

**Supporting Information for
“What controls variations in aftershock productivity?”**

Kelian Dascher-Cousineau¹, Emily E. Brodsky¹, Thorne Lay¹ Thomas H. W. Goebel²

¹University of California Santa Cruz, Santa Cruz, California, USA.

²University of Memphis, Memphis, Tennessee

Contents

| | | |
|-------------|---|----|
| S1 | Critical examination of relative productivity..... | 1 |
| S1.1 | Scaling effects | 1 |
| S1.2 | Catalog completeness..... | 6 |
| S1.3 | Alternative clustering method | 16 |
| S1.4 | Background Seismicity | 24 |
| S2 | Relative productivity as a function of miscellaneous parameters . | 26 |
| S3 | Additional Supporting Information (Files uploaded separately) .. | 28 |
| S3.1 | Trained SVM model..... | 28 |
| S3.2 | Catalogued source attributes | 28 |
| S3.3 | Alternative clustering code | 28 |

S1 Critical examination of relative productivity

In this section we examine potential biases to relative productivity that may arise from our counting method, its parameterization and catalog artifacts.

S1.1 Scaling effects

We first assess whether any biases are introduced by the definition of relative productivity as the number of aftershocks referenced to the number of aftershocks expected for a given magnitude. If the global trend with magnitude is incorrectly modelled, a systematic trend of relative productivity with magnitude would be observed. Figure S1 shows that no such trend exists and thus definition of relative productivity appears reasonable.

Next, we examined whether there may be issues related to the definition of the scaling relationship we used in space windowing. Figure S2 compares the measurement of relative productivity utilizing two different inferred scalings between rupture length and magnitude [Brengman *et al.*, 2019; Wells and Coppersmith, 1994]. We find that 79% of sequences have the same aftershock counts regardless of the scaling used and that the excursions tend to be small.

Using a the group of earthquakes with $M_W < 7.5$ reveals more statistically significant subgroups for attributes which do not require finite fault inversion (Figure S4). Conversely, those attributes derived from finite fault inversion become less stable and, as a result, their relative productivity statistics are more varied (Figure S3). The attributes defining the largest and smallest relative productivity are the same in both cases. Note that within for the set of smaller mainshocks, the median value of relative productivity can be censored, implying that the median mainshock of the subgroup had no measured aftershock.

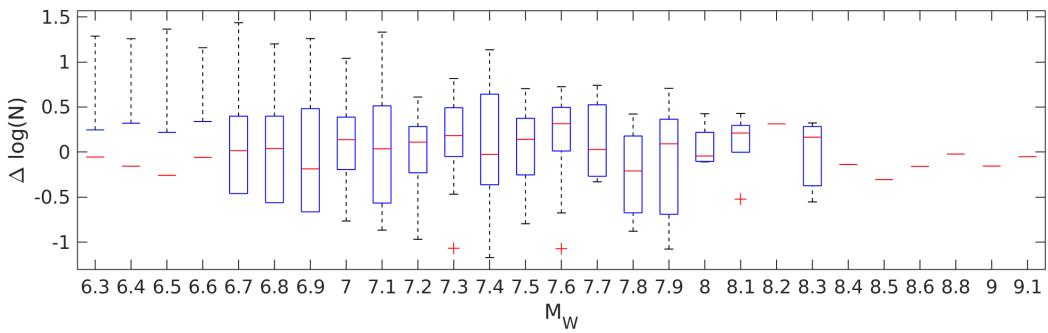


Figure S1. Box and whiskers plot of relative productivity as a function of magnitude. Boxes outline the interquartile range. Whiskers outline the range of the data and +'s indicate outliers. Censoring, $\Delta \log(N) = -\inf$, at the lower magnitudes ($M_W < 6.7$) and limited data at highest magnitudes ($M_W > 8.1$) prevent proper assessments of the interquartile range. Nonetheless, we do not observe a systematic decrease in variance that may introduce bias in subsequent analysis.

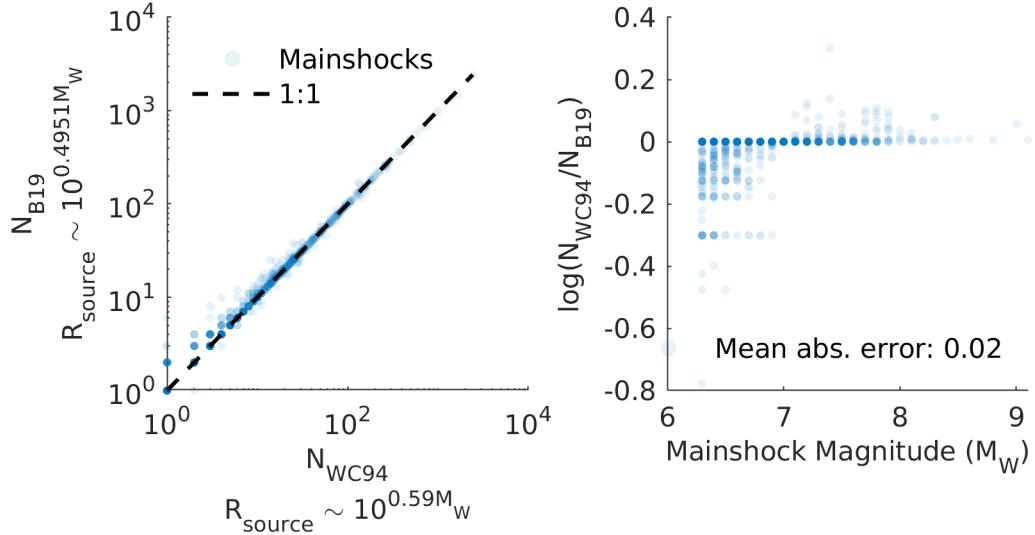


Figure S2. Comparison between relative productivity statistic as obtained from two different scaling relationships: 1) WC94 following [Wells and Coppersmith, 1994] and 2) B19 following [Brengman et al., 2019]. Left: Direct comparison of inferred number of aftershocks. Right: Discrepancy in measurements as a function of mainshock magnitude. We find that 79% of sequences have that exact same aftershock counts. Larger ruptures yield slightly smaller a aftershock counts the B19 scaling, whereas smaller ruptures have the converse relationship. Generally, the effect is subtle with a mean absolute error of 0.02. Note that the discrepancies, $\log(N_{\text{WC}}/N_{\text{B19}})$, are directly equal to differences in measurements of relative productivity.

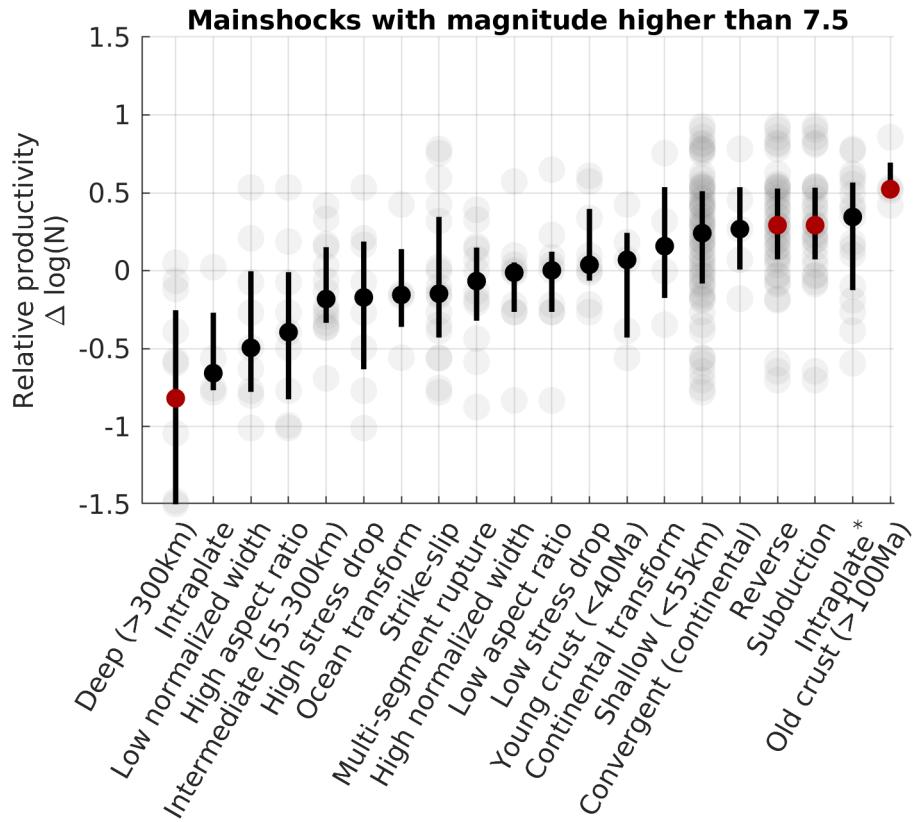


Figure S3. Major results of this study presented for mainshocks with M_W magnitude greater than 7.5. Attributes with red markers are more consistent with the hypothesis that they are sampled from a different continuous distribution than the overall population of earthquakes using a 2-sample Kolmogorov-Smirnov test at a 5% significance threshold.

