Extending the results of biodiversity - ecosystem function experiments to regional species collections

Kevin Gross[1,3] and Bradley J. Cardinale[2]

[1]Biomathematics Graduate Program
North Carolina State University
CB 8203
Raleigh, NC 27695
ph: 919-513-4690
fax: 919-515-1909
gross@stat.ncsu.edu

[2]Department of Zoology
University of Wisconsin–Madison
Madison, WI 53706
bjcardinale@wisc.edu

[3] to whom correspondence should be addressed

Abstract

Rapid loss of global biodiversity has motivated a number of recent experiments that seek to understand how species richness affects ecosystem functioning. In the first round of studies, researchers have formed local communities differing in richness by selecting species at random from an experimental species pool, which is a subset of the 'regional' collection of species. Often, the effects of local community diversity on ecological processes (e.g., nutrient cycling or primary production) are used to draw more general conclusions about how extinction from the regional pool might influence these and other ecosystem functions. Here, we investigate the extent to which such extrapolations are warranted. We use a simple numerical model to make the intuitive but important point that whenever the experimental species pool is a non-random sample of the regional pool, the results of the typical biodiversity experiment do not lead to valid conclusions about the broader SRF relationship for the regional collection of species. As experimental pools are seldom a random sample of species (most studies focus on taxa that are abundant, commercially available, or that have biological traits of particular interest), more work is needed to assess whether conclusions to date represent generalities. We go on to outline the limited set of circumstances under which current experimental designs can draw broader conclusions safely. Most promisingly, we show that if the experimental pool can be considered a random subset of the regional collection, re-sampling techniques enable inference about the broader SRF relationship on the basis of a single study.

Introduction

One of the central tasks confronting ecologists is the need to articulate the ecological consequences of global biodiversity loss. In recent years, a growing number of experiments have manipulated biodiversity explicitly in hopes of elucidating causal relationships between biodiversity and community or ecosystem functioning (Naeem et al., 1994; Tilman et al., 1996; McGrady-Steed et al., 1997; Hooper & Vitousek 1997; Wardle et al., 1997; van der Heijden et al., 1998; Hector et al., 1999; Mulder et al., 2001; Cardinale et al., 2003). While each experiment's design is unique, most of these experiments have involved combining species randomly selected from an experimental species pool into communities of pre-designated richness levels (Loreau and Hector, 2001; Schmid et al., 2002; Giller et al., 2004). The experimental species pool is often a subset of those species occurring regionally. Patterns between ecosystem function and richness that emerge are thought by some to suggest ways in which richness and function are related in natural communities, at least to the extent that natural communities are randomly assembled (Tilman, 1999).

An important question for interpreting random assembly experiments is: how do the conclusions depend on the particular set of species chosen for the experimental pool? As an example, consider the canonical species richness – function (SRF) system of grasslands. Suppose we were interested in understanding how loss of herbaceous plant species would impact the primary productivity of grassland ecosystems. Typically, we would compile a species list, or pool, of grasses, forbes, etc., for which sufficient propagules could be obtained to perform an experiment. Specific taxa that have particular biological traits of interest might also be added to this pool — for example, nitrogen fixing legumes. Then, we could conduct an SRF experiment in which communities differing in richness would be assembled by randomly selecting species from the pool. These communities would be sown in plots of some size, and biomass accrual in the plots measured some time later. Our question is: Would this experiment provide

meaningful information about how extinction of a species from the regional collection of grassland taxa would impact grassland productivity, or do the results only inform us about the effects of species loss on communities comprised exclusively of those species included in the experiment? Answering this question is important for assessing whether conclusions of modern experiments provide generalities.

The goal of this paper is to explore the relationship between experimental and regional SRF patterns using simple numerical models. The initial results are not surprising — we show that SRF patterns depend strongly on the particular group of species chosen for the experimental pool and, therefore, different experimental species pools can lead to different conclusions about the broader regional SRF relationship. What may be surprising, perhaps, is the frequency and degree by which an experimental SRF relationship may differ from the regional one. We explain why this is so, and identify one circumstance where experimental results can be generalized to regional species collections with the proper statistical tools. These results serve as both a reminder and a warning that the conclusions of SRF studies to date are limited in scope, and suggest ways in which future experimental designs can be used to support broader inference.

