Supporting Information

Bettencourt and Kaur 10.1073/pnas.1102712108

SI Text

Search Terms and Paper and Author Counts. We have assembled a comprehensive collection of sustainability science publications (including journal papers and conference proceedings). All items in the collection are publications written in English between 1973 and the end of 2009. Although considering publications in English only is a limitation, it represents by far the largest component of the scholarly literature, ensures consistency of records, and facilitates automatic text parsing. We searched the Institute for Scientific Information (ISI) Web of Science for publications in other major languages (as described in the *Discussion* section of the main text) and found only a relatively small number of items, accounting for perhaps at most a few percent of our total. Nevertheless, completeness is always extremely hard to establish or guarantee in any scientometric study, so that the corpus analyzed here is to be taken as representative but not as the totality of the literature on the subject.

We used the query ⟨sustainability⟩ in title, abstract, and key words (subject) on the databases Science Citation Index Expanded, Social Sciences Citation Index, and Arts & Humanities Citation Index for all years. Several other data sources were also considered; namely, Scopus and Los Alamos National Laboratory collections, but ISI Web of Science was chosen because its records are more consistent and standardized, allowing us to extract (automatically and with extensive manual inspection and correction in some cases) title text and author names and affiliations (including place names that could be geolocated).

Though journal and institutional names are becoming reported with increasing consistency, their use still poses major challenges of text processing, matching, and associated identification accuracy. For this reason, we have parsed manually over a thousand journal names and have chosen to report geographic location only at the city level. We hope to be able to analyze more detailed institutional affiliations in the near future, but this requires creating dictionaries of possible names for each institution worldwide, including variations in usage and language.

Following this procedure, each publication was geolocated from city and country (and often postal code) given in each record. Each publication can be associated with more than one city if there are multiple addresses listed. For this reason, geographic maps refer to number of publications (not authors) per city.

Corpus of Sustainability Science Publications. We assembled several corpora of scholarly publications (articles and conference proceedings) on sustainability science through key word searches using ISI Web of Science, the Los Alamos National Laboratory's library databases, and Scopus. The corpus discussed here comes from key word searches on "sustainability" in title, subject, or abstract for all years and databases: Science Citation Index Expanded, Social Sciences Citation Index and Arts, and Humanities Citation Index. The data contain 23,211 records (June 2010), with 20,376 until the end of year 2009. Similar searches for "sustainable development" yielded a subset of this corpus that was correct but substantially more incomplete, with 16,647 records through the end of 2009. We checked that properties of this alternative corpus are similar and, as discussed above, corresponds to a very similar disciplinary makeup. The Web of Science records were adopted here because they were found to have more complete records, easy downloadable data, including available addresses for publications from which we extracted (via parsing of text addresses) city and nation of authors' institutions. Years of publication and journals, which we matched to ISI disciplines, were also extracted from each record. A small number of journals in our corpus were not present in the maps of science classification and were excluded from the scientific discipline analysis.

Population Models and Parameter Estimates. It is worth elaborating briefly on the meaning of each of the terms in Eq. **1** (compare Fig. 1*B*). The first term ΛN is responsible for population growth and adds new individuals that are susceptible to the idea. The second term is responsible for the progression of these individuals to a state of exposure to the idea, through interaction (training, teaching, publications) with a community of publishing researchers I. The community of exposed individuals in turn can progress to practitioners via an incubation period (ϵE) or via continued contact ($\kappa E \frac{I}{N}$), which is atypical of population biology but was found to be important in previous analyses and ethnographic studies of other scientific fields where formal training programs, meetings, etc., were essential to guarantee that individuals initially exposed could become authors. Active researchers may then leave the field at a rate γ (γI). Terms of this form could be added to other classes to account for possible exit rates in the S and E classes, but these do not change the dynamics qualitatively. Estimates of model parameters from data were obtained using a stochastic ensemble method (see refs. 11 and 12) and are given in Table S1.

Maps of Authors and Citations. Author numbers (nonunique = number of addresses) and citations were extracted from ISI Web of Science records and assigned to cities and nations. Whenever a publication has several authors, it is counted and assigned to each location. We disregarded the possibility of differential credit assignment by order of authors, because this is a subjective measure that varies from field to field. The maps of Fig. 2 *A* and *B* were created using Google charts.

