

Appendix 1

Insights from a formal analogy with group selection

In this appendix I highlight ecological insights revealed by the formal analogy between my approach and models of group selection in the evolutionary literature. It is perhaps not obvious that there are any insights to be gained, since the analogy between group selection in evolution and ‘group selection’ in my approach is a purely formal one. However, it is widely-recognized that the Price equation has provided conceptual and empirical insights into problems in many fields outside evolutionary biology, even though the analogy between evolutionary processes and the processes of interest often is only a purely formal one (Loreau and Hector 2001, Andersen 2004, Fox 2005, 2006, Fox and Harpole 2008, Jäger 2008, Collins and Gardner 2009). Indeed, evolution by natural selection itself can be viewed as merely one example of the formal ‘selection process’ described by the Price equation (Price 1972, 1995). More broadly, many empirical and conceptual advances in ecology have arisen from the development of formal mathematical frameworks which classify the effects of complex underlying biological mechanisms. Chesson’s formal classification of coexistence mechanisms into those based on “equalizing” versus “stabilizing” effects is one example (Chesson 2000, Loreau 2010).

The concept of ‘group selection’ has been the subject of several long-standing debates in evolutionary biology, including the proper definition of the term (reviewed by Okasha 2004, 2006). Not all of these debates are relevant to the approach developed here, because the analogy between group selection in evolution and ‘group selection’ in the present context is purely formal. Conceptual debates concerning group selection in evolutionary biology primarily concern the appropriate mathematical description of biological reality. Because ‘groups’ in the present context exist only in a formal sense, debates about the appropriate description of real biological groups are moot.

One aspect of the group selection debate which is relevant in the present context is the distinction between two different classes of multi-level selection models. In evolutionary biology, the partitioning of group and individual selection used here comprises a “multi-level selection 1” model (MLS 1; Damuth and Heisler 1988, Okasha 2006). In MLS 1 models, the properties of groups are “aggregate” properties: sums or averages of the properties of the individuals comprising the groups (Damuth and Heisler 1988, Okasha 2006). For instance, in multi-level selection models of the evolution of altruism, individuals are either “selfish” or “altruistic”, and the “group phenotype” on which group selection acts is the average altruism of the group members (i.e. the proportion of altruists in the group). Note that the properties of individual group members might be determined by all sorts of complex biological factors (e.g. interacting genetic and environmental factors). But once the properties of the individuals comprising the group are

specified, the property of the group is specified simply by summing or averaging the properties of the individuals comprising the group. In contrast, in “multi-level selection 2” models (MLS 2 models), group properties are “emergent”, meaning that they bear some non-additive (and often unspecified) relationship to the properties of the individuals comprising the group. In evolutionary biology, an example of MLS 2 is provided by models of species selection (Jablonski 2008). For instance, geographic range size is a species-level property that may affect the probability of speciation (“giving birth”) and extinction (“dying”), so that geographic range size can evolve via species-level selection. Geographic range size presumably depends in some complex fashion on the phenotypic traits of the individual organisms comprising the species, but it is not a simple sum or average of any individual-level property.

In MLS 1 models, the fact that group-level properties are aggregates of individual-level properties means that individual-level evolution necessarily has group-level consequences, which the multi-level Price equation describes (Okasha 2006). In contrast, in MLS 2 models it is impossible to describe the group-level consequences of individual-level evolution (if indeed there are any at all) without first specifying the complex, non-additive connections between group-level properties and individual-level properties.

In the present context, the distinction between MLS 1 models and MLS 2 models is relevant because some commonly-used measures of community ‘stability’ are aggregate properties, while some are emergent properties. As discussed in the main text, the sum of the elements of the variance-covariance matrix is an aggregate property which has been taken as an index of community stability. However, stability measured as the largest eigenvalue of the community matrix, or return time to equilibrium following a perturbation, is an emergent property (Pimm 1992). Eigenvalues and return times are not simple sums or averages of any properties of the species comprising the community. Species loss will affect these emergent measures of stability, but it will do so via effects that are fundamentally different from those defined by Eq. 4–6. In order to partition the effects of species loss on an emergent measure of stability, it would be necessary to first specify precisely how that measure of stability emerges from the properties of the individual species comprising the community. Different partitionings would be necessary for measures of stability that ‘emerge’ from species-level properties in different ways.

References

- Andersen, E. S. 2004. Population thinking, Price’s equation, and the analysis of economic evolution. – *Evol. Inst. Econ. Rev.* 1: 127–148.
- Chesson, P. 2000. Mechanisms of maintenance of species diversity. – *Annu. Rev. Ecol. Syst.* 31: 343–366.
- Collins, S. and Gardner, A. 2009. Integrating physiological, ecological, and evolutionary change: a Price equation approach. – *Ecol. Lett.* 12: 744–757.
- Damuth, J. and Heisler, I. L. 1988. Alternative formulations of multi-level selection. – *Biol. Philos.* 3: 402–430.