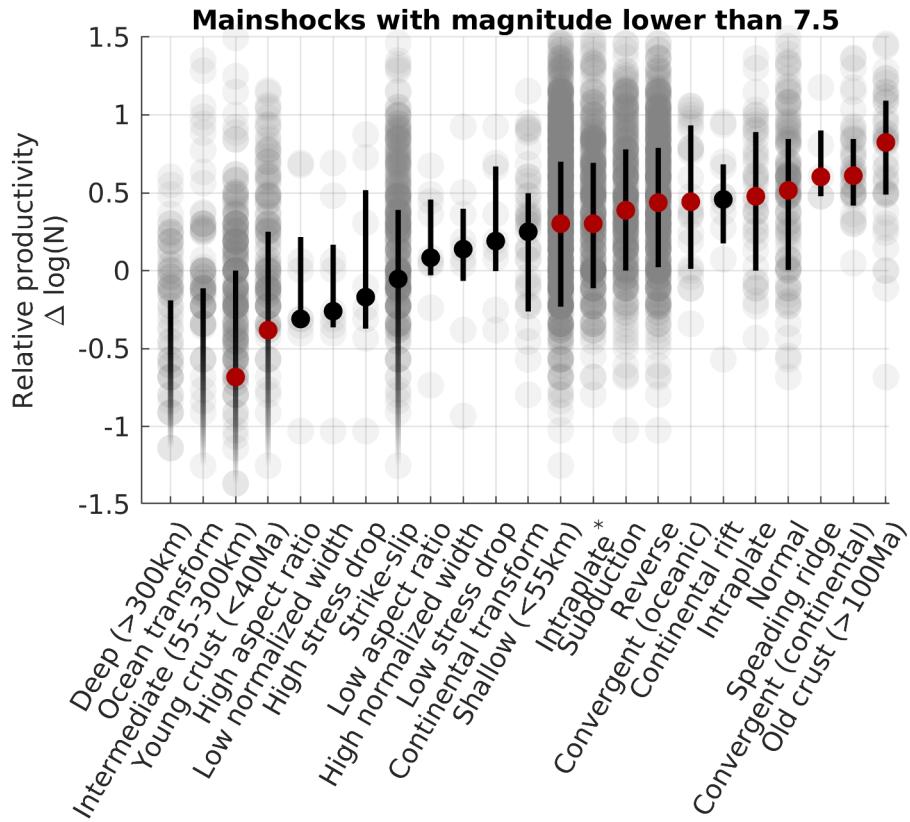


Figure S4. Major results of this study presented for mainshocks with M_W magnitude less than 7.5. Attributes with red markers are consistent with the hypothesis that they are sampled from a different continuous distribution than the overall population of earthquakes using a 2-sample Kolmogorov-Smirnov test at a 5% threshold.

S1.2 Catalog completeness

We test the robustness of our primary findings by testing their validity with a more conservative magnitude of completeness of M_w5 instead of $M_w4.5$ as done in the main text. In these results, there are fewer aftershocks detected and statistical significance is therefore decreased while the chance of inclusion of any background events in the counts is lowered because of the reduced rate for the larger magnitude cut-off, but the results do not violate any of the major findings of the main text. We reproduce Figures 1–11 of the main text with the alternative completeness in Figures S5–S15, which preserve the same order as the main text.

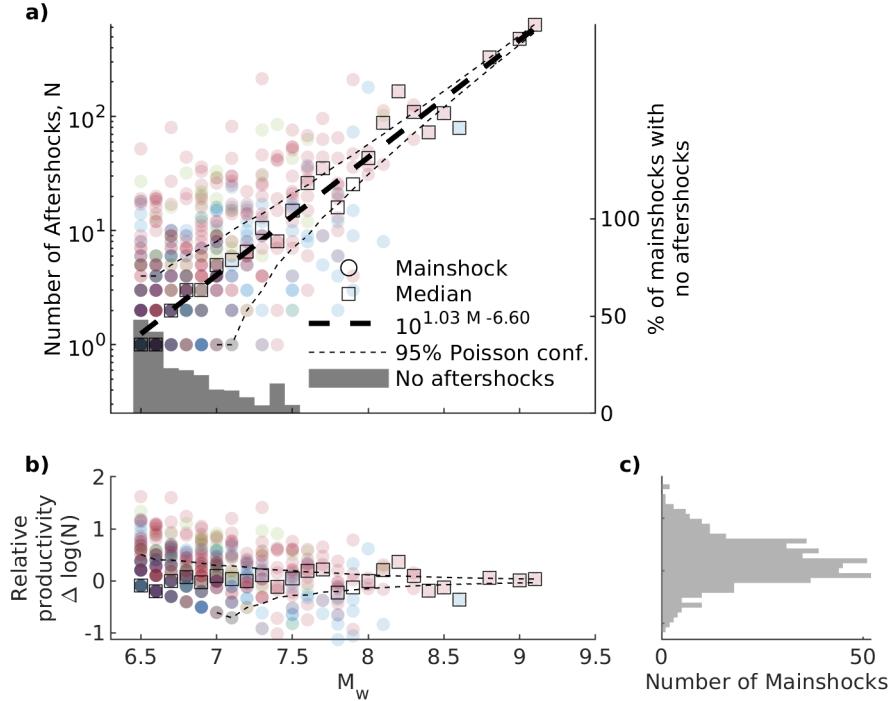


Figure S5. The number of aftershocks of $M_w \geq 5$ within three source dimensions and 60 days as a function of mainshock magnitude identified in the global ISC and NEIC catalogs from 1990 to 2019. Colors indicate faulting style of the mainshock; blue, green and red points correspond to earthquake sequences for which the mainshock was respectively strike-slip, normal or reverse. The global productivity law (dashed line) is fit using a least squares regression through the median log-number of aftershocks for each 0.1 magnitude bin (black squares). The median number includes mainshocks with no aftershocks which are not shown on the plot. Note the individual earthquake sequences (circles) exhibit significant scatter around the productivity law. b) Relative productivity as a function of mainshock magnitude. The relative productivity distribution does not show events with no aftershocks and thus the lower left corner of the plot is underpopulated. c) Histogram of the relative productivity of mainshocks considered in this study.

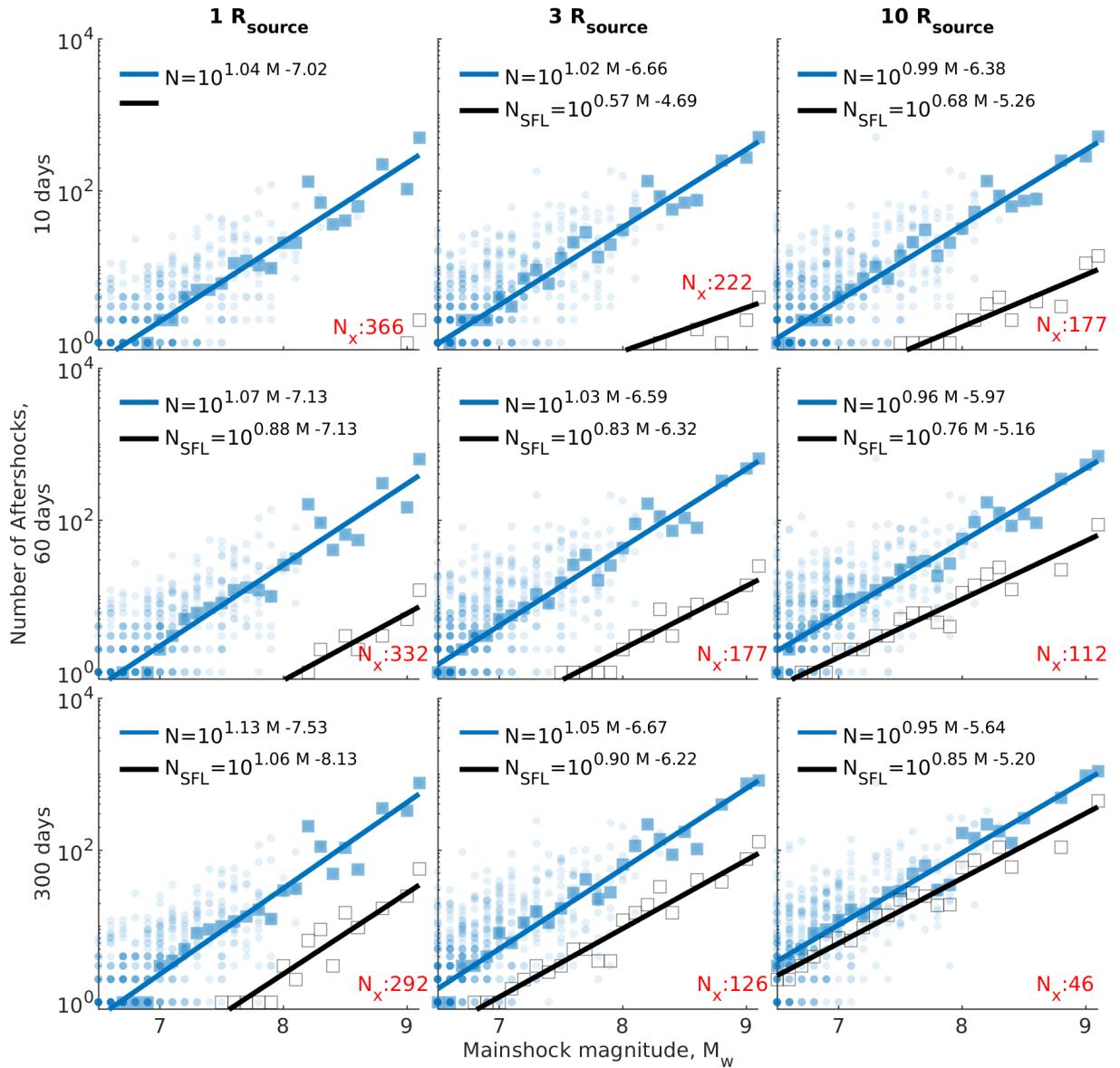


Figure S6. Sensitivity analysis of space-time windows. Time windows of 10, 60, and 100 days and spherical space with radii of 1, 3, and 10 source dimensions (R_{source}) are considered. Blue data are mainshocks identified through our hierarchical declustering routine. Circles are individual mainshocks. Squares are median values for each 0.1 magnitude bins. Regressions are computed using least squares through the median log-number of aftershocks for each 0.1 magnitude bin. For reference, we computed the median productivity relationship (grey squares) for 100 time-shuffled catalogs and the corresponding scaling relationship (black line). For each space-time window, we indicate the number of mainshocks with no aftershocks in red (N_x). Note that as space and time windows increase, more mainshocks have measurable aftershock counts. However, the likelihood of counting background productivity and significantly affecting subsequent parameterization becomes an increasing concern.

