Species loss drives ecosystem function in experiments, but in nature the importance of species loss depends on dominance

Mark A. Genung1,2 | Jeremy Fox3 | Rachael Winfree1

1Department of Ecology, Evolution and Natural Resources, Rutgers University, New Brunswick, NJ, USA
2Department of Biology, University of Louisiana at Lafayette, Lafayette, LA, USA
3Department of Biological Sciences, University of Calgary, Calgary, AB, Canada

Correspondence
Mark Genung,
Department of Biology, University of Louisiana at Lafayette,
410 E. St. Mary Blvd.,
Billeaud Hall, Room 108,
Lafayette, LA 70503
Email: mark.genung@louisiana.edu

Funding information
Division of Environmental Biology, Grant/Award Number: DEB-0620652, DEB-1234162, DEB-1355629 and DEB-1754657; Wildlife Conservation Society; National Institute of Food and Agriculture, Grant/Award Number: #2009-65104-5782; Division of Ocean Sciences, Grant/Award Number: OCE-0620276; Smithsonian Institution; Conservation International; Missouri Botanical Garden; Gordon and Betty Moore Foundation

Editor: Irena Simova

Abstract
Aim: Decades of experimental research have conclusively shown a positive relationship between species richness and ecosystem function. However, authoritative reviews find no consensus on how species loss affects function in natural communities. We analyse experimental and observational data in an identical way and test whether they produce similar results.

Location: North America and Europe (experimental communities); global (natural communities).


Major taxa studied: Experimental communities: temperate grassland plants; natural communities: temperate grassland plants, tropical forest trees, kelp forest producers and native bees.

Methods: We used an approach inspired by the Price equation to analyse 129 datasets from experimental and natural communities worldwide. We tested how the effects of species loss on ecosystem function varied with dominance and the non-randomness of species loss and, in turn, how these two factors differed between experiments and observations.

Results: Studies carried out in experimental and natural communities reached different conclusions regarding the effects of species loss. First, species loss had greater effects on ecosystem function in experiments than in nature. Second, the importance of species loss was negatively correlated with dominance in nature because as dominance increased, lost species were increasingly those contributing little to ecosystem function. Although experimental and natural communities exhibited similar levels of dominance, an analogous relationship was not possible in experiments because the order of species loss was randomized by design.

Main conclusions: Species loss was sometimes, but not always, the major driver of loss of function in nature. Variation in the importance of species loss was not messy and context dependent; instead, it was predicted by functional dominance. Although results from experimental and natural communities were similar in several key ways, they differed in that species loss was a consistent predictor of ecosystem function in experiments and not in nature.
1 | INTRODUCTION

Global loss of biodiversity doubtless threatens essential ecosystem functions (Cardinale et al., 2012; Hooper et al., 2012; Isbell et al., 2017), but despite decades of experimental and observational research, there is still no predictive understanding of how species loss will impact ecosystem function in nature (Cardinale et al., 2012). Studies in experimental and natural communities often find different results, with experiments pointing to a major role for species richness (e.g., Isbell et al., 2011; Reich et al., 2012; Tilman et al., 2001), whereas studies in natural communities variably find that species richness (Duffy, Goodwin, & Cardinale, 2017; Duffy, Lefcheck, Stuart-Smith, Navarrete, & Edgar, 2016; Grace et al., 2016; Mora et al., 2011), the order of species loss (Larsen, Williams, & Kremen, 2005), dominant species (Genung et al., 2017; Winfree, Fox, Williams, Reilly, & Cariveau, 2015) or aggregate abundance (Smith & Knapp, 2003) drive function. Furthermore, experimental and observational studies define “species loss” differently, consistent with their study designs.

Most experiments have defined species loss as a decrease in species richness (i.e., the number of species) and have isolated the effects of richness by randomizing the species composition of experimental communities at different levels of richness (Schmid et al., 2002). Thus, experiments measure the effect of losing n species, given that those n species were chosen at random. Generally, there is no literal species loss (i.e., species removals, but see e.g., Lyons & Schwartz, 2001; Smith & Knapp, 2003; Wardle & Zackrissron, 2005; Zavaleta & Hulvey, 2004) but instead the comparisons are made among plots that were established with different numbers of species. The results of > 600 biodiversity-ecosystem function experiments have convincingly established that decreasing richness decreases function (Cardinale et al., 2012), with effects of richness being comparable to environmental drivers such as drought, nitrogen fertilization and invasive species (Hooper et al., 2012). Nevertheless, it remains unclear whether the effect of reduced richness, as measured in experiments, mirrors the effect of species loss from natural communities.

Observational studies of the biodiversity–function relationship likewise have rarely studied literal species losses. Instead, they have measured the effects of species loss based on one of two designs: either space-for-time substitutions, in which sites with different levels of anthropogenic effects are compared (Duffy et al., 2016; Grace et al., 2016); or comparisons of sites that do not necessarily differ in levels of anthropogenic change but that do differ in levels of both biodiversity and ecosystem function (Duffy et al., 2017; Genung et al., 2017; Winfree et al., 2015, 2018). The key distinction between species loss in experiments and natural communities is that changes in richness and composition are confounded in natural communities (Larsen et al., 2005; Smith & Knapp, 2003; Suding et al., 2005; Winfree, Williams, Dushoff, & Kremen, 2014), but not in experiments. Thus, observational studies measure the functional effect of losing the n species that were in fact lost, rather than the expected effect of losing n species at random. This is a limitation in that it is difficult to separate the effects of richness and composition as drivers of function (Fridley, 2002; Mulder, Jumpponen, Högberg, & Huss-Danell, 2002; Tilman & Wardle, 1997). However, it is a strength in that it captures any association between the identity of species most likely to be lost as richness declines and the contribution of these species to function.

Dominance, or the tendency of communities to contain many rare and few common species (McCII et al., 2007), is an important feature of ecological communities that differs between experimental and observational studies. Experiments investigating the biodiversity–function relationship tend to equalize the initial abundances of species (but see e.g., Lamb, Kennedy, & Siciliano, 2011; Wilsey & Potvin, 2000; Wittebolle et al., 2009) to isolate the effects of species richness (Schmid et al., 2002). However, it is unclear how species abundances and contributions to function change over time in experiments. Do they, either quickly or over many years, start to mimic the higher functional dominance (an analogue of numerical dominance, in which contributions to function replace abundance) seen in natural communities (Schleuning, Fründ, & Garcia, 2015)? This question is important, because systematic differences in functional dominance between experimental and natural communities could lead to predictable differences in the biodiversity–function relationship. The basic prediction is that high functional dominance makes species richness less important to function, because a few common species could provide most of the function (Dangles & Malmqvist, 2004; Grime, 1998; Smith & Knapp, 2003; Winfree et al., 2015). In contrast, when communities are even, it is more likely that species richness will be important. This broad prediction is not system specific and provides some reason to expect that dominance could mediate the effects of species loss in a similar way across ecosystem functions.

In sum, to advance research on biodiversity-ecosystem functioning in real-world communities we need to know, first, whether experimental and natural communities differ in functional dominance, and second, whether functional dominance mediates the consequences of species loss for ecosystem function. We analysed 129 datasets from across the globe, of which 36% came from biodiversity–function experiments and 64% from natural communities, and answered the following questions. First, does species loss, as measured through changes in species richness and species composition, have similar effects on ecosystem function in experimental and natural communities? Second, do experimental and natural communities have similar levels of functional dominance, and how does functional dominance mediate the effect of species loss on function?

KEYWORDS
biodiversity, composition, dominance, ecosystem function, experiment, Price equation, species loss, species richness
Third, are species lost at random in experimental and natural communities, and how does non-randomness in the order of species loss affect function?

2 | METHODS

2.1 | Price equation partition

The Price equation was first developed to partition the drivers of microevolutionary change in mean phenotype (Price, 1972). The Price equation partition used here reinterprets and builds on the same mathematics to partition the difference in ecosystem function between two sites (a higher-function “baseline” site and a lower-function “comparison” site) into three additive terms: richness, composition and context dependence (Fox, 2006; Fox & Kerr, 2012; for details, see Supporting Information Appendix S1). Hereafter, we refer to our partition as the “ecological Price equation” for simplicity, although we are not claiming that any one definitive “ecological Price equation” exists. The ecological Price equation divides a between-site decline in function into components attributable to different drivers. Thus, the ecological Price equation reverses the question traditionally asked by biodiversity–function experiments, which is, “Given a change in the number of species, how does function change?”, and instead asks, “Given a between-site difference in function, how much can be assigned to changes in the number of species?”. Both questions are interesting, and our use of the ecological Price equation does not advocate for one over the other.

The ecological Price equation compares pairs of sites, and in each case asks why one has higher function than the other. The mathematics of the ecological Price equation shows that three, non-exclusive answers are possible, each corresponding to one ecological Price equation term. First, the higher-function site might simply have many more species. This would be captured by richness (RICH), which is the expected change in function if species loss is random with respect to function. We emphasize that this is richness in a strict, literal sense, that is, the number of species present. It is not inclusive of the identities of those species, nor any positive complementarity resulting from higher richness. Second, the higher-function site might have species that contribute, on average, more function. This would be captured by composition (COMP), which adjusts the expectation set by RICH because species are almost never lost exactly at random with respect to function. For example, if species lost between the higher- and lower-function site had above-average contributions to function, COMP would augment RICH because the effects of species loss were greater than the random expectation. Third, species present at both sites might contribute more to function at the higher-function site. This would be captured by the context dependence effect (CDE), which includes all between-site differences in function not attributable to between-site differences in species richness or composition. The CDE captures any compensatory (or depensatory) responses of the remaining species to species loss, effects of between-site differences in environmental conditions, and any other factors causing the remaining species to function differently at different sites. In this paper, we use the term “species loss” to refer the sum of the richness and composition effects (Supporting Information Appendix S1). This can be considered the direct or immediate effects of a species no longer being present at a site. Box 1 shows a general framework for interpreting ecological Price equation results, based on the signs and relative magnitudes of species loss and context dependence.