- Fox, J. W. 2005. Interpreting the 'selection effect' of biodiversity on ecosystem function. – *Ecol. Lett.* 8: 846–856.
- Fox, J. W. 2006. Using the Price equation to partition the effects of biodiversity loss on ecosystem function. – *Ecology* 87: 2687–2696.
- Fox, J. W. and W. S. Harpole. 2008. Revealing how species loss affects ecosystem function: the trait-based Price equation partition. – *Ecology* 89: 269–279.
- Jablonski, D. 2008. Species selection: theory and data. – *Annu. Rev. Ecol. Evol. Syst.* 39: 501–524.
- Jäger, G. 2008. Language evolution and George Price's "general theory of selection". – In: Cooper, R. and Kempson, R. (eds), *Language in flux: dialogue coordination, language variation, change and evolution*. College Publications, pp. 53–80.
- Loreau, M. 2010. Linking biodiversity and ecosystems: towards a unifying ecological theory. – *Philos. Trans. R. Soc. B.* 365: 49–60.
- Loreau, M. and Hector, A. 2001. Partitioning selection and complementarity in biodiversity experiments. – *Nature* 412: 72–76.
- Okasha, S. 2006. *Evolution and the levels of selection*. – Oxford Univ. Press.
- Pimm, S. L. 1992. *Food webs*. – Univ. of Chicago Press.
- Price, G. R. 1972. Extension of covariance selection mathematics. – *Ann. Human Genet.* 35: 485–489.
- Price, G. R. 1995. The nature of selection. – *J. Theor. Biol.* 175: 389–396.

Appendix 2

Additional results from the Biodiversity II experiment

Here I show results of using Eq. 4–6 to compare post-loss plots in the Biodiversity II experiment to the most- and least-variable pre-loss plots. In the main text I present results using a typically-variable pre-loss plot. Comparison of the results presented here with those presented in the main text illustrates that the sign of the group- and individual-level SRE on the summed covariance can differ depending on the pre-loss plot to which the post-loss plots are compared. Other results are qualitatively similar to those presented in the main text. Quantitative and qualitative differences in results, depending on the pre-loss plot to which the post-loss plots are compared, reflect variation in species dynamics among pre-loss plots.

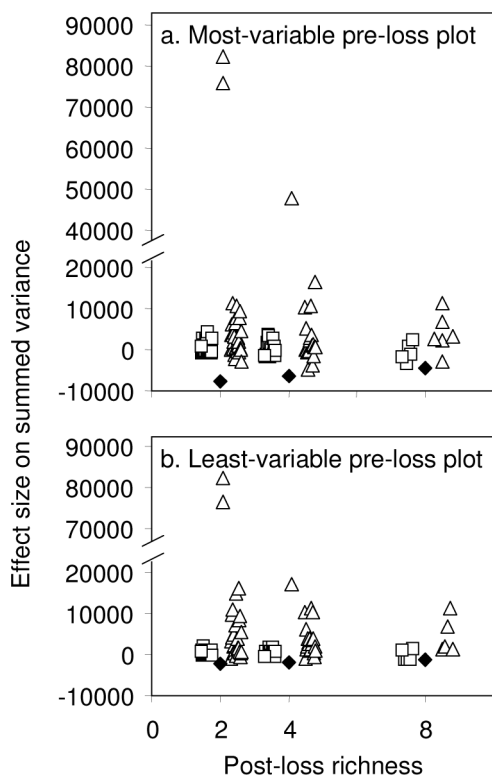


Figure A2.1. Effects of species loss on the summed variance in the Biodiversity II experiment when comparing post-loss plots to either (a) the most-variable or (b) the least-variable pre-loss plot. Points show the SRE (filled diamonds), SCE (open squares), and CDE (open triangles) for each post-loss plot, as a function of post-loss richness. Points are slightly jittered horizontally to improve visibility.

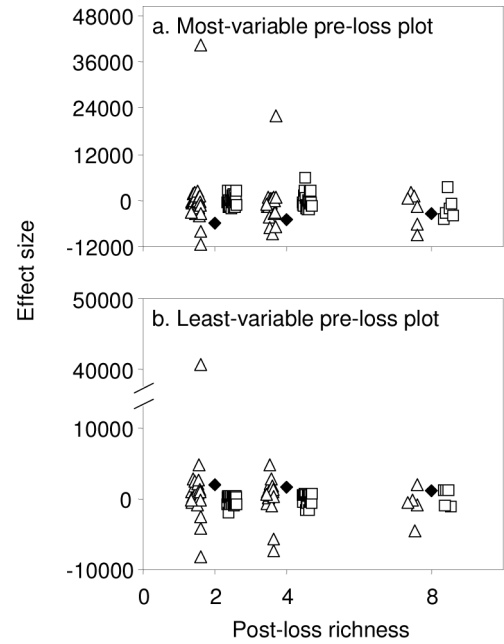


Figure A2.2. Group-level effects of species loss on the summed covariance in the Biodiversity II experiment when comparing post-loss plots to either (a) the most-variable or (b) the least-variable pre-loss plot. Points show the SRE (filled diamonds), SCE (open squares), and CDE (open triangles) for each post-loss plot, as a function of post-loss richness. Points are slightly jittered horizontally to improve visibility.

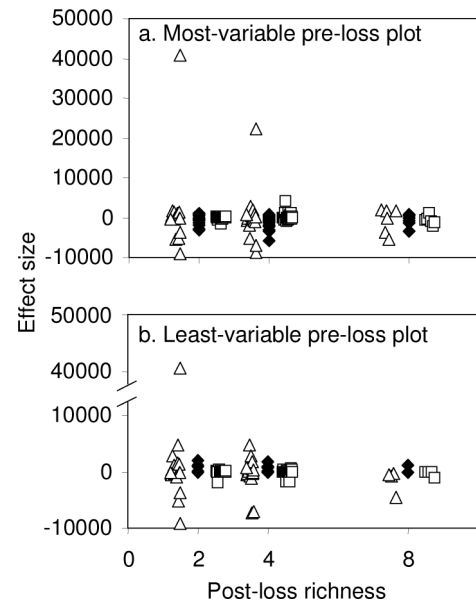


Figure A2.3. Individual-level effects of species loss on the summed covariance in the Biodiversity II experiment when comparing post-loss plots to either (a) the most-variable or (b) the least-variable pre-loss plot. Points show the SRE (filled diamonds), SCE (open squares), and CDE (open triangles) for each post-loss plot, as a function of post-loss richness. Points are slightly jittered horizontally to improve visibility.

