Within this supplement, we provide more details about analysis methods described in the main manuscript, to facilitate an easier reproducibility of our results. Data and analysis codes are available from the first author. We start by describing experimental procedures (Section DR1) and the determination of the correlation dimension, $D_2$, from hypocenter distributions which includes a sensitivity analysis on hypocentral uncertainty (Section DR2). We then describe approaches to resolve focal mechanism heterogeneity (Section DR3) and a maximum likelihood method for the determination of Gutenberg-Richter $b$-values (Section DR4).
DR1  Experimental methods and seismic analysis

We present results from triaxial compression experiments conducted on oven-dried, cylindrical (diameter = 50 mm, height = 105 mm) Westerly granite samples (Figure DR1). Rock fracture and frictional sliding were achieved within a hydrostatic pressure vessel with individual servo-controls for axial load and confining pressure ($P_c$=120–150 MPa). The samples were separated from the confining oil by an elastic rubber jacket and loaded axially at a displacement rate of $3 \cdot 10^{-4}$ mm/s up to a maximum vertical displacement of $u \sim 7$ mm or less depending on the characteristics of stick-slips in each experiment. An external load cell and two vertical and horizontal strain gauges measured axial force and strain with a sampling rate of 10 Hz.

To simulate a range of fault roughness, we used both cut and fractured specimens. We fractured two samples at $P_c = 75$ MPa and cut the other samples at a 30° angle to the vertical sample axis. The planar, cut surfaces were polished with an abrasive silicate-powder (grid: #290) or roughened with coarser powders (grid: #60 - #80). Fault structure and roughness were determined by X-ray imaging and white-light interferometry scans. The latter confirmed the expected increase in roughness from polished to roughened to fractured surfaces expressed in higher power-spectral-density across the examined wavelengths (see inset in Figure 3 in the main text). The corresponding Hurst exponents, which describe the relative distribution of spectral power at short vs. long wavelengths, showed little variations in direction of slip ($H \sim 0.4–0.5$). For the planar surfaces, we observed a pronounced rollover in power-spectral-density due to the absence of long-wavelength topography beyond $\sim 0.8$ mm. The fractured surfaces appear largely self-affine across the resolved wavelengths.
We analyze tiny AE events recorded and located by using a 16 channel, high-speed data acquisition system (for details, see Stanchits et al., 2006). Accurate AE locations were possible due to high-sampling rates (10 MHz) and anisotropic, layered velocity models from active velocity measurements in 30s intervals with hypocentral uncertainty estimates of 1 to 3 mm (Goebel et al., 2014). We determined relative AE event sizes in our experiments by averaging peak amplitudes across the laboratory array of piezo-transducers and correcting for source-station distances (Zang et al., 1998). Based on AE amplitudes, $A$, we assigned relative magnitudes $M=\log_{10} A$ to each event on an experiment-specific scale. The corresponding AE rupture dimensions are commonly small so that even the largest AEs are expected to be well below 10 mm (c.f. McLaskey et al., 2014). Thus, AE events rupture only part of the laboratory faults and seismicity statistics are not expected to be biased by the finiteness of the available fault area.

The analysis of spatial and magnitude distributions was limited to events that
Analyzed Time
Differential Stress
AE-rate
Magnitude
Polished
Relative Magnitude
Relative Magnitude

Figure DR2: Acoustic emission rate (green curve), magnitudes (black dots) and differential stress (orange curve) during the first stick-slip cycle on a polished (left), roughened (center) and fractured (right) fault. The analyses of spatial and magnitude distributions were limited to events within a time window close to peak stress (analyzed time). The onsets of slip events are highlighted by black, down-facing arrows. The onsets of slip events occurred at stresses close to failure stresses (Figure DR2). While we generally observed higher AE rates for fractured than for roughened and polished surfaces, the corresponding loading curves are complex. Initial stick-slip periods can be comparatively short (i.e. 2800 sec) or extend over more than 10,000 sec with no clear dependency on fault roughness. Similarly, stress drops vary significantly for initial stick-slip events with some indication that polished surfaces, on average, promote higher stress drops. The spatial and magnitude distributions, however, showed a clear correlation with initial fault roughness as described in the main text.