The ecological Price equation approach offers two important advantages. First, it creates a natural contrast between effects directly attributable to changes in the number and identity of species (i.e., species-level effects; RICH and COMP) and those that are not (CDE). Abundance, in particular, is often a confounding factor in observational studies, whereas experiments control the initial abundance (often by seeding species at equal densities, e.g., Reich et al., 2012; Tilman et al., 2001; Weigelt et al., 2010). The ecological Price equation partitions abundance effects into the CDE and can, therefore, be used to make comparisons among studies that did, or did not, control abundance. Second, the ecological Price equation separates the random (RICH) and non-random (COMP) effects of species loss. Thus, it can compare among studies that did, or did not, enforce random species loss by design.

2.2 | Datasets used

We searched for datasets using the following four criteria. First, the dataset had to include a measurement of ecosystem function expressible as a sum of species contributions. Second, the dataset had to include replicate sites (or plots, in the case of experiments) at which species composition and function were measured. These are basic requirements for using the ecological Price equation. Third, the dataset had to include a second level of sampling, either temporal (e.g., sampling the same collection of sites in a subsequent year) or spatial (e.g., measurements of function replicated with the same design in different regions of the world). This was to generate a range of functional dominance values for each ecosystem function. Fourth, to make our results more comparable with experiments, we focused on datasets in which ecosystem function providers belonged to the same trophic level. Fifth, we specifically searched for datasets that would represent a diversity of functions: aquatic and terrestrial, delivered by plants and animals, and spanning multiple continents.

2.3 | Data from experimental communities

We analysed 46 datasets from three long-running biodiversity–ecosystem function experiments, all of which measured aboveground biomass of grassland plant species: Biodiversity II, BioCON and the Jena experiment (n = 46 datasets from experiments in total). Biodiversity II manipulated plant species richness in one-, two-, four-, eight- and 16-species plots. BioCON manipulated species richness in one-, four-, nine- and 16-species plots. BioCON also included a two-by-two factorial manipulation of CO₂ (ambient and elevated) and nitrogen (unfertilized and
fertilized), yielding four global change treatments (e.g., ambient CO₂ and nitrogen fertilization). We analysed each global change treatment separately and found the same patterns across all four treatments. The Jena experiment is known for high maximum richness (it contains one-, two-, four-, eight-, 16- and 60-species plots) and large plot size (20 m × 20 m). For all three biodiversity-ecosystem function experiments, species were seeded at equal densities when establishing plots. For all experimental data, we excluded the one-species plots to make experiments more comparable with natural communities. Each year of each experiment was a dataset \[ n = 12 \text{ for Biodiversity II, } n = 28 \text{ (7 years × 4 treatments) for BioCON, and } n = 6 \text{ for Jena} \], and the ecological Price equation partitioned variation in function among plots.

### 2.4 Data from natural communities

There were 83 datasets from natural communities, spread across four ecosystem functions. Our first ecosystem function was crop pollination by wild bees (Winfree et al., 2018). We analysed pollination provided to blueberry \( n = 3 \) years of data), watermelon \( n = 5 \) years) and cranberry \( n = 2 \) years) crops by wild bee species. Each crop-year combination was a dataset, and replicate farms within crop-years were sites \( n = 10 \) datasets in total across three crops). Our second ecosystem function, collected by the Santa Barbara Coast LTER (Long-Term Ecological Research) group, was producer biomass in kelp forests at nine sites off the California coast. Each year of kelp forest surveys was a dataset \( n = 19 \) datasets in total). Our third ecosystem function was aboveground carbon storage in tropical forests. We used tree abundance data from four different continents, collected by the Tropical Ecology Assessment and Monitoring (TEAM) Network. Datasets were TEAM Network locations in different parts of the world \( n = 8 \) datasets in total), and sites were 1-ha forest plots within each location. We used allometric equations from (Chave et al., 2005) and wood density estimates (Zanne et al., 2009) to translate abundances to carbon storage. Our fourth ecosystem function was aboveground

---

### BOX 1 Simplified interpretations of Price equation terms

In broad terms, ecological Price equation results can be placed in four categories (the top four rows above; the fifth row is included only for completeness). These categories are based on the sign, and in some cases the relative magnitudes, of the species loss and context dependence terms. The first and second rows correspond to results from natural and experimental communities, respectively. Interpretations and example communities provide a likely explanation for each result. “Shared” means shared between the higher- and lower-function sites, that is, species that are present at both sites. “Smaller” and “larger” describe the relative magnitude of terms within rows. In the right columns, shapes are different species, sizes represent contributions to function, and open shapes with dotted borders indicate the absence of species (i.e., a species loss).

<table>
<thead>
<tr>
<th>Sign of Species Loss</th>
<th>Sign of Context Dependence</th>
<th>Likely Interpretation</th>
<th>Higher-function Site</th>
<th>Lower-function Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative (Smaller)</td>
<td>Negative (Larger)</td>
<td>Higher-function sites tend to have more species, but most species that turn over between sites have low function. Shared species have highly variable contributions to function across sites.</td>
<td><img src="#" alt="Shape1" /> <img src="#" alt="Shape2" /> <img src="#" alt="Shape3" /></td>
<td><img src="#" alt="Shape4" /> <img src="#" alt="Shape5" /></td>
</tr>
<tr>
<td>Negative (Larger)</td>
<td>Negative (Smaller)</td>
<td>Higher-function sites tend to have more species, and species turnover includes species with substantial contributions to function. Shared species have slightly variable contributions to function across sites.</td>
<td><img src="#" alt="Shape6" /> <img src="#" alt="Shape7" /> <img src="#" alt="Shape8" /></td>
<td><img src="#" alt="Shape9" /> <img src="#" alt="Shape10" /></td>
</tr>
<tr>
<td>Negative</td>
<td>Positive</td>
<td>Higher-function sites tend to have more species. After species loss, persisting species increase contributions to function, but cannot completely offset the effects of species loss. This could be because declines in species richness reduce complementarily.</td>
<td><img src="#" alt="Shape11" /> <img src="#" alt="Shape12" /> <img src="#" alt="Shape13" /></td>
<td><img src="#" alt="Shape14" /> <img src="#" alt="Shape15" /></td>
</tr>
<tr>
<td>Positive</td>
<td>Negative</td>
<td>Higher-function sites tend to have fewer species. This requires that shared species contribute more function when richness is lower. Perhaps, in high-richness sites, competition from low-functioning species limits the performance of functionally-dominant community members.</td>
<td><img src="#" alt="Shape16" /> <img src="#" alt="Shape17" /> <img src="#" alt="Shape18" /></td>
<td><img src="#" alt="Shape19" /> <img src="#" alt="Shape20" /></td>
</tr>
<tr>
<td>Positive</td>
<td>Positive</td>
<td>This result is not mathematically possible.</td>
<td><img src="#" alt="Shape21" /> <img src="#" alt="Shape22" /> <img src="#" alt="Shape23" /></td>
<td><img src="#" alt="Shape24" /> <img src="#" alt="Shape25" /></td>
</tr>
</tbody>
</table>
biomass of temperate grassland plants in Minnesota, USA, collected by the Cedar Creek LTER group. There were 23 years of data and two habitat types. Each year–habitat type combination was a dataset (n = 46 total datasets).

Data sources are listed in the Appendix. See the Supporting Information Appendix S2 for a full descriptions of datasets, functions and how the data were used. Although the experimental and observational communities were sampled at different spatial scales, our results appear to be robust to this difference (Supporting Information Appendix S3).

### 2.5 Environmental variation and interpretation of the ecological Price equation

As much as possible, datasets were selected to minimize environmental variation (e.g., habitat type, time of year, precipitation, temperature and elevation) among sites. However, there will doubtless be remaining environmental variation. Although the ecological Price equation does not have a term for environmental variation, the effects of environmental variation are still captured and attributed to the component of community structure through which they act. For example, higher precipitation at a focal site could increase carbon storage by trees indirectly, but it must do so by: (a) increasing focal site species richness; (b) shifting focal site community composition towards higher-function species; or (c) increasing the focal site abundance or per-capita function of species present at both sites.

### 2.6 Methods of analysis

For all 129 datasets, we first applied the ecological Price equation to all pairwise comparisons of sites, producing values for RICH, COMP and CDE for each comparison. We then calculated the mean, across pairwise comparisons, of each ecological Price equation term (Figure 1), yielding: effects of changes in species richness ($\overline{\text{RICH}}$); effects of changes in species composition ($\overline{\text{COMP}}$); or context-dependent changes in the function provided by species that are present at both sites ($\overline{\text{CDE}}$). Overbars indicate that ecological Price equation terms are means (e.g., $\overline{\text{RICH}}$), averaged across all pairwise comparisons of sites within the dataset. To determine whether species loss has similar effects on ecosystem function in experimental and natural communities, we compared the mean ecological Price equation terms (see above) between experimental and natural communities. Each dataset had one value for functional dominance, which was the mean functional dominance across all sites (or plots) in the dataset. We explored five dominance indices, all in terms of diversity (Hill numbers) of order $q$, where $0 < q \leq 2$ (Chao & Ricotta, 2019). Results were qualitatively consistent across indices and values of $q$ (Supporting Information Appendix S4); in the Results, we use the third class of indices with $q = 1$. A simplified equation for this form of dominance, specific to $q = 1$, is:

$$1 - \left( \frac{e^H - 1}{S - 1} \right),$$

where $H$ is the Shannon entropy index and $S$ is species richness.