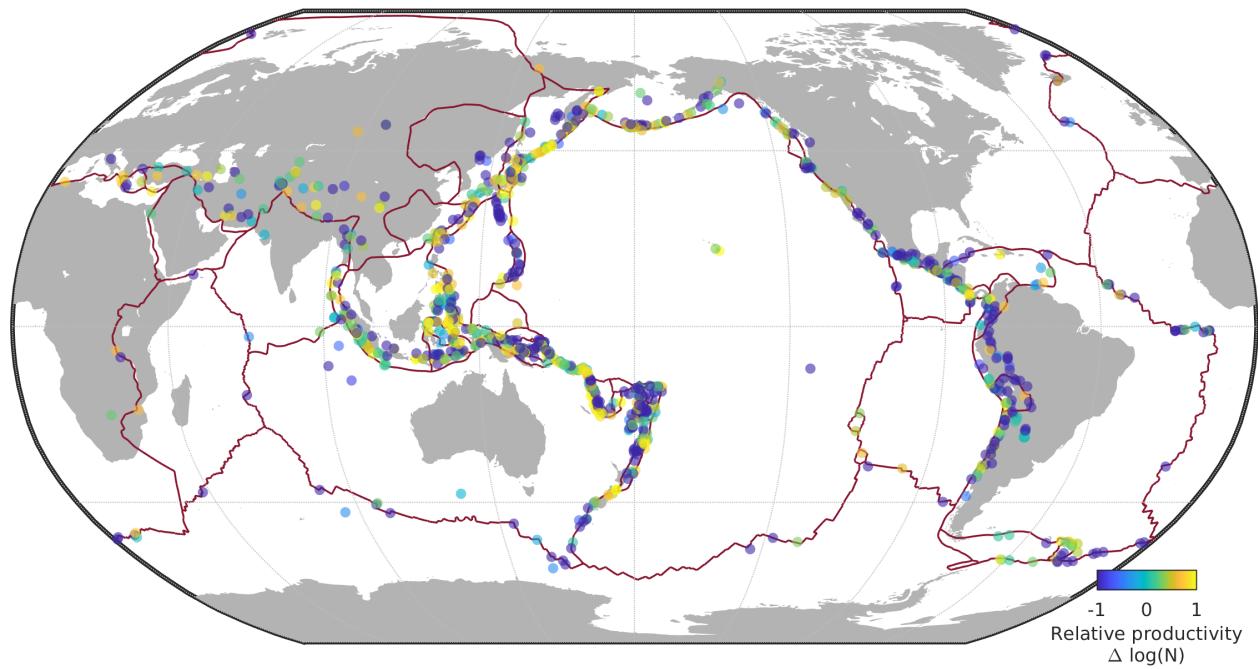


Figure S7. Global map of earthquake productivity as measured with a catalog completeness threshold of $M_w 5.0$. Red lines indicate the surface trace of the tectonic boundaries. Mainshocks with $M_w \geq 6.5$ color-coded according to their relative productivity.

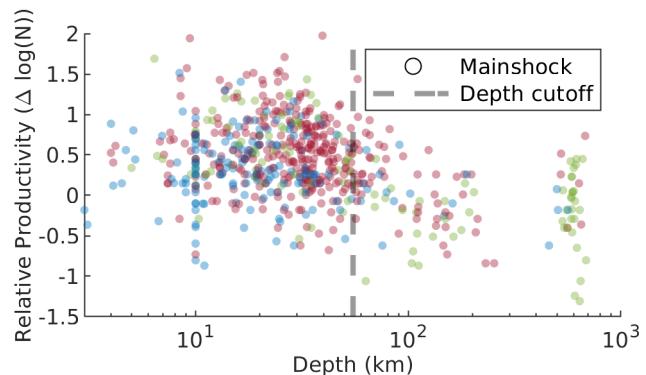


Figure S8. Relative aftershock productivity measured with a catalog completeness threshold of $M_w 5.0$ as a function of depth. Subsequent analysis will only consider earthquakes shallower than the 55 km cutoff (dashed line). Sequences are color-coded according to faulting style of the mainshock (blue: strike-slip, green: normal and red: reverse).

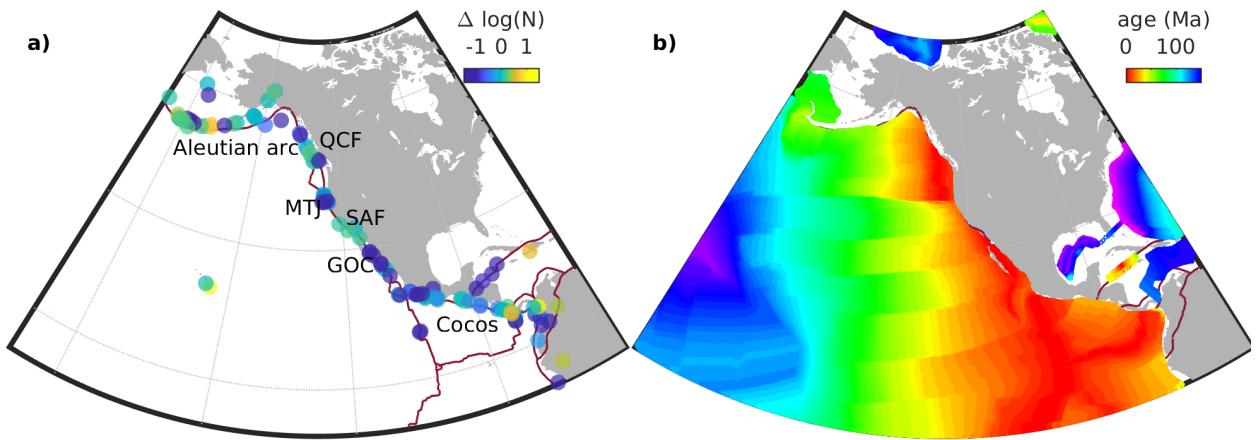


Figure S9. a) Aftershock productivity measured with a catalog completeness threshold of $M_W 5.0$ along the North American coastline. Individual mainshocks (points) are color-coded according to their relative aftershock productivity ($\Delta \log(N)$). The Aleutian arc, the Queen Charlotte Fault (QCF), the Mendocino Triple Junction (MTJ), San Andreas Fault (SAF), the Gulf of California (GOC) and the Cocos plate subduction include areas with coherent productivity. Red line indicates major plate boundaries [Bird, 2003]. b) Seafloor crustal age estimates from Müller *et al.* [2008].

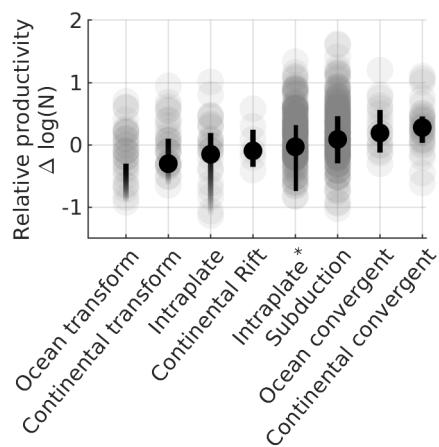


Figure S10. Earthquake productivity measured with a catalog completeness threshold of $M_W 5.0$ by tectonic boundary. Points indicate the relative productivity of individual sequences. Solid markers and error bars indicate the median and the interquartile range. A faded lower error bar implies that mainshocks with no aftershocks are within the interquartile range. Intaplate* indicates earthquakes within 400km from a plate boundary but with a faulting mechanism discordant with the plate boundary (e.g., outer rise events).