DR2 Pair correlation function and fractal dimensions

The spatial distribution of AE hypocenters in our experiments varies substantially as a function of roughness. To quantify this variability, we compute the Pair-Correlation-Function, $C(r)$, at all scales and for all AE event pairs, $N$, with sep-
aration distance, $s$, less than $r$ (Kagan and Knopoff, 1980; Grassberger, 1983):

$$C(r) = N(s < r) / N_{\text{tot}}^2,$$  \hspace{1cm} (1)

where $N_{\text{tot}}$ is the total number of AE events within each experiment. Here we use only AE events in narrow time windows prior to the first stick-slips which are most representative of the initial surface roughness. Through the selection of relatively narrow time windows, we can avoid events that repeatedly rupture the same fault patch which may be problematic for reliable estimates of fractal dimensions (Main, 1992). After log-transformation, the Pair-Correlation-Function in eq. 1 is approximately linear between 1 to 10 mm for our datasets, indicating that the spatial distributions of AEs are fractal within this distance-range (Main et al., 1992; Main, 1992; Henderson et al., 1992; Wyss et al., 2004). The corresponding correlation dimension, $D_2$, was estimated by a least-square fit of the linear portion of logarithmically-binned Pair-Correlation-Functions:

$$D_2 = \log(C(r)) / \log(r).$$  \hspace{1cm} (2)

The distance, $r$, in eq. 2 is bounded by a minimum ($r_{\text{min}}$) and a maximum cut-off ($r_{\text{max}}$) due to finite event density and finite sample size. The former is simply a function of the overall point-density relative to the size of the considered area, i.e. $r_{\text{min}} \approx 2 \cdot \text{height} \cdot \text{diameter} \cdot (1/N)^{1/D}$, where $N$ is the total number of events and $D$ is the dimension of the embedding Euclidean space (Kagan, 2007a). In practice, this minimum distance approximately coincides with $r$ at $C(r) \sim 2 / N_{\text{tot}}^2$, that is the distance at which the function is comprised of more than a single event. The maximum cut-off is determined by maximizing the goodness-of-fit between obser-
Figure DR3: Example results of synthetic test for $D_2$ estimates based on the algorithm described in the text using point clouds with known dimensions. True dimensions are indicated by dashed lines and estimated values are shown as a function of point density as colors markers.

vation and model. Because of the gradual rollover of $C(r)$ at large distances, we test different goodness-of-fit measures and obtained robust results for an adjusted $R^2$ value applied to the windowed derivative of the pair-correlation-function. The robustness of the described algorithms was tested for point clouds with known fractal geometry (Figure DR3) and uncertainties in $D_2$ can be determined by bootstrap resampling of the underlying spatial distributions of AEs. We select relatively narrow time windows to compute $C(r)$ and the corresponding correlation dimensions to avoid AE events that repeatedly rupture the same fault patch which may be problematic for reliable estimates of fractal dimensions (Main et al., 1992).
Kagan (2007a) showed that a number of factors such as location errors, geometric and boundary effects as well as spatial and temporal inhomogeneity may bias estimates of the correlation dimension of natural seismicity. In the laboratory, we encounter a favorable situation compared to natural seismicity because both array size and geometry remain approximately constant throughout the experiments so that only small variations in catalog quality and hypocentral uncertainty are expected. Moreover, both sample volume and fault orientations are similar in all experiments, which simplifies the comparison of $D_2$ estimates between experiments.

We performed a series of tests by adding Gaussian uncertainty to seismic event locations and by resampling subsets of the data (Figure DR4 & DR5). Location uncertainties are generally expected to inflate $D_2$-values due to the loss of fractal clustering at the smallest scales. For an AE location uncertainties of 1–3 mm (Goebel et al., 2014), we expect $D_2$ to increase by less than 0.1 given the point density and stable seismic array geometry. Since we hypocentral uncertainties remain constant between experiments, we expect systematic changes in $D_2$ so that relative differences can be resolved reliably.
Figure DR4: Estimates of correlations dimensions converge rapidly to the true dimension within a reasonable range of point densities which are comparable to the AE densities in our laboratory experiments.

Figure DR5: Influence of location uncertainty on correlation dimensions for a fractal point cloud with known dimension of $D_2 = 1.58$. Location uncertainties are expected to influence fractal dimensions in a systematic fashion for all experiments. For the here resolved AE location uncertainties of 1–3 mm, we expect $D_2$ to change by less than 0.1 given the point density and stable seismic array geometry.
Focal mechanisms and stress-field heterogeneity

Focal mechanisms orientations and their overall variability can provide insight into stress field heterogeneity at the scale of seismic events. In our laboratory experiment, we use AE focal mechanism heterogeneity as a proxy for stress heterogeneity within the laboratory fault zones.

We determine full moment tensors of AEs from first P-wave amplitudes and their rise times using the hybridMT package (Kwiatek et al., 2016). Before the inversion, P-wave amplitudes were corrected for coupling effects between AE sensors and rock surfaces, which change as a function of confining pressure (Kwiatek et al., 2014). The corrected amplitude data were inverted for six independent moment tensor components using a least squares approach (for details see Kwiatek et al., 2016).

