Price Equations for Understanding the Response of Ecosystem Function to Community Change

Tina Harrison,1 Rachael Winfree,2 and Mark A. Genung1,⁎

1. Department of Biology, University of Louisiana at Lafayette, Lafayette, Louisiana 70503; 2. Department of Ecology, Evolution, and Natural Resources, Rutgers University, New Brunswick, New Jersey 08901

Submitted January 15, 2021; Accepted January 21, 2022; Electronically published June 14, 2022

Online enhancements: supplemental PDF.

Abstract: The relationship between biodiversity and ecosystem function (BEF) remains unclear in many natural ecosystems, partially for lack of theoretical and analytical tools that match common characteristics of observational community data. The ecological Price equation promises to meet this need by organizing many different species-level changes into a few ecologically meaningful categories that sum to total ecosystem function change. Current versions of the ecological Price equation focus on species richness and presence-absence. However, abundance and relative abundance are better estimated in samples and are likely showing a stronger response to global change. Here, we present a novel, abundance-based version of the ecological Price equation in both discrete and continuous forms and explain the similarities and differences between this method and a related, previously developed richness-based method. We also present new empirical techniques for applying the Price equation to ecological data. Our two demonstration analyses reveal how additive effects of increasing abundance on total function are modified by concurrent selection effects due to shifts in species’ composition as well as intraspecific change in species’ per capita function. The ecological Price equations derived here complement existing approaches and together offer BEF researchers analytical tools and a unifying framework for studying BEF in observational community data.

Keywords: abundance, additive, biodiversity, composition, ecosystem function, Price equation.

Introduction

Understanding the effects of biodiversity and ecosystem function (BEF) has been central to ecological research for more than two decades (Tilman et al. 1996; Cardinale et al. 2012; Hooper et al. 2012; Jochum et al. 2020). While species richness clearly improves ecosystem function in experiments, understanding how much richness matters for ecosystem function in nature has always been challenging (Cardinale et al. 2012). Composition and abundance change nonrandomly alongside species richness, and environmental gradients have combined direct and indirect effects on species’ function (Brose and Hillebrand 2016; Spaak et al. 2017). Adaptations of the Price equation from evolutionary biology (Price 1970; Gardner 2008) have been promising for disentangling the effects of richness from these other ecological drivers of ecosystem function (Fox 2006; Fox and Kerr 2012). In large part, the BEF work described above has been motivated by a sense that declines in species richness would threaten ecosystem function in natural communities. However, recent meta-analyses suggest that the focus on richness may be misplaced. While species richness is clearly in decline at global scales, composition and abundance are more commonly changing at local scales (Vellend et al. 2013; Dornelas et al. 2014; McGill 2015; but see Cardinale et al. 2018). An expanded theoretical framework is needed to understand how declines in abundance, independent of any change in richness, affect ecosystem function in natural communities. Here, we present such a framework, using a novel Price equation partition to explain change in ecosystem function based on changes in species’ absolute and relative abundances.

All ecological adaptations of the Price equation, including the one we present below, can be used to analyze variation in any ecosystem functions that can be expressed as a sum across species—for example, biomass, pollination, or carbon storage (Fox 2006). Under controlled experimental conditions, the ecological Price equation isolates function due to complementarity among species in diverse mixtures (Loreau et al. 2001; Fox 2005). In observational data, it has been used to separate the effects of declining richness from changes in species composition, both of which are common dynamics in natural communities experiencing global change (Larsen et al. 2005; Winfree
The Price equation itself makes no distinction between specific ecological drivers of change; for example, community change due to species invasion versus human disturbance would not be treated differently.

Current applications of the ecological Price equation focus on the relationship between ecosystem function and species richness (Fox and Kerr 2012), in accordance with a historical emphasis on richness as the key component of biodiversity (Cardinale et al. 2012). This focus on richness is challenged by a current consensus, however, that emphasizes the arbitrary nature of richness as a metric (e.g., its dependence on sampling effort; Larsen et al. 2018; Roswell et al. 2021), the potential for abundance changes to outpace the effects of changing richness (Winfree et al. 2015), and the greater sensitivity of other components of biodiversity, such as abundance, to global change (McGill et al. 2015). For example, at local scales declines in sensitive species are often compensated by increases in tolerant species, buffering species richness loss while producing large changes in species’ total and relative abundance (Mayfield and Daily 2005; Vellend et al. 2013; Supp and Ernest 2014; Dornelas et al. 2014; Elahi et al. 2015; Hillebrand et al. 2018; Komatsu et al. 2019). When global change pressures are particularly strong, local richness loss is accompanied by sharp abundance declines and strong changes in species composition (Karp et al. 2012; Newbold et al. 2015). Biodiversity experiments are not generally designed to test the combined effects of changes in richness and abundance, but the Price equation is well suited for this task. Adapting the ecological Price equation to study the functional consequences of abundance change is important for increasing the relatively few theoretical and empirical studies focusing on abundance-function relationships (Vellend et al. 2013; Winfree et al. 2015; Wardle 2016; Spaak et al. 2017) and understanding the impacts of biodiversity loss as it occurs in natural communities.

Here, we work toward one central goal and then make two supporting points. Our central goal is the derivation of a novel, abundance-based version of the ecological Price equation. Supporting this central goal, in box 1 and figure 1 we discuss how our abundance-based version differs from the existing richness-based version (Fox 2006; Fox and Kerr 2012). Second, we discuss practical considerations for applying the ecological Price equation to specific research questions, both in a general sense and through demonstration

---

**Figure 1:** A simple scenario for comparing the abundance- and richness-based Price ecological equations. Each shape is one individual, with triangles and circles indicating different species. All shapes are the same size to represent per capita equivalence in function (i.e., $z_i = z' = z = 1$ for all $i$). From the abundance-based perspective, total function change is fully explained by abundance loss, since per capita function contribution is constant. From the richness-based perspective, total function change is due to both richness loss and change in mean per-species function (driven by change in abundance). In general, we recommend choosing either the richness- or abundance-based version, depending on the relevant data and hypothesis, rather than using both versions and trying to compare the results. We give more detail on the abundance- and richness-based Price equations in the supplemental PDF, both analytically (supplement B) and using the first demonstration analysis (watermelon pollination; supplement C).
analyses. Taken together, the new derivation and supporting points provide a flexible framework for analyzing BEF relationships in observational data.

**Previous Use of Ecological Price Equations to Understand Changes in Species Properties**

We first present the ecological Price equation in an abstract form, borrowing notation from Frank (2012). This version lacks the covariance term that readers may expect in the Price equation, but it is mathematically equivalent to the canonical Price equation (see supplement A) and still captures the same selection and transmission bias effects, albeit transferred to an ecological context. We then show that the ecological version can likewise be written to describe discrete (Price 1970) or continuous (Price 1972) change. The general form of the ecological Price equation partitions a change in any community mean property. The remainder of the text includes many variables, which are defined in table 1. Imagine we have observed species’ relative frequencies \( \sum_{i=1}^{n} p_i = 1 \) and some species property \( z_i \) for a community of \( i = 1, 2, \ldots, n \) species. Species relative frequency may be based on any type of data (e.g., presence-absence, counts, percent area, biomass), and the species property may be any trait, ecosystem function, or other quantity that can be measured on a per-individual basis. The mean property value across individuals (community-weighted mean) is then \( \bar{z} = \sum_{i=1}^{n} p_i z_i \).

We are interested in how change in this community-level mean, \( \Delta \bar{z} = \bar{z'} - \bar{z} \), is related to species-level changes in the constituent factors, relative frequencies \( \Delta p_i = p'_i - p_i \), and property values \( \Delta z_i = z'_i - z_i \). The Price equation achieves this decomposition with a simple algebraic rearrangement of the difference between two means (derivation in supplement A):

\[
\Delta \bar{z} = \sum_{i=1}^{n} (\Delta p_i z_i) + \sum_{i=1}^{n} (p'_i \Delta z_i). \tag{1}
\]

This expression states that all mechanisms that affect a mean property in ecological communities can be organized into two processes: interspecific changes in species relative frequencies \( \Delta p \), weighted by the original property distribution \( z \), plus intraspecific changes in species properties \( \Delta z \), weighted by species’ final relative frequencies \( p' \). Generality means that \( p \) and \( z \) are abstract placeholders for the relative frequencies and properties (examples given above), which can be replaced by any metric judged reasonable for an application (Frank 2018). All ecological Price equations are extensions or modifications of this general form. The general form of the Price equation unfortunately precludes analysis of functions that cannot be summed across individuals, as is common in ecosystem ecology (decomposition, nutrient cycling, etc.).