Discipline Mapping. We mapped each publication in our corpus of sustainability science to a traditional discipline and subdiscipline using Thomson–Reuters Journal Citation Reports and Web of Science commercial products. This scheme is the standard in disciplinary analysis and provides the classification of journals into 554 subdisciplines and 13 major disciplines. This is the same procedure used to generate maps of science, which provide the standard color assignments used in Fig. 3.

Collaboration Network and Analysis. Publications and authors form a bipartite graph. We projected this graph onto the space of authors, assigning links between them if they have coauthored at least once. This network was created each year between 1975 and 2010, and analyzed in terms of a variety of metrics, including number of edges, number of nodes, clustering, diameter, the fraction of edges in the largest cluster P, and the cluster susceptibility S. S is defined as $S = [\sum_i n_i^2 - [\max_i(n_i)]^2]/n^2$, (we adopted a normalization by n^2 for visualization), where the sum is over all disconnected clusters, n_i is the size of each cluster (in terms of number of nodes), and n is the total size of the system, over all clusters; $\max_i(n_i)$ is the size of the largest cluster. P and S are analogous to percolation cumulants where they suffice to define a second-order transition in the infinite system size limit. Network analysis was performed using the python package NetworkX (available online at http://networkx.lanl.gov/)

Fig. S1. Word cloud showing the relative frequency of most frequent words in publication titles. The multidisciplinarity of the field is apparent, addressing themes that range from economics and the social sciences to ecology, climate, and engineering. The figure was generated using Wordle (available online at http://www.wordle.net).

Fig. S2. Word cloud for ISI key words for collection of papers in sustainability science. Key words are assigned to publications by ISI Web of Knowledge, a standard scientific collection and search engine. As in Fig. S1, the multidisciplinarity of the field is clear from the diversity of themes expressed in these key words. The figure was generated using Wordle (available online at http://www.wordle.net).

Fig. S3. Word cloud for paper title bigrams. Bigrams (consecutive two-word combinations, obtained after frequent stop words were removed) often give a better sense of subjects addressed than the single words of Figs. S1 and S2. We reduced the frequency of "sustainable development" by 35%, so that other terms would be more visible. The figure was generated using Wordle (available online at http://www.wordle.net).

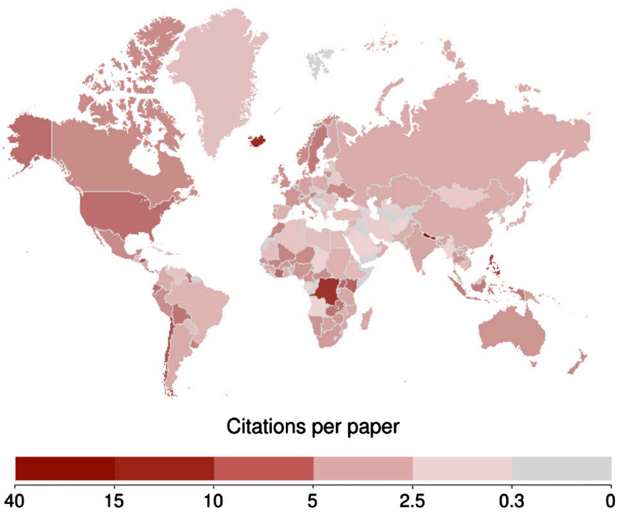

Fig. S4. World map of citations per paper in sustainability science. Most of the nations with highest citations per paper are small and have generated relatively few manuscripts. For this reason, we opted to present total citations in Fig. 2. The top 10 nations by this ranking are the Dutch Antilles (44 citations per paper, 2 papers), Gambia (29.2 citations per paper, 6 papers), Nepal (18.4 citations per paper, 45 papers), Iceland (14.4 citation per paper, 18 papers), Mauritius, (14.3 citation per paper, 4 papers), Democratic Republic of the Congo (14.2 citations per paper, 17 papers), the Philippines (13.3 citations per paper, 131 papers), Honduras (11.7 citations per paper, 6 papers), Chile (11.3 citation per paper, 107 papers), and Kenya (10.2 citation per paper, 206 papers). The United States appears in 14th place with 9.3 citations paper (9,435 papers).