To determine whether experimental and natural communities have similar functional dominance, we compared the distributions of functional dominance between experimental and natural communities. Our choice to use functional rather than numerical dominance reflected data limitations, because most datasets reported species-level function, but not abundance. Experiments seeded species at equal densities when establishing plots, but nonetheless the functional dominance could still be high in experimental communities for two reasons: (a) changes in species abundances from their initially equal values; and (b) interspecific variation in per-capita function.

We fitted three models using “lm” in R v.3.5.1 (R Core Team, 2018). We fitted each model to experimental and natural communities...
separately. All three included ecosystem function "type" as a categorical predictor and functional dominance as a continuous predictor, fitting separate slopes for dominance within each ecosystem function "type". Type means a specific ecosystem function (e.g., pollination) for natural communities and experiment name (e.g., Biodiversity III) for experiments. The response variable for the first model was the effect of species loss, $\text{RICH} + \text{COMP}$, which was the effect of changes in the number and identity of species. The response variable for the second model was $\text{CDE}$, which accounted for changes in the abundance and per-capita function of species present at both sites. The first two models were parallel in that they tested how functional dominance mediated the effect of either species loss or context dependence on function. The third model examined the non-randomness of species loss, using $\text{COMP} / \text{RICH}$ as the measure of non-randomness. This works because $\text{COMP}$ is zero when species are lost at random with respect to function, and there is no need to adjust $\text{RICH}$. As $\text{COMP}$ moves away from zero, species found at one site but not the other have lower ($\text{COMP} / \text{RICH} < 0$) or higher ($\text{COMP} / \text{RICH} > 0$) than average function.

For all three models, a significant effect of dominance nested within "type" would show that the response ($\text{RICH} + \text{COMP} , \text{CDE}$ or $\text{COMP} / \text{RICH}$) became more (or less) important to function as function became more concentrated in a few species, as opposed to being spread more evenly among species. Averaging across functions, we expected increasing dominance to decrease the effect of species loss because of a higher proportion of species having small contributions to function. We had no a priori expectation for the effect of dominance on the CDE. In the main text, we report $p$-values and semi-partial correlation coefficients for functional dominance nested within type, rather than for the whole model. This is because we were most interested in the amount of variance specifically described by functional dominance versus any given response variable (effect of species loss, effect of context dependence or the non-randomness of species loss), rather than variance between different "types". We had no a priori expectation for how mean functional dominance, or any response variable, would vary across "types". Full model results are available in the Supporting Information Appendix S5.

3 | RESULTS

3.1 | Does species loss, as measured through changes in richness and composition, have similar effects on ecosystem function in experimental and natural communities?

The total change in function between sites ($\text{RICH} + \text{COMP} + \text{CDE}$) did not differ between the experimental and natural communities included in our analysis ($r^2 = .013, p = .104$; Figure 2). If species had been lost at random, species loss would have accounted for a slightly greater decline in function in experimental communities than in natural communities. This is shown by a lower (i.e., larger magnitude) value for $\text{RICH}$ for experimental communities in Figure 2 ($r^2 = .149, p = 3.76 \times 10^{-6}$). Species were lost at random in experiments, as required by design. However, species were not lost at random in natural communities; instead, low-function species were more likely to be lost. This contrast is shown in Figure 2 by a positive $\text{COMP}$ for natural communities and a significantly lower, near-zero $\text{COMP}$ for experimental communities ($r^2 = .304, p = 7.60 \times 10^{-12}$). Thus, in natural communities, $\text{COMP}$ partly cancelled $\text{RICH}$, because these terms were of opposite signs. As a result, the effect size of species loss on function was greater in experimental communities than in natural communities ($r^2 = .444, p = 4.09 \times 10^{-18}$, shown by a more negative $\text{RICH} + \text{COMP}$ in Figure 2). If high-function species had been lost, $\text{COMP}$ would have augmented $\text{RICH}$ and increased the effects of species loss, but this result was rare in our data. Furthermore, we stress that knowing that low-function species were generally lost does not guarantee a low effect size of species loss as defined by the ecological Price equation, because: (a) species loss incorporates both the number and the identity of lost species ($\text{RICH}$ and $\text{COMP}$, respectively); and (b) for each pairwise comparison, the effects of species loss are scaled by baseline site function (see Methods and Supporting Information Appendix S5). Changes in factors that do not

![Figure 2](colour-figure-can-be-viewed-at-wileyonlinelibrary.com)
3.2 | Do experimental and natural communities have similar levels of functional dominance, and how does functional dominance mediate the effect of species loss on function?

Mean functional dominance was slightly higher in experimental communities than in natural communities, but the distributions were broadly overlapping (Figure 3). However, similarity in the distributions of functional dominance does not mean that the consequences of varying functional dominance were the same in experimental and natural communities.

The effect of species loss (\( RICH + COMP \)) did not vary with dominance in experiments (semi-partial \( r^2 = .080, p = .151; \) Figure 4a). However, in natural communities, species loss had little effect on ecosystem function at high-dominance sites, whereas it decreased ecosystem function at low-dominance sites (semi-partial \( r^2 = .366, p = 2.49 \times 10^{-10}; \) Figure 4b). Thus, lost species made either major or insignificant contributions to ecosystem function, depending on functional dominance.

Likewise, the effect of context dependence (\( CDE \)) on function did not change with functional dominance in experiments (semi-partial \( r^2 = .020, p = .562; \) Figure 4c). However, the effect of context dependence was positively correlated with functional dominance in natural communities (semi-partial \( r^2 = .182, p = 7.25 \times 10^{-6}; \) Figure 4d). Thus, in natural communities, as functional dominance increased and species loss explained less variation in function, context dependence explained more.

3.3 | Are species lost at random in experimental and natural communities, and how does non-randomness in the order of species loss affect function?

Species loss was, by design, random in experiments (semi-partial \( r^2 = .076, p = .305; \) Figure 4e). In natural communities, as functional dominance increased, species lost between sites increasingly made small contributions to function (semi-partial \( r^2 = .299, p = 1.03 \times 10^{-6}; \) Figure 4f). This helps to explain why the importance of species loss decreased with dominance (see section 3.2 above).

4 | DISCUSSION

Understanding the consequences of species loss for ecosystem function is a great challenge in ecology. Species loss consists of two components: the number of lost species and the identity of those species. In other words, species loss includes a decline in richness irrespective of the identity of the species lost (a random component) and any pattern that might exist in the identity of the species lost (a non-random component). Here, we found that the importance of species loss to ecosystem function can be predicted by two factors: (a) whether species loss was random or not; and (b) the extent of functional dominance in the ecological community. In experiments, and in natural communities with low functional dominance, the identity of lost species was random, and the effects of species loss were important. However, in natural communities with high functional dominance, the effect of species loss was weak, because the species that contribute less to function (often, the rare species; Supporting Information Appendix S6) were more likely to be lost. In these natural communities, changes in function were driven by shifts in the abundance and per-capita function of the persistent, functionally dominant species.

Our findings might help to reconcile a discrepancy between the results of biodiversity–ecosystem function experiments, which overwhelmingly find that species richness is a strong driver of ecosystem function (Cardinale et al., 2012), and studies done in natural communities, which have variously implicated species richness (Duffy et al., 2016, 2017; Grace et al., 2016; Mora et al., 2011), the order of species loss (Larsen et al., 2005), dominant species (Genung et al., 2017; Winfree et al., 2015) or aggregate abundance (Smith & Knapp, 2003) as important to function. The contributions of lost species were more important when functional dominance was low, whereas spatial variation in function provided by common species was more important when functional dominance was high. This finding is consistent with
recent work indicating that species richness and evenness can drive function through distinct mechanisms (Sonkoly et al., 2019).

Instead of linking species richness with function without accounting for community composition and abundance, in which case richness is an implicit surrogate for changes in composition and abundance that co-vary with richness, the ecological Price equation assigns effect sizes to all three (Fox, 2006; Fox & Kerr, 2012; shifts in abundance are a component of the context dependence effect). This is an important strength and allowed us to explore how changes in functional dominance affected the relative importance of species loss and context dependence. However, the ecological Price equation also has limitations. For example, it provides less information about the shape of the richness–function relationship, and it is not a tool for identifying the effects of complementarity (Fox, 2006). For two reasons, it is possible that the ecological Price equation could find a weak effect of species loss, whereas a conventional analysis with generalized linear models finds species loss to be correlated with function. First, it is possible that even when dominant species drive changes in function, positive effects of species richness on function are still strong enough to detect statistically. Second, the ecological Price equation defines the effects of “richness” differently from the way in which it is defined in the generalized linear models usually used to analyse biodiversity experiments (Fox, 2006). The ecological Price equation compares sites with one another in pairwise fashion, allowing it to isolate the effect of species richness per se, defined as the effect of changing species richness independent of any changes in mean function per species. Generalized linear models of biodiversity experiments estimate the often-nonlinear association between species richness and function, averaging over all sites and over the other predictor variables included in the model. There is no straightforward mapping between the terms in the ecological Price equation and the terms estimated by a generalized linear model. Despite these differences, our ecological Price equation analysis is, for experimental communities, in agreement with decades of studies showing that species loss can be a major driver of declines in function (Cardinale et al., 2012). Perhaps the more pressing issue to resolve is the difference between our results and other studies using real-world, observational data that have used sophisticated methods to control for abiotic differences between sites but did not investigate the role of shifts in composition and abundance that occur alongside changes in richness. These studies have found that species richness is an important driver of function (Duffy et al., 2017; Grace et al., 2016). Here, when we separate the effects of composition from richness but cannot account directly for the role of abiotic variation, we find that the importance of species loss is negatively correlated with functional

**FIGURE 4** (a,c) In experimental communities, functional dominance does not predict how ecosystem function changes as a result of either (a) species loss or (c) context dependence. (b,d) In natural communities, functional dominance (b) decreases the effects of species loss but (d) increases the effects of context dependence, which occurs independent of any changes in richness or composition. (e) In experiments, species are lost at random regardless of dominance. (f) In natural communities, as dominance increases the lost species are increasingly those with low function. The $p$ and semi-partial $r^2$ values are for the effect of functional dominance nested within different experiments or ecosystem functions. Symbols show experiments (a,c,e; all grassland plant biomass, Minnesota, USA; Germany) or functions (b,d,f; grassland plant biomass, Minnesota, USA; kelp forest biomass, near California, USA; tropical forest carbon storage, tropics worldwide; and crop pollination, eastern USA) [Colour figure can be viewed at wileyonlinelibrary.com]
dominance. Whether richness, composition, abundance or abiotic effects drive ecosystem functioning in nature is a key question for future studies.