Appendix 3

Illustrative application to a theoretical model

Here I provide an illustrative application of my approach to data generated by a theoretical model. This application complements the application to the Biodiversity II data, by showing that my approach can recover the effects of known underlying mechanisms. Loreau and de Mazancourt (2008) presented a stochastic Lotka-Volterra competition model:

$$N_{i,t+1} = N_{i,t} \exp \left[r \left(1 - \frac{N_{i,t} + \alpha \sum_{j \neq i} N_{j,t}}{K'} \right) + \varepsilon_{i,t} + \frac{d_{i,t}}{\sqrt{N_{i,t}}} \right], \quad (\text{A3.1})$$

where $N_{i,t}$ is the abundance of species i at time t ($I = 1, \dots, S$ where S is species richness), r is intrinsic rate of increase (assumed for simplicity to be the same for all species), α is the competition coefficient (assumed the same for all interspecific interactions). K' is the carrying capacity, assumed to be the same for all species, and is given by $K' = K \frac{1 + \alpha(S-1)}{S}$ so that the mean total abundance of

all species is independent of S . The random variable $d_{i,t}$ gives the effect of demographic stochasticity on the per-capita growth of species i during timestep t , and is normally distributed with zero mean and standard deviation σ_d . The random variable $\varepsilon_{i,t}$ gives the effect of environmental stochasticity on the per-capita growth of species i during timestep t , and is normally distributed with zero mean and standard deviation σ_e .

I used this model to simulate the dynamics of a pre-loss community of $S = 10$ species which differed in their response to environmental variation, and so differed in their variances and covariances. Species $i = 1, \dots, 5$ responded identically to environmental fluctuations, so that $\varepsilon_{i,t} = \varepsilon_{1,t}$ for all $I = 1, \dots, 5$. Species $i = 6, \dots, 10$ responded identically to environmental fluctuations that were independent of those experienced by species $i = 1, \dots, 5$. Species 1–5 and 6–10 comprised two ecologically-distinct groups, with group members covarying more positively with one another than with members of the other group.

I then simulated the dynamics of a series of post-loss communities, the first of which lacked species 1, the second of which lacked species 1 and 2, the third of which lacked species 1, 2 and 3, etc. Species loss was thus confined to the species in the first ecological group until all five of these species were lost. The most species-poor community had only two species ($i = 9, 10$). I simulated the dynamics of each community for 1000 timesteps, and discarded the first 500 timesteps in order to eliminate transient dynamics. I calculated the variance-covariance matrix for each community and used Eq. 4–6 to partition the difference in community variability between the pre-loss community and each post-loss community. All species initially present persisted for the entire duration of every simulation. I conducted 10 replicate simulations of this experiment. Parameter values were $r = 1$, $\alpha = 0.5$, $K = 2000$, $\sigma_e = 0.3$, and $\sigma_d = 0.09$.

Figure A3.1a shows that species loss has little effect on community variability until the loss of the last remaining species in the first ecological group. Loss of this species dramatically increases total community variability, which changes little under further species loss. However, application of Eq. 4–6 reveals that this simple pattern in overall community variability actually arises from a more complicated set of effects, some of which cancel one another out.

Figure A3.1a also shows the approximate predicted effect of species loss on community variability derived by Loreau and de Mazancourt (2008; their Eq. B8). The first-order linear approximation derived by Loreau and de Mazancourt (2008) states that the species loss modeled here increases community variability in two ways. First, species loss increases species' carrying capacities (recall that carrying capacities vary with S so that total community size is independent of S). Second, species loss increases the average pairwise correlation between species' responses to environmental variation (up to a maximum), because as species are lost an increasing proportion of the remaining species are in the second ecological group and so respond identically to environmental variation. The approximation derived by Loreau and de Mazancourt (2008) provides an alternative, complementary way to partition the effects of species loss on community variability. Because the approach developed here makes no system-specific assumptions, it does not attempt to separate effects mediated by changes in carrying capacity from effects mediated by changes in the correlation of species' environmental responses. Conversely, the approximation derived by Loreau and de Mazancourt (2008) does not attempt to separate the total effect of loss of species richness per se from effects of changes in species composition or post-loss changes in species' variances and covariances. The two approaches thus provide different, complementary ways to partition the same total effect, which may be useful for different purposes. The approximation derived by Loreau and de Mazancourt (2008) provides mechanistic insight into the behavior of their model, but does not provide insight into the behavior of different models incorporating different ecological mechanisms (e.g. models in which competition is not described by the Lotka–Volterra model). The approach developed here is not itself mechanistic (although applying the approach often suggests mechanistic insights), but allows comparison of the effects of different mechanisms within a common framework. Further, the approximation derived by Loreau and de Mazancourt (2008) is not equally accurate at all levels of species loss simulated here (Fig. A3.1a). The exact partitioning provided by my approach can reveal subtle effects of species loss not captured by the approximation.

Figure A3.1b shows that the summed variance declines linearly with increasing species loss due to an SRE, as it must. However, this trend is partially counterbalanced (and at high levels of species loss, more than counterbalanced) by the CDE. Loss of species from the first ecological group slightly increases the variances of the remaining species, in part because the mean abundances of the remaining species increase, leading to variance-mean scaling (an effect which could be quantified using the approach described in Appendix 4). However, once all species from the first ecological group are lost, further species loss dramatically increases the variances of the remaining species. The results in Fig. A3.1b illustrate that species' variances depend on interspecific interactions, just as do species' covariances.