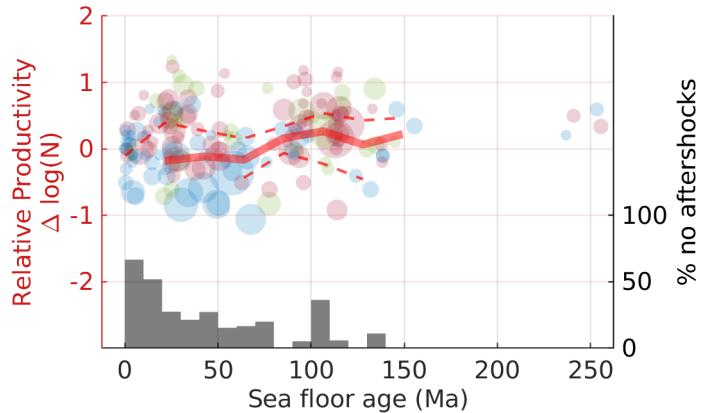


Figure S11. Relative productivity measured with a catalog completeness threshold of $M_W 5.0$ increases as a function of the age of the oceanic lithosphere. Each point indicates an individual earthquake sequence. Sequences are color-coded by faulting style of the mainshock (blue: strike-slip, green: normal and red: reverse). The red line indicates the median average for 20Ma crustal age bins. Dashed lines indicate the corresponding interquartile ranges. Bars indicate the fraction of earthquakes with no aftershocks within each 10 Ma crustal age bin.

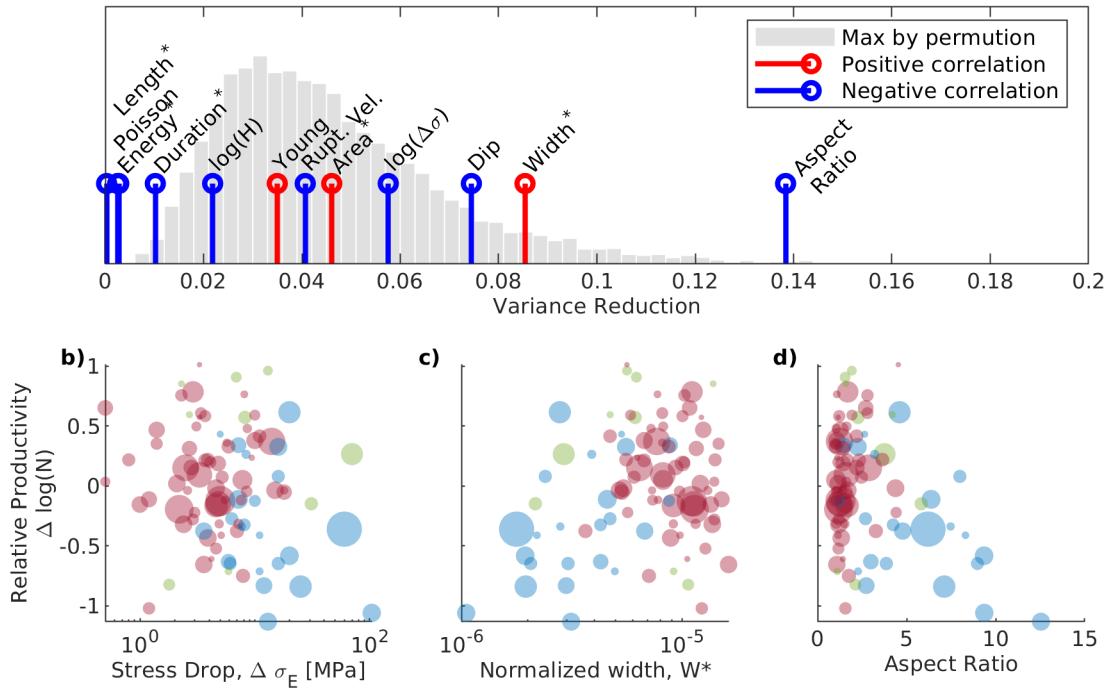


Figure S12. Inspection of the relationship between source attributes and relative productivity measure with a catalog completeness of M_{W5} . a) Goodness of fit of linear regressions for each source attribute in our combined catalog. Top and bottom axes respectively represent the p-value and goodness of fit of each attribute (stems). The probability distribution function in the backdrop indicates the maximal variance reduction outcome of 10000 permutation test of the entire data set. The probability of obtaining a spurious correlation by chance over the whole family of attributes we tested for is derived from the number of random shuffles exceeding the measured variance reduction and normalized to the overall sample (10000). Asterisks indicate scaled and log-transformed variables. The scaled energy, length, duration and area, material properties, velocity, dip, and stress drop ($\Delta\sigma$) of the mainshock rupture all do not yield a statistically significant ($p = 0.05$) linear fit to the relative productivity; the normalized rupture width and aspect ratio of the rupture yield the best fitting linear regressions. Stems are color-coded to indicate whether the source attribute is positively (red) or negatively (blue) correlated with relative productivity. b-d): Relative earthquake productivity as a function of mainshock stress drop, normalized rupture width, and aspect ratio. Individual mainshocks are color-coded according to faulting style as in Figure S5.

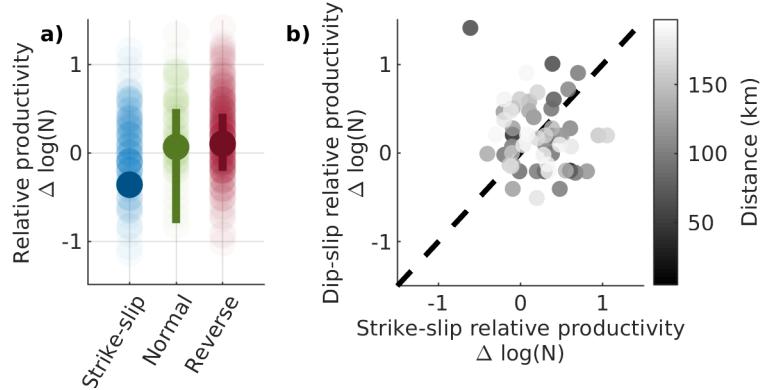


Figure S13. Inspection of the relationship between focal mechanism and relative productivity as measured with a catalog completeness threshold of $M_w 5$. a) Relative aftershock productivity ($\Delta \log(N)$) by focal mechanism. b) Relative aftershock productivity for pairs of earthquake sequences with strike-slip and dip-slip mainshocks within 200 km from each other. Each pair is shaded according to its relative distance. Dashed line indicates a 1:1 relationship, the expectation for a purely site dominated control on relative productivity. Co-located mainshocks pairs generally follow this 1:1 trend, but exhibit considerable scatter.

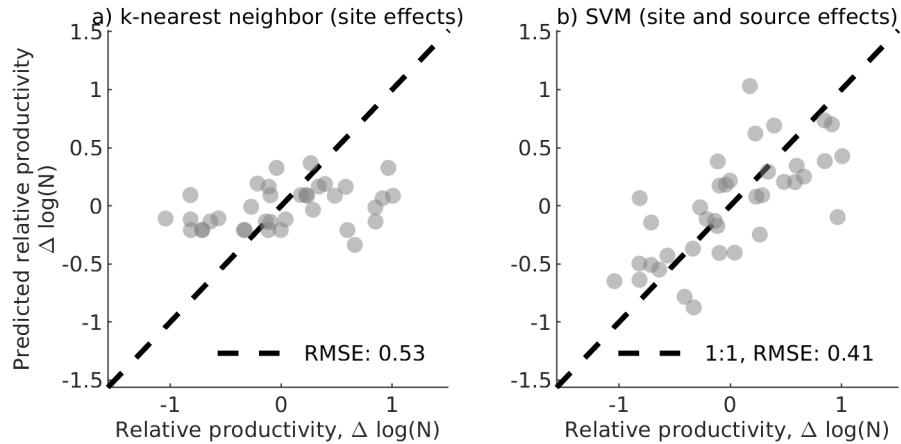


Figure S14. Sensitivity of two forecasting approaches to catalog completeness. Measurements here of relative productivity are computed using a catalog completeness of $M_W 5.0$. Response plots (prediction versus observation) for the k-nearest neighbor algorithm and SVM models. Each point is an individual earthquake sequence. A perfect prediction would place all values on the 1:1 line. The SVM model outperforms the k-nearest neighbor model. Hold-one-out cross-validation ensures that the data for model calibration is separate from the prediction data. Combining both contextual information about the setting (crustal age) and the source (dip and normalized area) yields a root mean square value of 0.39. In particular, the SVM model better predicts extreme cases (highly productive or unproductive).