Our findings also show that biodiversity–ecosystem function experiments are more similar to natural communities than expected. First, it has been suggested that experiments that equalize the initial abundances of species (Schmid et al., 2002) do not provide an accurate reflection of ecological communities (Kwak et al., 2007; Lamb et al., 2011; Schwartz et al., 2000), which tend to have a few common and many rare species (McGill et al., 2007). However, we found that the mean functional dominance in experiments was consistent over time (Supporting Information Appendix S7) and slightly higher than functional dominance in natural communities (Figure 2). This is a crucial point; the average decline in function between higher- and lower-function sites, at least for the datasets included in our analysis, was similar in experimental and natural communities (Figure 2). This is a crucial point; the larger effects of species loss in experiments was largely driven by the identity of lost species (COMP, Figure 2), not because there was a greater reduction in function between experimental plots relative to between sites in natural ecosystems (see also Supporting Information Appendix S8).

In conclusion, it has often been questioned whether biodiversity–ecosystem function experiments, which are based on random community assembly, provide a good model for the functional effects of species loss in nature, where extinction risk varies among species (Gross & Cardinale, 2005; Loreau et al., 2001; Schlüpf, Pfisterer, & Schmid, 2005; Srivastava, 2002; Symstad & Tilman, 2001). Here, we used the ecological Price equation to compare the roles in ecosystem function of species loss, in its random and non-random components, with changes in the functional contributions of persistent species, which take place in the absence of species loss. We found that experiments provide a good model system in several important ways. However, the design that allowed experiments elegantly to isolate the effect of richness (i.e., random assignment of species to plots) prevented detection of the non-random loss of low-functioning species, a key factor determining whether species loss matters for ecosystem function in nature.

Continuing to develop an understanding of when, and how often, it is necessary to retain many species to sustain ecosystem function is an ongoing challenge for ecologists and will help to clarify the conditions in which conservation based on ecosystem function extends to the conservation of rare species (Adams, 2014; Kleijn et al., 2015).

ACKNOWLEDGMENTS

M.A.G. was funded by U.S. National Science Foundation grant DEB-1556629 to R.W. Tropical forest carbon storage data were provided by the Tropical Ecology Assessment and Monitoring (TEAM) Network, a collaboration between Conservation International, the Missouri Botanical Garden, the Smithsonian Institution and the Wildlife Conservation Society, and partially funded by these institutions, the Gordon and Betty Moore Foundation and other donors. Pollination data were funded by a United States Department of Agriculture Agriculture and Food Research Initiative grant #2009-65104-5782 to R.W. Kelp forest datasets were provided by the Santa Barbara Coastal Long-Term Ecological Research (LTER) Program and funded by the U.S. National Science Foundation (OCE-0620276). We thank the Jena experiment for making data publicly available through Ecological Archives. Collection of Cedar Creek data (Biodiversity II, BioCON and observational grassland plant biomass data) was supported by grants from the U.S. National Science Foundation LTER Program, including DEB-0620652 and DEB-1234162. Further support was provided by the Cedar Creek Ecosystem Science Reserve and the University of Minnesota. The work presented here is a product of U.S. National Science Foundation grant DEB-1754657 to M.A.G. and R.W. Helene Muller-Landau and James Grace provided useful feedback on earlier drafts.

DATA AVAILABILITY STATEMENT

Sources, links and data use licences to publicly available data are given in the Supporting Information Appendix S2. Pollination data is available on Dryad and the appropriate link is given in Appendix S2. Tropical forest data can be accessed by emailing info at wildlifein sights.org.

CODE AVAILABILITY STATEMENT

The R code used to generate the results of this study is available in Figshare at https://figshare.com/s/981b7ad2ab845ca3ea18

ORCID

Mark A. Genung https://orcid.org/0000-0002-9383-667X
Jeremy Fox https://orcid.org/0000-0002-3432-5247

REFERENCES

Duffy, J. E., Lefcheck, J. S., Stuart-Smith, R. D., Navarrete, S. A., & Edgar, G. J. (2016). Biodiversity enhances reef fish biomass and resistance...


**BIOSKETCHES**

**Mark A. Genung** studies the community and evolutionary ecology of plant–pollinator interactions.

**Jeremy Fox** studies population, community and evolutionary ecology and is especially interested in population dynamics.

**Rachael Winfree** studies the importance of biodiversity for real-world ecosystem functioning. All three authors are interested in the relationship between biodiversity and ecosystem function and especially in the analysis of observational biodiversity–function data.

---

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Genung MA, Fox J, Winfree R. Species loss drives ecosystem function in experiments, but in nature the importance of species loss depends on dominance. *Global Ecol Biogeogr*. 2020;29:1531–1541. [https://doi.org/10.1111/geb.13137](https://doi.org/10.1111/geb.13137)

**APPENDIX: DATA SOURCES**


Figure S1.1. Consider two communities, at a baseline (higher-function) and a comparison (lower-function) site. Within sites, each circle is a different species, with colours indicating species identity and size indicating contribution to function. Now, consider the familiar axes of richness and function. The units are not important for the explanation, although it may be helpful to think of the y-axis values as being standardized by baseline site function, because this is the approach we use in the main text. The red and blue dots show the richness and function of the baseline and comparison sites, respectively. The Price equation partitions the difference in function between the baseline and comparison sites into three components: richness, composition, and context dependence. This figure above shows a simple example of the ecological Price equation in which the comparison site includes a nested subset of species found at the baseline site. This is sufficient for explaining the core approach of the Price equation. Our
analyses use a more complex version that does not require the comparison site to be a nested subset (Fox and Kerr 2012).

18 **Basic mathematics of the ecological Price equation**

19 The richness term is $\bar{z}\Delta s$, which is the expected decline in function if species are lost at random, and before any compensatory responses by persisting species. We define $\bar{z}$ as mean function across species at the baseline site and $\Delta s$ as the decline in species richness between the baseline and comparisons sites. Thus, we display richness as a red arrow along a line connecting the baseline site and the origin. The length of the arrow depends on the number of species lost between the baseline and comparison sites. The composition term is $\sum_i^s (w_i - \bar{w}) (z_i - \bar{z})$, where overbars are means, $s$ is the number of species at the baseline site, $w_i$ is a binary variable which is one if species $i$ is present at both the baseline and comparison sites and zero otherwise, and $z_i$ is species $i$’s contribution to function at the baseline site. The composition term is analogous to selection in the evolutionary Price equation and accounts for the fact that species almost certainly were not lost at random. We display the composition term as a pink arrow connecting the richness expectation with a pink circle. The pink circle is defined by the number and identity of persisting (i.e. present at both the baseline and comparison sites) species and the amount of function they provide. The context dependence term is $\sum_i^s w_i \Delta z_i$, where $\Delta z_i$ is the change in function provided by species $i$ between the baseline and comparison sites. It is analogous to imperfect transmission in the evolutionary Price equation and compares the amount of function provided by persisting species at the baseline and comparison sites. We display the context dependence term as a purple arrow connecting the pink and blue dots.
Following the three arrows helps show the logic of the ecological Price equation: first, we account for how much function would decline if species were lost at random (richness, red); second, we account for non-randomness with respect to function (composition, pink); third, we account for how the same set of species provides different levels of function at the baseline and comparison site (context dependence, purple).

Composition terms are only one component of the effects of species loss

When lost species have below- (or above-) average contributions to function, the composition term decreases (or increases) the importance of species loss. But this is only one of several factors that determine the importance of species loss. First, and most obviously, we define species loss as the sum of the ecological Price equation’s richness and composition terms. Second, for each pairwise comparison, all ecological Price equation terms are standardized by baseline site function. Thus, knowing whether lost species had below- or above-average contributions to function is insufficient to know the effects of species loss. The pairwise comparisons of sites in the above figure illustrates these points, as pair A shows a larger effect of species loss (richness + composition) than pair B, despite the loss of low-function species in pair A and high-function species in pair B.

Use of the term “species loss” in this manuscript.

In the main text we use the term “species loss” to mean the changes in richness and composition that accompany a decrease in EF. A decrease in EF usually, but not always, is accompanied by a decrease in richness. In 83% of the 129 datasets, it was more common for the higher-function site to have higher richness, than for the higher-function site to have lower richness. Thus, we
think the term “species loss” is appropriate when looking at the dataset-level results that are the focus of our analyses.