We use several measures to determine the overall heterogeneity at the scale of polished, roughened and fractured faults. First, we analyze the variations in P-axis orientations of focal mechanisms which show increased scatter and deviation from the applied far-field stresses for roughened and fractured surfaces (see main manuscript, Figure 1). Second, we determine the distance angle between AE moment tensors as measure of difference of micro-scale stresses at failure. For this purpose, we determine distance-angle between full moment tensors and double-couple components using the quaternion convention and inner-tensor product (e.g. Kagan, 2007b; Tape and Tape, 2012; Cesca et al., 2014). Both P-axes orientations and rotation angles between tensor pairs were determined from deviatoric components of the decomposed moment tensor (Knopoff and Randall, 1970).

While, the distribution of distance angle may change depending on the corresponding approach, the overall tendencies for minimum rotations angles and
inner tensor product are consistent. Our results suggest that smoother faults and larger magnitude AEs show less variations in focal mechanism and smaller distance angle between tensor pairs which may originate from smoother underlying stress fields compared to rougher faults and small-magnitude AEs (Figure DR6 & DR7).
Figure DR6: Distribution of minimum rotation angles between focal mechanism pairs for polished, roughened and fractured surfaces. The median rotation angles increase systematically from polished to roughened and fractured surfaces in agreement with observed $P$-axis scatter (see main manuscript). Moreover, rotation angles decrease if only larger magnitude events (>5.8) are considered.
Figure DR7: Minimum rotation angles of focal mechanism pairs as a function of AE magnitude. Solid lines show the median for different magnitude bins and shaded area is the 95% confidence interval. Large angles highlight a high degree of heterogeneity in focal mechanisms. This heterogeneity decreases significantly from polished, to roughened to fractured surfaces. Moreover, polished and roughened surfaces exhibit less heterogeneity at the scale of larger magnitude AE events. The resolved focal mechanism heterogeneity is apparent in both minimum rotation angles (Kagan, 2007b) and inner tensor products (Tape and Tape, 2012).
DR4  Magnitude distributions and $b$-value estimates

The AE magnitude distribution in the our experiments can commonly be described by a power-law with an exponent, $b$. This $b$-value was determined by using a maximum-likelihood approach (Aki, 1965).

$$ b = \frac{1}{M - M_c} \log(e), $$

(3)

where $M_c$ is the magnitude of completeness, corrected for bin-size (Utsu et al., 1965; Guo and Ogata, 1997) and $\bar{M}$ is the mean magnitude above completeness. For reliable $b$-value estimates, we require distributions to contain at least 300 AE events. Moreover, eq. 3 shows that $b$-value estimates are sensitive to the estimated magnitude of completeness, $M_c$. To avoid biases, we use an objective approach to invert for both $M_c$ and $b$ through minimizing the misfit between observed and modeled power-law distributions (Clauset et al., 2009).

The statistical uncertainty of $b$-value estimates can be determined from (Shi and Bolt, 1982):

$$ \sigma_b = \frac{b^2 \log(e)}{\sqrt{\frac{\sum_{i=1}^{n}(M_i - \bar{M})^2}{n(n-1)}}}, $$

(4)

where $n$ is the number of events above $M_c$, $M_i$ is the magnitude of individual AE events and $\bar{M}$ is the mean magnitude for events greater or equal than $M_c$.

While the AE magnitude distributions for polished, roughened and fractured surfaces can commonly be described by a power-law, this power-law breaks down for polished surfaces after several successive stick-slip events (Figure DR8). This marks a potential transition from Gutenberg-Richter-like to Characteristic magnitude distributions as a function of fault smoothing and surface roughness, suggesting that magnitude distributions are indeed controlled by surface heterogeneity.
Figure DR8: Distribution of event magnitudes for AEs that occurred during latter stick-slip cycles on a polished surface. Error bars highlight 95% confidence level assuming Poissonian distributed event numbers. The distribution substantially deviates from the commonly observed exponential decrease in event number for large magnitude events, showing a pronounced second peak above $M$-5. This second peak and connected break-down of Gutenberg-Richter-type scaling may suggest that magnitude distributions become more characteristic and are controlled by a dominant length scale which lead to the preferred occurrence of events within a narrow range of magnitudes (e.g. Wesnousky, 1994). The transition from Gutenberg-Richter to characteristic magnitude distribution may be a result of fault evolution and surface smoothing.
References


McLaskey, G.C., Kilgore, B.D., Lockner, D.a., and Beeler, N.M., 2014, Laboratory Generated M -6 Earthquakes: Pure