Relationship between experimental and regional SRF relationships

Before proceeding, we introduce some notation. Let $\mathcal{R}$ be a regional collection of species, and let $\mathcal{P}$ be an experimental species pool drawn from $\mathcal{R}$. To formalize our notion an SRF relationship, consider some measure of ecosystem function for communities that could be assembled randomly from either $\mathcal{R}$ or $\mathcal{P}$. For sake of illustration, we call this measure of ecosystem function "productivity", although nothing about our model is specific to productivity, as opposed to other measures of ecosystem function (Giller et al., 2004). Let μ_r denote the average productivity for a community with r species randomly drawn from $\mathcal{R}$ or $\mathcal{P}$. For this study, we define a SRF relationship to be the (log-)linear relationship

between μ_r and r. Other SRF relationships are certainly plausible (e.g., relationships between variance

and r; Huston (1997); McGrady-Steed et al. (1997)), but are not the focus of this study. For the species

in $\mathcal{R}$, we write the slope of the best-fitting log-linear relationship between μ_r and r as $\theta_\mathcal{R}$, and for the

species in $\mathcal{P}$, we write this slope as $\theta_\mathcal{P}$.

To explore the relationship between $\theta_\mathcal{R}$ and $\theta_\mathcal{P}$, we have devised the following model. The model

is not intended to be an exact portrayal of community function, in the sense that it is not a detailed

model of how species interact in a community. Instead, we have aimed for a simple model that retains

the salient features of random assembly experiments, without getting bogged down in too much detail.

Our model supposes that the productivity of multispecies communities is determined by competitive

sorting ("selection" in the terminology of Loreau and Hector (2001), but see comparable terminology

of Huston (1997); Wardle (1999); Aarssen et al. (2003)). We will argue later that our results hold for

other mechanisms of determining community productivity as well.

In the model, each species i is described by two quantities: a score used to rank competitive ability

X_i and a value of productivity in monoculture Y_i. In communities, the strength of competitive sorting

(i.e., dominance) is controlled by a parameter ϕ $(0 \leq \phi \leq 1)$. Specifically, mean community productivity

is equal to $\phi \times$ the productivity of the competitive dominant, plus $(1 - \phi) \times$ the average productivity

of all species in the community, so that larger values of ϕ correspond to increased dominance by the

strongest competitor. The SRF relationship in $\mathcal{R}$ ($\mathcal{P}$) is entirely determined by ϕ and the correlation

between competitive ability and productivity among the species in $\mathcal{R}$ ($\mathcal{P}$). When $\phi = 0$, expected

productivity is the same for all richness levels, so there is no SRF relationship. When $\phi > 0$, the

SRF relationship will be positive (negative) if and competitive ability and productivity are positively

(negatively) correlated among the species in $\mathcal{R}$ or $\mathcal{P}$. All simulations in the first part of this paper have

$\phi = 0.75$.

In the first simulations, regional species collections were formed by drawing X_i and Y_i independently from normal distributions with zero mean and unit variance. The observed productivity for each community was equal to mean productivity plus a mean-zero noise term with standard deviation 0.25. For all simulations, we considered communities consisting of 1, 2, 4, and 8 species. The results do not depend on these details.

Fig. 1 illustrates the connections between $\mathcal{R}$, $\mathcal{P}$, and experimental data. Here, the regional collection has 20 species. For the regional collection shown here, competitive ability and productivity are weakly negatively correlated, and so the regional SRF relationship is also weakly negative ($\theta_\mathcal{R} = -0.05$). From this regional collection, we have drawn 3 different experimental pools of 10 species each. These might represent different subsets of species chosen by different researchers for separate experiments. If these researchers had infinite resources and could assemble and replicate all possible combinations of species from their experimental species pool, researchers 1 ($\theta_\mathcal{P} = 0.07$) and 3 ($\theta_\mathcal{P} = 0.15$) would find a positive relationship between species richness and productivity while researcher 2 would find a negative relationship ($\theta_\mathcal{P} = -0.27$). With limited resources, each researcher can only assemble some of the possible communities at each richness level. Here, we simulate experiments with 20 communities per richness level. For each simulated experiment, the experimental SRF pattern was estimated using weighted least-squared regression, using the observed variation within each richness level to calculate weights. (Weighted least-squares accounts for the variance reduction effect.) Confidence intervals were calculated for each SRF pattern as well, assuming independent errors among the data. As these results show, these intervals are valid for $\theta_\mathcal{P}$, but not for $\theta_\mathcal{R}$. (To emphasize this, we write the experimental SRF pattern as $\widehat{\theta_\mathcal{P}}$.) This makes sense, because the communities included in a random assembly experiment are randomly selected from the population of all communities that could be formed from species in $\mathcal{P}$, not $\mathcal{R}$.

The results in fig. 1 are not peculiar to this particular regional collection of species. Fig. 2 shows regional SRF relationships for four different regional collections (chosen by different random selections of X_i and Y_i) and distributions of $\theta_\mathcal{P}$ and $\widehat{\theta_\mathcal{P}}$ for experimental pools drawn from each regional collection. For each regional collection, the mean of the $\theta_\mathcal{P}$'s is nearly equal to $\theta_\mathcal{R}$. However, the SRF relationship for any one experimental pool need not resemble the regional SRF relationship, and may in fact be far from it.