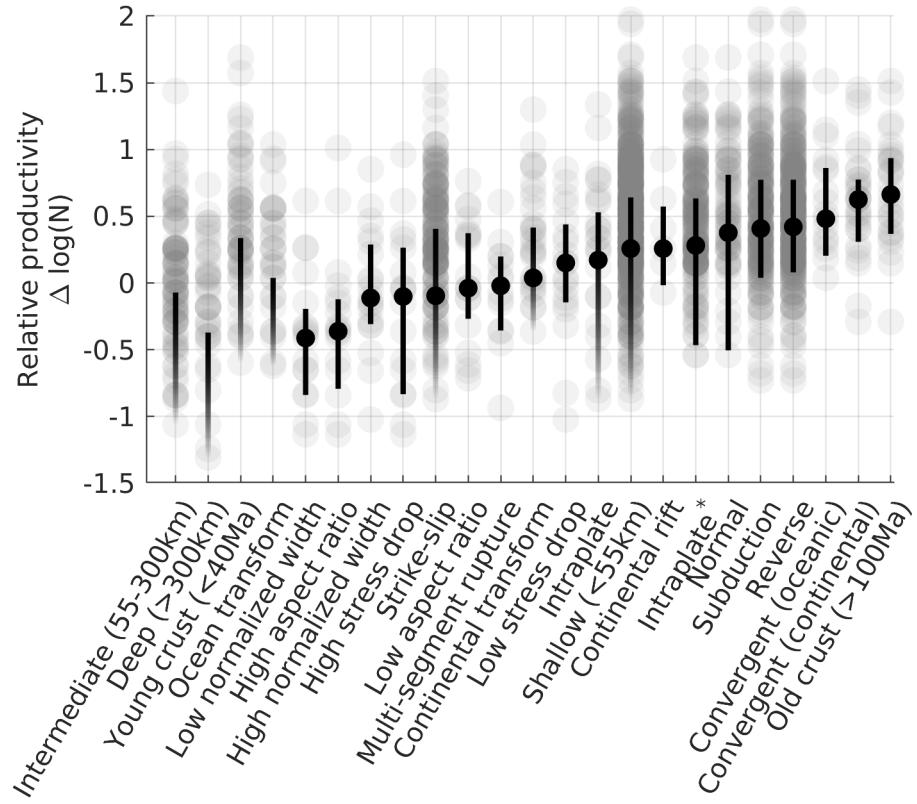


Figure S15. Synthesis of relative productivity measured with a catalog completeness threshold of $M_W 5.0$ according to catalog subsets. The group considered here are the short list which best distinguished relative productivity based on our different lines of investigation. ‘High’ and ‘low’ subsets respectively refer to > 80 th and < 20 th percentile ranges of the data. Grey circles are individual mainshocks. Black points and error bars respectively indicate the median and interquartile range of the subset. Fading error bars imply that mainshock sequences with no aftershocks are within the interquartile range of the data.

S1.3 Alternative clustering method

In this section we present the salient results of our study using an alternate clustering approach. The main manuscript presents results obtained using a hierarchical space-time windowing scheme. We preferred this method because it is more readily reproducible on an event by event basis and therefore provides transparency. A reader is readily capable of inspecting individual past or future earthquakes to see where they fall with respect to the global average, and whether they corroborate or challenge our conclusions without having to perform extensive additional calculations. However, it is important to also check the result robustness with more mathematically rigorous aftershock detection methods.

Here we present our major results using aftershock counts obtained following *Zaliapin et al.* [2008]. This approach seeks to build earthquake families by linking earthquakes to parent events. Separation of clusters is achieved by finding a decision boundary between background events and clustered events. See *Zaliapin et al.* [2008] for a detailed overview of the method, distance metrics, and theoretical connections to other schemes (ETAS). Though there are discrepancies, aftershock counts obtained using this method correspond well to counts obtained using space-time windowing (see Figure S16). Additionally, we find few differences to the major results of the study (see Figure S17–S26). The workflow and detailed explanations of the clustering routine are available here: <https://github.com/tgoebel/cluster-analysis>.

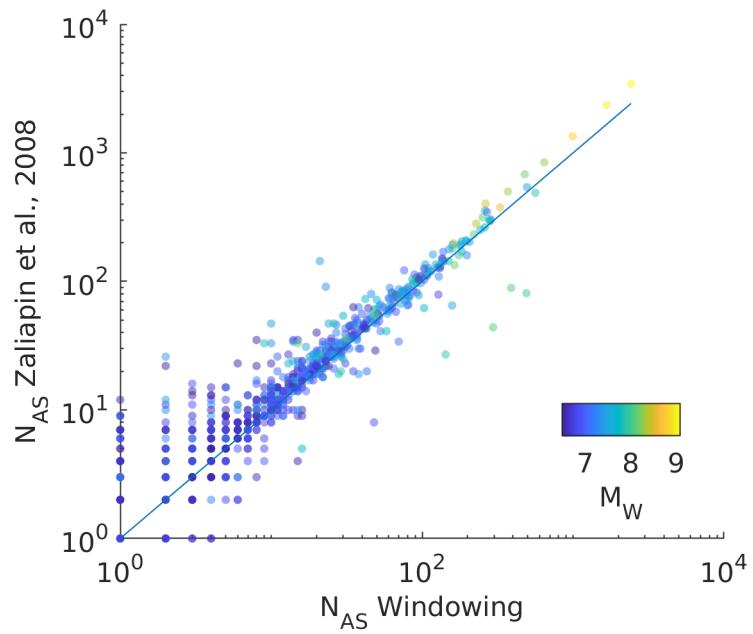


Figure S16. Comparison of aftershock counts obtained using space time windowing and those obtained using nearest neighbor clustering [following Zaliapin et al., 2008]. Each point represents an individual mainshock. Mainshocks are colored according to moment magnitude. Blue line indicates a 1 to 1 correspondence.

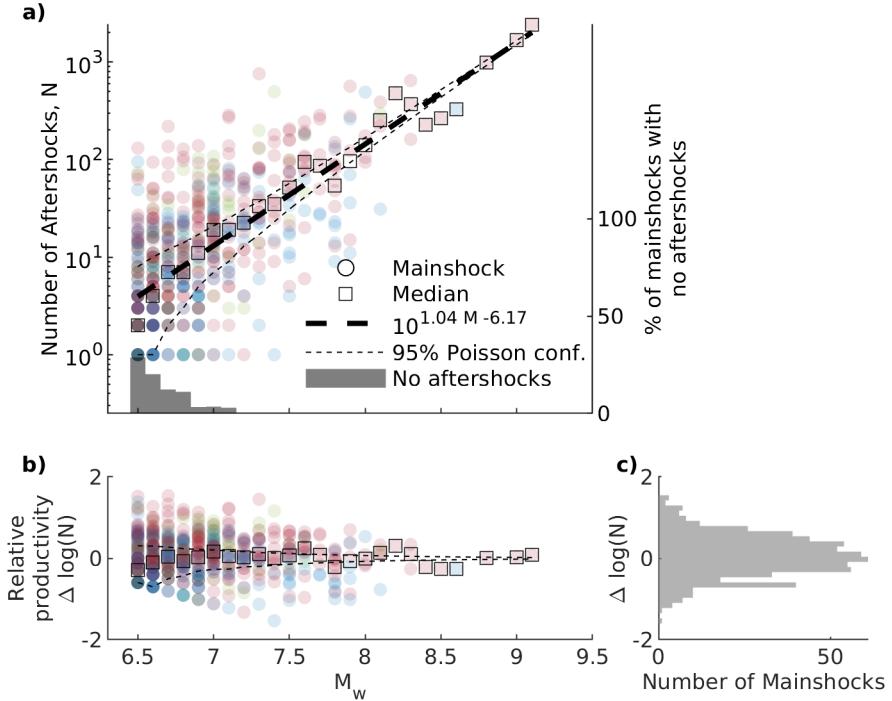


Figure S17. Results using Zaliapin *et al.* [2008] clustering. a) The number of aftershocks of $M_w \geq 4.5$ within three source dimensions and 60 days as a function of mainshock magnitude identified in the global ISC and NEIC catalogs from 1990 to 2019. Colors indicate faulting style of the mainshock; blue, green and red points correspond to earthquake sequences for which the mainshock was respectively strike-slip, normal or reverse. The global productivity law (dashed line) is fit using a least squares regression through the median log-number of aftershocks for each 0.1 magnitude bin (black squares). The median number includes mainshocks with no aftershocks which are not shown on the plot. Individual earthquake sequences (circles) scatter significantly above and below the productivity law. b) Relative productivity (Eq. 5) as a function of mainshock magnitude. The relative productivity distribution does not show events with no aftershocks and thus the lower left corner of the plot is underpopulated. c) Histogram of the relative productivity of mainshocks considered in this study.

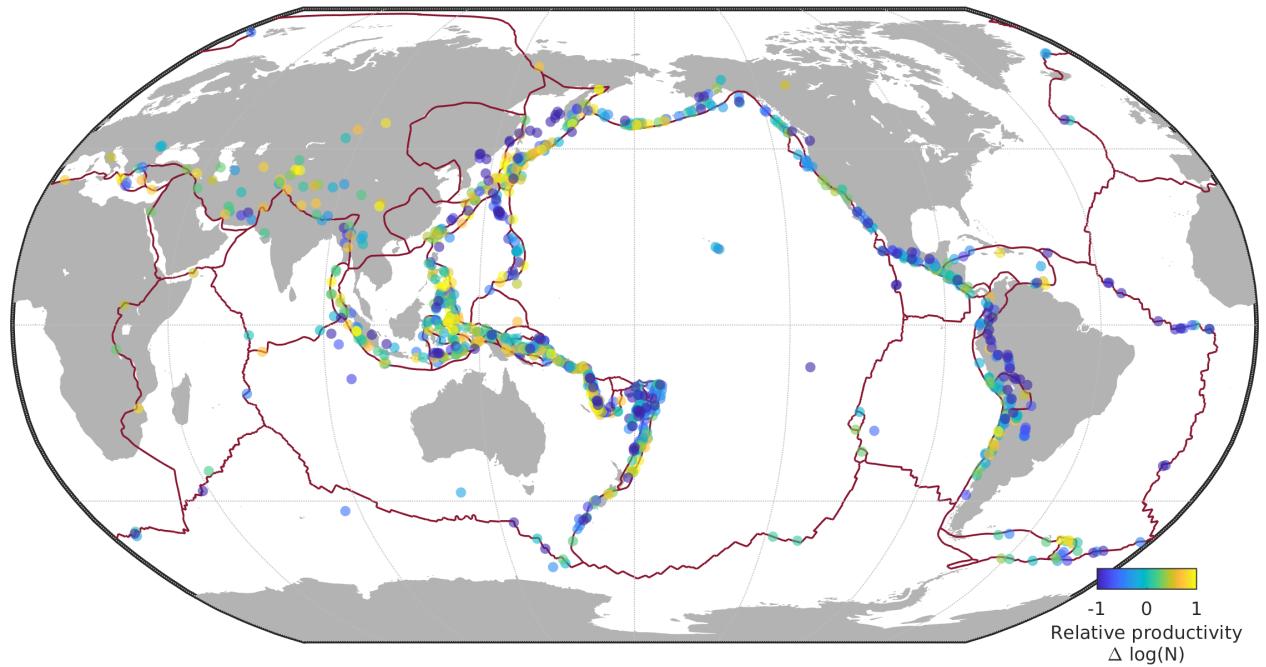


Figure S18. Results using Zaliapin *et al.* [2008] clustering. Global map of earthquake productivity. Red lines indicate the surface trace of the tectonic boundaries. Mainshocks with $M_W \geq 6.5$ color-coded according to their relative productivity (Equation 5)

