triggering associated with the target events

[4] To determine earthquake triggering properties, an ideal study would consider individual event sequences in isolation to avoid contamination from overlapping seismicity from other sequences. In practice, however, this is only possible for larger main shocks with substantial and well-defined aftershock sequences. Because smaller earthquakes may generate, on average, less than one aftershock each, it is only

¹Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California, USA.

Corresponding author: P. M. Shearer, Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093-0225, USA. (pshearer@ucsd.edu)

^{©2012.} American Geophysical Union. All Rights Reserved. 0148-0227/12/2012JB009471



Figure 1. (a) Event locations from the LSH catalog of *Lin et al.* [2007], windowed to only include events with good station coverage. Events of M > 6 are shown as black circles. (b) The number of events as a function of minimum magnitude, plotted at 0.1 magnitude intervals. The best-fitting line for $2 \le M \le 5$ is plotted and has a slope of -1.00.

possible to study their triggering properties by averaging results from many events. These results will always be contaminated to some extent by unrelated background seismicity and their aftershocks, particularly when studying the full time and range dependence of triggering. My approach here is to accept that real earthquake catalogs will contain a complex pattern of both background and triggered activity, and attempt to model this behavior by generating synthetic catalogs based on earthquake triggering models, which can then be processed in the same way as the real data.

[5] A fundamental characteristic of aftershock sequences is the time asymmetry that they introduce into earthquake catalogs. There are many more aftershocks than foreshocks, and aftershock rates typically exhibit a power law decay with time described by Omori's law. These properties can be use to help determine which parts of observed seismicity clustering can be explained by triggering and which parts may reflect underlying physical driving mechanisms, such as the fluid flow, dyke intrusion, or creep events that are often observed to drive seismic swarms [e.g., *Toda et al.*, 2002; *Hainzl and Ogata*, 2005; *Lohman and McGuire*, 2007; *Llenos et al.*, 2009].

2. Relocated California Earthquake Catalogs

[6] To explore the distance dependence of seismicity at the shortest possible scale lengths, it is helpful to use earthquake catalogs from major networks that have been relocated using high-resolution methods, including waveform crosscorrelation to improve the timing accuracy among closely spaced events. Two such catalogs have recently been produced for large numbers of California earthquakes: the Lin et al. [2007] southern California catalog of 430,000 events from 1981 to 2005 (here termed LSH), and the Waldhauser and Schaff [2008] northern California catalog of 311,000 events from 1984 to 2003 (here termed NCA). These catalogs achieve relative location accuracy of tens of meters for the majority of the earthquakes that can be cross-correlated with other events. Because location accuracy and catalog completeness suffer at the edges of the network, I first window the catalogs to well-resolved subsets of the data. Figure 1 shows a selected subset of the LSH catalog. It includes 294,502 events located within and near the network, but excludes offshore earthquakes and others outside the network that have relatively poor locations. To estimate the minimum magnitude of catalog completeness, I also plot the cumulative number of earthquakes versus magnitude. The results are well-fit with a b value of one in the Gutenberg-Richter magnitude-frequency relation (agreeing with Hutton et al. [2010]). Our windowed LSH catalog appears nearly complete down to M 1.5.

[7] For northern California, Figure 2 shows a selected subset of the NCA catalog. It includes 113,003 events of $M \ge 1$ within the western part of the network, but excludes offshore events, Mendocino seismicity, the Central Valley, and eastern California seismicity, including large numbers of events near Mammoth Lakes in eastern California. These event magnitudes are also well-fit with a *b* value of one and appear nearly complete down to about M 1.5. For subsequent analysis, I will only use earthquakes of $M \ge 1.5$ from both catalogs. For the LSH catalog I also exclude events identified as quarry blasts by the network operators (no such flag is available for the NCA catalog). This results in 173,058 LSH events and 53,361 NCA events.

3. Quantifying Earthquake Clustering

[8] We consider the earthquakes as point sources, defined by their origin times, (x, y, z) coordinates, and their



Figure 2. (a) Event locations from the northern California catalog of *Waldhauser and Schaff* [2008], windowed to only include events with good station coverage from the western part of the network. Events of M > 6 are shown as black circles. (b) The number of events as a function of minimum magnitude, plotted at 0.1 magnitude intervals. The best-fitting line for $2 \le M \le 5$ is plotted and has a slope of -1.01.

magnitudes. Now imagine an earthquake catalog with completely random locations and times, and magnitudes randomly drawn from the G-R relation. In this case the average seismicity rate (events per volume per time) would be constant within any space-time window in the catalog. In particular, the presence or absence of an event at a particular location would have no effect on the seismicity rate for other events nearby. However, real earthquake catalogs are strongly clustered in both space and time. The seismicity rate is much higher for events close in space and time to a given target event than the rate at random locations and times. Thus, one way to measure the clustering in a catalog is to calculate the average seismicity rate as a function of the time and distance separation from other events. However, the event magnitudes provide an additional constraint on the clustering and are key to resolving earthquake triggering in the catalog. Earthquake triggering is most obvious in the large aftershock sequences that follow major earthquakes, but it occurs to some extent for smaller earthquakes as well, and earthquake triggering models generally assume that all earthquakes increase the chances of subsequent events. Because M 2–3 earthquakes may individually trigger on average less than one aftershock above the catalog minimum magnitude, it is necessary to average (i.e., stack) results from many such small events to resolve their average aftershock productivity, which is the approach used by *Helmstetter* [2003] and Felzer and Brodsky [2006] in their analyses of southern California seismicity.