Two basic results appear in these simulations. First, SRF relationships among experimental pools (the $\theta_\mathcal{P}$'s) need not resemble the SRF relationships among the regional collection ($\theta_\mathcal{R}$) from which the experimental pool was drawn. Second, with customary statistical techniques, the data from a single random assembly experiment provide a basis for drawing inferences about $\theta_\mathcal{P}$, but not $\theta_\mathcal{R}$. These results raise several questions. Why does $\theta_\mathcal{P}$ vary dramatically among experimental pools drawn from the same regional collection? To what extent do these results depend on the construction of this particular model? And must inference from experimental data always stop with $\theta_\mathcal{P}$, or are there cases where inference can be extended to the regional SRF pattern?

The variability among $\theta_\mathcal{P}$'s drawn from the same regional pool can be best understood by contrast with the following thought experiment. Suppose we were to construct many communities of various richness levels by drawing species directly from $\mathcal{R}$. These communities, and the data they would yield, would be independent draws from the populations of all communities that could be assembled from $\mathcal{R}$. Consequently, statistical analyses of these data that assume independent errors would provide valid inferences about $\theta_\mathcal{R}$.

Now consider a similar experiment, but with communities randomly assembled from an experimental pool. These communities are also a sample from the population of communities that could be assembled from $\mathcal{R}$, and the SRF pattern among these communities provides an estimate of $\theta_\mathcal{R}$. However, these

communities are not independent samples from the populations of all communities that could be assembled from $\mathcal{R}$, because they all share dependence on the same experimental pool. The resulting overlap in species composition overlap makes the ecosystem response (here, productivity) positively correlated both within and among richness levels when viewed as a sample from the population of all communities that could be assembled form $\mathcal{R}$. Positive correlations among data points in a regression inflates the variability of the slope, and so the variance of the SRF pattern observed in communities assembled from $\mathcal{P}$ is inflated, relative to the variance of an SRF pattern observed by assembling communities directly from $\mathcal{R}$. Others have already discussed the effects of shared dependence on $\mathcal{P}$ on community responses *within* a richness level (Huston, 1997; Fukami et al., 2001; Morin and J.McGrady-Steed, 2004), but the variability in $\theta_{\mathcal{P}}$ described here is caused by correlations among richness levels as well.

Thus, inflated variability among $\theta_{\mathcal{P}}$'s is a consequence of shared dependence on an experimental pool, and the resulting overlap in species composition among observed communities. Importantly, this explanation does not rely on the particular mechanism by which species interactions affect community productivity. In addition to competitive sorting, the same phenomenon could arise with any of a number of alternative mechanisms (e.g., facilitation, inhibition, complementary resource use, ecosystem engineering, etc.; Fridley (2001); Loreau and Hector (2001); Cardinale et al. (2002); Naeem (2002)); all that is required is that compositional overlap results in correlated productivities across richness levels. The only biologically feasible example we can envision where $\theta_{\mathcal{P}}$'s would not vary among subsets of $\mathcal{R}$ occurs when mean community productivity is exactly the average of all member species' productivities in all communities (see scenario C in the simulation study that follows).

What can be done? A resampling method for extending inference to regional SRF relationships

Are there alternative statistical methods that would permit inference about $\theta_\mathcal{R}$ on the basis of a single experiment from a single experimental pool? If we could model the effect of shared dependence on $\mathcal{P}$ on the variability in $\widehat{\theta_\mathcal{P}}$, then we could estimate the variability in $\theta_\mathcal{P}$'s and use this estimate to draw statistical inferences about $\theta_\mathcal{R}$. Unfortunately, it is difficult to envision a statistical model that could accommodate the many ways in which species may interact to determine community function. However, it may be possible to use resampling methods (ie, jackknife and bootstrap (Efron and Tibshirani, 1993)), which are designed to estimate the sampling variability of an estimator without making parametric assumptions about the distribution from which the data were drawn. Importantly, any inference about $\theta_\mathcal{R}$ made on the basis of such a method is only valid if the pool was selected by randomly drawing species from $\mathcal{R}$. If the species in $\mathcal{P}$ were not selected at random from the regional collection, then there may be a systematic relationship between $\theta_\mathcal{P}$ and $\theta_\mathcal{R}$ that cannot be accommodated by purely statistical inference.