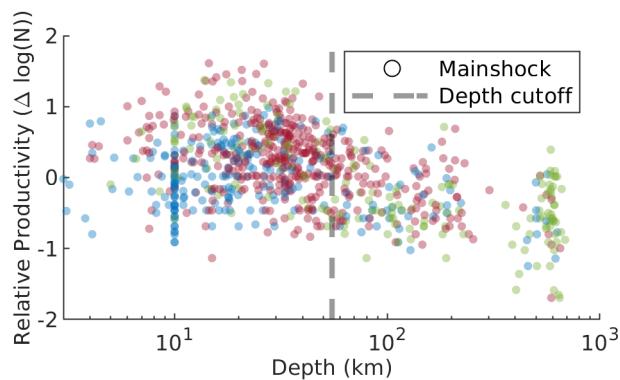


Figure S19. Results using Zaliapin *et al.* [2008] clustering. Relative aftershock productivity as a function of depth. Subsequent analysis will only consider earthquakes shallower than the 55 km cutoff (dashed line). Sequences are color-coded according to faulting style of the mainshock (blue: strike-slip, green: normal and red: reverse). Note: Discretization of depth is apparent in this plot as some events have default values. Depths of 33 km, 5 km, 10 km and 15 km are reported for 6%, 1%, 10% and 0.7%, respectively, of the catalog.

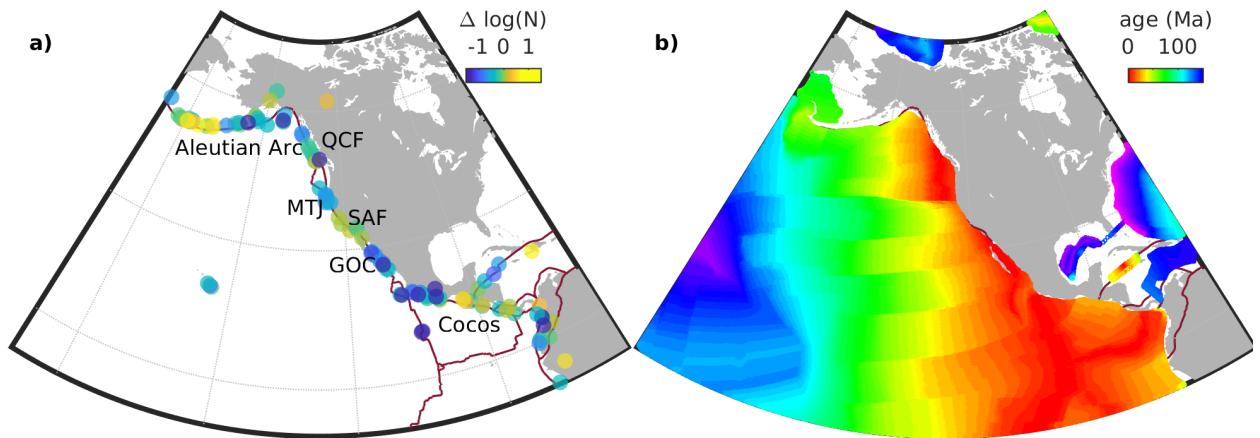


Figure S20. Results using *Zaliapin et al.* [2008] clustering. a) Aftershock productivity along the North American coastline. Individual mainshocks (circles) are color-coded according to their relative aftershock productivity. The Aleutian Arc, Queen Charlotte Fault (QCF), Mendocino Triple Junction (MTJ), San Andreas Fault (SAF), Gulf of California (GOC) and Cocos Plate Subduction Zone include areas with coherent productivity. Red line indicates major plate boundaries [*Bird*, 2003]. b) Seafloor crustal age estimates from *Müller et al.* [2008].

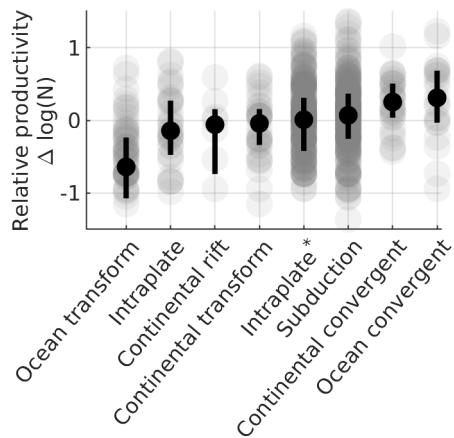


Figure S21. Results using *Zaliapin et al.* [2008] clustering. Earthquake productivity by tectonic boundary. Circles indicate the relative productivity of individual sequences. Solid markers and error bars indicate the median and the interquartile range. A faded lower error bar implies that mainshocks with no aftershocks are within the interquartile range. Intraplate* indicates earthquakes within 400 km of a plate boundary but with a faulting mechanism discordant with the plate boundary (e.g., outer rise events).

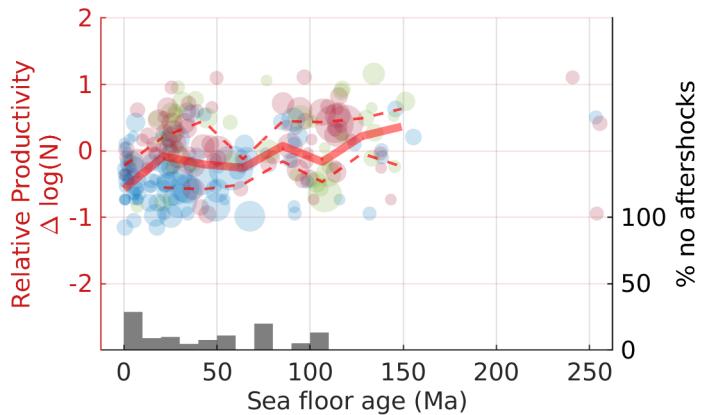


Figure S22. Results using Zaliapin *et al.* [2008] clustering. Relative productivity increases as a function of the age of the oceanic lithosphere. Each circle indicates an individual earthquake sequence. Sequences are color-coded by faulting style of the mainshock (blue: strike-slip, green: normal and red: reverse). The red line indicates the median average for 20 Ma crustal age bins. Dashed lines indicate the corresponding interquartile ranges. Bars indicate the fraction of earthquakes with no aftershocks within each 10 Ma crustal age bin and correspond to the right-hand axis.

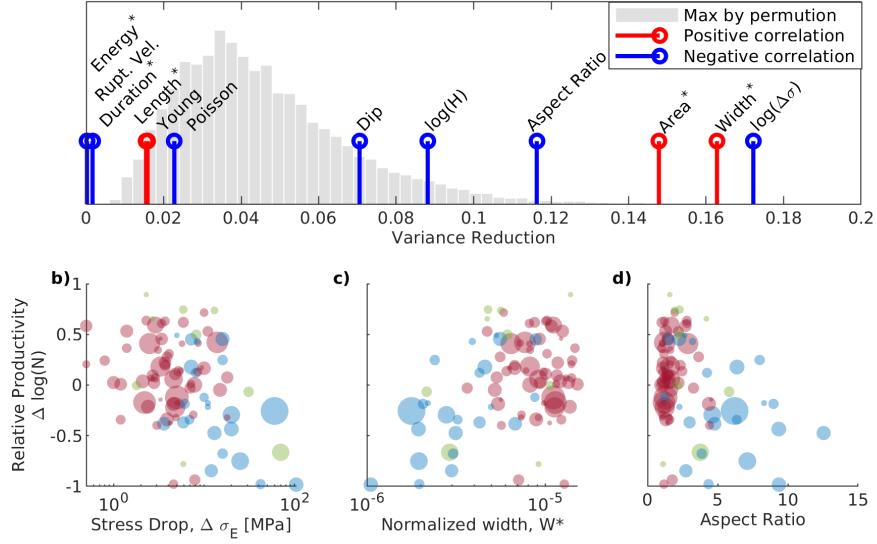


Figure S23. Results using Zaliapin *et al.* [2008] clustering. a) Goodness of fit of linear regressions for each source attribute in our combined catalog. Top and bottom axes respectively represent the p-value and goodness of fit of each attribute (stems). The probability distribution function in the backdrop indicates the maximal variance reduction outcome of 10000 permutation test of the entire data set we tested. Asterisks indicate scaled and log-transformed variables. The scaled energy, length, duration and area, material properties, velocity, dip, and log-stress drop ($\Delta\sigma$) of the mainshock rupture all do not yield a statistically significant ($p = 0.05$) linear fit to the relative productivity the normalized rupture width and aspect ratio of the rupture yield the best fitting linear regressions. Stems are color-coded to indicate whether the source attribute is positively (red) or negatively (blue) correlated with relative productivity. b-d): Relative earthquake productivity as a function of mainshock stress drop, normalized rupture width, and aspect ratio.