Figure A3.1c shows the group-level effects of species loss on species' covariances. Total covariance decreases linearly with increasing species loss due to an SRE, as it must because species covary positively on average. However, this effect is more than counterbalanced by a CDE. Loss of species from the first ecological group increases the summed covariances of the remaining species, with the biggest increase associated with loss of the last species from the first ecological group. Further species loss actually decreases the summed covariances of the remaining species.

To fully understand context-dependence of species' covariances, we need to partition the group-level CDE into its individual-level components (Fig. A3.1d). The individual-level SRE is negative because, on average, the persisting species have positive pre-loss pairwise covariances with other species. Random species

loss therefore reduces the summed covariances of the persisting species, all else being equal. This effect varies in absolute magnitude depending on how many species are lost because it comprises a sum across different species having different pre-loss summed covariances. The individual-level SRE is small, and is swamped by the other two effects. At an individual level, species were lost non-randomly with respect to their pairwise covariances with the remaining species: loss of species from the first ecological group represents non-random loss of species that covary negatively with the majority of the remaining species (those in the other ecological group). This individual-level SCE increases the summed covariances of the remaining species. After all the species in the first ecological group are lost, further species losses come from the second group, making the cumulative species loss less non-random and reducing the individual-level SCE. Finally, the pairwise covariances between the remaining species increase as more species are lost, with the largest increases associated with loss of the last species in the first ecological group and the first species in the second group. Species loss generates this individual-level CDE in part by increasing species' mean abundances, which reduces the importance of desynchronizing demographic stochasticity (Loreau and de Mazancourt 2008).

References

Loreau, M. and de Mazancourt, C. 2008. Species synchrony and its drivers: neutral and nonneutral community dynamics in fluctuating environments. – *Am. Nat.* 172: E48–E66.

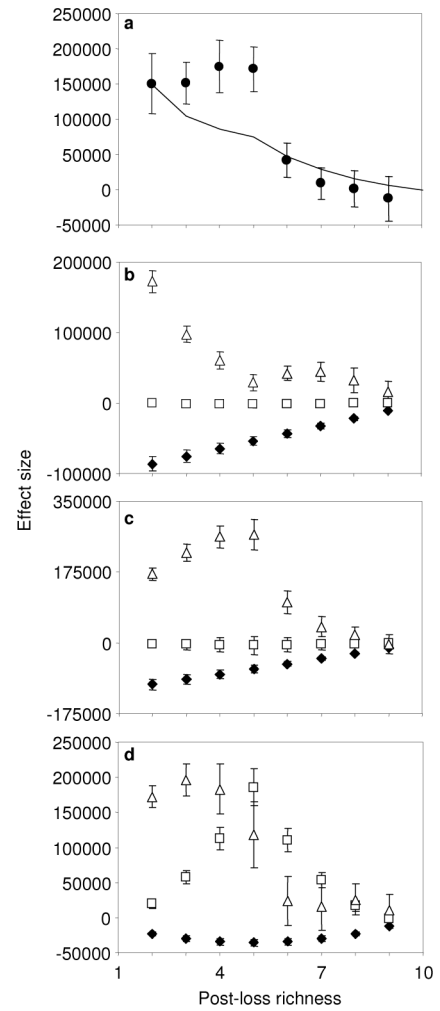


Figure A3.1. Illustrative application of Eq. 4–6 to simulated data from Eq. A3.1. (a) Observed difference in total community variability (points) and the predicted difference from the first-order approximation derived by Loreau and de Mazancourt (2008). (b) Effects on the summed variance. (c) Group-level effects on the summed covariance. (d) Individual-level effects on the summed covariance. Effect sizes are means of 10 replicate simulated experiments, \pm SD. Some error bars are hidden by plot symbols.

Appendix 4

Incorporating mean-variance scaling

Here I demonstrate two ways to incorporate species' mean biomasses into the approach developed in the main text. One way is to analyze a scaled measure of variability such as the variance-mean ratio:

$$\frac{V}{B} = \frac{\sum_{i=1}^s \text{var}(B_i) + \sum_{i=1}^s \sum_{j \neq i} \text{cov}(B_i, B_j)}{\sum_{i=1}^s \bar{B}_i} \quad (\text{A4.1})$$

where \bar{B}_i is the temporal mean biomass of species i . In practice, many workers take the square root of the numerator of Eq. A4.1 before dividing by total biomass, thereby obtaining the coefficient of variation (CV; Lehman and Tilman 2000, Petchey et al. 2002). However, this nonlinear transformation of the numerator is made for purposes of obtaining a dimensionless, scale-invariant ratio, and is inessential for present purposes. An analogous equation can be written for the ratio of post-loss variance to post-loss total biomass:

$$\frac{V'}{B'} = \frac{\sum_{i=1}^s w_i \text{var}(B'_i) + \sum_{i=1}^s w_i \sum_{j \neq i} w_j \text{cov}(B'_i, B'_j)}{\sum_{i=1}^s w_i \bar{B}'_i} \quad (\text{A4.2})$$

Lehman and Tilman (2000) suggested that the effects of diversity on stability can be understood by considering the effects of diversity on each of the three components of Eq. A4.1: the summed variance $\sum_{i=1}^s \text{var}(B_i)$, the summed covariance $\sum_{i=1}^s \sum_{j \neq i} \text{cov}(B_i, B_j)$, and the total biomass $\sum_{i=1}^s \bar{B}_i$. At a broad level, this is clearly true: if species loss (or gain) is to affect the variance-mean ratio (or the CV), it must affect one or more of the three components of Eq. A4.1. Several empirical studies interpret experimental data on species loss and variability by considering the separate effects of species loss on each of the three components of Eq. A4.1 (Petchey et al. 2002). Further, the Price equation partition of Fox (2006) can be used to describe the effects of species loss on total biomass, thereby partitioning the difference between the denominators of Eq. A4.1 and Eq. A4.2. This partitioning of the difference in denominators could be used together with the approach developed in the main text to interpret the effects of species loss on all three components of Eq. A4.1.