In the remainder of this paper, we introduce a bootstrap method for constructing confidence intervals for $\theta_\mathcal{R}$ on the basis of a single experiment and study the performance of this bootstrap in a simulation study. To set the problem up, we divide the random assembly experiment into a two-step process (fig. 1). In the first step, $\mathcal{P}$ is drawn randomly (without replacement) from $\mathcal{R}$. In the second step, experimental communities species are formed from $\mathcal{P}$ and an SRF relationship is estimated from these data. For the remainder of this section, we call this estimate $\hat{\theta}$, dropping the $\mathcal{P}$ subscript to emphasize that $\hat{\theta}$ can also be used as an estimate of $\theta_\mathcal{R}$ if $\mathcal{P}$ is randomly drawn from $\mathcal{R}$. We have examined several resampling methods for using a single data set to construct confidence intervals for $\theta_\mathcal{R}$ (namely, bootstrap, jackknife, and infinitesimal jackknife estimators and their variations), and have found that bootstrap-based confidence intervals perform best. Consequently, we present only the bootstrap here.

To create a bootstrap mimic of the data-generating process, we first use the species in $\mathcal{P}$ to create a bootstrap regional collection, $\mathcal{R}^\star$. (We follow the convention of using $\star$'s to denote bootstrap analogs.) Assuming that the number of species in $\mathcal{R}$ is known, we populate $\mathcal{R}^\star$ with copies of the species in $\mathcal{P}$ until $\mathcal{R}^\star$ has (nearly) the same number of species as $\mathcal{R}$. We then draw $\mathcal{P}^\star$ from $\mathcal{R}^\star$ by sampling from $\mathcal{R}^\star$ without replacement the same number of species as were in $\mathcal{P}$. We then eliminate duplicate species from $\mathcal{P}^\star$, leaving fewer species in $\mathcal{P}^\star$ than in $\mathcal{P}$. The bootstrap data set was formed by retaining those communities from the original data set that were subsets of $\mathcal{P}^\star$. If the bootstrap data set contained only monocultures, a new $\mathcal{P}^\star$ was selected. For each bootstrap data set, the log-linear slope of the SRF pattern, $\hat{\theta}^\star$, was estimated using weighted least squares, with weights calculated from the original data set. We used percentile-based confidence intervals for $\theta_\mathcal{R}$, so that (for example) a 95% CI would be formed by the 2.5th and 97.5th empirical percentile of the bootstrap distribution of $\hat{\theta}^\star$ (Efron and Tibshirani, 1993, §13.3).

We tested the bootstrap for three different simulation scenarios (fig. 3). The scenarios differed either in the degree of competitive sorting within the communities, or in the correlation between competitive ability and productivity among species. In scenario A, community productivity was determined primarily by the productivity of the competitive dominant ($\phi = 0.75$), but the competitive abilities and productivities of species in the regional collection were determined independently of one another. In scenario B, competitive sorting was again strong ($\phi = 0.75$), but competitive ability and productivity of species were drawn from a bivariate normal normal distribution with correlation $\rho = 0.75$. Finally, in scenario C, community productivity was simply the average of the productivities of the species in the community $\phi = 0$, and competitive ability and productivity were determined independently. Scenario C was a worst-case scenario for the bootstrap, because $\theta_\mathcal{P} = \theta_\mathcal{R} = 0$ for all $\mathcal{P}$'s and all $\mathcal{R}$'s.

In all scenarios, the following simulation was repeated 1000 times. First, 20 species were selected to

form the regional collection, and the regional collection's SRF pattern ($\theta_\mathcal{R}$) was determined analytically.

10 species were then selected from $\mathcal{R}$ to form $\mathcal{P}$. From $\mathcal{P}$, 20 communities were formed at random at

each of four richness levels (communities with 1, 2, 4, and 8 species), and data were simulated for this

collection of communities. 500 bootstrap replications were run for the simulated data set, and these

were used to calculate a 95% CI for $\theta_\mathcal{R}$. This CI was compared to the actual value of $\theta_\mathcal{R}$.

The percentile-based bootstrap CI covered the actual $\theta_\mathcal{R}$ in 92.6%, 97.7%, and 97.0% of the simu-

lations from scenarios A, B, and C, respectively (fig. 3). For comparison, 95% CIs for $\theta_\mathcal{P}$ covered $\theta_\mathcal{R}$ in

55.8%, 60.5%, and 93.7% of the simulations. (Coverage rates are close to the nominal rate for scenario

C because $\theta_\mathcal{P} = \theta_\mathcal{R} = 0$ for every simulation.) These coverage rates are encouraging, although it is

difficult to determine if the CIs are as small as they could be for data sets of this size, or if alternative

methods of interval construction could provide similar coverage with smaller intervals.