In the derivation above, change is described as a difference between discrete baseline and comparison states. This is an intuitive way to present the Price equation (Gardner 2008; Frank 2012), and it corresponds naturally to some real-world problems, such as communities before and after a disturbance event (Fox 2006). However, it is just as valid to describe change as a continuous process (Robertson 1966; Price 1972), which corresponds to other scenarios, such as change through time or along environmental gradients (Elner et al. 2011; Norberg et al. 2012).

The continuous ecological Price equation describes continuous change in a mean property \( \bar{z} \) along some gradient \( t \), assuming each species’ relative frequency \( p \) and property \( z \) is also a continuous and differentiable function of \( t \) (derivation in supplement A):

\[
\frac{d\bar{z}}{dt} = \sum_{i=1}^{n} \frac{dp_i}{dt} \cdot z_i + \sum_{i=1}^{n} p_i \frac{dz_i}{dt}. \tag{2}
\]

This expression, conceptually equivalent to equation (1), states that all mechanisms that drive change in a community mean property over gradient \( t \) can be organized into two processes: the rates of change in species relative frequencies over the gradient \( dp_i/dt \) relative to the species’ property distribution \( z \), plus rates of intraspecific changes in species’ properties over the gradient \( dz_i/dt \), weighted by species’ relative frequencies \( p \). Equation (2) has previously been used to understand how community mean traits

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A )</td>
<td>Total abundance, summed across species within a community</td>
</tr>
<tr>
<td>( E )</td>
<td>Statistical expectation, used in the context of the expected site</td>
</tr>
<tr>
<td>( i )</td>
<td>Indicates a specific species for summations</td>
</tr>
<tr>
<td>( n )</td>
<td>Richness (i.e., the total number of species in a community)</td>
</tr>
<tr>
<td>( p )</td>
<td>Relative abundance (i.e., proportion of total abundance)</td>
</tr>
<tr>
<td>Indicator symbol</td>
<td>Refers to comparison sites (corresponding variables without prime notation refer to baseline sites)</td>
</tr>
<tr>
<td>( t )</td>
<td>Time, used in the context of trait values or ecosystem function changing over time</td>
</tr>
<tr>
<td>( T )</td>
<td>Total ecosystem function, summed across species within a community</td>
</tr>
<tr>
<td>( z )</td>
<td>A species property, such as pollen deposition or biomass</td>
</tr>
</tbody>
</table>
change over gradients (Collins and Gardner 2009; Norberg et al. 2012; Govaert et al. 2016). For example, earlier work focused on understanding community-level phenotypic change in response to climate change developed an eco-evolutionary model in which the first term of equation (2) captured ecological effects and the second term captured evolutionary effects (Norberg et al. 2012). The continuous Price equation has seen use in ecology as well. One prominent example comes from Ellner et al. (2011), who combined the continuous Price equation with the approach of Hairston et al. (2005) to understand how genotype, phenotype, and the environment interact to affect a chosen response variable.

**Novel Ecological Price Equations for Changes in Ecosystem Function**

Adapting the above approaches to incorporate changes in abundance, as shown below, is the key advance presented here. To use the ecological Price equation to analyze changes in total ecosystem function, mean function must be scaled by total community size, introducing a third partition term (Fox 2006). The derivation proposed here is an abundance-based partition of total ecosystem function. Thus, abundance is our measure of community size, and total ecosystem function \( T \) is total abundance \( A \) times the community weighted mean \( \bar{z} \). We treat \( \pi_i \) as species’ relative abundance in a community and \( z_i \) as each species’ per capita contribution to ecosystem function. Discrete change in total function, \( \Delta T = T' - T \), can then be decomposed into changes in total abundance \( \Delta A = A' - A \) and community mean per capita function \( \Delta \bar{z} = \bar{z}' - \bar{z} \), which can be further decomposed using equation (1) into changes in species relative abundance and per capita function (derivation in supplement A):

\[
\Delta T = \Delta A \cdot \bar{z} + A' \sum_{i=1}^{n} (\Delta \pi_i z_i) + A' \sum_{i=1}^{n} (\pi_i' \Delta z_i) \tag{3}
\]

This equation is a novel three-term partition for analyzing ecosystem function change due to changes in species’ absolute and relative abundances. The first term, \( \Delta A \cdot \bar{z} \), equals the total amount of function change due to total abundance change, holding mean per capita function fixed at its original value. The second and third terms partition the change in total function due to the change in mean per capita function (eq. [1]), holding abundance fixed at its final value \( A' \). In the second term, \( \Delta \pi_i = \pi_i' - \pi_i \). As in equation (1), there is no explicit covariance term in equation (3), but \( \Delta \pi_i z_i \) is equivalent to a covariance (supplement A). To understand why, it may be helpful to consider that for \( \Delta \pi_i z_i \) to be large and positive, large positive \( \Delta \pi_i \) values must coincide with large \( z_i \) values. A parallel extension to the continuous Price equation (eq. [2]) can be used to partition continuous change in total function along a gradient (derivation in supplement A).

**Ecological Price Equations and Models of Community Change**

When two communities are compared, each may have species that are absent in the other. One should decide, for a given data set, whether such species gains are more likely statistical (e.g., due to incomplete sampling) or ecological (e.g., the species can persist at only some sites). If one considers statistical effects more likely, as we did in the demonstration analyses below, no modifications of the above equations are needed. Given that most species are rare (McGill et al. 2007), it is difficult to fully sample communities, and we consider that an increase in abundance from 0 to 1 often holds similar information to an increase in abundance from 1 to 2. However, if one considers an ecological effect more likely, as may be the case when species absences are clearly controlled by different processes than the relative abundance of present species, an alternative is available because equation (3) may be extended to separate the functional effect of adding new species from the effects of frequency changes among original residents. This extension was first developed in evolutionary biology by Kerr and Godfrey-Smith (2009) and then adapted for richness-based ecological Price equations by Fox and Kerr (2012); it is equally valid for abundance-based analysis (supplement B).

Allowing \( \pi_i \) to equal zero is a departure from evolutionary versions of the Price equation, where such a decision would be incoherent. It would suggest that parental group \( i \) had zero relative frequency and that offspring had descended from zero parents. While offspring may belong to a different subgroup than their parents, they still represent their parents’ productivity. Therefore, the new relative frequencies \( \pi_i' \) must be calculated as the fraction of offspring descended from parents in the \( i \)th subgroup (Frank 2012). However, in transferring the Price equation to ecology, “descent” metaphorically refers to the shared identity between members of the same species in two different communities. There is no useful sense in which members of one species may be descended from a different species, so we can calculate species’ new relative frequencies \( \pi_i' \) directly from species’ frequencies in the new community.

In the existing richness-based partition (equation in box 1), the magnitude of the selection effect (composition effect) is closely constrained by the observed change in richness (Fox 2006), especially when treating species loss and gain separately (Fox and Kerr 2012; see also eq. [S9] in supplement B). Therefore, the selection effect cannot be directly compared with the richness effect but can be
compared with the range of possible selection effects given
the observed change in richness (Fox 2006). In our parti-
tion, the magnitude of the selection effect is less constrained
because there are many more ways to lose a proportion of
individuals than to lose a proportion of species from a com-
munity, especially when many species are abundant. On the
other hand, when the site-species matrix is sparse (as may
occur in degraded environments or with undersampling),
adding or losing individuals will force changes in composi-
tion even if the ecological processes driving abundance loss
and gain is random with respect to species identity (Karp
et al. 2012). This will induce dependence between the abun-
dance and composition terms, which may be explored with
further analysis (Blüthgen et al. 2006) or null models (Karp
et al. 2012).