However, the suggestion of Lehman and Tilman (2000) has a technical limitation. When community variability is measured by the variance-mean ratio, the difference in variability between a post-loss and a pre-loss community, $\Delta \frac{V}{B}$, is given by Eq. A4.2 minus Eq. A4.1:

$$\Delta \frac{V}{B} = \frac{V'}{B'} - \frac{V}{B}. \quad (\text{A4.3})$$

However, the expression for the separate effects of species loss on each of the three components of Eq. A4.1 is

$$\frac{(\sum \text{var}' - \sum \text{var}) + (\sum \text{cov}' - \sum \text{cov})}{(B' - B)} = \frac{V' - V}{B' - B} \quad (\text{A4.4})$$

where $\sum \text{var}$ and $\sum \text{var}'$ respectively are the pre- and post-loss summed variances and $\sum \text{cov}$ and $\sum \text{cov}'$ respectively are the pre- and post-loss summed covariances. Equations A4.3 and A4.4 are

not equal. In general, the difference of two ratios does not equal the ratio of the difference in numerators to the difference in denominators. So while the Price equation partition can be applied to both the numerator and denominator of Eq. A4.4, the resulting expression is not a partition of the difference between pre- and post-loss variability, which is given by Eq. A4.3. More broadly, the difference between pre- and post-loss variability, Eq. A4.3, cannot be completely understood by considering the difference between the pre- and post-loss values of each of its components Eq. A4.4. Nor is it possible to partition the between-site difference in variance-mean ratio, Eq. A4.3, into interpretable subcomponents via the Price equation partition, because the variance-mean ratio of a community does not equal the sum of the variance-mean ratios of the species comprising it. The same points hold true if community variability is measured in terms of CV rather than the variance-mean ratio.

If the variance-mean ratio (or the CV) is the investigator's preferred measure of community variability, then substantial but incomplete insight into the between-site difference in variability can be gained by separately partitioning the effects of species loss on the numerator and the denominator. Alternatively, instead of scaling community variability by total biomass, the effects of species' biomasses on species' variances can be quantified via mean-variance scaling.

Incorporating mean-variance scaling

Mean-variance scaling refers to the observation that, for species' biomasses (and many other random variables), the temporal variance often is related to the mean via a power law:

$$\text{var}(B_i) = a \bar{B}_i^b + \varepsilon_i \quad (\text{A4.5})$$

where \bar{B}_i is the mean biomass of species i , $\text{var}(B_i)$ is the temporal variance of species i , ε_i is residual error, and a and b are parameters, the values of which depend on the nature of the processes controlling the mean and the variance. For instance, for a Poisson process the variance equals the mean ($a = 0$, $b = 1$). For ecological populations, theoretical arguments and most empirical evidence indicate that $1 \leq b \leq 2$ (Hanski 1982, Taylor and Woiwood 1982, Murdoch and Stewart-Oaten 1989, Tilman 1999, Lepš 2004, Steiner 2005, Steiner et al. 2005, Romanuk et al. 2006, Vogt et al. 2006).

Mean-variance scaling will affect the diversity-variability relationship if species' mean biomasses vary with species richness. In general, mean biomasses should decline with increasing species richness due to interspecific competition, causing species' temporal variances to decline via mean-variance scaling (Doak et al. 1998, Tilman et al. 1998). This can result in a reduction in the summed variance (Doak et al. 1998, Tilman et al. 1998). However, species' variances also can vary with species loss for reasons independent of mean-variance scaling (Petchey et al. 2002), necessitating a way to isolate the role of mean-variance scaling from other factors affecting species' variances.

The approach of Fox and Harpole (2008) can be used to isolate the effect of mean-variance scaling, and provide a more refined picture of how species loss affects community variability via mean-variance scaling. Fox and Harpole (2008) use linear regression to describe how the contributions of individual species to pre- or post-loss ecosystem function depend on their underlying 'traits'. A trait is simply any measurable property of a species that predicts its functional contribution. This approach is formally analogous to Fisher's (1958) regression approach describing how the phenotypes of individual organisms depend on their underlying genotypes. While much insight into evolution can be gained without

knowledge of the underlying genetics of individual phenotypes, additional insight can be gained by explicitly specifying the relationship between genotypes and phenotypes. This allows evolution by natural selection to be re-described in terms of selection operating on the underlying genotypes rather than on phenotypes, and transmission bias operating at the level of genotypes rather than phenotypes (Fisher 1958).