These results suggest that it is possible to generalize the results of SRF studies to a broader collection

of species, as long as the species chosen for the experimental pool were drawn randomly from that

broader collection. Although we have been equating this broader collection with the regional pool, these

methods can be used to extend inference to any collection of species from which the experimental pool

was randomly selected. Thus, even if logistical constraints prevent some regionally occurring species from

being considered for a random assembly experiment, bootstrapping could be used to extend inference to

the collection of species from which the experimental pool was randomly drawn. Of course, the surest

way to make inferences about the SRF relationship among a collection of species is to include all those

species in the experimental pool, if that is feasible.

Conclusions

SRF patterns observed in random assembly experiments depend strongly on the particular choice of species in the experimental pool. Hence, SRF patterns observed in experiments do not necessarily provide strong evidence of similar SRF patterns among larger collection of species from which the experimental pool was drawn. This strong dependence on the experimental pool results from inflated species-composition overlap of communities both within and among richness levels, relative to the overlap of communities drawn directly from the regional collection. Hence, if the goal of an experiment is to elucidate a general SRF relationship for a regional collection of species — for example, the effect of grassland vegetation on production, or insect detritivores on carbon processing — then the experimental species pool should include as many species from the regional collection as possible. Alternatively, if the experimental pool can be drawn randomly from a subset of the regional collection (or better yet, from the entire regional collection), then the bootstrap described here can be used to draw statistical inference about the collection of species from which the experimental pool was drawn. If random selection is not possible, we must simply be aware that inferences from random assembly experiments are limited, and cannot automatically be extended to larger collections of species.

Acknowledgments

References

Aarssen, L., Laird, R. and Pither, J. 2003. Is the productivity of vegetation plots higher or lower when there are more species? Variable predictions from interaction of the 'sampling effect' and 'competitive dominance effect' on the habitat templet. – Oikos 102: 427–432.

Cardinale, B., Harvey, C., Gross, K. and Ives, A. 2003. Biodiversity and biocontrol: emergent impacts of a multi-enemy assemblage on pest suppression and crop yield in an agroecosystem. – Ecol. Lett. 6: 857–865.

Cardinale, B., Palmer, M. and Collins, S. 2002. Species diversity enhances ecosystem functioning through interspecific facilitation. – Nature 415: 426–429.

Efron, B. and Tibshirani, R. 1993. An introduction to the bootstrap. – Chapman & Hall.

Fridley, J. 2001. The influence of species diversity on ecosystem productivity: how, where, and why?. – Oikos 93: 514–526.

Fukami, T., Naeem, S. and Wardle, D. A. 2001. On similarity among local communities in biodiversity experiments. – Oikos 95: 340–358.

Giller, P., Hillebrand, H., Berninger, U.-G., Gessner, M. O., Hawkins, S., Inchausti, P., Inglis, C., Leslie, H., Malmqvist, B., Monaghan, M., Morin, P. and O'Mullan, G. 2004. Biodiversity effects on ecosystem functioning: emerging issues and their experimental test in aquatic environments.. – Oikos 104: 423–436.

Hector, A. et al. 1999. Plant diversity and productivity experiments in European grasslands. – Science 286: 1123–27.

van der Heijden, M., Klironomos, J., Ursic, M., Moutoglis, P., Streitwolf-Engel, R., Boller, T., Wiemken, A. and Sanders., I. 1998. Mycorrhizal fungal diversity determines plant productivity, ecosystem variability, and productivity. – Nature 396: 69–72.

Hooper, D. and Vitousek, P. 1997. The effects of plant composition and diversity on ecosystem processes. – Science 277: 1302–1305.

Huston, M. 1997. Hidden treatments in ecological experiments: re-evaluating the ecosystem function of biodiversity. – Oecologia 110: 449–460.

Loreau, M. and Hector, A. 2001. Partitioning selection and complementarity in biodiversity experiments. – Nature 412: 72–75.

McGrady-Steed, J., Harris, P. and Morin, P. 1997. Biodiversity regulates ecosystem predictability. – Nature 390: 162–165.

Morin, P. and J.McGrady-Steed 2004. Biodiversity and ecosystem functioning in aquatic microbial systems: a new analysis of temporal variation and species richness-predictability relations. – Oikos 104: 458–466.

Mulder, C., Uliassi, D. and Doak, D. 2001. Physical stress and diversity-productivity relationships: The role of positive interactions. – PNAS 98: 6704–6708.

Naeem, S. 2002. Ecosystem consequences of biodiversity loss: The evolution of a paradigm. – Ecology 83: 1537–1552.

Naeem, S., Thompson, L., Lawler, S., Lawton, J. and Woodfin, R. 1994. Declining biodiversity can alter the performance of ecosystems. – Nature 368: 734–737.

Schmid, B., Hector, A., Huston, M., Inchausti, P., Nijs, I., Leadley, P. and Tilman, D. 2002. The design and analysis of biodiversity experiments. – In: Loreau, M., Naeem, S. and Inchausti, P. (eds.), Biodiversity and Ecosystem Functioning. Oxford University Press, pp. 61–75.