Applying the Ecological Price Equation to Data

Previous applications of the ecological Price equation par-
titioned pairwise differences in function among replicate
samples and then summarized a subset of the pairwise com-
parisons—for example, comparisons for which ΔT is neg-
ative (Fox 2006; Winfree et al. 2015; Mateo-Tomás et al.
2017; Genung et al. 2020) or comparisons with one site
selected to represent the extreme end of an ecological gra-
dient (Bloom et al. 2019). This approach is simple to imple-
ment because equation variables are represented by raw
data. However, the resulting summaries (means or med-
ians) of subsets of pairwise partitions are difficult to in-
terpret. A post hoc standardization is needed to compare
partition values from many different baseline sites, and the
meaning of variability in partition values calculated from
many nonindependent site pairs is unclear.

Instead of summarizing a subset of pairwise compari-
sions, an alternative approach begins by imagining that ob-
served communities are random samples from a regional
species abundance distribution. Then, we propose an
“expected community” in which each species’ abundance,
relative abundance, and per capita function is equal to its
mean across all observed communities. We use the adjecti-
ve “expected” because this is the community we expect
if n individuals are drawn from the regional species abun-
dance distribution at random, where n is the average abun-
dance across all observed communities. We then use the
Price equation to partition deviations from or changes in
the expected community, which takes on the role of the
baseline community (Ellner et al. 2011; Okasha and Otsuka
2020). In this analytical framework, partition values can be
calculated for single communities, representing comparison sites
that deviate from the baseline. Adding this expected com-
nunity’s expected ecosystem function to any community’s par-
tition value produces the absolute function that would be
observed at that community if only one attribute of the
community differed between that community and the bas-
eline. We think that this interpretation, together with a clearly
defined role for community-level replication, is intuitive and
powerful enough to justify the larger up-front analytical ef-
fort, relative to pairwise comparisons of raw data. For this
approach, we show a demonstration analysis that uses the
discrete Price equation to study variability in crop pollina-
tion across replicated farms, unconditioned by spatial or en-
vironmental gradients.

The expected community approach is less reasonable if
communities occur along a clearly identified gradient in
time or space (e.g., seasonal variation, elevation, and lati-
itude). If communities are changing deterministically along
an identifiable gradient in either space or time, the contin-
uous Price equation may be more appropriate. To demon-
strate the continuous Price equation approach, we present
a second demonstration analysis to study total invertebrate

Box 1: Comparing abundance- and richness-based ecological Price equations

Equation (3) factors total function into total abundance and the per capita function of the average individual in
the community, \( T = \lambda x \). It is equally valid to factor total function into species richness \( s \) and the total functional
contribution of the average species in the community \( T = \bar{s} \bar{x} \), where \( \bar{x} \) is the total function provided the average
species in the community. Applying the Price equation to analyze the contributions of change in each of these
factors to change in total function results in Fox’s previously published (2006) partition:

\[
\Delta T = \Delta s \cdot \bar{x} + \bar{s} \sum_{i=1}^{n} \Delta w_i \cdot x_i + \bar{s} \sum_{i=1}^{n} w_i' \cdot \Delta x_i.
\] (4)

In the equation above, \( w \) is the vector of species’ presence-absence in the community relativized (i.e., divided) by total
richness, \( x_i \) is each species’ total functional contribution, and \( n \) is the number of unique species across both sites. The
partitions in equations (3) and (4) must sum to the same \( \Delta T \), so changes assigned to a term under the abundance-
based partition may be assigned to a different term, or combination of terms, under the richness-based partition.

To demonstrate, we present parallel calculations for abundance- and richness-based approaches in figure 1.
How Does Bee Abundance and Composition Drive Variability in Watermelon Crop Pollination across Farms?

To obtain pollinator visitation rates to crop flowers, we net-collected wild pollinator specimens from standard areas of flowering watermelon crop at 16 replicated farms in 2012. We also collected single-visit pollen deposition data to get a point estimate for the number of watermelon pollen grains deposited on a watermelon stigma by an average visit by a member of each bee genus. Crop pollination rates were estimated by multiplying pollinator visitation rates with their per-visit pollen deposition. Our data summarize per-visit effectiveness as a genus-level property; we do not consider variability in per-visit effectiveness within species across farms. This is common in landscape pollination studies, since collecting single-visit pollen deposition data is labor intensive and pollinator effectiveness is strongly correlated with species traits such as hairiness (Stavert et al. 2016). Additional study design and data collection details are published in Winfree et al. (2015).

Our question is whether variability in crop pollination is driven by variability in bee abundance or composition (selective change in relative abundance across genera with different per-visit function). Variability is generally measured as a summary of the deviation between individual sample values and their collective mean. We use the Price equation to partition the difference in function between each observed community and the expected community. Crop pollination at each farm is the product of total bee abundance and the expected community weighted mean per capita function across all farms. The second right-hand term of equation (1) captures an important truth: the extent to which two factors covary is exactly the extent to which we cannot learn about their independent effects on a third outcome (total function). In experimental studies, this covariance is deliberately minimized by good study design. In observational studies, this covariance can be partitioned out (as in eq. [S14], supplement C), allowing us to learn from the residual orthogonal variation. In the watermelon crop pollination data set, covariance between abundance and mean per capita function across farms was very low (1.7% of mean total function $E(T)$), indicating that our Price equation partition terms carry a lot of information about the relative importance of the two factors (fig. 2A). Data underlying figure 2 have been deposited in the Dryad Digital Repository (https://doi.org/10.5061/dryad.8kpr4xqf; Genung et al. 2022).

We can use equation (3) to partition the difference in function between each farm and the uncorrelated component of average function across farms, $E(A)E(z)$ (detailed derivation in supplement C):

$$T - E(T) = [A - E(A)]E(z) + A[z - E(z)] - \text{cov}(A, z).$$

The first right-hand term gives the deviation in pollen deposition at a farm if its per capita function is equal to the mean per capita function across all farms. The second right-hand term gives the deviation in pollen deposition due to the farm having a different per capita function. In these data, species’ per capita function is fixed across farms, so the values of the second partition will be driven by how much a farm’s species relative abundance distribution differs from composition of the regional species pool. If comparing equation (5) to equation (3), note that the second term of equation (5) combines the second and third terms of equation (3), because species per capita function does not vary across sites. If species per capita function did vary across sites, we could further partition $z - E(z)$ using the derivation for equation (1).

The final term in equation (5), the covariance between abundance and mean per capita function across farms, captures an important truth: the extent to which two factors covary is exactly the extent to which we cannot learn about their independent effects on a third outcome (total function). In experimental studies, this covariance is deliberately minimized by good study design. In observational studies, this covariance can be partitioned out (as in eq. [S14], supplement C), allowing us to learn from the residual orthogonal variation. In the watermelon crop pollination data set, covariance between abundance and mean per capita function across farms was very low (1.7% of mean total function $E(T)$), indicating that our Price equation partition terms carry a lot of information about the relative importance of the two factors (fig. 2A). Data underlying figure 2 have been deposited in the Dryad Digital Repository (https://doi.org/10.5061/dryad.8kpr4xqf; Genung et al. 2022).

We added the mean total function $E(T)$ to each partition value to calculate the absolute pollen deposition expected at each farm if only abundance or only per capita function deviated from the mean community. We present these alternative scenario pollination levels plotted against bee abundance at each farm (fig. 2B). We found that across the watermelon farms, the addition of an individual bee increased pollination service to a farm by a total of 128 (±25) pollen grains (slope of gray line in fig. 2B, with 95% confidence interval). Of this increase, 113 pollen grains were contributed via abundance effects (the slope of the red line is equal to mean per capita function in the expected community; thus, the fit is perfect, and a confidence interval is
not appropriate) and 15 (±25) pollen grains via selection effects (slope of blue line). It is no coincidence that the slopes of the blue and red lines must sum to the slope of the gray line; this will always be the case. The expected community method, proposed in the “Applying the Ecological Price Equation to Data” section, provided key benefits for this analysis. By calculating the expected function \( E(T) \), we were able to plot the three alternative scenarios described above (fig. 2) and calculate slope estimates showing how much function was provided by abundance versus selection effects, neither of which would have been possible with methods used by previous applications of the ecological Price equation.

We have focused on changes in abundance instead of changes in per capita function because changes in abundance are more predictive of changes in function. However, one could instead ask how a one-unit increase in per capita function affects total function (fig. 2C). We include this panel to fully illustrate the proposed method. For the watermelon analysis, however, this relationship is not significant, and the panel thus contains little useful information.