To apply the approach of Fox and Harpole (2008) to mean-variance scaling, we first log-transform Eq. A4.5 in order to linearize the mean-variance relationship:

$$\log[\text{var}(B_i)] = \log(a) + b \log(\bar{B}_i) + e_i \quad (\text{A4.6})$$

where e_i is residual error on the log-transformed scale. Equation A4.6 is a linear regression of species' phenotypes (their log-transformed variances) on their "genotypes" (their log-transformed mean biomasses). Equation A4.6 applies to the pre-loss site; the corresponding equation for the post-loss site is

$$\log[\text{var}(B'_i)] = w_i \log(a') + w_i b' \log(\bar{B}'_i) + w_i e'_i \quad (\text{A4.7})$$

where primes denote attributes of the post-loss site. Fox and Harpole (2008) show that the difference between the summed log-transformed variances can be written as

$$\begin{aligned} & \sum_{i=1}^s w_i \log[\text{var}(B'_i)] - \sum_{i=1}^s \log[\text{var}(B_i)] = \overline{\log[\text{var}(B)]} \Delta s \\ & + \left[\sum_{i=1}^s b \log(\bar{B}_i) (w_i - \bar{w}) + \sum_{i=1}^s e_i (w_i - \bar{w}) \right] \\ & + \left[\sum_{i=1}^s w_i b \Delta \log(\bar{B}_i) + \left(s' \Delta a + \sum_{i=1}^s w_i \Delta b \log(\bar{B}_i) \right) + \sum_{i=1}^s w_i \Delta b \Delta \log(\bar{B}_i) + \sum_{i=1}^s w_i \Delta e_i \right] \end{aligned} \quad (\text{A4.8})$$

where

$$\Delta \log(\bar{B}_i) = \log(\bar{B}'_i) - \log(\bar{B}_i), \quad \Delta a = a' - a, \quad \Delta b = b' - b, \quad \text{and} \quad \Delta e_i = e'_i - e_i.$$

Equation A4.8 partitions the difference between the pre- and post-loss summed log-transformed variances into three additive components. The first component of Eq. A4.8, $\overline{\log[\text{var}(B)]} \Delta s$, is the SRE, isolating the effect of random loss of species richness on the summed log-transformed variances. The second component of Eq. A4.8,

$$\sum_{i=1}^s b \log(\bar{B}_i) (w_i - \bar{w}) + \sum_{i=1}^s e_i (w_i - \bar{w}),$$

is the SCE, isolating the effect of non-random loss of species. The third component of Eq. A4.8,

$$\sum_{i=1}^s w_i b \Delta \log(\bar{B}_i) + \left(s' \Delta a + \sum_{i=1}^s w_i \Delta b \log(\bar{B}_i) \right) + \sum_{i=1}^s w_i \Delta b \Delta \log(\bar{B}_i) + \sum_{i=1}^s w_i \Delta e_i,$$

is the CDE, isolating the effect of post-loss changes in the properties of the remaining species.

The novel feature of Eq. A4.8 is that it partitions the SCE and CDE into subcomponents that identify the sources of mean-variance scaling and separate the effects of mean-variance scaling from other factors. The SCE in Eq. A4.8 has two additive subcomponents. The first, $\sum_{i=1}^s b \log(\bar{B}_i) (w_i - \bar{w})$, gives the effect on the summed log-transformed variance of non-random loss of species with respect to their pre-loss log-transformed biomasses. For instance, non-random loss of high-biomass species will reduce the mean biomass of the remaining species, all else being equal. This

reduction in mean biomass per species will reduce the mean variance of the remaining species, thereby reducing their total variance via mean-variance scaling. The second subcomponent of the SCE, $\sum_{i=1}^s e_i (w_i - \bar{w})$, is the residual SCE attributable to non-random loss of species with respect to other, unidentified traits that affect species' log-transformed variances. The residual SCE will be small in absolute magnitude if interspecific variation in log-transformed biomass explains most of the interspecific variation in log-transformed variance at the pre-loss site.

The CDE in Eq. A4.8 has four additive subcomponents. The first, $\sum_{i=1}^s w_i b \Delta \log(\bar{B}_i)$, is that part of the CDE attributable to between-site variation in the log-transformed biomasses of the persisting species. For instance, if the persisting species all exhibit lower biomasses at the post-loss site than at the pre-loss site, this term will be negative, because these reduced biomasses will lead to reduced variances via mean-variance scaling. Context dependence of species' biomasses is a second source of mean-variance scaling, distinct from non-random loss of species with respect to

their pre-loss biomasses. The second subcomponent of the CDE, $s' \Delta a + \sum_{i=1}^s w_i \Delta b \log(\bar{B}_i)$, is that part of the CDE attributable to between-site differences in the slope and intercept of the mean-variance scaling relationship. For instance, imagine that the species at the post-loss site exhibit log-transformed variances that increase very steeply with their log-transformed biomasses, while the species at the pre-loss site do not. In this case, Δb will be positive, which will increase the summed log-transformed variance at the post-loss site. Total post-loss variance is increased because the persisting species exhibit higher post-loss variances than they would have had the mean-variance scaling relationship retained its pre-loss properties. This subcomponent is not an effect of mean-variance scaling per se, because this subcomponent does not arise from changes in species' mean biomasses. Empirical studies report a range of slopes and intercepts for the mean-variance scaling relationship (Hanski 1982, Taylor and Woiwood 1982, Tilman 1999, Petchey et al. 2002, Steiner 2005, Steiner et al. 2005, Romanuk et al. 2006),

but the drivers of this variation are largely unknown. The third subcomponent of the CDE, $\sum_{i=1}^s w_i \Delta b \Delta \log(\bar{B}_i)$, is an interaction term between the effects of post-loss changes in species' log-transformed mean biomasses, and post-loss change in the slope of the mean-variance scaling relationship. This term is a third source of mean-variance scaling, but one which is non-zero only when the slope of the mean-variance scaling relationship differs between sites (so that $\Delta b \neq 0$). This term is an interaction term between mean-variance scaling itself and any factor affecting the slope of the mean-variance scaling relationship. The fourth subcomponent of the CDE, $\sum_{i=1}^s w_i \Delta e_i$, is the residual context dependence in species' log-transformed variances attributable to unidentified factors other than post-loss changes in species' biomasses and the properties of the mean-variance scaling relationship.