[9] The triggered event need not be smaller than the initiating event; in cases where it is larger, the 'aftershock' is usually reclassified as a 'main shock' and the initiating event becomes a 'foreshock.' However, foreshocks are much less common than aftershocks in real catalogs and thus a defining characteristic of triggered seismicity is that it introduces a time asymmetry-large events are followed by many more small events than those that precede them. To see this asymmetry around target events, one simply compares the average number of events that precede the target events with the numbers that follow the targets. However, this approach only works if the count is limited to events smaller than the target events. Otherwise every event pair A-B (B occurring after A) would be counted twice—both as B following target event A, and as A preceding target event B. In this case, in which magnitude information is ignored, the pre- and posttarget event density would be exactly the same. However, if A is larger than B and we require that the counted events be smaller than the target event, then only B is counted as following A.

[10] Although all events in an aftershock sequence can contribute to triggering (spawning second and later generation aftershocks), triggering from an individual target event will be most obvious if it is larger than any other nearby events within a given time window. This makes it more likely that the target event is a 'main shock' rather than the 'aftershock' of an earlier event or the 'foreshock' of a subsequent event. *Felzer and Brodsky* [2006] required that their target events be larger than any events occurring within 150 km and both 3 days before and 0.5 days after the target. I adopt a similar approach but require no larger event within ± 3 days from the target to preserve time symmetry in the analysis. In addition, because my focus is mainly on the clustering within 20 km of the targets,



Figure 3. Target event selection criteria: (a) All events are excluded for one to three months following large earthquakes, such as the 1992 M 7.3 Landers earthquake. In addition, smaller magnitude events within 3 days and 50 km of a larger earthquake are excluded. (b) An example of the selection method for a set of nearby earthquakes (i.e., all within 50 km of each other). Only the events shown in red qualify as target earthquakes. Note that the events between the 3rd and 4th boxes are not included because they have larger events within 3 days, even though these larger events are not themselves target events.

I only require that the targets be larger than earthquakes within 50 km for the ± 3 day interval. Because catalog completeness often suffers following major earthquakes due to the high seismicity rate [e.g., *Kagan*, 2004], I also exclude target events during certain specified time periods. For southern California, events are excluded for one month following the 1987 *M* 6.2/6.6 Elmore Ranch/Superstition Hills and 1992 *M* 6.1 Joshua Tree earthquakes, 2 months following the 1994 *M* 6.7 Northridge earthquake, and 3 months following the 1992 *M* 7.3 Landers and 1999 *M* 7.1 Hector Mine earthquakes. For northern California, events are excluded for one month following the 1984 *M* 6.2 Morgan Hill earthquake and for two months following the 1989 *M* 6.9 Loma Prieta earthquake. My intention is not to remove all of the aftershocks of these major earthquakes or to 'decluster' the data, but simply to exclude time periods where the seismicity rates are very high due to aftershocks of a single large initiating event and it is likely more difficult to measure triggering from smaller earthquakes. Figure 3 illustrates how the target event selection procedure works.

[11] For the windowed LSH catalog, this event selection criteria results in 8569 M 2-3 target events, 1533 M 3-4 target events, and 162 M 4-5 target events. For the windowed NCA catalog, this yields 4628 M 2–3 target events, 1174 M 3–4 target events, and 110 M 4–5 target events. Note that there are noticeably fewer M 2–3 target events than would be expected based on the observed G-R b-value of one. This results from the target event selection criteria, which rejects events in which a larger nearby event occurs within 3 days, a rejection that is more likely for smaller target events. Smaller earthquakes are then summed within 100 total distance-time bins with respect to the target event location and time, equally spaced at ten log time intervals between 0.001 and 1000 days and ten log distance intervals between 0.01 and 100 km. Note that events within at least the shortest distance bin (10 to 25 m) are likely affected by location error. Figure 4 contours the resulting estimates of pre- and post-target activity as a function of time before/after and distance from target events in the LSH catalog. The contours are unit spaced in \log_{10} density, to accommodate the fact that the observed event rates span a huge range, from less than 10^{-5} events per day per km³ to more than 100 events per day per km³. The event density is greatest at small times and distances, reflecting the strong space-time clustering of the seismicity.

[12] A comparison between the pre-target activity shown in Figure 4a with the post-target activity in Figure 4b shows a very similar pattern in the event rates. However, by subtracting the pre-target activity from the post-target activity, a region of enhanced post-target activity can be identified at roughly ± 1 day and up to 1–10 km from the targets (see Figure 4c). This time asymmetry is strongest and extends to greater distances and times for the M4-5 target events and is weakest for the M 2–3 targets. Within this region the rate falls off with both time and distance, i.e., the enhanced posttarget activity is most pronounced at small times and distances. Notice that the results are most complete for the M 2–3 target events because of the much greater number of target events for small magnitudes. The lack of events closer than 100 m to the M 4–5 targets likely reflects the finite rupture dimensions of these events and the fact that they are located less accurately in the catalog because their waveforms typically do not cross-correlate well with smaller earthquakes. Analogous results for the windowed NCA catalog are plotted in Figure 5. In general the pre- and posttarget clustering appears quite similar to that resolved in southern California, as does the surplus of post-target events observed at short times and distances.

[13] Evenly spaced log contours in density on a log plot are indicative of a power law relationship. Thus, the roughly evenly spaced contours versus time at distances less than 1 km could be interpreted as Omori's Law for the post-target activity (and the differenced activity in part c in each plot), and inverse Omori's Law for the pre-target activity. Some studies [e.g., *Ouillon and Sornette*, 2005; *Ouillon et al.*, 2009] have found that the Omori Law decay steepens for



Figure 4. Space/time behavior of seismicity in southern California with respect to target events of varying magnitude. (a) Average event rate prior to target earthquakes of M (left) 2–3, (middle) 3–4, and (right) 4–5, at times from 0.001 day (86 s) to 1000 days prior to the target events at distances from 10 m to 100 km. Contours are uniform in log event density (per day per cubic kilometer). White shows regions of no data. (b) Corresponding post-target event rates. (c) The difference between the pre- and post rates, i.e., the extra events observed after the targets compared to before the targets. White indicates no data or negative values.

larger main shocks. Although our focus here is mainly on the distance-dependence of triggering, it is interesting to note that comparing the post-target contour spacing with time at 1-km distance among the different sized target events in Figures 4b and 4c supports this apparent magnitudedependence in Omori's law. However, notice that this decay in event density with time is very gentle to non-existent at distances larger than about 10 km, suggesting that the spatial clustering observed at these scales is not primarily caused by triggering.