For comparison, we used the previously published richness-based Price equation (Fox and Kerr 2012) to conduct a parallel analysis of the same watermelon pollination data. We found that higher richness was associated with higher total function and that this increase was closely matched by a scenario in which only per species function varied (holding richness and composition function fixed). However, in the scenario where only composition varied, higher richness was associated with function loss, since additional species were likely to have low baseline function due to rarity (a negative selection effect; supplement C).

**Figure 2**: A. Crop pollination at farms can be factored into total bee abundance (x-axis) and mean per capita function at each farm (y-axis). The total pollen deposition at each farm is therefore the product of the two axes and corresponds to the size of each gray point \((N = 16)\). Since the correlation between total abundance and per capita function is low \((r = 0.16)\), total function in the expected community can be approximated as the product of average abundance (red line) and average per capita function (blue line) across farms. We use the Price equation to partition the difference in total function between each farm and an “expected community” (open circle) into components due to deviation in abundance (red arrow) and per capita function (blue arrow). B, C. These panels, which have different x-axes, plot the total function for each site under scenarios of abundance variation only (red points) or per capita function variation only (blue points). Each of these scenarios can be compared with the observed total pollen deposition for the 16 watermelon farms (solid gray points). There are only two lines because species’ per capita function did not vary across sites, reducing the number of terms in equation (5) to two. In data where the per capita function also varies across sites, a complete analysis would produce a corresponding third line. Of the two, B is more informative because its x-axis (abundance) is the better predictor of total function.
How Do Abundance, Composition, and Body Size Drive Change in Total Stream Invertebrate Biomass across a Pollution Gradient?

We adapted the continuous Price equation to study changes in total community biomass in a previously published publicly available data set of stream invertebrates collected from 25 sites along a strong, continuous pollution gradient (Pomeranz et al. 2019a). Specimens were individually measured and assigned to one of six functional groups. Finer taxonomic resolution is available in the original study, but we use functional groups here to get enough data to effectively model. Additional study details are found in Pomeranz et al. (2019a). Acidic mining pollution strongly reduced stream invertebrate abundance and richness, so communities in the most polluted sites were dominated by a single group of larval midges (Diptera: Orthocladiinae). Since we are using biomass as a proxy for ecosystem function, a functional group’s per capita contribution to function is its average body size across specimens collected at a site. We excluded functional groups with no observed specimens at a site from the calculation for that site’s average body size; an alternative would be to use zero as the body size for missing functional groups, which we do not recommend because it would confound changes in composition with change in body size. We want to know the extent to which the loss in total biomass across the pollution gradient is explained by loss of total abundance, selective shifts in functional group composition, and within-group changes in body size.

We first modeled each of these factors (abundance, each functional group’s relative abundance and body size) as continuous linear functions of the pollution gradient $g$. For example, the relative abundance of the $i$th functional group is modeled as $\hat{a}_i(t) = \alpha_{i,a} + \beta_{i,a} \cdot g$. The slope of this linear function captures the rate of change in the group’s relative abundance over the pollution gradient, that is, a derivative in the continuous Price equation (eq. [2]). Using all model slopes $\beta_{i,a}$, $\beta_{i,p}$, $\beta_{i,s}$ and intercepts $\alpha_{i,a}$, $\alpha_{i,p}$, $\alpha_{i,s}$, together with a simplifying linear approximation, we parameterized a continuous Price equation that partitions the rate of change in total biomass along the gradient into three additive components, each capturing the rate of change in only one factor (abundance, composition, or body size) while holding the remaining factors fixed at intercept values (supplement D):

$$\frac{dT}{de} \approx \beta_{i,a} \sum_{i=1}^{n} (\alpha_{i,a} \cdot \alpha_{i,s}) + \alpha_{i,p} \sum_{i=1}^{n} (\beta_{i,a} \cdot \alpha_{i,s}) + \alpha_{i,s} \sum_{i=1}^{n} (\alpha_{i,p} \cdot \beta_{i,s}).$$

(6)

(Note that $n$ represents functional groups, which have taken the place of species in this analysis.) To interpret the analysis in terms of absolute function rather than change in function, we can substitute each partition value for the slope in the equation for total biomass $T$ over the pollution gradient $g$, $T = \alpha T + \beta \cdot e$. After some simplifying algebra (details in supplement D), we can express total biomass given a change in one factor—for example, relative abundance—as $T|\beta_{a} = \alpha_{a} \sum_{i=1}^{n} \beta_{a,i}$. This is the total biomass that would have been observed along the pollution gradient if species’ relative abundances changed at their modeled rates while total abundances and per capita biomass remained fixed at the values observed at midpoint sites (i.e., intercept values). The corresponding counterfactuals for change in total abundance and per capita function are $T|\beta_{a} = \alpha_{a} \sum_{i=1}^{n} \alpha_{a,i} \cdot \alpha_{i,s}$ and $T|\beta_{a} = \alpha_{a} \sum_{i=1}^{n} \alpha_{a,i} \cdot \beta_{i,s}$, respectively. Plotting all three counterfactuals together (fig. 3) shows that loss of total abundance is the primary driver of loss in total biomass (red line, slope = 0.023). Data underlying figure 3 have been deposited in the Dryad Digital Repository (https://doi.org/10.5061/dryad.v6g985s; Pomeranz et al. 2019b). Change in composition, however, has barely any effect (blue line, slope < 0.001). Finally, loss of body size within functional groups explains a modest amount of biomass loss (green line, slope = 0.008).

Discussion

The abundance-based ecological Price equation describes a linear relationship between abundance and function, a simple model that is nevertheless essential for understanding how function responds to global change in real-world ecosystems. Additive abundance effects are a common, well-validated assumption in the wider BEF literature (Vázquez et al. 2005; Kleijn et al. 2015; Spaak et al. 2017; Sonkoly et al. 2019), but their potential to explain ecosystem function loss is limited in common BEF study designs that restrict the range of abundance to isolate nonadditive biodiversity effects (Cardinale et al. 2011; Wardle 2016). Positive linear relationships between abundance and function might be expected wherever abundance is highly variable and interactions are weak. These conditions describe competitive plant communities under strong environmental stress (Spaak et al. 2017) as well as many mobile animal communities (Kremen et al. 2007). Accordingly, our demonstration analyses of two invertebrate communities both found that total ecosystem function loss was closely matched by a linear effect of abundance loss. By identifying strong abundance-function relationships in observational data, the abundance-based ecological Price equation can help resolve a long-standing debate about the importance of additive biodiversity effects in natural systems (Cardinale et al. 2011; Hooper et al. 2012; Vellend et al. 2013; Winfree et al. 2013).

The ecological Price equation is an exact partition of total ecosystem function change and therefore must include the net effect of all drivers, including the species complementarity mechanisms that are central to both BEF theory and many empirical BEF relationships (Hooper et al. 2012; Duffy et al. 2017). Species complementarity may affect ecosystem function via per capita function (Loreau and Hector 2001) or abundance (Cardinale et al. 2011). In our demonstration analysis, direct effects of pollution on stream invertebrate body size may have been augmented (or mitigated) by shifting biotic interactions—for example, loss of predators (Pomeranz et al. 2019a)—and the strong additive abundance effects observed in the watermelon pollination data are partially supported by complementarity in species responses to daytime temperatures (Rader et al. 2013). However, there is no direct mapping between Price equation partitions of observational data and partitions of experimental data, such as Loreau and Hector’s (2001) complementarity term.