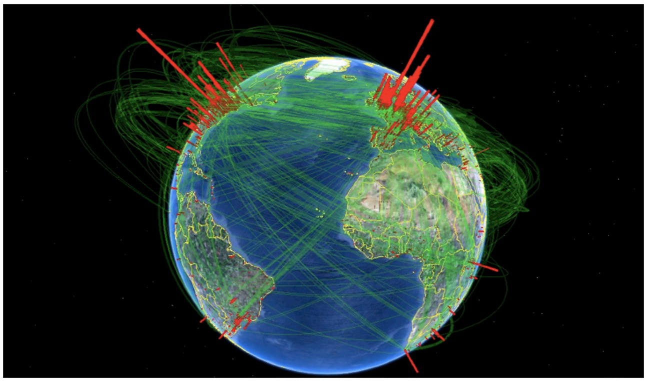

Fig. S5. Global collaboration network of sustainability science. Atlantic view. The map shows number of authors in cities worldwide (red columns) and their coauthorship networks (green lines). Thicker lines indicate a greater number of collaborations between places. The interactive Google Earth map is available at http://www.santafe.edu/~bettencourt/sustainability/.

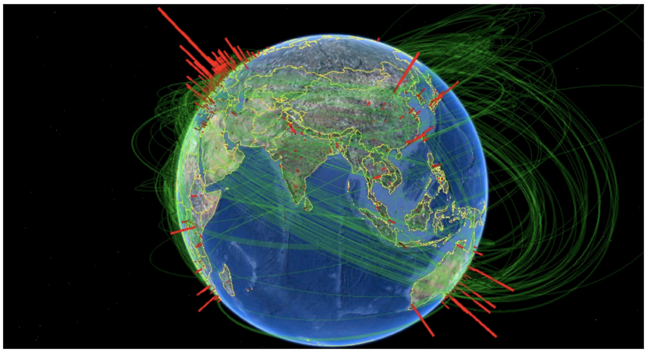

Fig. S6. Global collaboration network of sustainability science. Indian Ocean view. The map shows number of authors in cities worldwide (red columns) and their coauthorship networks (green lines). Thicker lines indicate a greater number of collaborations between places. The interactive Google Earth map is available at http://www.santafe.edu/~bettencourt/sustainability/.

Fig. S7. Collaboration network of sustainability science of Washington, DC. Washington, DC, is the leading city worldwide in terms of numbers of authors in sustainability science (with over 4,000 authors publishing between 1973–2009). As before, the map shows number of authors in cities worldwide (red columns) and their coauthorship networks (green lines). Thicker lines indicate a greater number of collaborations between places. The interactive Google Earth map is available at http://www.santafe.edu/~bettencourt/sustainability/.

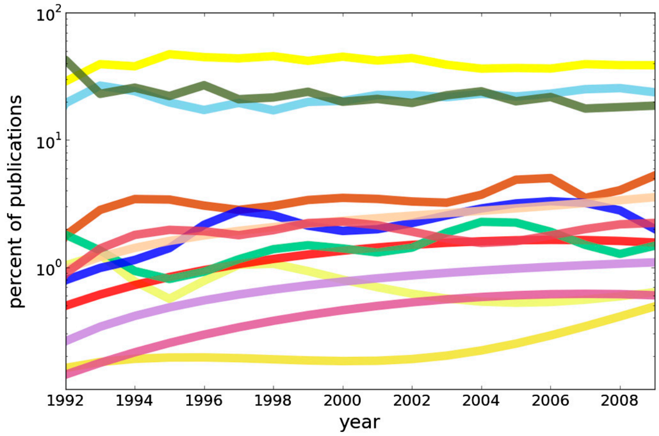

Fig. S8. The temporal evolution of disciplinary contributions to sustainability science obtained from a sustainable development query. The contributions from the several disciplines are similar to those of Fig. 3, with the social sciences contributing slightly more (39.5% of all manuscripts) here. The contributions of biology (19.8%) and chemical, mechanical, and civil engineering (23.0%) are slightly smaller than in Fig. 3, but within the same range of variation over the total time period.

Table S1. Best-fit estimates and standard deviation for population model of sustainability science authors (see Fig. 1*B* and *Materials and Methods*)

Parameter	Best-fit estimate	Standard deviation
S_0	3.45	0.27
E_0	1.83	0.10
I_0	1.00	0.07
β	2.04	0.28
Λ	0.46	0.04
κ	1.35	0.08
ε	0.89	0.02
γ	0.94	0.01
β_0	1.50	0.02
R_0	2.17	0.03