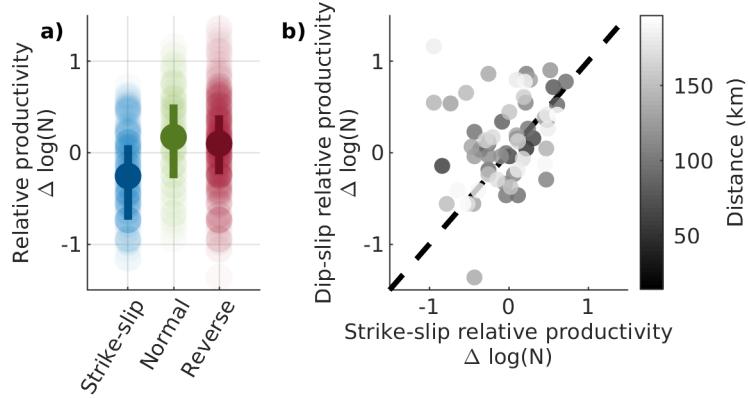


Figure S24. Results using Zaliapin *et al.* [2008] clustering. a) Relative aftershock productivity ($\Delta \log(N)$) by focal mechanism (Equation 5). b) Relative aftershock productivity for pairs of earthquake sequences with strike-slip and dip-slip mainshocks within 200 km from each other. Each pair is shaded according to its relative distance. Dashed line indicates a 1:1 relationship, the expectation for a purely site dominated control on relative productivity. Co-located mainshocks pairs generally follow this 1:1 trend, but exhibit considerable scatter.

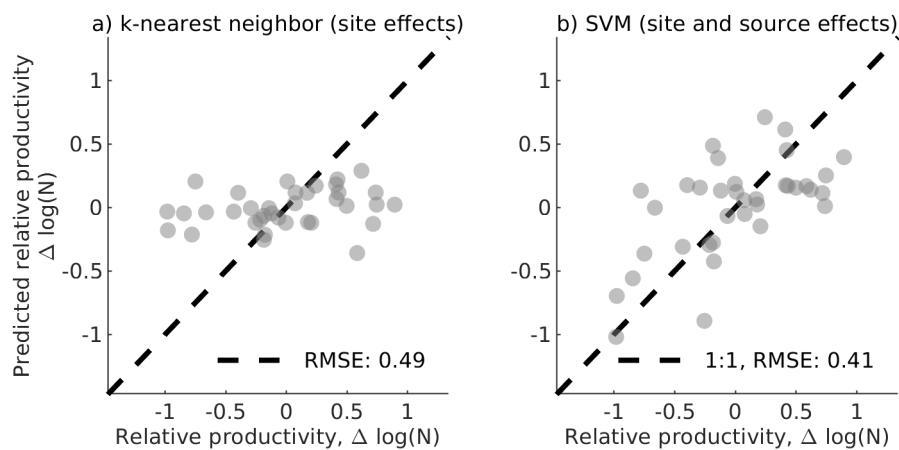


Figure S25. Results using Zaliapin *et al.* [2008] clustering. Response plots (prediction versus observation) for the k-nearest neighbor algorithm (a) and SVM models (b). Each point indicates prediction of relative productivity relative to that which was observed for individual earthquake sequences. A perfect prediction would place all values on the 1:1 line.

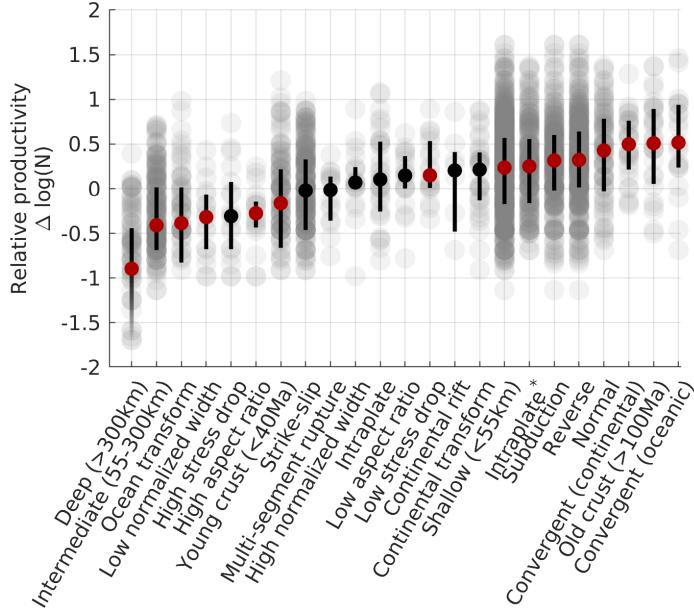


Figure S26. Results using Zaliapin *et al.* [2008] clustering. Synthesis of relative productivity according to catalog subsets. The group considered here are the short list which best distinguished relative productivity based on our different lines of investigation. ‘High’ and ‘low’ subsets respectively refer to >80 th and <20 th percentile ranges of the data. Grey circles are individual mainshocks. Opaque points and error bars respectively indicate the median and interquartile range of the subset. Fading error bars imply that mainshock sequences with no aftershocks are within the interquartile range of the data. Attributes with red markers are consistent with the hypothesis that they are sample from a different continuous distribution than the overall population of earthquakes using a 2-sample Kolmogorov-Smirnov test at a 5% significance threshold.

S1.4 Background Seismicity

Since our selection criteria for aftershocks is magnitude dependent, there is a concern that aftershock counts will be biased by background seismicity. In this section, we show that the background rate of seismicity is in fact negligible and within counting error of a Poisson process.

Measuring background seismicity is challenging and an ongoing topic of research. Approaches to measure this quantity include modelling the earthquake process as an epidemic type sequence with a stationary background rate and various declustering algorithms. However, these methods suffer from instability, subjectivity and are subject to strong trade-off between precision (granularity) and accuracy. In the extreme, event-wise calibration

of background seismicity is particularly under constrained. Background seismicity is also critically dependent on the duration of an aftershock sequence (if we let aftershock sequences continue to infinity, background vanishes). Fundamentally, background estimation is sensitive to the definition of a declustering method.

Here, we circumvent these problems by considering all events (including aftershocks) to produce an apparent background rate. We then compare this apparent background rate of events to aftershock counts. This is a conservative approach since the apparent background rate may include aftershocks that would not abide to a strict definition of background seismicity.

Figure 2 in the manuscript shows how the aftershock-counts compare to a catalog where temporal information has been shuffled. The spatial structure, however, is conserved. This test is both a means to optimize the choice of space-time window but also an absolute upper-bound of the effect background seismicity as a function of magnitude. The effect is indeed scale dependent. However, comparing this effect to the aftershock counts reveals that the bias is small, typically 1 to 2 orders of magnitude below the aftershock counts. Figure S27 demonstrates that the bias from background seismicity lies within one standard deviation of a Poisson counting error for 97% of mainshocks. Again, this calculated bias is an absolute upper-bound which includes aftershocks and likely a significant overestimate of background activity. The occurrence of $M_W 4.5$ events in a specific small space and time window is very rare.