Equation A4.8 has several novel advantages as a framework in which to analyze the effects of mean-variance scaling on community stability. First, it isolates that portion of the between-site difference in the summed log-transformed variances attributable to mean-variance scaling per se from between-site differences in the summed log-transformed variances arising from other factors. Second, it identifies a novel interaction between mean-variance scaling and factors affecting the slope of the scaling relationship.

Previous discussions of mean-variance scaling have not recognized the interaction term (Doak et al. 1998, Tilman et al. 1998, Tilman 1999, Petchey et al. 2002, Valone and Hoffman 2003, Lepš 2004, Steiner 2005, Steiner et al. 2005, Romanuk et al. 2006, Vogt et al. 2006). Third, Eq. A4.8 separates two distinct sources of mean-variance scaling: non-random loss of high- or low-biomass species, and post-loss changes in the biomasses of the remaining species. Many causes of species loss will tend to eliminate rare, low-variance species. However, species loss also can lead to increases in the biomasses, and thus the variances, of the remaining species (e.g. if species compete). Equation A4.8 separates these two countervailing effects on the summed variance, both of which operate via mean-variance scaling.

References

- Doak, D. F. et al. 1998. The statistical inevitability of stability-diversity relationships in community ecology. – *Am. Nat.* 151: 264–276.
- Fisher, R. A. 1958. *The genetical theory of natural selection* (2nd ed.). – Dover.
- Hanski, I. 1982. On patterns of temporal and spatial variation in animal populations. – *Ann. Zool. Fenn.* 19: 21–37.
- Lepš, J. 2004. Variability in population and community biomass in a grassland community affected by environmental productivity and diversity. – *Oikos* 107: 64–71.
- Romanuk, T. N. et al. Nutrient enrichment weakens the stabilizing effect of species richness. – *Oikos* 114: 291–302.
- Steiner, C. F. 2005. Temporal stability of pond zooplankton assemblages. – *Freshwater Biol.* 50: 105–112.
- Steiner, C. F. et al. 2005. Temporal stability of aquatic food webs: partitioning the effects of species diversity, species composition and enrichment. – *Ecol. Lett.* 8: 819–828.
- Taylor, L. R. and Woiwood, I. P. 1982. Comparative synoptic dynamics. I. Relationships between inter- and intraspecific spatial and temporal variance/mean population parameters. – *J. Anim. Ecol.* 51: 879–906.
- Tilman, D. 1999. The ecological consequences of changes in biodiversity: a search for general principles. – *Ecology* 80: 1455–1474.
- Tilman, D. et al. 1998. Diversity–stability relationships: statistical inevitability or ecological consequence? – *Am. Nat.* 151: 277–282.
- Valone, T. J. and Hoffman, C. D. 2003. Population stability is higher in more diverse annual plant communities. – *Ecol. Lett.* 6: 90–95.
- Vogt, R. J. et al. 2006. Species richness–variability relationships in multi-trophic aquatic microcosms. – *Oikos* 113: 55–66.

Appendix 5

Assumptions and interpretation of the Price equation

Here I discuss the assumptions and interpretation of the Price equation and its extension by Fox (2006). I first discuss issues pertaining to both the original Price equation and its extension by Fox (2006), and then issues specific to the extension by Fox (2006). For further discussion see Rice and Papadopoulos (2009), who emphasize that the assumptions of the Price equation are axiomatic, in contrast to the simplifying assumptions which comprise the starting point for conventional theoretical models.

The Price equation, and its extension by Fox (2006), is an additive partition. It is important to recognize that the additivity of the SRE, SCE, and CDE is not an assumption of the Price equation partition. Rather, the additivity of the partitioned effects emerges naturally from the definition of the problem. The problem is to partition the difference between two populations of entities in some aggregate property of those entities (Price 1972, Frank 1997, Okasha 2006). Additivity is assumed in the definition of the property of interest: an aggregate property is an average or sum of the properties of the individual entities comprising the population. Additivity also is assumed in the definition of the between-population comparison: it is the difference between populations that is considered, not the ratio or the product or etc.

The assumptions of additivity involved in defining the problem determine the domain of applicability of the Price equation. The approach developed here applies to community variability because the variance of a sum of random variables (the population-level property of interest) is an aggregate property by definition: it is the sum of the elements of the variance-covariance matrix. If the variance of a sum of random variables was defined as, e.g. the determinant of the variance-covariance matrix, it would be an emergent property and the Price equation would not apply (see Appendix 1 for more on aggregate vs emergent properties). By identifying important distinctions between different problems, such as the effects of species loss on aggregate versus emergent properties, the Price equation prevents the investigator from comparing “apples to oranges” (Fox and Harpole 2008).

The Price equation’s assumptions of additivity are not simplifying assumptions, and do not limit its generality within its domain of applicability (Rice and Papadopoulos 2009). The approach developed here makes no assumptions about the nature of the underlying ecological mechanisms determining species’ variances and covariances, or about the mechanisms causing species loss. The approach developed here therefore is consistent with the operation of any complex, nonadditive, nonlinear processes affecting species’ variances and covariances. Analogously, in evolutionary biology concepts like ‘selection’ and ‘heritability’ are defined without making any assumptions about the complex biological mechanisms that determine their values. This is what gives the original Price equation its generality (Frank 1997). For instance, the original Price equation applies to both natural and artificial selection. The distinction between natural and artificial selection is a distinction between different underlying mechanisms that generate selection, and the Price equation makes no assumptions about these mechanisms. Analogously, the approach developed here describes the effects of both natural species loss, and artificial species loss such as in the Biodiversity II experiment.