This same issue exists at the level of pairwise comparisons (there are hundreds or thousands of these within datasets). In experimental communities, the higher-function site had lower richness 28% of the time. In natural communities, the higher-function site had lower richness 41% of the time. This could happen for any number of reasons; the only requirement is that species at the lower-richness, but higher-function, site have a high enough mean contribution to function to offset the difference in richness. We acknowledge that the term “species loss” is not precise. We still use the “species loss” term to place our work in the broader realm of studies investigating the biodiversity-ecosystem function relationship, none of which (to our knowledge) measure species loss in a literal sense.

Although finding data that can inform the effects of species loss, strictly defined, is difficult, our approach does not confront fundamentally new issues. The reality of the data is that higher-richness sites or plots sometimes have lower function. In traditional biodiversity analyses, this somewhat weakens the importance of species loss, which these studies define as declines in plot-level function with declines in species richness. In ecological Price equation analyses, the effect of higher-richness sites having lower function is the same: namely, it weakens the importance of species loss, which in our ecological Price equation analysis is defined as the proportion of between-site changes in function attributable to the RICH and COMP terms.
Appendix S2. Data sources and description of use.

Within each type of ecosystem function (e.g. pollination) or experiment ID (e.g. Biodiversity II at Cedar Creek), we have multiple datasets that serve as replicates. Within each dataset there are multiple sites (or plots), and the ecological Price equation partitions the between-site (or between-plot) difference in ecosystem function (Fig. S2.1). We use multiple years of data as replicates, except for tropical forest carbon storage, which did not allow for the use of multiple years because of temporal autocorrelation (see Tropical forests section). For example, there are 19 years of data from kelp forests, and nine kelp forest sites were sampled in each year. Within each year, we apply the ecological Price equation to understand differences in ecosystem function between kelp forest sites. Thus, in Fig. 4 of the main text, each year has its own value for mean functional dominance, $\overline{\text{RICH} + \text{COMP}}$, and $\overline{\text{CDE}}$, and the overbars indicate means across the nine sites.

Fig S2.1 Ecological Price equation comparisons (arrows) are made between sites within datasets (each box represents one dataset). Comparisons are generally made between communities within
replicate years (e.g. “2011” and “2012”, above), but for tropical forest carbon storage are made between sites and the replicates are different countries (e.g. “Panama” and “Brazil”, see below).

The numbers rank the communities in order of ecosystem function, and the arrows point from the baseline site (higher ecosystem function) to the comparison site (lower ecosystem function).

Each dataset is represented by one point in Fig. 4 of the main text (a simplified version is shown here).

Natural communities: Wild bee pollination of crops in New Jersey and Pennsylvania, USA

Data source: We analysed data collected by the Winfree laboratory group, which constitutes one of the most complete data sets worldwide on the pollination function (i.e. grains of pollen deposited to stigmas) provided to crops by entire communities of wild bee species. These data sets are described in more detail in Winfree et al. (2018) and the references therein. We analysed data from farms that of grow either blueberry (N=13-16 farms per year, 3 years of data), watermelon (N=17-23 farms per year, 5 years of data), or cranberry (N=16 farms per year, 2 years of data) in New Jersey and Pennsylvania, USA. For consistency across different ecosystem functions, we hereafter call the farms “sites”. Comparisons are made only within sites growing the same crop. We excluded a small number of uncommon and rare bee species for which per-visit pollen deposition data was unavailable (1 of 71 species for watermelon; 7 of 44 species for blueberry; 1 of 57 for cranberry). All the crop pollination sites have high numerical dominance and all abundant species had available pollen deposition data.
Community providing ecosystem function: The community providing the ecosystem function consisted of the wild bees collected from a standard-area and -time transect of crop flowers at a given farm. The specific measure of ecosystem function is total grains of pollen deposited. Honey bees, which are a non-native, domesticated species in our systems, were excluded because spatial variation in function provided by honey bees is determined by human activity (e.g. hive placement) rather than ecological factors.

Defining datasets and sites: There were 10 datasets (one per crop-year combination, each yielding one point in Fig. 4 of the main text), each with 13-23 sites per year. In total, the datasets contained 10,563 records of individual bees identified to 137 species, and 1,665 records of the number of pollen grains deposited in a single visit to a crop flower by a particular type of bee. Within each dataset, the ecological Price equation partitioned between-site variation in function. Watermelon sites were in a 90 x 60 km area in New Jersey and Pennsylvania, blueberry and cranberry sites were in the same 35 x 55 km area in southern New Jersey. All farms were at least 1 km apart.

Use of multiple years of data: Multiple years of data are available for each crop. Because bees are annual organisms with highly variable population dynamics, we used years as replicates. However, because the same farms were sampled in each year, we checked for potential temporal autocorrelation. We repeated the analysis from the main text but allowed each site to be used in only one year. For example, we sampled at 16 blueberry farms over three years, and we recreated each year’s dataset as follows: in year 1 (2010), we kept data from farms 1 through 6; in year 2 (2011) we kept data from farms 7 through 11; in year 3 (2012) we kept data from farms 12
Combining crops: Each crop has spatial replicates with multiple years of data. We decided to lump the three crops as one ecosystem function (this only affects the analyses shown in main text Figure 4, where there are best fit lines for each ecosystem function) because: 1) this allows us to explore a broader range of dominance values (blueberry has lower dominance than the other crops); 2) lumping the crops gives a more similar number of data points, and thus similar statistical power, as for the other ecosystem functions; and 3) if all three crops were presented separately in Fig. 4 of the main text, each crop shows a slope consistent with the lumped data (i.e. a negative slope), suggesting that our results are due to variation in dominance and the effects of species loss within each crop, rather than a result of lumping crops.

Colour and shape in Fig. 4: Black triangles

Link: Dataset is available on Dryad with the DOI https://doi.org/10.5061/dryad.7m0cfxpr4.

Natural communities: Biomass of kelp and algae near Santa Barbara, California, USA

Data source: The Santa Barbara Coast Long Term Ecological Research Site has sampled kelp forest communities along underwater transects at kelp forest sites since 2000. We excluded sites
near the Channel Islands and used only the nine kelp forest sites near the US mainland, to keep
sites as environmentally similar as possible. For all sites, we excluded transects which were
surveyed in only one year, resulting in a minimum of two transects per site. For sites with more
than two transects each sampled in more than two years, we chose the two lowest-numbered
transects to equalize transect number across sites. We then summed data across transects within
each site, and the ecological Price equation partitions between-site differences. We included data
with genus but not species identity; each such genus is treated as one species by the analysis. We
excluded data in the kingdom Animalia (invertebrates were counted on the same transects),
because these do not contribute to the function of interest (primary production of biomass).

*Community providing ecosystem function:* Algae (which includes giant kelp) detected within
quadrats along transects at each site. Surveyors measure the abundance of each species, which is
then converted to function (biomass, specifically ash-free dry biomass) using published
allometric equations.

*Defining datasets and sites:* There were 19 datasets (one per year, each yielding one point in Fig.
4 of the main text); all datasets had nine sites except the first year (2000), which had three. In
total, the datasets contained 15,839 biomass records identified to 48 taxa (genus or species).
Within each dataset, the ecological Price equation partitioned between-site variation in function.
Sites were along an east-to-west line off the California coast near Santa Barbara, at least 2 km
and at most 70 km apart.
Use of multiple years of data: Many years of data are available. There were too many years (19), and too few sites (9), to test the effects of temporal autocorrelation by using each site in only one year (we consider this the stronger method, which is why we used it for pollination). Instead, we compared temporal similarity of species’ contributions to function across years (i.e., we calculated the mean correlation coefficient for all pairwise comparisons of years within a site) with spatial similarity in the analogous matrices across sites (i.e., across all sites within a year). Here, spatial similarity serves as a null for how similar we might expect kelp forest sites to be, absent temporal autocorrelation. We found that spatial similarity was nearly the same (mean correlation coefficient: 0.745) as temporal similarity (mean correlation coefficient: 0.761), indicating that indicating a lack of site-specific temporal autocorrelation in site composition. In other words, we expect kelp forests to have a certain degree of similarity because they are all kelp forests, and the temporal similarity did not meaningfully exceed that. Therefore, we used the site-year as the replicate and calculated the ecological Price equation across all sites within each year.

Colour and shape in Fig. 4: Light blue squares

Link: https://portal.edirepository.org/nis/mapbrowse?scope=knb-lter-sbc&identifier=50&revision=7 downloaded 12/12/19, licenced under the Creative Commons Attribution 4.0 International License

Natural communities: Aboveground carbon storage in tropical forests
Data source: Tree abundance data comes from The Tropical Ecology Assessment and Monitoring (TEAM) Network, which established and maintains long-term surveys of tropical forests worldwide. We analysed eight tropical forest datasets from four different continents (Barro Colorado-Soberania, Bukit Barisan, Caxiuana, Coca Cashu, Manaus, Nouabalé Ndoki, Ranomafana, and Yasuni).