Tilman, D. 1999. The ecological consequences of changes in biodiversity: A search for general principles. – Ecology 80: 1455–1474.

Tilman, D., Wedin, D. and Knops, J. 1996. Productivity and sustainability influenced by biodiversity in grassland ecosystems. – Nature 379: 718–720.

Wardle, D. 1999. Is "sampling effect" a problem for experiments investigating biodiversity – ecosystem function relationships?. – Oikos 87: 403–407.

Wardle, D., Bonner, K. and Nicholson, K. 1997. Biodiversity and plant litter: experimental evidence which does not support the view that enhanced species richness improves ecosystem function. – Oikos 79: 247–258.

Figure Legends

Figure 1. Relationships among $\theta_\mathcal{R}$, $\theta_\mathcal{P}$, and random assembly experiments. Here, letters indicate species. The top panel shows simulated productivities for 1600 communities randomly assembled from a regional pool $\mathcal{R}$ of 20 species — 400 communities for each of four richness levels ($r = 1, 2, 4, 8$). From $\mathcal{R}$, three experimental pools have been drawn with ten species each, and for each $\mathcal{P}$ 200 simulated communities are shown for each richness level. Simulated experiments are also shown for each $\mathcal{P}$, with 20 communities at each richness level. Values of $\theta_\mathcal{R}$ and $\theta_\mathcal{P}$ are the slopes of the displayed least-squares regression lines of mean productivity (for all communities, not just those shown) against log richness. Lines shown for the experimental data are weighted least-squares regression lines for the data shown, and $\widehat{\theta_\mathcal{P}}$ values are the slope of the regression line and 95% confidence intervals calculated assuming independent errors.

Figure 2. Relationships between regional and experimental SRF relationships for four different regional collections. Panels on the left show regional SRF relationships for separate regional collections, histograms in the center column show $\theta_\mathcal{P}$ values for 500 experimental pools drawn from each regional collection, and histograms on the right show $\widehat{\theta_\mathcal{P}}$ values for 500 experiments, one from each experimental pool. Regional collections were formed by drawing (X_i, Y_i) pairs from bivariate normal distributions with either no correlation (upper two rows), a correlation of 0.5 (second row from the bottom), or a correlation of -0.5 (bottom row). The top row is the same regional collection that produced the data in fig. 1.