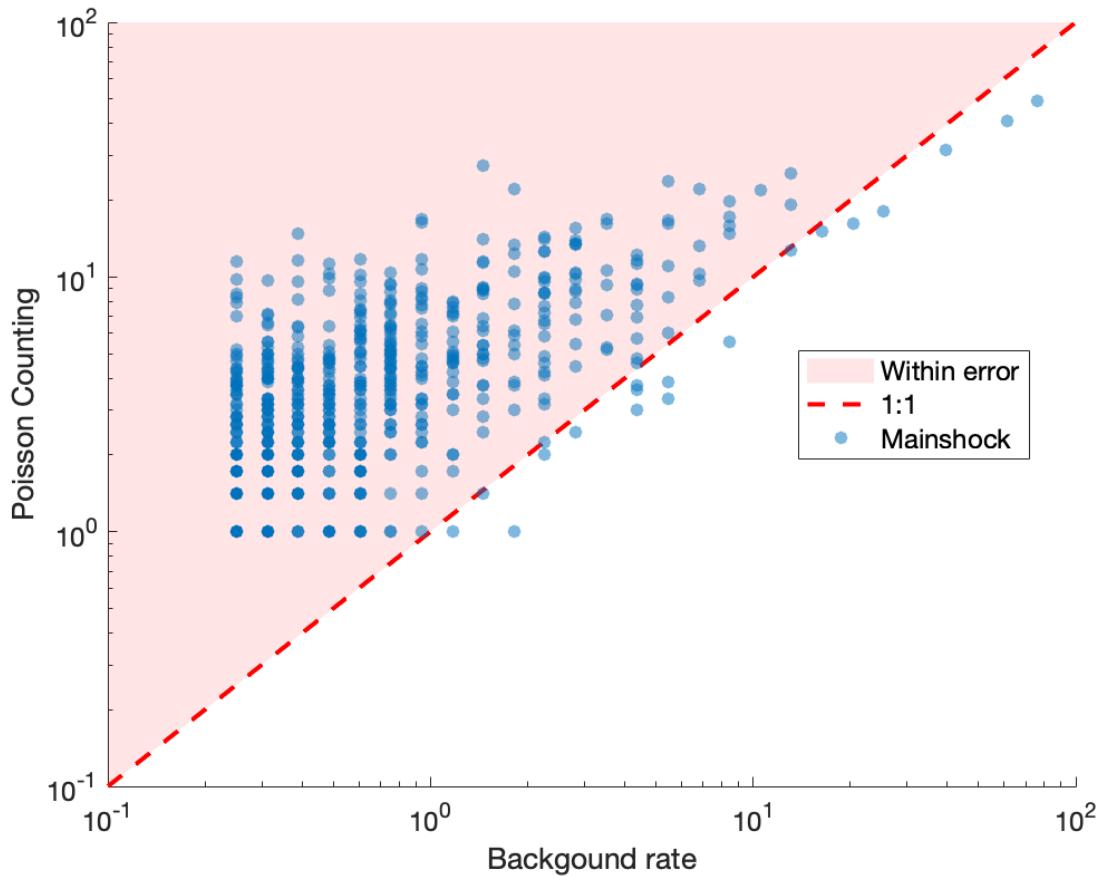


Figure S27. Comparison of the apparent background seismicity to one standard deviation of a Poisson counting error. Dashed line indicates a one-to-one relationship. 97% of events fall in the upper half of the plot highlighted in red indicating that they are within the error of the aftershock count.

S2 Relative productivity as a function of miscellaneous parameters

Figure S28 presents how the relative productivity changes as a function of each variable we tested for in this study.

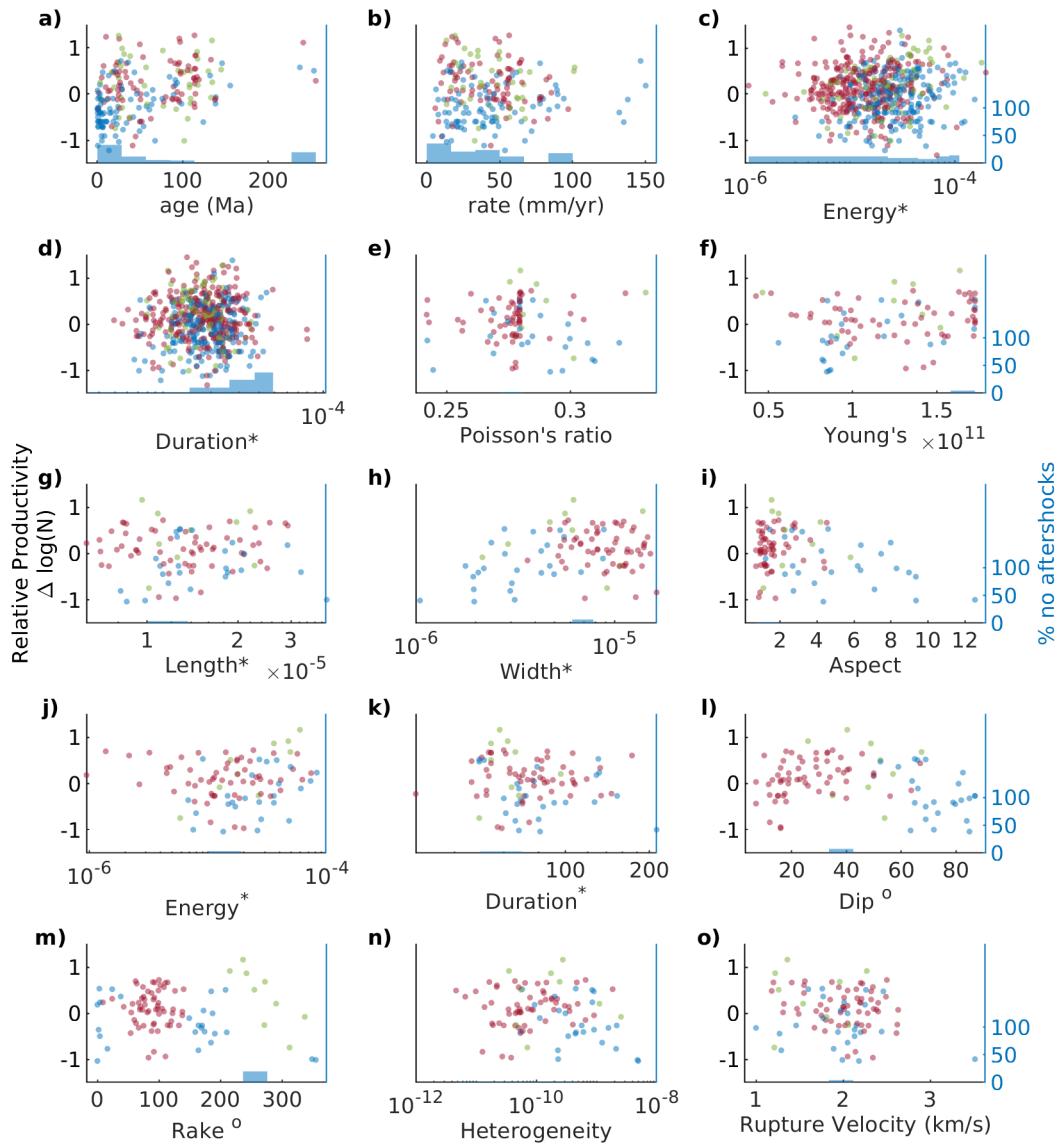


Figure S28. Miscellaneous setting and source attributes as a function of relative productivity (N^*). Source attributes that scale with magnitude are normalized (norm.) to eliminate any such dependence. Note that for the normalized rupture durations and the normalized radiated energy, we show their relationship to relative productivity for both the *Convers and Newman* [2011] and *Hayes* [2017] estimates. Mainshocks are color-coded by focal mechanism (strike-slip: blue, normal: green, and reverse: red). Note that a linear relationship has its limitations for more non-linear relationships (e.g. dip and rake).

S3 Additional Supporting Information (Files uploaded separately)

S3.1 Trained SVM model

We provide as a separate file, *train_SVM.m*, the function used to train our SVM mode using leave-one-out cross-validation. The function includes hyperparamaters used for our analysis. Given 1) a 2-D array of predictors where rows represent individual earthquakes and columns represent predictors (in our case dip, normalized area and plate age); and 2) a response vector (in our case the relative productivity), the function returns the predictions, *validationPredictions*, the root means squared error, *validationRMSE*, and the trained model, *regressionSVM*.

S3.2 Catalogued Source Parameters

We provide separately a table (Table S1), *FSPcat.csv*, for all the source attributes we consider. The table includes earthquake date (in *matlab datenum* format), latitude, longitude, depth (km), moment magnitude, moment (Nm), dip, rake, rupture velocity (km/s), rupture duration (s), Young's modulus (Pa), and Poisson's ratio as reported by *Hayes* [2017]. We also include the focal mechanism, width (km), length (km), aspect ratio (length over width), slip heterogeneity and stress drop (MPa) that we derived.

S3.3 Alternative clustering method code

The full description of the alternative aftershock detection routine is available in a Python Jupyter Notebook at the following GitHub page: <https://github.com/tgoebel/clustering-analysis>. In addition, this repository contains all the functions and utilities used in the process in addition to the earthquake catalog used in this study.

References

- Bird, P. (2003), An updated digital model of plate boundaries, *Geochemistry, Geophysics, Geosystems*, 4(3), doi:10.1029/2001GC000252.
- Brengman, C. M. J., W. D. Barnhart, E. H. Mankin, and C. N. Miller (2019), EarthquakeScaling Relationships from Geodetically Derived Slip Distributions, *Bulletin of the Seismological Society of America*, 109(5), 1701–1715, doi:10.1785/0120190048.

- Convers, J. A., and A. V. Newman (2011), Global evaluation of large earthquake energy from 1997 through mid2010, *J. Geophys. Res.*, 116, doi:10.1029/2010JB007928.
- Hayes, G. P. (2017), The finite, kinematic rupture properties of great-sized earthquakes since 1990, *Earth and Planetary Science Letters*, 468, 94–100, doi:10.1016/j.epsl.2017.04.003.
- Müller, R. D., M. Sdrolias, C. Gaina, and W. R. Roest (2008), Age, spreading rates, and spreading asymmetry of the world's ocean crust, *Geochemistry, Geophysics, Geosystems*, 9(4), 1–19, doi:10.1029/2007GC001743.
- Wells, D. D. L., and K. J. K. Coppersmith (1994), Empirical relationships among magnitude, rupture length, rupture width, rupture area and surface displacements, *Bulletin of the Seismological Society of America*, 84(4), 974–1002, doi:10.1785/bsa.84.4.974.
- Zaliapin, I., A. Gabrielov, V. Keilis-Borok, and H. Wong (2008), Clustering analysis of seismicity and aftershock identification, *Physical Review Letters*, 101(1), doi:10.1103/PhysRevLett.101.018501.