The abundance-based ecological Price equation and the previously developed richness-based ecological Price equation (Fox 2006) are conceptually analogous partitions of the same total function change, but they are not easily compared. Lost richness represents lost individuals, but these same individuals may be alternatively interpreted as lost abundance. Therefore, we can find that total variation in watermelon pollination is explained by the richness and per species function terms under the richness-based partition or by the abundance term under the abundance-based partition. When an entire community is lost at once, as in a cleared forest, ascribing function loss to richness or abundance is a distinction without a difference. In less extreme scenarios, it may be helpful to distinguish systems where richness loss is primary and abundance loss secondary (highly selective drivers and biogeographic spatiotemporal scales) from the reverse (drivers shared by many species and smaller scales). To our knowledge, the

Figure 3: A. This panel shows the pattern we are trying to explain, namely, the total biomass of stream invertebrates declining over a pollution gradient. We use the Price equation to partition the change in function (biomass) along the pollution gradient into components due to deviation in abundance, composition (relative abundance), and per capita function. B. Total biomass can be subdivided into six functional groups (each color indicates one functional group). We modeled the intercept ($\alpha$) and linear slope ($\beta$) of biomass for each functional group ($\alpha_1$ and $\beta_1$ shown as examples) as functions of the pollution gradient. C. Using equation (6), the intercepts and slopes are used to calculate biomass under three counterfactuals scenarios of community response to the pollution gradient: abundance loss only (red line), change in composition (i.e., relative abundance) only (blue line), and change in per capita function only (green line).
conceptual overlap between additive richness and abundance effects has not been resolved in the BEF literature, and this may not be possible since it arises from the close mathematical interdependence between richness and abundance (Gotelli and Colwell 2001).

The general ecological Price equation conceptually unites ecological selection effects that have been defined throughout the BEF literature. Just as evolutionary selection effects capture covariance between changes in traits and fitness, ecological selection effects capture covariance between changes in species’ frequencies and species’ baseline function. Thus, ecological selection effects include the effects of higher function species being selected by chance into a local community (Huston 1997), displacing low-performing species during community growth (Loreau and Hector 2001), disproportionately benefiting from complementarity (Fox 2005), or being particularly robust or sensitive to environmental stress (Larsen et al. 2005; Mateo-Tomàs et al. 2017; Spaak et al. 2017). In these examples, composition change occurs in response to different drivers (random assembly, growth, environmental change) and may be correlated with species’ baseline function for various, largely unknown reasons (Carroll et al. 2011; Funk et al. 2017). But these various ecological processes have one shared consequence for ecosystem function: altered relative proportions of species with different baseline function.

The Price equation shows that a selection effect can be a component of every change in a community mean (in some cases the selection effect will be zero), although isolating it in a Price partition may or may not be important for a particular research question. We found weak selection effects in both of our demonstration analyses because composition changes were not strongly correlated with species’ per capita function. In other words, farms that received high pollination simply had high bee abundance, rather than individual bees belonging to especially efficient species. Similarly, biomass decreased along the pollution gradient because fewer individuals were present, not because the individuals at high-pollution sites were small bodied. However, if communities have limited size, even random composition changes must eventually result in selective loss of interspecific variability in function (ecological drift; Vellend 2010). In the following concluding paragraphs, we explore other ways the ecological Price equation invites analogy between community ecology and evolution in populations.

Field norms in BEF research and wider community ecology place high value on describing, explaining, and predicting empirical patterns. Therefore, having adapted the Price equation for BEF questions, it is tempting to immediately apply the equation to data analysis (Fox 2006), as we have done in this article. In contrast, the evolutionary Price equation’s most important contributions have been theoretical, such as the proof that kin selection and group selection are equivalent (Gardner 2008). The power of evolutionary theory derives from the tautological nature of the Price equation and natural selection itself (Frank 2012). The trade-off loss of ability to predict real-world systems is an ongoing challenge for evolutionary biology (Hunt 2014; Okasha and Otsuka 2020). Adapting the Price equation to empirical questions is therefore a work in progress that benefits both evolutionary biology and BEF research. For example, our analysis of biomass loss is inspired by an analysis of trait change over environmental gradients (Ellner et al. 2011) and the partitioning of trait change into ecological and evolutionary components (Norberg et al. 2012). On the other hand, the Price equation is already a proven theoretical tool, so its best use in BEF research may be for developing the theoretical structure needed to organize the field’s overlapping hypotheses and contingent empirical patterns.

One way this may be achieved is by using the Price equation to formalize some of the verbal parallels between evolution and ecology (Vellend 2010). Trait distributions are subject to any number of interacting biotic and abiotic drivers, which evolutionary theory reorganizes into a small, logically complete set of universally relevant processes: mutation, selection, drift, and dispersal. This conceptual compression is powerful because a pattern (“finch beaks are bigger”) has both a specific contingent explanation (“more hard-shelled seeds”) and an abstract generalizable explanation (“natural selection”). Similarly, the BEF pattern “species richness increases total biomass” has specific explanations that can be demonstrated only in experimental settings (“reduced competition,” “increase in limiting resources”) and general explanations that can be applied in any context (“selection for species with higher per capita function”). By reframing the BEF hypothesis in just three terms, the ecological Price equation offers a highly generalizable framework for describing how changes in species’ distributions and function result in changes in total community function.

**Acknowledgments**

This work was funded by National Science Foundation Division of Environmental Biology grant 1915938 to M.A.G. We thank the editor (Erol Akçay), associate editor (Mark Vellend), and two anonymous reviewers for comments that greatly improved the manuscript.

**Statement of Authorship**

T.H. and M.A.G. conceptualized the study. M.A.G. and R.W. acquired funding. T.H. developed the new abundance-based partitions with assistance from M.A.G. All authors wrote the original draft and contributed to revisions.
Data and Code Availability


Literature Cited


Appendices to:

Price equations for understanding the response of ecosystem function to community change

*The American Naturalist*

Tina Harrison¹, Rachael Winfree², Mark A. Genung¹*

Contact Information:

1 – Department of Biology, University of Louisiana at Lafayette, Lafayette, LA, USA

2 – Department of Ecology, Evolution and Natural Resources, Rutgers University, New Brunswick, NJ, USA

* – corresponding author
Supplement A. Deriving ecological Price equations

These derivations produce the general equations (Equations 1-3 in main text, plus the continuous partition of a community total). They are not specific to abundance-based analysis, or to analysis of ecosystem function. We used notation based on change in relative frequencies (Frank 2012), because we find it more accessible and useful to ecologists than the traditional Price equation notation based on growth rates and covariances.

_Discrete ecological Price equation: difference in means (Equation 1)_

Assume we have observed the frequencies and functions (or any additive trait value) of species in two communities, one representing a baseline state and one representing the end state after some process of change (hereafter and in the main text, “comparison site”). Pooled richness across the communities is \( n \), and some arbitrary measure of species’ absolute quantity (abundance, biomass, percent cover, presence-absence, etc.) is represented by vectors \( a_i, a_i' \) indexed \( i = 1, 2, \ldots, n \). The prime indicates values at the comparison site. Species’ frequency in each community is then \( p = a / \sum a \) so that \( \sum p = \sum p' = 1 \). The corresponding vectors of each species’ function in each site are \( z, z' \). If \( a \) represents abundance, \( z \) represents per-capita function; if \( a \) represents presence-absence, \( z \) represents per-species function. In other contexts, such as community weighted mean trait analysis, \( z, z' \) may represent any individual or species-level trait.

The weighted mean function for each community is the product of each species’ relative quantity and per-unit function, summed across all species: \( \bar{z} = \sum_{i=1}^{n} p_i z_i \) and \( \bar{z}' = \sum_{i=1}^{n} p_i' z'_i \). Change in mean function is then:

\[
\Delta \bar{z} = \bar{z}' - \bar{z} = \sum_{i=1}^{n} p_i' z'_i - \sum_{i=1}^{n} p_i z_i
\]
There are many summations in the appendices, and in all cases the summation subscripts would be from $i = 1, 2, \ldots, n$, so from this point we have dropped the subscripts to keep the equations neater. Our goal is to partition change in mean function into additive terms that separate the effects of $\Delta p = p' - p$ and $\Delta z = z' - z$. If species’ relative quantities changed, but the function distribution is fixed so $z = z'$, then the mean function in the second community will equal $\sum p'z$.

We can both add and subtract this quantity to the expression above, regroup:

$$\Delta z = \sum p'z - \sum pz + (\sum p'z - \sum p'z) = (\sum p'z - \sum pz) + (\sum p'z - \sum p'z)$$

And emerge with the ecological Price equation:

$$\Delta \bar{z} = \sum \Delta p \ast z + \sum p' \ast \Delta z \quad (A1)$$

This expression states that all mechanisms that affect mean function in ecological communities can be organized into two processes: interspecific changes in species relative quantities $\Delta p$ weighted by the original function distribution $z$, plus intraspecific changes in species per-capita function $\Delta z$ weighted by species’ final relative quantities $p'$. This form is equivalent to a form of the evolutionary Price equation used to study evolutionary trait changes (Frank 2012; Price 1970), but indexes species in communities rather than individuals in populations. The notation is that of Frank (2012), rather than the statistical notation of Price (1970).