Within each TEAM Network dataset (e.g. Manaus, Yasuni) there are six, 1-ha plots which are separated by at least 2 km. For consistency across different ecosystem functions, we hereafter call the 1-ha plots “sites”. The TEAM Network protocols requires at least 100 meters between sites and any trail access points. Sites are randomly placed within the dominant vegetation type, are required to be on homogenous soils, and cannot include any bodies of water. To further control between-site differences, we excluded TEAM Network datasets with significant elevation gradients (i.e. those where the maximum difference in elevation between sites was higher than 400 m). This elevation-based control excluded TEAM Network datasets at Bwidi, Korup, Udzungwa, Virunga Massif, and Volcán Barva. Additional TEAM Network datasets (e.g. Central Suriname, Nam Kading) largely lacked genus and species identities and were excluded. Pasoh Forest Reserve in Malaysia was excluded because it was unclear whether data were publicly available. In addition to the exclusions described above, we excluded any individual trees if the genus or species was listed as any variation on “unknown” or if a question mark was included in the genus or species name. In total these removed individuals accounted for ~10% of all individuals in the raw data set. However, we kept species that were given a consistent identity, even if it was not taxonomically resolved (e.g. “ma sp. 1”). We removed “cf.” designations and assumed the species in question had been identified correctly. Excluding
individuals with “cf.”, or maintaining “cf.” as a unique species, had no meaningful effect on the results. We did not remove “aff.” designations. Again, this decision had no meaningful effect on the results (i.e. removing “aff.” designations, or removing species with “aff.”, did not change the results). We removed subspecies and “var.” designations because “species” was our focal level of biodiversity.

Community providing ecosystem function: Within each site, carbon storage was estimated for all individual trees with a trunk more than ten cm in diameter at breast height. We used allometric equations from Chave et al. 2005 to translate abundances to carbon storage. These allometric equations rely on genus-level wood density estimates from Zanne et al. 2009. We excluded genera without genus-level wood density estimates (about 10% of individuals).

Defining datasets and sites: There were eight datasets (each yielding one point in Fig. 4 of the main text), each with 6 sites. In total, the datasets contained >27,000 records of individual trees identified to ~2200 species. Within each dataset, the ecological Price equation partitioned between-site variation in function. Datasets are separated by between 2,066 and 18,673 km (mean = 10,478 km).

Declining to use multiple years of data: Many years of data are available, but sites are permanent, and the function is provided by long-lived organisms. As for the kelp data, were too many years and too few sites to use the method we used for pollination. Using the method described above for kelp forests, temporal correlation (0.988) far exceeded spatial correlation (0.537). Thus, completing separate analyses for each year would lead to pseudoreplication, and
we analysed only the first year of available data for each dataset. For this function, replication
comes from the eight datasets from different parts of the world, rather than replicate years.

Colour and shape in Fig. 4: Green circles

Link: Ownership of the TEAM Network is in transition, and data is currently unavailable online.
Data can be obtained by emailing info ‘at’ wildlifeinsights.org.

Natural communities: Grassland plant biomass (Minnesota, USA)

Data source: Plant biomass was measured in plots in four “fields” from 1982-2004. Fields A, B, and C were successional grassland fields and field D was a prairie savanna. Fields A, B, and C had 54 plots, each 4x4 meters, and none were burned before 2005. Field D had 45 plots, each 1.5x4 meters, and these plots were burned in 1981, 1982, and frequently since 1987. Within each field, plots differed in N fertilization rates, and we chose to use only unfertilized plots because they are the most natural, unmanipulated treatment. Data is available for years after 2004, but sampling was less consistent, so we used only data collected between 1982 and 2004. We lumped fields A, B, and C together because they were the same burn treatment, habitat type (successional grasslands) and had the same size plots (4x4 meters). We kept field D separate because it was a different burn treatment, habitat type (prairie savanna) and had smaller plots. “Lumped” means that we looked at all pairwise comparisons of plots within fields A, B, and C. Some of these decisions are obvious (use of unfertilized plots for realism) and some are more subjective. To ensure the subjective decisions did not bias our results, we re-ran the analysis
using all available years (instead of stopping at 2004) and without lumping fields A, B, and C

together. Neither alternate analysis substantially affected our conclusions.

Community providing ecosystem function: Within each plot, aboveground biomass was sampled
by clipping biomass 1-2 cm above the soil surface in a 10cm x 3m strip. Each field is sampled at
a different time between early July and late August, chosen to match peak productivity of that
field.

Defining datasets and sites: There were 23 years of data, which translates to 46 datasets because
fields A, B, and C were lumped together within each year, but field D was kept separate. Within
each dataset, the ecological Price equation partitioned between-site variation in function for all
pairwise comparisons of plots.

Use of multiple years of data: 23 years of data are available. Biomass data was collected yearly
in 10cm x 3m strips, which were located in different sections of the plot each year. Because of
the long duration of the experiment, some strips were eventually reused, but this was at least 6
years later. For this reason, we treat each plot’s biomass data as independent across years.

Colour and shape in Fig. 4: Pink diamonds

Link: https://portal.edirepository.org/nis/mapbrowse?packageid=knb-lter-cdr.14.8, downloaded
12/10/2019, licenced under the Creative Commons Attribution 4.0 International License
Experiments: General data use methods applied to all experiments

Use of multiple years of data in experiments: For experiments, “plot” has the same meaning as “site” had for natural communities. Multiple years of data are available for all experiments. Experimental plant species are perennials and site composition is maintained across years, but biomass data is collected yearly in a small subset of each plot. Thus, we used all available years of the data, because the biomass sampled in plot \( m \) in year \( n \) is not directly related (i.e., not the same subsampled area) to the biomass sampled in plot \( m \) in year \( n+1 \).

Exclusion of monocultures: For all experimental data, to make comparisons with natural communities more valid, we excluded monoculture plots.

Exclusion of non-target species: We excluded biomass measurements of non-target species (i.e., species not planted as part of the experiment), because in all experiments the plots were regularly weeded to exclude these species.

Observed vs. planted richness in experiments: Experimental plots are established with a certain combination of species, but over time some of these species disappear from the plot. In our analyses, we must decide whether to account for knowledge of which species were planted (“planted richness”) or, as we do in the main text, ignore the species that were no longer present when function was measured (“observed richness”). To use “planted richness” with the ecological Price equation we would have to manually input a function value of ~0 (we used 1e-07) for species that were planted but did not persist. We re-ran our analyses doing just that and
found no substantial differences in our results. Our results are robust to this decision because, to
the extent that using “planted richness” matters at all, it weakens the importance of species loss
and increases the importance of context dependence.

Experiments: Biodiversity II at Cedar Creek

Data source: Biodiversity II is an experiment at the Cedar Creek Ecosystem Science Reserve. Plant species richness is manipulated in 1-, 2-, 4-, 8-, and 16-species plots (as for all other experiments, we excluded the 1-species plots and used the rest). Species in each plot were randomly chosen from a pool of 18 species and were seeded at equal rates (see https://www.cedarcreek.umn.edu/research/experiments/e120 for details). We analysed only the biomass produced by those 18 species (i.e. we excluded other species that colonized the plots and unsorted/unidentified biomass) because including other species without experimentally controlled planting would confound our efforts to compare the results of experimental and natural communities. Further, the other species were occasionally weeded and not always identified to the species level. We use all years of publicly available data for aboveground biomass (2001-2008, 2010-2013; species-specific aboveground biomass data was not included in the data download of aboveground biomass for the years 2009 or 1996-2000).

Community providing ecosystem function: The function is biomass production by plant species in experimental plots.
Defining datasets and plots: There were 12 datasets (one per year, each yielding one point in Fig. 4 of the main text). In total, the datasets contained biomass measurements for 15 to 18 species in 121 to 127 plots, depending on the year. Within each dataset, the ecological Price equation partitioned between-plot variation in function. More specifically, function was measured at the sub-plot level, in 10 cm x 6 m strips, harvested from the 9 m x 9 m plots.

Colour and shape in Fig. 4: Orange circles

Link: [http://www.cedarbrook.umn.edu/research/data/dataset?ple120](http://www.cedarbrook.umn.edu/research/data/dataset?ple120), downloaded 10/06/17, licenced under the Creative Commons Attribution 4.0 International License

Experiments: BioCON at Cedar Creek

Data source: BioCON explores how global change factors such as CO₂ and nitrogen fertilization affect plant communities. Plant species richness is manipulated in 1-, 4-, 9-, and 16-species plots (as for all experiments, we excluded the 1-species plots). As with Biodiversity II, we analysed only the biomass produced by the 16 originally planted species (i.e. we excluded other species that colonized the plots). We limit our analysis to the years 1998-2004 because the amount of “unsorted biomass” substantially increases after 2004. We re-ran our analyses include all years of available data (through 2012) and it does not substantially change our results. We chose to stick with ending with 2004 because including all years weakens the effect size of dominance for experimental data, as may be expected when more biomass is uncategorized. Later years of
BioCON include fully factorial manipulations of CO₂, nitrogen, temperature, and precipitation; however, up to 2004 only CO₂ and nitrogen were manipulated.

Community providing ecosystem function: The function is biomass production by plant species in experimental plots.

Defining datasets and plots: There were 28 datasets (4 per year, one per “global change treatment”, 7 years of data; each yielding one point in Fig. 4 of the main text). The BioCON experiment includes a two by two factorial manipulation of CO₂ (elevated CO₂ and ambient) and nitrogen (nitrogen fertilization and ambient), yielding the four global change treatments. To control for between-plot environmental variation, we compared plots only within treatments, and there were no obvious among-treatment differences in terms of the relationship shown in Fig. 4 in the main text. In total, the datasets contained biomass measurements for 15 to 16 species in 54 to 59 plots. Within each dataset, the ecological Price equation partitioned between-plot variation in function. More specifically, function was measured at the sub-plot level, in 10 cm x 1 m strips.

Colour and shape in Fig. 4: Blue squares

Link: http://www.cedar creek.umn.edu/research/data/dataset?ple141, downloaded 10/06/17, licenced under the Creative Commons Attribution 4.0 International License

Experiments: The Jena Experiment (Main Experiment)
Data source: The Main Experiment explores the relationship between species richness and plant productivity (among other response variables) with high maximum richness and large plot size (20 m x 20 m). Plant species richness is manipulated in 1-, 2-, 4-, 8-, 16-, and 60-species plots (as for all experiments, we excluded the 1-species plots). We analyse data from 2003-2008, because these are publicly available through an Ecological Archives Data Paper. Data from 2002 is also available, but less data was collected and in a different month.