[14] At short times, the nearly evenly spaced contours versus distance also indicate a power law distribution, which has been related to a fractal dimension for the seismicity. However, because events also cluster in time, the computed fractal dimension will vary as a function of the time interval over which the measurements are made, as noted by *Kagan* [2007]. The curved and non-parallel contour lines in Figures 4 and 5 indicate clustering behavior fundamentally different from that expected for separable time and distance power laws, i.e., an equation where the event rate, D, is given by

$$D = \kappa (t+c)^{-p} r^{-q} \tag{1}$$

where κ is a scaling constant, t is time, r is distance, c and p are the Omori's law parameters, and q is the distance decay



Figure 5. Space/time behavior of seismicity in northern California with respect to target events of varying magnitude. (a) Average event rate prior to target earthquakes of M (left) 2–3, (middle) 3–4, and (right) 4–5, at times from 0.001 day (86 s) to 1000 days prior to the target events at distances from 10 m to 100 km. Contours are uniform in log event density (per day per cubic kilometer). White shows regions of no data. (b) Corresponding post-target event rates. (c) The difference between the pre- and post rates, i.e., the extra events observed after the targets compared to before the targets. White indicates no data or negative values.

parameter. This equation predicts uniform time decay regardless of distance and uniform distance decay regardless of time.

[15] To further explore the surplus of post-target events at short time intervals and its distance dependence, Figure 6 (left) plots the linear event density (the average number of events, in any direction, per distance interval Δr from the target) versus distance for ± 1 hour from the LSH target events. Plots of linear event density versus distance are useful for strongly clustered seismicity because they flatten out the much stronger falloff with distance that would be seen in plots of event density per unit volume versus distance [see *Richards-Dinger et al.*, 2010, supplemental Figure 1]. A bootstrap resampling method is used to estimate one standard error bars. For comparison, 'background' events rates at ± 900 to 1000 days from the target events are also plotted. Note that the pre- and post-target event densities exhibit a similar distance dependence and that enhanced post-target activity is resolvable to less than 1 km for M 2–3 targets, to less than 4 km for M 3–4 targets, and to about 10 km for M 4–5 targets. In addition, the background seismicity at long times merges with the ± 1 hour seismicity at about 10 km and accounts for the flattening of the ± 1 hour curves at longer distances. The apparent flattening of the falloff with distance at short ranges is likely due to location error. The M 3–5 target events have poorer relative location



Figure 6. Linear event density versus distance for the windowed southern and northern California catalogs, comparing results for M 2–3, M 3–4, and M 4–5 target earthquakes. Average pre- and post-target event densities are computed in ± 1 hour windows from the target event times. For comparison, a 'back-ground' rate estimated for between 900 and 1000 days from the target events is also plotted. The pre-target densities are shown as dashed lines. One standard error bars are estimated from bootstrap resampling of the target events. Reference slopes of distance $r^{-1.5}$ to $r^{-2.5}$ are also plotted.

accuracy than the smaller events because they typically do not have waveforms that cross-correlate with nearby events. Thus, the steepest parts of these curves are likely the best resolved and exhibit falloff rates of roughly r^{-2} to $r^{-2.5}$. Figure 6 (right) shows analogous results for the northern California NCA catalog. Results are similar, except the surplus of post-target events compared to pre-target events can be resolved to somewhat greater distances. In addition, the falloff with distance for the post-target M 2–4 events is somewhat shallower, i.e., roughly $r^{-1.5}$ to r^{-2} .

[16] For both catalogs, both the pre- and post-event densities grow with target event magnitude, as does the surplus of post-target events. This behavior is consistent with the fact that the counted events must be smaller than the targets (so that there is a greater range of event magnitudes available in the case of larger magnitude targets) and with some degree of triggering associated with the target event (larger

magnitude targets will trigger more aftershocks). In the next section, I will use computer-generated synthetic catalogs to test whether standard triggering models can account for the observed behavior. Before doing so, however, it is useful to summarize what can be reasonably inferred from the catalog data alone. First, the spatial patterns of seismicity observed in Figures 1 and 2 define a stable long-term tendency for earthquakes to occur close to each other. This clustering can be seen in the falloff in event density with distance at ± 1000 days in Figures 4 and 5 and in the approximately constant linear event density for $\pm 900-$ 1000 days in Figure 6 (in three-dimensions the corresponding falloff is r^{-2}). This clustering behavior is greatly enhanced at short timescales, leading to a surplus of events at ± 1 hour out to distances of 5 to 20 km from the target events. This seismicity exhibits time asymmetry, with a greater number of post-target events compared to pre-target events. However, the surplus of post-target events is only resolvable to distances of 1 to 10 km. At distances beyond about 10 km, the observed clustering is not primarily caused by aftershock triggering. This can be seen both in the merging of the preand post-target event densities beyond this distance and the lack of an Omori's law time decay in the event densities. These results thus do not support the claim of *Felzer and Brodsky* [2006] that aftershocks from M 2–4 target events can be resolved to a distance of 50 km in southern California.