Figure 3. Left panels: Regional SRF relationships ($\theta_\mathcal{R}$) and experimental SRF patterns ($\widehat{\theta_\mathcal{P}}$) for 1000 different regional collections and experiments. Right panels: Bootstrap confidence intervals for $\theta_\mathcal{R}$, based on $\widehat{\theta_\mathcal{P}}$, for 50 of these simulated experiments. Vertical ordering of the CI's is not meaningful. ×'s show the actual value of $\theta_\mathcal{R}$ for the regional collection from which the experimental pool was drawn. Upper panels: simulation scenario A. Middle panels: scenario B. Lower panels: scenario C.
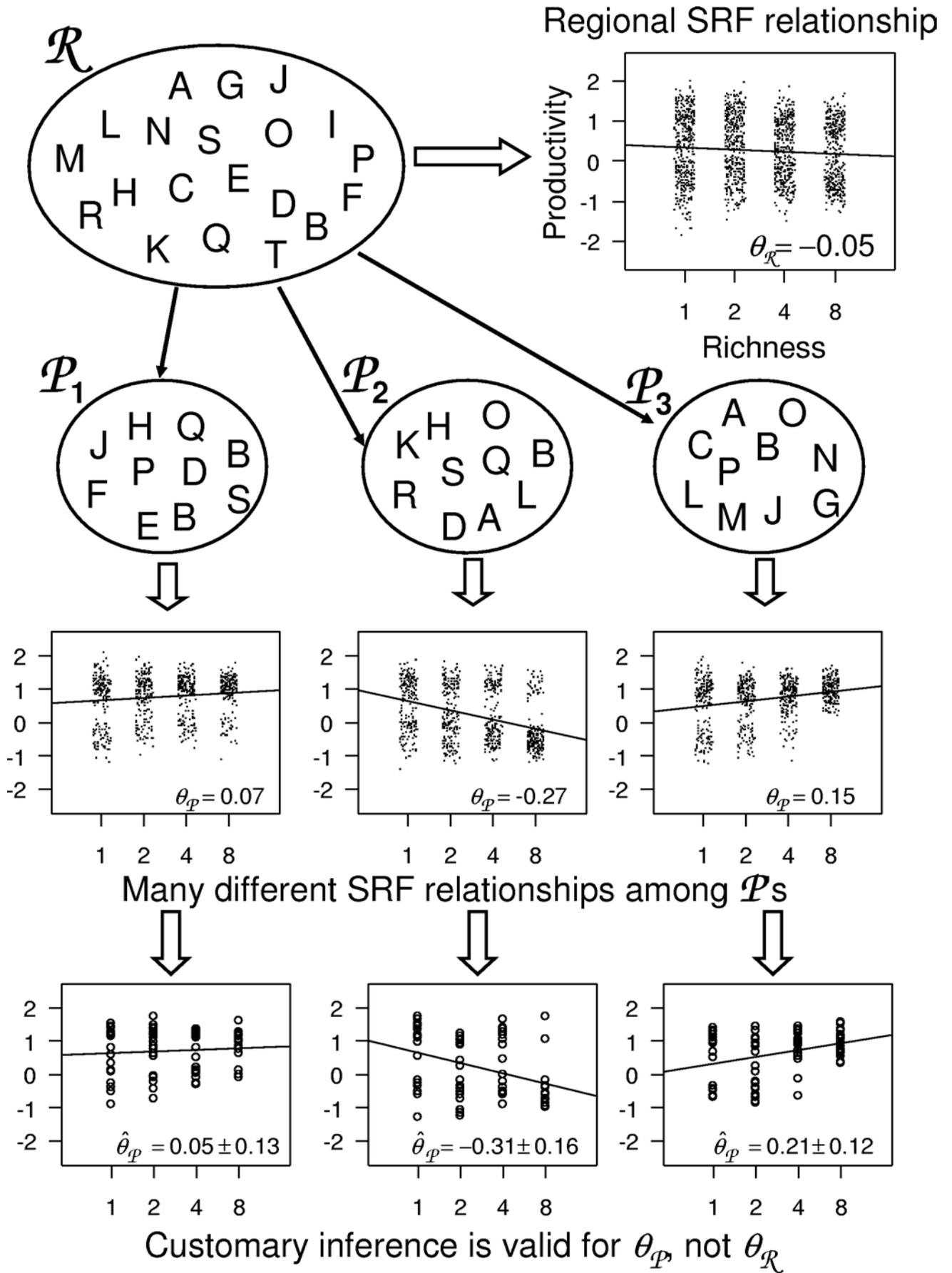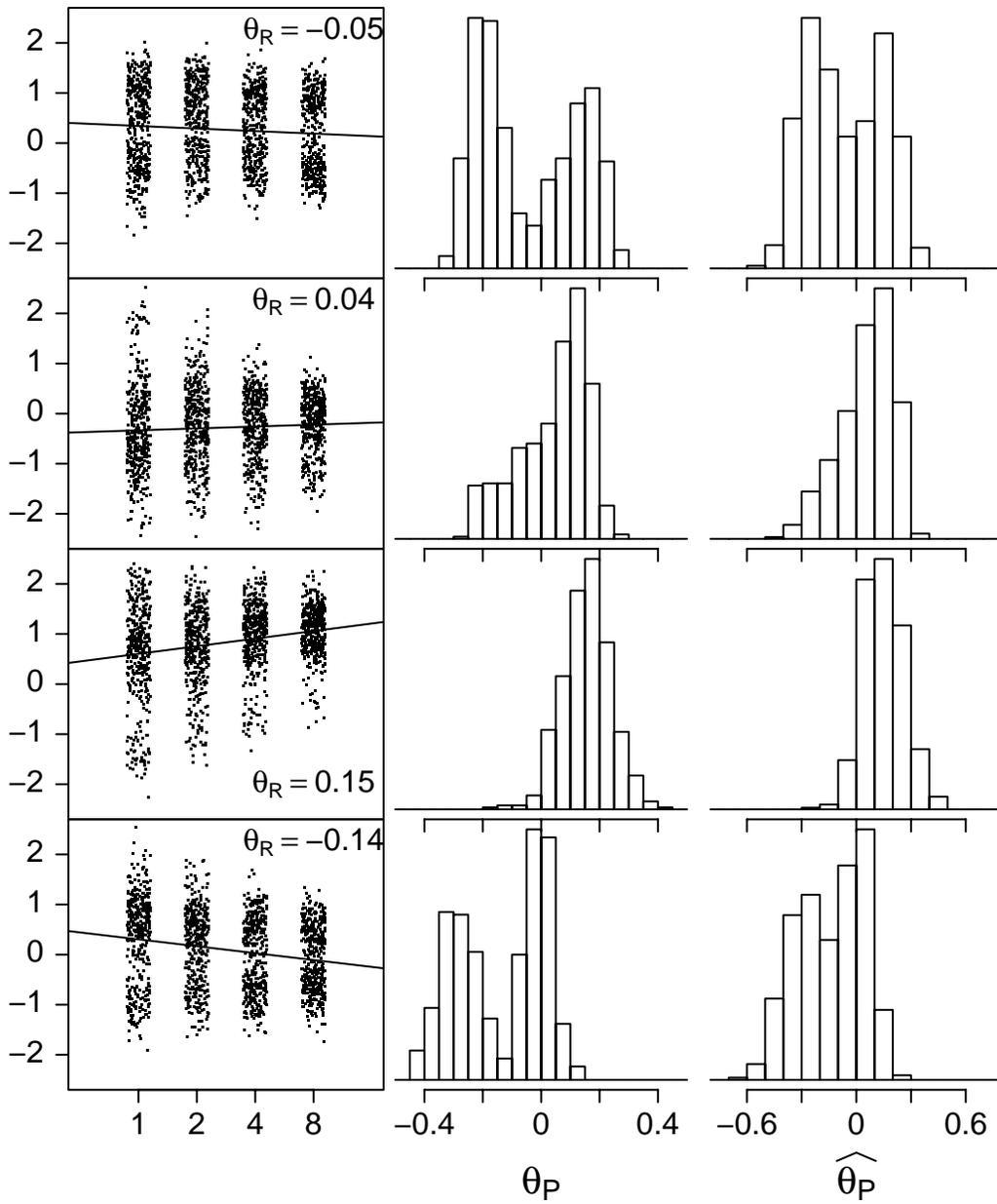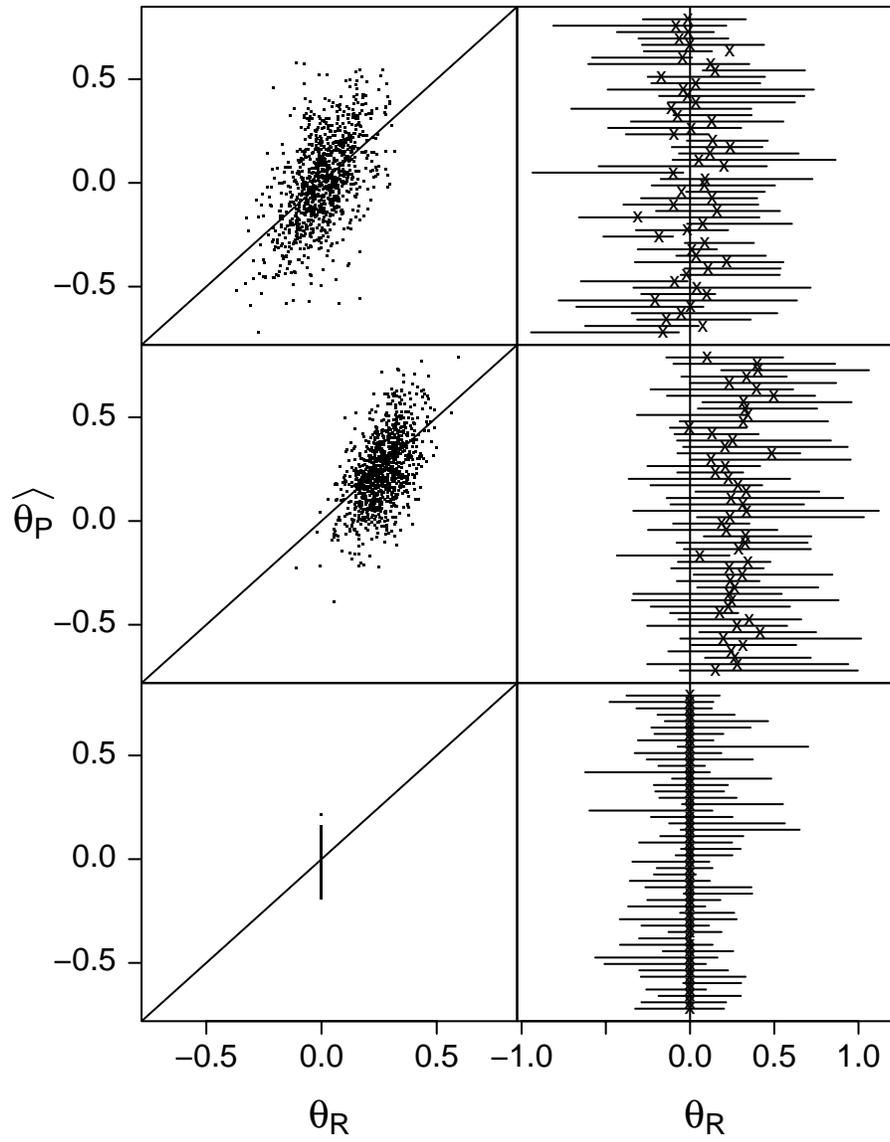
Figure 1

Figure 2

Figure 3