Community providing ecosystem function: The function is biomass production by plant species in experimental plots.

Defining datasets and plots: There were six datasets (one for each year 2003-2008, each yielding one point in Fig. 4 of the main text). In total, the datasets contained biomass measurements for 56-60 species in 62-66 plots. Within each dataset, the ecological Price equation partitioned between-plot variation in function. More specifically, function was measured based on harvesting 3 or 4 0.2 m x 0.5 m subplots from 20 m x 20 m plots. (The number of subplots (3 or 4) is consistent within years, meaning we never compare two plots in which function was determined from a different number of subplots.)

Colour and shape in Fig. 4: Yellow triangles

Link: http://esapubs.org/archive/ecol/E091/066/default.htm, downloaded 3/24/18, from a published, peer-reviewed data paper
Harvest dates in experiments: Biodiversity II plots were almost always harvested in late July and early August, except in 2001 when there was a June harvest (we excluded the June data). BioCON plots were harvested in June and again in August. Jena plots were harvested in May and August, except in 2006 when they were harvested in June and August. For consistency across experiments, we aimed to limit our analysis to collections that occurred in late July and August. However, the Jena dataset was missing species-specific biomass values for the 2004 August collection. Thus, when analysing the Jena dataset, we used the May/June collection dates. In years with both May/June and August data, the results are not substantially affected by choosing May/June harvests vs. August harvests.
Appendix S3. Effects of increasing spatial scale.

Figure S3.1. It is not immediately clear whether increasing the spatial scale of data collection will increase or decrease the effects of species loss as measured by the ecological Price equation. If increases in scale increase the similarity in composition between the two sites being compared, the effect of species loss will decrease relative to context dependence. However, if increases in scale result in an accumulation of unique rare species at each site or reduce variation in the functional contributions of shared species, the effect of species loss will increase relative to context dependence. The tropical forest data included the coordinates of individual trees within...
each plot, allowing us to artificially reduce the spatial scale of data collection. We reduced the
spatial scale in increments of 10% and examined the size of ecological Price equation terms each
time (Fig. S3.1). We found that increasing spatial scale of data collection increased the between-
site difference in function ($\Delta EF$). $\text{COMP}$ was unchanged, suggesting that the non-randomness of
species loss was robust to changes in spatial scale. Similarly, $\text{CDE}$ was unchanged, suggesting
that the relative contribution of changes in abundance and per-capita function was robust to
changes in spatial scale. In contrast, $\text{RICH}$ increased with spatial scale, suggesting that random
species loss is expected to reduce function more as the two sites being compared become larger.
By extension this increased the effect of species loss ($\text{RICH} + \text{COMP}$). How the effect of species
loss changes across spatial scales is an open and important question beyond the scope of this
paper. However, this analysis does help convince us that the smaller effect of species loss in
natural systems is not driven by a larger spatial scale of data collection.
Appendix S4. Consistency of results across dominance indices.

**Figure S4.1.** We explored how five different dominance indices, described by Chao and Ricotta (2019), affected our results. For clarity, we emphasise that Chao and Ricotta’s indices were presented as evenness indices, all varying from 0 to 1. We use them as dominance indices by subtracting each index from 1. Each index is in terms of diversity (Hill numbers) of order \( q > 0 \).

Fig. S4.1 shows the effect size (semi-partial \( r^2 \), see main text) of “dominance vs. effect of species loss” (A), “dominance vs. effect of CDE” (B), and “dominance vs. non-randomness of species loss” (C) relationships from \( q = 0.25 \) to \( q = 2 \). Here, dominance specifically means “dominance nested within type”, as in Fig. 4 of the main text. \( Q \) controls the relative weighting of species’ contributions to function, with increasing values of \( q \) increasing the weight of high-function species. Thin lines are the five different dominance indices, and the thicker line is the average across the five indices. Results are qualitatively similar across indices and across values of \( q \). In the main text, we use the third class of Chao & Ricotta’s (2019) indices, with \( q = 1 \). This class of indices was generalized across \( q \) by Jost (2010) and the case of \( q = 1 \) was developed by Heip (1974).
Appendix S5. Full results for statistical models.

(1) Does species loss, as measured through changes in richness and composition, have similar effects on ecosystem function in experimental and natural communities?

All data (experimental and natural communities), predictor = experimental vs. natural, response = ecological Price equation terms (Fig. 2)

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>Difference</th>
<th>SD Diff.</th>
<th>t</th>
<th>p</th>
<th>FDR p</th>
<th>Adjusted r²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delta(Function)</td>
<td>127</td>
<td>0.030</td>
<td>0.018</td>
<td>1.64</td>
<td>0.104</td>
<td>0.104</td>
<td>0.013</td>
</tr>
<tr>
<td>RICH</td>
<td>127</td>
<td>0.079</td>
<td>0.016</td>
<td>4.84</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>0.149</td>
</tr>
<tr>
<td>COMP</td>
<td>127</td>
<td>0.109</td>
<td>0.014</td>
<td>7.55</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>0.304</td>
</tr>
<tr>
<td>CDE</td>
<td>127</td>
<td>-0.158</td>
<td>0.021</td>
<td>-7.47</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>0.300</td>
</tr>
<tr>
<td>RICH+COMP</td>
<td>127</td>
<td>0.188</td>
<td>0.019</td>
<td>10.16</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>0.444</td>
</tr>
</tbody>
</table>

“FDR p” indicates p-values adjusted with false discovery rate of 0.05.

(2) Do experimental and natural communities have similar levels of functional dominance, and how does functional dominance mediate the effect of species loss on function?

Experimental communities, response = effects of species loss (Fig. 4a)

ANOVA table

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>Sum of Sq.</th>
<th>Mean Sq.</th>
<th>F</th>
<th>p</th>
<th>Adjusted r²</th>
</tr>
</thead>
<tbody>
<tr>
<td>EF Type</td>
<td>2</td>
<td>0.098</td>
<td>0.049</td>
<td>12.399</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>dominance within EF Type</td>
<td>3</td>
<td>0.022</td>
<td>0.007</td>
<td>1.867</td>
<td>0.151</td>
<td></td>
</tr>
<tr>
<td>Residuals</td>
<td>40</td>
<td>0.157</td>
<td>0.004</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Model</td>
<td>5/40</td>
<td>6.08</td>
<td>&lt;0.001</td>
<td>0.361</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Model Summary

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.422</td>
<td>0.381</td>
<td>-1.107</td>
<td>0.275</td>
</tr>
<tr>
<td>Experiment ID (BioCON)</td>
<td>-0.017</td>
<td>0.394</td>
<td>-0.044</td>
<td>0.965</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------</td>
<td>-------</td>
<td>--------</td>
<td>-------</td>
</tr>
<tr>
<td>Experiment ID (Jena)</td>
<td>-0.597</td>
<td>0.547</td>
<td>-1.091</td>
<td>0.282</td>
</tr>
<tr>
<td>dominance within</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment ID</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Biodiversity II)</td>
<td>0.189</td>
<td>0.641</td>
<td>0.296</td>
<td>0.769</td>
</tr>
<tr>
<td>dominance within</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment ID</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(BioCON)</td>
<td>0.330</td>
<td>0.189</td>
<td>1.747</td>
<td>0.088</td>
</tr>
<tr>
<td>dominance within</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment ID</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Jena)</td>
<td>1.003</td>
<td>0.639</td>
<td>1.569</td>
<td>0.124</td>
</tr>
</tbody>
</table>

Natural communities, response = effects of species loss (Fig. 4b)

ANOVA table

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>Sum of Sq.</th>
<th>Mean Sq.</th>
<th>F</th>
<th>p</th>
<th>Adjusted r²</th>
</tr>
</thead>
<tbody>
<tr>
<td>EF Type</td>
<td>3</td>
<td>0.252</td>
<td>0.084</td>
<td>16.192</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>dominance within</td>
<td>4</td>
<td>0.370</td>
<td>0.092</td>
<td>17.842</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Residuals</td>
<td>75</td>
<td>0.388</td>
<td>0.005</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Model</td>
<td>7/75</td>
<td>17.13</td>
<td>&lt;0.001</td>
<td>0.579</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Model Summary

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.528</td>
<td>0.072</td>
<td>-7.356</td>
<td>0.000</td>
</tr>
<tr>
<td>EF Type (Pollination)</td>
<td>-0.145</td>
<td>0.124</td>
<td>-1.175</td>
<td>0.244</td>
</tr>
<tr>
<td>EF type (Kelp)</td>
<td>-0.421</td>
<td>0.259</td>
<td>-1.628</td>
<td>0.108</td>
</tr>
<tr>
<td>EF type (Forest)</td>
<td>0.435</td>
<td>0.398</td>
<td>1.092</td>
<td>0.278</td>
</tr>
<tr>
<td>dominance within EF type (Grassland)</td>
<td>0.613</td>
<td>0.102</td>
<td>6.013</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>dominance within EF type (Pollination)</td>
<td>0.941</td>
<td>0.202</td>
<td>4.650</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>dominance within EF type (Kelp)</td>
<td>1.192</td>
<td>0.324</td>
<td>3.681</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>dominance within EF type (Forest)</td>
<td>-0.113</td>
<td>0.566</td>
<td>-0.200</td>
<td>0.842</td>
</tr>
</tbody>
</table>
Experimental communities, response = effects of context dependence (Fig. 4c)