Specific ecological assumptions are only required to incorporate mean-variance scaling via the method of Fox and Harpole (2008). It is assumed that the true mean-variance scaling relation-

ship is reasonably well-described by Eq. A4.5. If it is not, Eq. A4.8 remains mathematically valid, but provides little insight because the residual SCE and CDE will be large in magnitude relative to the other terms in Eq. A4.8.

The Price equation is an exact partition. There is no possibility that any effects are omitted from Table 1, not even residual error. This allows the Price equation to solve the conceptual problem of how species loss affects community variability in a more complete and convincing way than would be possible by attempting to generalize from specific ecological models making specific ecological assumptions. The Price equation is exact because it is retrospective. The Price equation assumes that the observed properties of entities are known without error, which can never be the case for future (and therefore unobserved) values of these properties. It does not matter if the observed properties of entities (here, species’ variances and covariances) were affected by stochastic mechanisms, since the Price equation considers only the known, observed outcomes of past stochastic (and deterministic) processes. Recently, Rice (2009) derived a prospective equivalent of the Price equation which incorporates uncertainty arising from the fact that future property values have not yet been observed and so must be characterized by probability distributions.

In reality, species’ true variances and covariances are rarely known without error even in retrospect, because of sampling error. However, sampling error is a feature of the data input into the Price equation, not of the Price equation itself. The Price equation can be interpreted as averaging over sampling error and describing the expected change in the property of interest (Gardner and Grafen 2009).

The Price equation does not assume equilibrium. The values of its terms will depend on when the system is sampled if the system state varies over time. This is not a drawback of the approach, but instead reflects real temporal variation in the strength of different effects. For instance, community variability likely will change over time as the community relaxes to a new dynamical state following species loss.

Partitionings are not unique. The difference between two populations in some aggregate property can be partitioned in an infinity of ways. Many possible partitionings will be uninterpretable and so provide no insight. But there may be no single ‘best’ partitioning because different partitionings provide different, complementary insights. In several previous applications of the Price equation, alternative partitionings are the subject of ongoing discussion (Fox 2005, 2006, Okasha 2006, Collins and Gardner 2009). The partitioning of group-level and individual-level effects used here is widely used in evolutionary biology, which provides a motivation for choosing this partitioning (Okasha 2006).

The original Price equation partitions the difference between two populations in the mean of some property of the entities comprising the populations (Price 1972). Fox (2006) extended the original Price equation to partition the between-population difference in the sum, rather than the mean. By definition, the sum of the property values of the entities comprising a population always equals the mean property value multiplied by the total number of entities (e.g. in a population of s entities, where entity i

has property value z_i , $\sum_{i=1}^s z_i = s \frac{\sum_{i=1}^s z_i}{s} = s\bar{z}$). Populations that differ

in their summed property value necessarily differ in population size (number of entities), mean property value, or both. The portion of the between-population difference in the sum which is attributable to the difference in population size is the SRE (Fox 2006). It is important to recognize that this SRE is not an effect expected under some possibly-false null hypothesis such as that species do not interact. Rather, it simply isolates the effect on a

sum of removing some of the summands, without changing the mean value of those summands. The SRE could be regarded as a statistical fiction rather than a “real” biological effect, because it necessarily takes on a non-zero value whenever species are lost (unless the mean property value equals zero). If the SRE is viewed as a statistical fiction, it can be added to the SCE, with the sum of the two giving the total ‘direct effect’ of species loss on the sum of the property values (Fox 2006). However, viewing the SRE as a statistical fiction is tantamount to ignoring the distinction between effects of species loss that depend on the identity of the lost species vs. effects that don’t. The SRE isolates that portion of the total effect of species loss that would have occurred no matter which species were lost, and so cannot be uniquely attributed to the particular species that were lost (Fox 2006). Many experimental studies use statistical approaches such as ANOVA to separate effects of species richness from effects of species identity (Tilman et al. 2001). The Price equation partition is in the same spirit, but with the difference that in the Price equation partition the effects of species richness and species composition depend only on the biological properties of the species themselves, not on details such as the choice of experimental design or statistical model.

References

- Collins, S. and A. Gardner. 2009. Integrating physiological, ecological and evolutionary change: a Price equation approach. – *Ecol. Lett.* 12: 744–757.
- Fox, J. W. 2006. Using the Price equation to partition the effects of biodiversity loss on ecosystem function. – *Ecology* 87: 2687–2696.
- Fox, J. W. and W. S. Harpole. 2008. Revealing how species loss affects ecosystem function: the trait-based Price equation partition. – *Ecology* 89: 269–279.
- Frank, S. A. 1997. The Price equation, Fisher’s fundamental theorem, kin selection and causal analysis. – *Evolution* 51: 1712–1729.
- Gardner, A. and A. Grafen. 2009. Capturing the superorganism: a formal theory of group adaptation. – *J. Evol. Biol.* 22: 659–671.
- Okasha, S. 2006. *Evolution and the levels of selection*. – Oxford Univ. Press.
- Price, G. R. 1972. Extension of covariance selection mathematics. – *Ann. Human Genet.* 35: 485–489.
- Rice, S. H. 2009. A stochastic version of the Price equation reveals the interplay of deterministic and stochastic processes in evolution. – *BMC Evol. Biol.* 8: 262–277.
- Rice, S. H. and A. Papadopoulos. 2009. Evolution with stochastic fitness and stochastic migration. – *PLoS One* 4:e7130. doi:10.1371/journal.pone.0007130
- Tilman, D. et al. 2001. Diversity and productivity in a long-term grassland experiment. – *Science* 294: 843–845.