4. Comparisons to Triggering Models

[17] Because natural seismicity contains a complex overlapping mixture of background and triggered events, it is difficult to isolate triggering behavior directly from earthquake catalogs. Instead, I will attempt to model the behavior discussed above with computer-generated Monte Carlo simulations of seismicity. There is considerable literature on various aspects of ETAS-like earthquake triggering models [e.g., Ogata, 1999; Helmstetter and Sornette, 2002; Felzer et al., 2002; Helmstetter et al., 2005] in which all earthquakes, no matter how small, increase the probability of subsequent events. Here I will create synthetic catalogs based on methods discussed in Felzer et al. [2002] and used in the AftSimulator program distributed by Karen Felzer (http://pasadena.wr.usgs. gov/office/kfelzer/AftSimulator.html). We assume that the average number of direct (first generation) aftershocks, N_{asl} following an event of magnitude *m* follows a productivity law

$$N_{\rm as1} = Q \, 10^{\alpha(m-m_1)} \tag{2}$$

where m_1 is the minimum magnitude earthquake that triggers other earthquakes, Q is an aftershock productivity parameter (denoted k by Sornette and Werner [2005b]), and α is a parameter that determines the rate of increase in the number of aftershocks observed for larger main shock magnitudes. This type of triggering model assumes that every aftershock has a single parent and is sometime termed the branching model, in which each event independently triggers its own aftershock chain [e.g., Kagan, 1991; Felzer et al., 2002]. In contrast, in the ETAS modeling approach of Ogata [1998] the probability of aftershock triggering is governed by the sum of all previous activity. As discussed in Sornette and Werner [2005b], both approaches produce seismicity catalogs with the same statistical properties, although the branching model is much faster computationally.

[18] Note that equation (2) provides the average number of direct aftershocks for a given magnitude main shock and typically will have non-integer values. In computer simulations, the actual number of direct aftershocks for a specific event is computed as a random deviate drawn from a Poisson distribution of mean N_{asl} . The total number of aftershocks generated by a single initiating event will in general be larger than the number of direct aftershocks because each aftershock is capable of spawning its own aftershocks, which in turn can generate additional aftershocks. However, provided the aftershock productivity parameter Q is not set too large, this process will eventually converge to a finite total number of aftershocks.

[19] Once a model event occurs, its magnitude, time, and location must be assigned. Most triggering models assume that the magnitude of earthquakes, whether occurring as background or triggered events, is a random variable drawn from the Gutenberg-Richter (G-R) distribution. In this case,

$$N(\ge m) = 10^{a-bm} \tag{3}$$

where *m* is the magnitude, *a* is related to the total number of earthquakes, and *b* (the *b*-value) determines the relative number of large quakes compared to small quakes and is generally observed to lie between 0.8 and 1.2. In practice, for computer simulations a random event magnitude, m_r , can be computed as

$$m_r = m_1 - \log_{10} x_r$$
 (4)

where x_r is randomly and uniformly distributed between $10^{m_1-m_2}$ and 1. For the models presented here, I assume $m_1 = 0$ and $m_2 = 5.5$. The former is chosen to be well below most catalog thresholds, while not excessively increasing the computation cost of the computer simulation. The latter is chosen to avoid aftershock sequences from events substantially larger than m = 5, the maximum target event magnitude. Of course, the California earthquake catalogs contain larger events, such as the 1992 M 7.3 Landers earthquake, but they have extended ruptures that are not well-described by a point source assumption.

[20] Recent studies seem to be converging on $\alpha \approx 1$ in equation (2) [Felzer et al., 2004; Helmstetter et al., 2005]. The case of $\alpha = b$ has received special attention because it produces self-similar behavior in which the increased triggering caused by larger magnitude events is exactly compensated for by their decreased numbers in the G-R relation [Agnew and Jones, 1991; Felzer et al., 2002, 2004; Helmstetter et al., 2005; Shearer, 2012]. My simulations presented here assume $\alpha = b = 1$.

[21] Often the triggering productivity can be defined in terms of the branching ratio, n, given by

$$n = Qb\ln(10)(m_2 - m_1) \tag{5}$$

which gives the ratio of the average number of first generation aftershocks to the number of background events (simplified in this case by using an approximation for large m_2). This parameter is used by *Helmstetter and Sornette* [2003b], *Helmstetter et al.* [2003], and *Sornette and Werner* [2005a, 2005b]. For the G-R magnitude limits $m_1 = 0$ and $m_2 = 5.5$, following the method described in *Shearer* [2012] we obtain n = 0.39 in order to satisfy Båth's law [*Båth*, 1965], the observation that the largest aftershock is, on average, 1.2 magnitudes smaller than its main shock.

[22] The time delay following the triggering event can be determined from Omori's Law:

$$D \propto (t+c)^{-p} \tag{6}$$

where *D* is the aftershock rate, *t* is the time delay, *c* is a constant the defines the flattening of the aftershock decay curve at short times, and *p* is the decay constant, which is often assumed to be one. *Felzer et al.* [2002] describes how a random aftershock time can be drawn from this distribution using a computer-generated random number. For the simulations presented here, c = 0.001 day (86 s) and p = 1.



Figure 7. Space/time behavior of a computer-generated synthetic seismicity catalog with respect to target events of varying magnitude. (a) Average event rate prior to target earthquakes of M (left) 2–3, (middle) 3–4, and (right) 4–5, at times from 0.001 day (86 s) to 1000 days prior to the target events at distances from 10 m to 100 km. Contours are uniform in log event density (per day per cubic kilometer). White shows regions of no data. (b) Corresponding post-target event rates. (c) The difference between the preand post rates, i.e., the extra events observed after the targets compared to before the targets. White indicates no data or negative values.

[23] Following *Felzer and Brodsky* [2006], I assume that the distance, r, from the triggering event to the aftershock obeys the power law

$$D \propto r^{-q}$$
 (7)

where D is the aftershock rate and q is the distance decay constant. After a random r is drawn from this distribution, the aftershock location is assigned as a random location on a sphere of radius r centered on the triggering event, excluding any portions of the sphere that are above the surface or below an assigned maximum depth of seismicity (30 km for the simulations presented here).