*Discrete ecological Price equation: difference in totals (Equation 3)*

Total ecosystem function $T$ can be analyzed by multiplying community mean function by absolute community size $A = \sum a$. Then, the parallel logic used to derive equation 1 is applied to
partition change in total ecosystem function into change in total abundance, holding mean per-capita function fixed at its original value, plus change in mean per-capita function, holding total abundance fixed at its final value.

\[ \Delta T = \Delta(A\bar{z}) = \Delta A \ast \bar{z} + A' \ast \Delta \bar{z} \]  

(A2)

Equation A1 is then substituted for change in mean per-capita function \( \Delta \bar{z} \):

\[ \Delta T = \Delta(A\bar{z}) = \Delta A \ast \bar{z} + A' \sum \Delta p z + A' \sum \Delta p \ast z + A' \sum p' \ast \Delta z \]  

(A3)

This expression states that all mechanisms that affect total ecosystem function in ecological communities can be organized into three processes: changes in total community size \( \Delta A \), interspecific changes in species proportions \( \Delta p \), and intraspecific changes in species functions \( \Delta z \). Each term isolates the effect of change in one factor by holding the remaining two factors fixed.

Continuous ecological Price equation: difference in means (Equation 3)

The Price equation can also be used to describe continuous change in function resulting from continuous change in species abundance, relative abundance and per-capita function. A version of the Price equation already existed in continuous form before Price re-discovered it (Robertson 1966). For the continuous ecological Price equation, each factor \( A, p, z \) is considered a function of some independent variable \( g \), which might describe an environmental or temporal gradient. To describe how the mean function \( \bar{z}(t) \) changes across the gradient, we take the derivative with respect to \( t \). The derivative of a sum is the same thing as the sum of derivatives:
\[
\frac{d\bar{z}}{dt} = \frac{d}{dt} \left( \sum p z \right) = \sum \frac{d}{dt} (pz)
\]

To calculate the derivatives of the products \((pz)\), we apply the product rule:
\[
\frac{d\bar{z}}{dt} = \sum \left[ \frac{d(p)}{dt} \ast z + p \ast \frac{dz}{dt} \right]
\]  
\(\text{(A4)}\)

This is the continuous version of Equation A1; both describe changes in mean function as a weighted sum of changes in species’ relative abundances and changes in species’ per-capita function.

**Continuous ecological Price equation: difference in totals**

To analyze continuous change in total function \(T = A \ast \sum pz\), we apply the product rule twice:
\[
\frac{dT}{dt} = \frac{dA}{dt} \ast \bar{z} + A \ast \frac{d\bar{z}}{dt}
\]

\[
= \frac{dA}{dt} \sum pz + A \sum \frac{d(p)}{dt} z + A \sum p \frac{d(z)}{dt}
\]

Multiplying through the last line up yields the continuous version of Equation A3:
\[
\frac{dT}{dt} = \frac{dA}{dt} \ast \sum pz + A \ast \sum \frac{d(p)}{dt} z + A \ast \sum p \frac{d(z)}{dt}
\]  
\(\text{(A5)}\)

The models describing individual species’ abundance \(\hat{a}_i\) and per-capita function \(\hat{z}_i\) may be any functions of the gradient \(i\), as long as they are differentiable over the interesting portion of the gradient. From these species’ level models, we can calculate total community function across the
gradient, $\hat{T} = \hat{A} \sum \hat{p}_i \hat{z}_i$, and the derivative that describes the rate of change across the gradient $(d\hat{T})/dt$. Using the continuous form of ecological Price Equation, we can then partition the total derivative into additive components isolating the contributions of change in single factors, for example per-capita function: $\sum \hat{a}_i (d\hat{z}_i/dt)$. Like the discrete equations, the continuous equations can also accept any form of site-species data, including presence/absence data if modeled as continuous occurrence probabilities.

**Equivalence of Price’s statistical notation and Frank’s notation**

Here we demonstrate how the first term in Eq 1 can be rewritten as a covariance. For this demonstration, assume that $i$ indexes population subgroups rather than species. Then, the first term of Equation 1 ($\sum_{i=1}^{n} \Delta p_i z_i$) is the covariance between parent phenotype and relative fitness $[\text{Cov}(w/w^*, z)]$. Frank (1997) explained the identity of these superficially different notations. We present one of his later explanations here (Frank 2012, p 1006) but in reverse and using $p_i$ to indicate relative frequency rather than his $q_i$.

The commonly used form of the Price equation in evolutionary biology uses covariance notation (e.g., Gardner 2008). Often mean fitness $\bar{w}$ appears on the left-hand side, but this should still be familiar:

$$\Delta \bar{z} = \text{Cov}(\frac{w}{\bar{w}}, z) + E(\frac{w}{\bar{w}} \Delta z)$$

As usual, fitness $w_i$ is the number of offspring from an average individual in the $i$th ancestral group. Because groups vary in relative abundance $p_i$, we use a weighted mean to calculate fitness of the average individual in the whole population, $\bar{w} = \sum p_i w_i$. The proportion of the next generation descended from the $i$th subgroup is then $p'_i = p_i (w_i/\bar{w})$. 
Covariance between two vectors is the difference between the mean of their products and the product of their means, $\text{Cov}(x, y) = (1/n) \sum xy - \bar{x} \bar{y}$. The selection term can therefore be re-written as $\sum((w_i/\bar{w}) - 1)p_i z_i$. As above, summations are from $i = 1, 2, \ldots, n$, where $n$ is the number of species in the community, and subscripts are omitted to keep equations neater.

Substituting $p_i(w_i/\bar{w}) = p'_i$, we can write the selection term as

$$\sum (p'_i - p_i) z_i = \sum \Delta p z$$

In other words, the covariance between relative fitness and phenotype (canonical form of the selection term) is the interaction between the change in relative group frequencies and the parent groups’ (weighted) phenotype distribution (the same selection term in Frank’s notation, and the first term of Equation 1 in our text). Stepping back from the algebraic proof, that $\sum \Delta p z$ is a covariance can be appreciated by noticing that for it to be large and positive, large positive $\Delta p_i$ must coincide with large $z_i$ values. If large positive $\Delta p_i$ coincide with small $z_i$ values, that means large negative $\Delta p_i$ must coincide with large $z_i$ values (since $\sum \Delta p_i = 0$).

The transmission bias term is very simple to re-write:

$$E\left(\frac{w}{\bar{w}} \Delta z\right) = \sum p_i \frac{w_i}{\bar{w}} \Delta z = \sum p'_i \Delta z$$

In other words, the population-wide expectation for intra-group change in average phenotype is equal to each group’s intra-group phenotype change, weighted by the frequency of its offspring.

We could have used the usual covariance notation instead of Frank’s sum-product notation throughout the paper (with one exception, see below). In fact, Fox’s papers deriving the richness-
based partition use covariance notation (Fox 2006, Fox and Kerr 2012). However, we think
Frank’s sum-product notation is more accessible to ecologists.