ANOVA table

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>Sum of Sq.</th>
<th>Mean Sq.</th>
<th>F</th>
<th>P</th>
<th>Adjusted r²</th>
</tr>
</thead>
<tbody>
<tr>
<td>EF Type</td>
<td>2</td>
<td>0.087</td>
<td>0.043</td>
<td>31.161</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>dominance within EF Type</td>
<td>3</td>
<td>0.003</td>
<td>0.001</td>
<td>0.692</td>
<td>0.562</td>
<td></td>
</tr>
<tr>
<td>Residuals</td>
<td>40</td>
<td>0.056</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Model</td>
<td>5/40</td>
<td>0.056</td>
<td>0.001</td>
<td>12.88</td>
<td>&lt;0.001</td>
<td>0.569</td>
</tr>
</tbody>
</table>

Model Summary

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.052</td>
<td>0.227</td>
<td>-0.227</td>
<td>0.821</td>
</tr>
<tr>
<td>Experiment ID (BioCON)</td>
<td>-0.089</td>
<td>0.234</td>
<td>-0.381</td>
<td>0.705</td>
</tr>
<tr>
<td>Experiment ID (Jena)</td>
<td>0.310</td>
<td>0.325</td>
<td>0.951</td>
<td>0.347</td>
</tr>
<tr>
<td>dominance within Experiment ID (Biodiversity II)</td>
<td>-0.189</td>
<td>0.381</td>
<td>-0.496</td>
<td>0.623</td>
</tr>
<tr>
<td>dominance within Experiment ID (BioCON)</td>
<td>0.074</td>
<td>0.112</td>
<td>0.655</td>
<td>0.517</td>
</tr>
<tr>
<td>dominance within Experiment ID (Jena)</td>
<td>-0.450</td>
<td>0.380</td>
<td>-1.184</td>
<td>0.243</td>
</tr>
</tbody>
</table>

Natural communities, response = effects of context dependence (Fig. 4d)

ANOVA table

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>Sum of Sq.</th>
<th>Mean Sq.</th>
<th>F</th>
<th>p</th>
<th>Adjusted r²</th>
</tr>
</thead>
<tbody>
<tr>
<td>EF Type</td>
<td>3</td>
<td>0.661</td>
<td>0.220</td>
<td>27.510</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>dominance within EF</td>
<td>4</td>
<td>0.281</td>
<td>0.070</td>
<td>8.775</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Residuals</td>
<td>75</td>
<td>0.600</td>
<td>0.008</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Model</td>
<td>7/75</td>
<td>0.600</td>
<td>0.008</td>
<td>16.80</td>
<td>&lt;0.001</td>
<td>0.574</td>
</tr>
</tbody>
</table>

Model Summary

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.189</td>
<td>0.089</td>
<td>2.119</td>
<td>0.037</td>
</tr>
</tbody>
</table>
EF Type (Pollination) -0.181 0.154 -1.179 0.242
EF type (Kelp) 0.067 0.322 0.209 0.835
EF type (Forest) 0.423 0.495 0.855 0.395
dominance within EF type (Grassland) -0.610 0.127 -4.815 <0.001
dominance within EF type (Pollination) -0.572 0.251 -2.276 0.026
dominance within EF type (Kelp) -0.866 0.402 -2.152 0.035
dominance within EF type (Forest) -1.022 0.704 -1.451 0.151

(3) Are species lost at random in experimental and natural communities, and how does non-randomness in the order of species loss affect function?

Experiments, response = non-randomness of loss (Fig. 4e)

ANOVA table

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>Sum of Sq.</th>
<th>Mean Sq.</th>
<th>F</th>
<th>p</th>
<th>Adjusted r²</th>
</tr>
</thead>
<tbody>
<tr>
<td>EF Type</td>
<td>2</td>
<td>0.041</td>
<td>0.021</td>
<td>2.623</td>
<td>0.085</td>
<td></td>
</tr>
<tr>
<td>dominance within EF Type</td>
<td>3</td>
<td>0.029</td>
<td>0.010</td>
<td>1.249</td>
<td>0.305</td>
<td></td>
</tr>
<tr>
<td>Residuals</td>
<td>40</td>
<td>0.314</td>
<td>0.008</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Model</td>
<td>5/40</td>
<td>1.80</td>
<td>0.135</td>
<td>0.082</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Model Summary

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.357</td>
<td>0.538</td>
<td>0.662</td>
<td>0.512</td>
</tr>
<tr>
<td>Experiment ID (BioCON)</td>
<td>-0.018</td>
<td>0.556</td>
<td>-0.032</td>
<td>0.974</td>
</tr>
<tr>
<td>Experiment ID (Jena)</td>
<td>-0.242</td>
<td>0.773</td>
<td>-0.313</td>
<td>0.756</td>
</tr>
<tr>
<td>dominance within Experiment ID (Biodiversity II)</td>
<td>-0.579</td>
<td>0.905</td>
<td>-0.639</td>
<td>0.526</td>
</tr>
<tr>
<td>4dominance within Experiment ID (BioCON)</td>
<td>-0.487</td>
<td>0.267</td>
<td>-1.824</td>
<td>0.076</td>
</tr>
<tr>
<td>dominance within Experiment ID (Jena)</td>
<td>-0.096</td>
<td>0.902</td>
<td>-0.107</td>
<td>0.915</td>
</tr>
</tbody>
</table>
**Natural communities, response = non-randomness of loss (Fig. 4f)**

**ANOVA table**

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Sq.</th>
<th>Mean Sq.</th>
<th>F</th>
<th>p</th>
<th>Adjusted r²</th>
</tr>
</thead>
<tbody>
<tr>
<td>EF Type</td>
<td>3</td>
<td>8.752</td>
<td>2.917</td>
<td>19.565</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>dominance within EF Type</td>
<td>4</td>
<td>8.519</td>
<td>2.130</td>
<td>14.283</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Residuals</td>
<td>75</td>
<td>11.183</td>
<td>0.149</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Model</td>
<td>7/75</td>
<td></td>
<td></td>
<td>16.55</td>
<td>&lt;0.001</td>
<td>0.570</td>
</tr>
</tbody>
</table>

**Model Summary**

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.084</td>
<td>0.385</td>
<td>5.407</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>EF Type (Pollination)</td>
<td>-1.656</td>
<td>0.664</td>
<td>-2.494</td>
<td>0.015</td>
</tr>
<tr>
<td>EF type (Kelp)</td>
<td>-0.450</td>
<td>1.389</td>
<td>-0.324</td>
<td>0.747</td>
</tr>
<tr>
<td>EF type (Forest)</td>
<td>3.924</td>
<td>2.137</td>
<td>1.836</td>
<td>0.070</td>
</tr>
<tr>
<td>dominance within EF type (Grassland)</td>
<td>-3.590</td>
<td>0.547</td>
<td>-6.566</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>dominance within EF type (Pollination)</td>
<td>-1.745</td>
<td>1.085</td>
<td>-1.607</td>
<td>0.112</td>
</tr>
<tr>
<td>dominance within EF type (Kelp)</td>
<td>-3.399</td>
<td>1.737</td>
<td>-1.956</td>
<td>0.054</td>
</tr>
<tr>
<td>dominance within EF type (Forest)</td>
<td>-8.379</td>
<td>3.039</td>
<td>-2.757</td>
<td>0.007</td>
</tr>
</tbody>
</table>
Appendix S6. Rare species and changes in richness.

Figure S6.1. Regionally rare species (i.e., low abundance across multiple sites within a dataset) tend to occur at fewer sites (Brown, 1984) (a: tropical forest carbon storage; b: pollination). Regionally rare species also tend to have low local abundance at the sites where they do occur (c: forests; d: pollination). Regionally rare species tend to contribute little to function (Vázquez, Morris, & Jordano, 2005) (e: forests; f: pollination). However, rare species, when present, contribute to richness the same as any common species. Taken together, this means that rare species drive the difference in richness—but not the difference in function—between sites, leading to a decoupling of richness and function in natural communities (Winfree, Fox, Williams, Reilly, & Cariveau, 2015). Each data point represents one species in one dataset.
“Regional occurrence” is the proportion of sites at which a species was detected. “Regional abundance” is the sum of a species’ abundance across all sites within a dataset. “Local abundance” is abundance measured at each site. Thus, each species has multiple “local abundance” values for each “regional abundance” value. We are unable to include kelp, grasslands (observational) or any experiments in these analyses because these systems do not include abundance data.
Appendix S7. Consistency of dominance over time in experiments.

Figure S7.1. Dominance (measured in terms of species’ functional contributions, as in the main text), is mostly stable over time. For BioCON, there are four points for each year, one for each “global change treatment” (see Appendix S2).
Appendix S8. Extent of species loss in natural communities and experiments.

Figure S8.1. The extent of species loss can affect “the effects of species loss” as presented in the main text. In experiments, the effects of species loss are inferred by comparing randomly assembled plots with fixed levels of richness (e.g. 6-species mixtures vs. 3-species mixtures). Thus, the extent of species loss is determined by experimental design. In natural communities, stochastic and ecological processes account for the degree of species loss sites. Thus, there is no guarantee that the level of the species loss treatment will be similar between experimental and natural communities. In the figure, replicates are all pairwise comparisons of sites, within each analysis. “Proportion of Species Retained” is the proportion of species shared between sites (meaning neither lost nor gained). On average, the relative extent of species loss is similar between experimental plots and spatial replicates in natural communities. This comparison indicates it is highly unlikely that differences in the effects of species loss between experimental
and natural communities (as measured by the ecological Price equation) are due to differences in the extent of species loss. Further, the result that “species loss” has a greater impact in experiments is robust to this result, because the extent of species loss is slightly lower in experimental than natural communities.
References (Supplementary Material)