[24] Much of the distance dependence of seismicity clustering at long timescales is determined by the fault locations and regions of active seismicity. To model this as accurately as possible, I use the actual earthquake locations in the LSH catalog as a starting point for assigning background seismicity. My simulations will thus focus on the southern California observations, although as shown in Figures 4 and 5, results from northern California are reasonably similar. Each background event location is randomly drawn from the catalog, but its time is randomly assigned from a 9000 day interval. Aftershocks are then computed using the above relationships and the resulting synthetic catalog is processed in exactly the same way as the real catalog, assuming a



Figure 8. Linear event density versus distance for two different computer-generated synthetic seismicity catalogs, comparing results for M 2-3, M 3-4, and M 4-5 target earthquakes, for (a) a catalog that satisfies Båth's law, and (b) a catalog with greater aftershock productivity than Båth's law predicts. Average preand post-target event densities are computed in ± 1 hour windows from the target event times. For comparison, a 'background' rate estimated for between 900 and 1000 days from the target events is also plotted. The pre-target densities are shown as dashed lines. One standard error bars are estimated from bootstrap resampling of the target events. Reference slopes of distance $r^{-1.5}$ to $r^{-2.5}$ are also plotted.

cutoff catalog magnitude of 1.5. That is, target events are identified that have no larger events within 3 days and 50 km and then average pre- and post-target event seismicity rates are computed. Figure 7 shows the space/time pre- and post-target event behavior of a catalog obtained from 5,000,000 starting background events of M 0–5.5, using q = 1.37 for the distance decay parameter (the value estimated in *Felzer and Brodsky* [2006]). The synthetic calculation captures many features of the real catalog (see Figure 4), including the existence of a limited zone close in time and space to the targets, in which the post-target seismicity exceeds the pre-target seismicity. However, in general the post-target event rate contour lines within this zone are straighter than those seen in the real data, reflecting the separation of the time and distance dependence of triggering in the synthetic model.

[25] The pre- and post-target event rates versus distance within ± 1 hour for this model run can be seen in more detail

in Figure 8a, which should be compared with the southern California LSH catalog result of Figure 6. Again this captures many features of the real data, including the merging at longer distances of the enhanced seismicity seen at short times with the background seismicity seen at long times. However, there are several aspects in which the synthetic results differ substantially from the LSH catalog: (1) the synthetics do not reproduce the flattening in the data curves at short distances, which presumably occurs due to location error and/or non-zero fault dimensions, (2) the observed decay in seismicity rate with distance at ± 1 hour from the targets is steeper in the real data than the $r^{-1.37}$ slope of the synthetics. (3) The post-target rates are significantly greater for the real data than for the synthetics for M 2–4 target events at distances less than1-2 km. (4) the pre- and posttarget curves are much closer together for the real data than in the synthetics. For example, the post-target rates for the M 3–4 target events in the synthetics are about 10 times higher than the pre-target rates, while the corresponding ratio is about 3 in the LSH catalog. If the observed pre- and post-rates are interpreted entirely in terms of triggering models (i.e., as foreshocks and aftershocks), (3) and (4) imply that greater aftershock productivity is required in the synthetic calculation to match the observed foreshock-to-aftershock ratios in the real data. However, it should be noted that one cannot arbitrarily increase aftershock productivity to match one feature in a data set without affecting the fit to other features. In particular, the foreshock-to-aftershock ratio seen in the synthetic calculation plotted in Figure 8a is derived from an aftershock productivity the matches Båth's law. Increasing the aftershock productivity will necessarily produce a synthetic catalog in which the average aftershock magnitude is larger than that predicted by Båth's law. The validity of Båth's law was recently verified for $M \ge 5.5$ earthquakes in California by Shcherbakov and Turcotte [2004] who obtained an average magnitude difference between the main shock and the largest aftershock of 1.16. However, it should be noted that although self-similar earthquake triggering is often assumed, Båth's law has not been established for smaller magnitude main shocks [see Shearer, 2012].

[26] Figure 8b shows the result of modifying the synthetic model in an attempt to better fit the data by increasing the distance decay parameter to q = 2 and increasing the branching ratio to n = 0.60. This enhanced triggering model predicts that the largest aftershock is, on average, 0.95 magnitude units smaller than the main shock. This produces a much better fit to the distance decay at ± 1 hour in the real data, but the real data still have a higher post-target rate at short distances and the pre- to post-target event ratio is still larger than observed for the LSH catalog, particularly for the smaller magnitude target events. It is possible that an additional increase in the branching ratio would improve the fit somewhat more, but at the cost of a further degradation in the fit to Båth's law. However, the observed pre to post ratio for the M 2–3 targets is so small that it is doubtful that any subcritical branching ratio (n < 1) can explain the observations. For these magnitude target events, the pre- and posttarget seismicity is nearly time symmetric, which is difficult to explain with standard triggering models (for additional discussion, see Shearer [2012]).

[27] However, earthquake-to-earthquake triggering is not the only process that can give rise to seismicity clustering. Correlation does not prove causation. Richards-Dinger et al. [2010] showed that a power-low decay with distance beyond 5 km is observed for post-target events even at times before the P-wave arrival from the target event, where direct earthquake-to-earthquake triggering should not be occurring. Temporal clustering of events also does not necessarily require earthquake-to-earthquake clustering. For example, earthquake swarms of small events often occur that do not have an obvious triggering main shock, and are commonly thought to result from an underlying physical driving mechanism, such as fluid flow or slow slip. In this case the occurrence of a target earthquake increases the probability of nearby events not because they are triggered by the target earthquake, but because all the seismicity is responding to some change in conditions. How can we distinguish this type of behavior from earthquake-to-earthquake triggering?