The one exception is the part where we allow species abundances to increase from initial
abundance of 0, since $p'_i = p_i(w_i/\bar{w})$ cannot equal anything other than 0 if $p_i = 0$. Our
relaxation of this constraint will frequently be useful for ecological data but is not strictly
necessary for any of our further derivations, especially since cross-site mean abundances always
provides non-zero baseline relative abundances.
Supplement B. Species loss and gain in ecological Price equations

Derivation when all a=0 represent true absences

The effects of losing and gaining species (or other entities) on community mean function can be symmetrically partitioned into three terms: differences in mean function between lost species and the total original community, differences in mean function between persisting species, and differences in mean function between gained species and the total changed community (Equation 2 in Kerr and Godfrey-Smith 2009). We use an indicator vector $C \in \{1,0\}$ to select shared species, and calculate relative frequencies of shared species in the original and changed communities, $\hat{p} = aC/(\Sigma(aC))$, $\hat{p}' = (a' C)/(\Sigma(a' C))$. Then,

$$\Delta \bar{z} = \sum p'z' - \sum pz = \sum (p' + \hat{p}' - \hat{p}z' - \sum p + \hat{p} - \hat{p})z$$

Regrouping the terms,

$$\Delta \bar{z} = \sum (\hat{p} - p)z - \sum (\hat{p}' - p')z' + (\sum \hat{p}'z' - \sum \hat{p}z) \quad \text{(A6)}$$

This shows that the total difference in mean function can be partitioned into three terms: selective species loss from the original community, selective species gain into the new community, and change in mean function among persisting species. Change in mean function among persisting species $\Delta \bar{z}$ can be further be divided into changes in persisting species’ relative abundances and per-capita functions, yielding a four-term partition:

$$\Delta \bar{z} = \sum (\hat{p} - p)z - \sum (\hat{p}' - p')z' + \sum (\hat{p}' - \hat{p})z + \sum \hat{p}'\Delta z \quad \text{(A7)}$$
To write an expression for changes in total function (rather than mean function), we can repeat the same logic using total abundances $A, A'$, where $\hat{A} = \sum aw; \hat{A}' = \sum a'w$, to achieve another four-term partition:

$$\Delta T = (\hat{A} - A)\bar{z} - (\hat{A}' - A')\bar{z}' + (\hat{A}' - \hat{A})\bar{z} + \hat{A}'\Delta \bar{z}$$

Substituting Equation A7 for $\Delta \bar{z}$ yields a seven-term partition of total function:

$$\Delta T = (\hat{A} - A)\bar{z} - (\hat{A}' - A')\bar{z}' + (\hat{A}' - \hat{A})\bar{z} + \hat{A}'[\sum (\hat{p} - p)z - \sum (\hat{p}' - p')z' + \sum (\hat{p}' - \hat{p})z$$

$$\quad + \sum \hat{p}' \Delta z]$$

The interpretation is parallel with the three-term partition that doesn’t distinguish between lost, gained and persisting species (Equation A3, or Equation 3 main text). When applied to abundance and per-capita function data, the first three terms include all changes in total function due to changes in total community abundance, partitioned into loss of total abundance from species loss from the original community, increase in total abundance from species gain in the novel community, and changes in total abundance of shared species, holding fixed species’ original mean function. The next four terms include all changes in total function due to total changes in species’ mean function, holding fixed the total new abundance of shared species. These four terms (in the square brackets) partition total change in species’ mean function into change from selective species loss from the original community, selective species gain into the novel community, changes in relative abundance of shared species, and changes in shared species’ per-capita function.
Equivalency of Equation A8 and Equations in Fox & Kerr (2012)

A simple re-coding shows that Equation A8 is the same as Fox & Kerr’s 2012 five-part partition (Eq 1 and 2 in Fox & Kerr 2012). When \( a, a' \) is transformed to species presence-absence data \( b, b' \) species richness in the original community is \( s = \sum b \) and species richness in the novel community is \( s' = \sum b' \). The Price equation uses relativized frequencies, here presence-absence divided by total site richness \( w = b / s \) and \( w' = b' / s \). As before, we use an indicator vector \( C \in \{1,0\} \) to select shared species and define \( \tilde{w} = (bC) / (\sum (bC)) \), \( \tilde{w}' = (b'C) / (\sum (b'C)) \) to calculate relativized presence-absence within the subset of shared species in the original and changed communities. \( s_c \) will denote the number of shared species.

If we are using species richness as the scale between mean and total function, then we need to use the total function of each species at a site \( x, x' \) (instead of per-capita function \( z, z' \)), so \( T = s \sum z_i a_i = s \bar{x} \). In summary: to compare with the parallel abundance-based analysis presented in the main text, richness \( s \) corresponds to abundance \( A \), mean per-species function \( \bar{x} \) corresponds to mean per-capita function \( \bar{z} \), and species’ relativized presence-absence \( w_i \) corresponds to relative abundance \( p_i \). Then, the seven-term partition shown above becomes:

\[
\Delta T = (s_c - s) \bar{x} - (s_c - s') \bar{x}' + (s_c - s_c) \bar{x} +
\]

\[
s_c [(\bar{x_c} - \bar{x}) - (\bar{x'_c} - \bar{x}') + (\bar{x_c} - \bar{x_c}) + (\bar{x'_c} - \bar{x_c})]
\]

The third and sixth terms go to zero because, unlike abundance and relative abundance, species presence-absence patterns cannot affect distributions within shared species. We can also multiple through to eliminate the square brackets.
\[ \Delta T = (s_c - s)\bar{x} - (s_c - s')\bar{x}' + s_c(\bar{x}_c - \bar{x}) - s_c(\bar{x}_c' - \bar{x}') + s_c(\bar{x}_c' - \bar{x}_c) \] (A9)

Which is Fox & Kerr’s 2012 five-part partition for presence-absence data, showing (from left to right) effects of random richness loss, random richness gain, selective richness loss, selective richness gain, and changes in function within persisting species.
Appendix C. Analysis of watermelon pollination data

Here we first show the derivation for main-text Equation 5, which partitions difference between observed communities and a baseline calculated from the mean across communities. For comparison, we next show the same derivation for a richness-based analysis, and present the empirical results.

Abundance-based analysis

Total crop pollination at farm $k$ ($T_k$) is the product of total bee abundance at that farm ($A_k$) and the weighted mean per-visit pollination efficiency across species ($\bar{z}_k = \langle \sum p_i z_i \rangle_k$). Variability is generally measured as a summary of the difference between individual sample values and their collective mean, $T_k - \mathbf{E}(T)$. Since $T = A\bar{z}$, the mean function across sites is $\mathbf{E}(T) = \mathbf{E}(A\bar{z}) = \mathbf{E}(A)\mathbf{E}(\bar{z}) + \text{cov}(A, \bar{z})$. Thus, for each site,

$$T - \mathbf{E}(T) = A\bar{z} - \mathbf{E}(A)\mathbf{E}(\bar{z}) + \text{cov}(A, \bar{z})$$

The difference between the first two right-hand side terms can be partitioned using the chain rule (or equivalently following the derivation for Eq 3):

$$T - \mathbf{E}(T) = [A - \mathbf{E}(A)]\mathbf{E}(z) + A[\bar{z} - \mathbf{E}(\bar{z})] + \text{cov}(A, \bar{z}) \quad \text{(A10)}$$
We do not present the partition of \( \bar{z} - E(z) \) isolating the component \( p_l - E(p_l) \) and \( z_l - E(z_l) \), first because our particular data do not have variable \( z_i \) across sites, and second, because this derivation is mathematically equivalent to the derivation for this same partition in context of richness-based analysis, presented below.

**Richness-based analysis**

In the main text, we analyze total pollination service as a product of bee abundance, a proxy for visitation rate, and community weighted mean per-visit pollination efficiency. This same data set can be analyzed using the same approach, but instead analyzing total pollination service as a product of bee richness and mean per-species pollination service at a site. We present this alternative partition here, first to reinforce the equivalence of the abundance and richness-based partitions as explained in Appendix 2, and second in order to demonstrate an analysis where the second term of Equation A10 (main text Equation 5) is further partitioned into three separate terms, resulting in one additional counterfactual scenario. In context of a richness-based analysis, this third counterfactual gives the deviation of a site’s total function from the mean, assuming that richness and species composition (presence-absence) are fixed at mean values, but per-species function at a site changes (due to changes in abundance).

We use the same notation as in the re-derivation of Fox & Kerr’s 2012 five-part partition (Appendix 2). To compare with the parallel abundance-based analysis presented in the main text, richness \( s \) corresponds to abundance \( A \), mean per-species function \( \bar{x} \) corresponds to mean per-capita function \( \bar{z} \), and species’ relativized presence-absence \( w_i \) correspond to relative abundance \( p_i \).
Total pollination function at farm \( k \), \( (T_k) \) is the product of total bee richness at that farm \( (s_k) \) and the mean per-species pollination contribution, \( \bar{x}_k = \sum (w_i x_i)_k \).

Since \( T = s \bar{x} \), the mean function across sites is \( E(T) = E(s \bar{x}) = E(s)E(\bar{x}) + \text{cov}(s, \bar{x}) \). Thus, for each site:

\[
T - E(T) = s \bar{x} - E(s)E(\bar{x}) - \text{cov}(s, \bar{x})
\]

The difference between the first two right-hand side terms can be partitioned using the Price equation (following the derivation for Eq 3):

\[
T - E(T) = [s - E(s)]E(\bar{x}) + s[\bar{x} - E(\bar{x})] - \text{cov}(s, \bar{x})
\]

This is the richness based equivalent of Equation 5 in the main text. As noted above, mean per-species function at one site is the product of species’ relative quantity and quality at that site:

\( \bar{x}_k = (\sum w_i x_i)_k \) Its mean across sites, \( E(\bar{x}) \), can be re-written using definition of covariance to isolate a term where species relative quantity and quality are independent, so that

\[
\bar{x} - E(\bar{x}) = \sum w_i x_i - E \left( \sum w_i x_i \right) = \sum w_i x_i - \sum E(w_i)E(x_i) - \sum \text{cov}(w, x)_i
\]

This is the richness based equivalent of Equation 5 in the main text. The difference between the first two right-hand side terms can be partitioned using the Price equation (following the derivation for Eq 1):

\[
\bar{x} - E(\bar{x}) = \sum [w - E(w)]x + \sum E(w)[x - E(x)] - \sum \text{cov}(w, x)
\]
All summations are across species, while means denoted by the expectation symbol and covariance are across sites within species. The first right-hand side term gives the absolute deviation of a site’s function if species presence-absence differs from the mean site (where all species are present) but per-species contribution is fixed. The second right-hand side term gives the absolute deviation due to the site having different per-species contribution, which in this example is due to different species abundances across sites. The summed covariance captures the remaining deviation that cannot be partitioned by treating each species’ presence-absence and per-species function as varying orthogonally across sites.

Substituting Equation A13 into Equation A12 yields the complete partition of total function deviation $T - \mathbf{E}(T)$:

$$[s - \mathbf{E}(s)]\mathbf{E}(\bar{z}) + s \sum w - \mathbf{E}(w)z + s \sum \mathbf{E}(w)[x - \mathbf{E}(x)] - s \sum \mathbf{cov}(w, x) - \mathbf{cov}(s, \bar{x})$$ (A14)

The first three r.h.s. terms correspond to richness, composition, and context-dependent terms in Fox (2006) and to the abundance, composition, and per-capita function terms in Eq 2 (this paper). The two covariance terms appear because of treating richness, composition, and per-species function as independent factors, to partition their independent effects on a site’s total deviance from the mean site.

In the watermelon pollination function data set, covariance between richness and mean per-species function was 1.2% of the mean function across sites (see lack of clear trend in Figure S1 A). Data underlying Figure S1 are deposited in the Dryad Digital repository at
http://doi.org/10.5061/dryad.14nn1 (Genung et al. 2022). For individual species, covariance between presence-absence and per-species function ranged from -2% to 17% of mean function across sites. To plot and interpret the partitions, we added the mean total function $E(T)$ to each r.h.s. term of Equation A14 to calculate the total function at each farm in the counterfactual scenario. We present them plotted against bee richness at each farm (Figure S1 B). We found that across the watermelon farms, the addition of an individual bee species increased pollination service to a farm by $1674 \pm 1018$ pollen grains (slope of gray line with 95% confidence intervals calculated as $\pm$ two standard errors). Of this increase, 1406 pollen grains were contributed via richness effects (slope of red line; fit is almost perfect), plus an additional $367 \pm 258$ pollen grains via increases in per-species function (slope of green line). The addition of an individual bee species produced a countervailing loss of $342 \pm 90$ pollen grains via composition effects (negative slope of blue line), indicating that sites with higher richness select lower-functioning species into the composition mix. We could have alternatively put per-species function on the x-axis, as shown in Figure S1 C. We focus our interpretation on Figure S1 B because richness is the better predictor of total function.
Appendix 4. Analysis of stream invertebrate biomass

We want to know how the loss in total biomass across the pollution gradient is driven by loss of total abundance, intergroup changes in relative abundance across functional groups with different initial sizes, and intragroup changes in body size across the gradient. We begin by factoring biomass into total abundance and summed relative abundance times per-capita biomass $T = A \sum pz$ and consider each factor to be a linear function of the pollution gradient $t$. Fitting the linear models gives each factor a fit line $\hat{A}, \hat{p}_i, \hat{z}_i$ with intercepts $\alpha_A, \alpha_{p_i}, \alpha_{z_i}$ and slopes $\beta_A, \beta_{p_i}, \beta_{z_i}$. We can then use these slopes as the derivatives in equation A5:

$$\frac{dT}{dt} = \beta_A \sum \hat{p} \hat{z} + \hat{A} \sum \beta_{p_i} \hat{z} + \hat{A} \sum \beta_{z_i}$$  \hspace{1cm} (A15)

To calculate total biomass along the pollution gradient under different counterfactual scenarios of community change, we need to integrate each of the three terms. Of course, since each of its three factors is a function of the pollution gradient $t$, total biomass calculated from the linear models $T$ is cubic over $t$, and the derivatives are quadratic. However, we are going to approximate total biomass as linear over the pollution gradient, which works well in this case, but should be tested if this method is applied to other data. This linear approximation means that each of the three partitions of change in total biomass must be treated as constants, so we need to substitute point values in place of the conditional means in the equation above. Because we specifically want to know how pollution degrades ecosystem function from a relatively pristine state, we use the intercept values from the linear models (same as main-text Eq 6):

$$\frac{dT}{dt} \approx \beta_A \sum \alpha_p \alpha_z + \alpha_A \sum \beta_{p_i} \alpha_z + \alpha_A \sum \alpha_p \beta_{z_i}$$  \hspace{1cm} (A16)
Note that intercepts are usually estimated with less certainty than unconditional means $\hat{A}, \hat{p}_i, \hat{z}_i$ and that in other analyses the latter might be more appropriate.

It is now easy to integrate of each of the terms to calculate change in total biomass along the pollution gradient under different counterfactual scenarios of community change. For example, the counterfactual total biomass under a scenario where only abundance changes along the pollution gradient is:

$$\hat{T}|\beta_A = \int \left( \beta_A \sum \alpha_p \alpha_z \right) dt = \left( \beta_A \sum \alpha_p \alpha_z \right) * t + C$$

For our analysis describing function loss from a pristine state, the value of the integration constant C should be the intercept of all the counterfactual $\hat{T}, \alpha_{\hat{T}} = \alpha_A \sum \alpha_p \alpha_z$. Therefore,

$$\hat{T}|\beta_A = \left( \beta_A \sum \alpha_p \alpha_z \right) * t + \alpha_A \sum \alpha_p \alpha_z$$

Re-arranging the terms:

$$\hat{T}|\beta_A = (\alpha_A + \beta_A * t) \sum \alpha_p \alpha_z = \hat{A} \sum \alpha_p \alpha_z$$

This expression explains a very clear interpretation for the counterfactual: it is the total biomass that would have been observed at a point along the pollution gradient, if abundance had changed at its observed rate but species’ relative abundances and per-capita biomass remained at their values at unpolluted sites. The corresponding counterfactuals for change in relative abundance and per-capita function are $\hat{T}|\beta_{p_i} = \alpha_A \sum \hat{p} \alpha_z$ and $\hat{T}|\beta_{z_i} = \alpha_A \sum \alpha_p \hat{z}$, respectively.
Figure Legends.

Figure S1. (A) Total pollination function to farms (N = 16, gray points) can be factored into total bee richness (x-axis) and mean per-species' function (y-axis). The product of the mean richness and mean per-species function is a good estimate for function at the average farm, since their correlation across farms is low (r = 0.07). We use the Price equation to partition the difference in total function between each farm (gray point with arrows) and the average farm, into components due to deviation in richness (red arrow) and mean per-species function (blue arrow). Adding these decomposed differences to mean total function produces the counterfactual total function for the example site under a scenario where only richness varies or only mean per-species function varies. (B and C) These panels, which have different x-axes, plot the total function expected at each site under three counterfactual scenarios, where variability across sites is limited to richness (red points), species presence-absence (blue points), or the function of present species (green points). Each of these scenarios can be compared to the observed total pollen deposition for the 16 watermelon farms (solid gray points). Of the two, panel (B) is more informative because its x-axis (richness) is the better predictor of total function.
Figure S1.
References (Appendices)