One approach is to recognize that most aftershocks are smaller than their initiating events, which introduces a fundamental time asymmetry into earthquake catalogs. Temporal clusters caused by triggering tend to have their largest events earlier in the sequence, whereas many swarms are distinguished by a more random distribution in time of event sizes [e.g., *Vidale and Shearer*, 2006; *Roland and McGuire*, 2009]. Thus, we can remove this magnitude-dependent time asymmetry related to triggering in an earthquake catalog by replacing the actual magnitudes in the catalog with random magnitudes drawn from the G-R distribution.

[28] Figure 9a illustrates the result of this magnitude randomization for the LSH catalog. In this case, the LSH locations and times are retained, but the magnitudes are replaced with random G-R magnitudes between 1.5 and 5.5. This simulates a catalog in which all the temporal and spatial clustering in the LSH catalog is driven by external factors rather than earthquake-to-earthquake triggering. The ± 1 hour earthquake rates appear similar to the original catalog (see Figure 6a) except that the pre- and posttarget curves coincide within their estimated errors. This is expected because there is no added triggering and the random magnitudes imply that every event pair A-B has an equal chance of being counted as A preceding target event B (if B is larger than A) or as B following target event A (if A is larger than B). The ± 1 hour earthquake pre- and post-target rates in this magnitude-randomized catalog roughly agree with the observed pre-target rates in the real LSH catalog. This suggests that the bulk of the ± 1 hour pre-target seismicity rate in the observations is not caused by earthquake-to-earthquake triggering related to the target events, but has some other cause. However, explaining the enhanced post-target rates in the real catalog requires some amount of triggering.

[29] To model this, we use the magnitude-randomized LSH catalog to generate background seismicity and then generate aftershock sequences using the triggering model parameters used for Figure 8b (distance falloff $q = 2, m_1 = 0$, $m_2 = 5.5$, branching ratio n = 0.60). The results are plotted in Figure 9b, which shows the additional post-target event activity generated by the triggering. This model provides a much better fit to the LSH data (see Figure 6a) than the model in which the background events have random times (see Figure 8), particularly for the M 2–4 target events. However, there remains significant misfit for the M 4–5 target events, for which the real catalog has higher posttarget seismicity rates and a steeper falloff rate with distance than in the synthetics. This might indicate greater triggering for larger magnitude events than $\alpha = 1$ predicts, i.e., a breakdown in self-similarity. In addition, the inclusion of the background seismicity has the effect of reducing the distance fall-off below the q = 2 slope used in the triggering model (compare Figures 8b and 9b). This implies that the true falloff in triggering rate with distance is steeper than r^{-2} .

5. Discussion

[30] Much of this paper concerns the spatial clustering that is observed in seismicity within narrow time windows around target earthquakes of varying sizes, e.g., the ± 1 -hour windows used for Figures 6, 8, and 9. In an attempt to isolate triggering associated with the target events, the target events



Figure 9. (a) Linear event density versus distance for the LSH southern California catalog in which the actual event magnitudes have been replaced with random magnitudes from the G-R distribution, and (b) the same as Figure 9a but with additional computer-generated triggered seismicity. The different rows compare results for M 2-3, M 3-4, and M 4-5 target earthquakes. Average pre- and post-target event densities are computed in ± 1 hour windows from the target event times. For comparison, a 'background' rate estimated for between 900 and 1000 days from the target events is also plotted. The pre-target densities are shown as dashed lines. One standard error bars are estimated from bootstrap resampling of the target events. Reference slopes of distance $r^{-1.5}$ to $r^{-2.5}$ are also plotted.

are selected to be larger than any earthquakes within specified time and distance limits (3 days and 50 km). In general, we can distinguish among three different kinds of spatial clustering that might be observed using this approach, which are illustrated in Figure 10.

[31] (i) Direct triggering. The target event triggers the post target events and/or is triggered by the pre-target events. If the target event did not occur, the post-target events would not have occurred. If the pre-target events (e.g., in the 1-hour window) did not occur, the target event likely would not have occurred. This process should produce greater post-target than pre-target densities.

[32] (ii) Correlated triggering from a previous large event, which generates an aftershock sequence that lasts longer than the exclusion period around the target events. This causes spatially correlated seismicity that does not primarily represent triggering caused by the target event or its immediate precursors. Events within the time window are triggered either by the main shock or by subsequent aftershocks, but the triggering event is outside the target event exclusion time period. This process will tend to produce equal to greater pre-target event densities compared to posttarget event densities, given the gradual Omori's Law decay in main shock aftershocks with time.

[33] (iii) Correlations originating from some underlying physical process, such as a change in stress state caused by a nearby slow slip event or a change in pore pressure due to fluid migration. In this case, the event clustering does not involve earthquake-to-earthquake triggering.

[34] In real seismicity catalogs, all of these processes likely occur and it is difficult to completely separate them. For example, fluid migration might generate a seismicity swarm, but some of the events in the swarm might be triggered by previous swarm events rather than directly by the



Figure 10. A cartoon illustrating three possible sources of clustering in short time intervals around target earthquakes. The target event X is selected to be larger than nearby events within a narrow time window. In mechanism i, the target earthquake X triggers aftershocks b, c, and d. The target event itself may have been triggered by foreshock a. In mechanism ii, the target earthquake X and both post- and pre-target activity (a, b, and c) are triggered by an earlier main shock M outside the time window of interest. In mechanism iii, all the earthquakes are triggered by an external event, such as fluid migration or slow slip.

fluid migration. Similarly, chains of triggering can cause some blurring between categories (i) and (ii), although the distinction between these processes is reasonably distinct for short time windows around the target events.

[35] Felzer and Brodsky [2006] implicitly assumed in their analysis of southern California seismicity that their observed post-target event densities were driven by triggering related to the target event (i.e., mechanism (i) above). However, *Richards-Dinger et al.* [2010] showed that many of their (target event, aftershock) pairs were likely independent aftershocks of an earlier large main shock (i.e., mechanism (ii) above). This can explain at least some of the time symmetric behavior we observe in our ± 1 hour event densities because the aftershock sequences of large earthquakes like Landers and Hector Mine extend beyond the one to three month exclusion windows that we applied in our analysis. Mechanism (ii) predicts equal numbers of pre- and posttarget events (i.e., time symmetric behavior) when the preand post-target windows are short compared to the time since the major earthquake. However, for longer pre- and

post-target time windows, there will be more pre-target events than post-target events. This can be seen in the 900 to 1000 day windows around the target events used to define the 'background' rate in Figure 6, in which there are more pre-target than post-target events. This effect is more apparent for southern California because of the larger number of major earthquakes during the catalog time periods. Because this effect is independent of the magnitude of the aftershocks, it also appears in the magnitude-randomized LSH catalog, plotted in Figure 9a.

[36] Because neither effect (ii) or (iii) can produce a surplus of post-target events compared to pre-target events in short time windows, the key to isolating triggering related to the target events is to compare pre- and post-target rates. Figure 6 shows that a surplus of post-target events can only be reliably resolved to distances of 1 to 3 km for M 2–3 targets, and distances of 3 to 10 km for M 3–5 targets. Determining the aftershock density as a function of distance is complicated by the fact that the 1-hour pre-target density exhibits a similar falloff with distance as the 1-hour posttarget density. This could be interpreted as indicating that both are dominated by a similar earthquake-to-earthquake triggering process, i.e., the pre-target events are a foreshock sequence and the post-target events are aftershocks. Support for this model comes from the fact that both the pre- and post-target activity at close distances exhibit a power law time dependence, i.e., the foreshocks follow an inverse Omori's Law and the aftershocks follow Omori's Law (see Figure 4 and discussion in *Helmstetter and Sornette* [2003a] and *Brodsky* [2011]). However, a difficulty with this interpretation is that the observed foreshock-to-aftershock ratios for the M 2–4 targets are too large to be compatible with predictions of standard triggering models (as shown here and in Shearer [2012]). And, as discussed above, a power law decay in event density with distance does not necessarily imply earthquake-to-earthquake triggering, as such a decay can be observed even at time separations before the P-wave arrival [Richards-Dinger et al., 2010]. It is also not clear if the ± 1 hour pre- and post-target curves in Figure 6 have exactly the same slope. At least for the LSH data, it appears that the post-target falloff with distance is slightly steeper than the pre-target falloff. This plot, together with the synthetic calculations, show that the fall-off in linear event density in the triggered seismicity is about r^{-2} , although there is some variation in slope among the different target magnitude bins and between the southern and northern California catalogs. It is apparent, however, that the falloff is substantially steeper than the $r^{-1.37}$ obtained by Felzer and Brodsky.

[37] Resolving among these mechanisms is important for several reasons. First, as discussed by *Felzer and Brodsky* [2006] and *Richards-Dinger et al.* [2010], resolving the distance dependence of aftershocks provides a key constraint on the issue of static versus dynamic triggering. Static stress changes are very unlikely to trigger aftershocks at 50 km distance from a M 2–3 main shock. However, as we have shown, aftershocks in California from such small main shocks can only be reliably resolved to distances of 1 to 3 km and exhibit a relatively steep fall off with distance. Second, exploring the limitations of earthquake triggering models will help to identify non-random behavior in earthquake catalogs that may be associated with physical changes

that drive earthquakes. Swarms are a common example of such behavior, but they may be only the most obvious manifestation of changes in earthquake rates that produce clustering over a variety of spatial and temporal scales. Developing quantitative methods for identifying such changes and distinguishing them from aftershock clustering is challenging, but will provide a better understanding of both the processes that trigger aftershocks and the underlying changes in crustal properties that ultimately drive seismic activity.

[38] Acknowledgments. This research was supported by the USGS/ NEHRP program and by the Southern California Earthquake Center. Shinji Toda and an anonymous reviewer provided helpful comments. Karen Felzer provided a copy of her AftSimulator program, which was very helpful in understanding how computer simulations of triggering work, although I did not use AftSimulator itself for the results presented here.

References

- Agnew, D. C., and L. M. Jones (1991), Prediction probabilities from foreshocks, J. Geophys. Res., 96, 11,959–11,971.
- Båth, M. (1965), Lateral inhomogeneities of the upper mantle, *Tectonophysics*, 2, 483–514.
- Brodsky, E. E. (2011), The spatial density of foreshocks, *Geophys. Res. Lett.*, *38*, L10305, doi:10.1029/2011GL047253.
- Felzer, K., and E. Brodsky (2006), Decay of aftershock density with distance indicates triggering by dynamic stress, *Nature*, 441, 735–738.
- Felzer, K. R., T. W. Becker, R. E. Abercrombie, G. Ekstrom, and J. R. Rice (2002), Triggering of the 1999 Mw 7.1 Hector Mine earthquake by aftershocks of the 1992 Mw 7.3 Landers earthquake, *J. Geophys. Res.*, 107(B9), 2190, doi:10.1029/2001JB000911.
- Felzer, K., R. Abercrombie, and G. Ekstrom (2004), A common origin for aftershocks, foreshocks, and multiplets, *Bull. Seismol. Soc. Am.*, 94, 88–98, doi:10.1785/0120030069.
- Hainzl, S., and Y. Ogata (2005), Detecting fluid signals in seismicity data through statistical earthquake modeling, J. Geophys. Res., 110, B05S07, doi:10.1029/2004JB003247.
- Helmstetter, A. (2003), Is earthquake triggering driven by small earthquakes?, *Phys. Rev. Lett.*, 91, 058501.
- Helmstetter, A., and D. Sornette (2002), Subcritical and supercritical regimes in epidemic models of earthquake aftershocks, *J. Geophys. Res.*, *107*(B10), 2237, doi:10.1029/2001JB001580.
- Helmstetter, A., and D. Sornette (2003a), Foreshocks explained by cascades of triggered seismicity, J. Geophys. Res., 108(B10), 2457, doi:10.1029/ 2003JB002409.
- Helmstetter, A., and D. Sornette (2003b), Predictability in the Epidemic-Type Aftershock Sequence model of interacting triggered seismicity, *J. Geophys. Res.*, 108(B10), 2482, doi:10.1029/2003JB002485.
- Helmstetter, A., D. Sornette, and J.-R. Grasso (2003), Mainshocks are aftershocks of conditional foreshocks: How do foreshock statistical properties emerge from aftershock laws, *J. Geophys. Res.*, 108(B1), 2046, doi:10.1029/2002JB001991.
- Helmstetter, A., Y. Y. Kagan, and D. D. Jackson (2005), Importance of small earthquakes for stress transfers and earthquake triggering, *J. Geophys. Res.*, 110, B05S08, doi:10.1029/2004JB003286.
- Hutton, K., J. Woessner, and E. Hauksson (2010), Earthquake monitoring in southern California for seventy-seven years (1932–2008), *Bull. Seismol. Soc. Am.*, 100, 423–446, doi:10.1785/0120090130.

- Kagan, Y. Y. (1991), Likelihood analysis of earthquake catalogs, *Geophys. J. Int.*, 106, 135–148.
- Kagan, Y. Y. (2004), Short-term properties of earthquake catalogs and models of earthquake source, *Bull. Seismol. Soc. Am.*, 94, 1207–1228, doi:10.1785/012003098.
- Kagan, Y. Y. (2007), Earthquake spatial distribution: The correlation dimension, *Geophys. J. Int.*, 168, 1175–1194, doi:10.1111/j.1365-246X.2006.03251.x.
- Lin, G., P. M. Shearer, and E. Hauksson (2007), Applying a threedimensional velocity model, waveform cross correlation, and cluster analysis to locate southern California seismicity from 1981 to 2005, *J. Geophys. Res.*, 112, B12309, doi:10.1029/2007JB004986.
- Llenos, A. L., J. L. McGuire, and Y. Ogata (2009), Modeling seismic swarms triggered by aseismic transients, *Earth Planet. Sci. Lett.*, 281, 59–69, doi:10.1016/j.epsl.2009.02.011.
- Lohman, R. B., and J. J. McGuire (2007), Earthquake swarms driven by aseismic creep in the Salton Trough, California, J. Geophys. Res., 112, B04405, doi:10.1029/2006JB004596.
- Ogata, Y. (1998), Space-time point-process models for earthquake occurrence, *Ann. Inst. Stat. Math.*, 50, 379–402.
- Ogata, Y. (1999), Seismicity analysis through point-process modeling: A review, *Pure Appl. Geophys.*, 155, 471–507.
- Ouillon, G., and D. Sornette (2005), Magnitude-dependent Omori law: Theory and empirical Study, *J. Geophys. Res.*, 110, B04306, doi:10.1029/2004JB003311.
- Ouillon, G., E. Ribeiro, and D. Sornette (2009), Multifractal Omori law for earthquake triggering: New tests on the California, Japan and worldwide catalogs, *Geophys. J. Int.*, 178, 215–243.
- Richards-Dinger, K., R. S. Stein, and S. Toda (2010), Decay of aftershock density with distance does not indicate triggering by dynamics stress, *Nature*, 467, 583–586.
- Roland, E., and J. J. McGuire (2009), Earthquake swarms on transform faults, *Geophys. J. Int.*, *178*(3), 1677–1690, doi:10.1111/j.1365-246X. 2009.04214.x.
- Shcherbakov, R., and D. L. Turcotte (2004), A modified form of Båth's law, *Bull. Seismol. Soc. Am.*, 94, 1968–1975.
- Shearer, P. M. (2012), Self-similar earthquake triggering, Båth's law, and foreshock/aftershock magnitudes: Simulations, theory, and results for southern California, J. Geophys. Res., 117, B06310, doi:10.1029/ 2011JB008957.
- Shearer, P. M., and G. Lin (2009), Evidence for Mogi doughnut behavior in seismicity preceding small earthquakes in southern California, *J. Geophys. Res.*, 114, B01318, doi:10.1029/2008JB005982.
- Sornette, D., and M. J. Werner (2005a), Constraints on the size of the smallest triggering earthquake from the epidemic-type aftershock sequence model, Båth's law, and observed aftershock sequences, J. Geophys. Res., 110, B08304, doi:10.1029/2004JB003535.
- Sornette, D., and M. J. Werner (2005b), Apparent clustering and apparent background earthquakes biased by undetected seismicity, J. Geophys. Res., 110, B09303, doi:10.1029/2005JB003621.
- Toda, S., R. S. Stein, and T. Saglya (2002), Evidence from the AD 2000 Izu islands earthquake swarm that stressing rate governs seismicity, *Nature*, *419*, 58–61.
- Vidale, J. E., and P. M. Shearer (2006), A survey of 71 earthquake bursts across southern California: Exploring the role of pore fluid pressure fluctuations and aseismic slip as drivers, J. Geophys. Res., 111, B05312, doi:10.1029/2005JB004034.
- Waldhauser, F., and D. P. Schaff (2008), Large-scale relocation of two decades of Northern California seismicity using cross-correlation and double-difference methods, J. Geophys. Res., 113, B08311, doi:10.1029/ 2007JB005479.