

1 **An Empiricist's Guide to Modern Coexistence Theory for Competitive Communities**

2 **List of authors:** Casey M. Godwin^{1,2*}, Feng-Hsun Chang¹, Bradley Cardinale^{1,2}

3

4 ¹School for Environment and Sustainability, University of Michigan, 440 Church Street, Ann

5 Arbor, Michigan, USA

6 ²Cooperative Institute for Great Lakes Research, University of Michigan, 440 Church Street,

7 Ann Arbor, Michigan, USA

8

9 *Correspondence: cgodwin@umich.edu

10

11 Key-Words: competition experiments, modern coexistence theory, niche differences, relative

12 fitness differences

13 Running Title: Empiricist's guide to modern coexistence theory

14 Abstract

15 While most ecological theories have historically invoked niche differences as the primary
16 mechanism allowing species coexistence, we now know that species coexistence in competitive
17 communities actually depends on the balance of two opposing forces: niche differences (ND)
18 that determine how species limit their own growth rate versus that of their competitor, and
19 relative fitness differences (RFD) that establish competitive hierarchies among species. Several
20 different empirical methods have been proposed for measuring ND and RFD in order to make
21 predictions about coexistence of species, yet it remains unclear which method(s) are appropriate
22 for a given empirical study and whether or not those methods actually yield the same
23 information. Here we summarize and compare five different empirical methods, with the aim of
24 providing a practical guide for empiricists who want to predict coexistence among species. These
25 include two phenomenological methods that estimate ND and RFD based on observing
26 competitive interactions among species; two mechanistic methods that estimate ND and RFD
27 based solely on information about species' resource requirements; and a fifth method that does
28 not yield ND and RFD but describes the impacts of those forces within communities. Based on
29 the specific requirements, limitations, and assumptions of each approach, we offer a series of
30 decision steps that can be used to determine which method(s) are best for a given study system.
31 In particular, we show there are important tradeoffs between mechanistic methods, which require
32 detailed understanding of species niches and physiology but are more tractable experimentally,
33 and phenomenological methods which do not require this detailed information but can be
34 impractical for some study designs. Importantly, we show that although each method can be used
35 to estimate ND and RFD, the methods do not always yield the same values. Therefore we caution
36 against future syntheses that compile these estimates from different empirical studies. Finally, we

37 highlight several areas where modern coexistence theory could benefit from additional empirical

38 work.

39

40

41 **Introduction**

42 Throughout most of the history of community ecology, it has been assumed that niche
43 differentiation among species is primary biological mechanism that can offset the negative
44 impacts of interspecific competition on species coexistence (Volterra 1931, Gause 1934, Tilman
45 1982, Leibold 1995, Chase and Leibold 2003). This idea originated when Volterra (Volterra
46 1931) introduced a dynamic model of competition that became the foundation for the
47 competitive exclusion principle, which states that if two species have identical niche
48 requirements then one of them will inevitably become locally extinct (Gause 1934). The
49 competitive exclusion principle led to two conclusions about coexistence in competitive
50 communities: 1) species will coexist only if they are limited by different resources (or
51 consumers) at the same location and time, or if they partition resources (or consumers) in space
52 or time and, as a result, 2) ecosystems should contain only as many species as there are limiting
53 resources (or consumers) (Rescigno and Richardson 1965, MacArthur and Levins 1967,
54 MacArthur 1970, Tilman 1977, Leibold 1995). Nearly all subsequent hypotheses to explain
55 coexistence have argued that biodiversity exists because of such niche differences among
56 species.

57 While niche partitioning remained the theoretical basis for understanding species
58 coexistence, empiricists and theorists struggled to reconcile the rich biodiversity that exists in
59 many of the world's ecosystems with the prediction that competitive communities should contain
60 only as many species as there are limiting resources (or consumers) (Hutchinson 1961, Oksanen
61 et al. 1981, Huisman and Weissing 1999). But starting in 2000, theories of species coexistence
62 began to undergo a major revision. In 2001, Hubbell published The Unified Neutral Theory of
63 Biodiversity (Hubbell 2001), which argued that patterns of biodiversity in nature can be

64 explained by a simple model that does not invoke niche differences among species. According to
65 Hubbell's theory, species coexist not because they are different, but because their demographic
66 parameters are identical, or nearly so, such that the consequences of their interactions are
67 'neutral' (i.e. essentially equal among all species). Based on this neutral theory, Hubbell argued
68 that the biodiversity we observe in nature can be explained by a series of stochastic events that
69 cause some populations to become dominant while others exhibit random walks toward
70 extinction.

71 Even as Hubbell was developing his neutral theory, Chesson (2000) was completing a
72 ground-breaking theory of coexistence that would provide a framework for integrating the niche
73 and neutral perspectives on biodiversity. Chesson's coexistence framework was built on his
74 insight into the invasibility criterion: a pair of species can coexist only if each species is capable
75 of invading a steady-state population of its competitor. Chesson showed how a species' long-
76 term growth rate when invading a resident species can be decomposed into two general terms,
77 which he called stabilizing and equalizing forces. Stabilizing forces cause species to limit their
78 own growth rate more than they limit the growth rate of other species (intra > interspecific
79 competition). These stabilizing forces, also known as niche differences (ND), occur when
80 species partition limiting resources in space or time, or when they experience differential
81 consumption by consumers. In contrast, equalizing forces minimize differences in competitive
82 abilities among species. Equalizing forces, which have also been called relative fitness
83 differences (RFD), are the result of inherent variation in biological traits such as minimum
84 requirements for shared resources or consumers, differential resistance to consumers, or
85 differences in potential growth rates (Adler et al. 2007, Levine and HilleRisLambers 2009,
86 HilleRisLambers et al. 2012).

87 Chesson showed it is the balance of these two forces – RFDs that establish competitive
88 hierarchies, and NDs that prevent competitive exclusion – that ultimately determines whether
89 species maintain non-negative long-term growth rates in competitive communities (Chesson
90 2000). For a pair of species to coexist, ND must be sufficiently large to offset and stabilize the
91 competitive hierarchies generated by RFDs. When two species exhibit identical niches (ND
92 equals 0), their RFD alone determines the competitive hierarchy and which species will become
93 extinct. It has subsequently been shown that Hubbell's neutral theory represents a specific,
94 limiting case of Chesson's coexistence theory where NDs and RFDs are both zero, causing the
95 outcome of competition to be approximated by a random walk toward extinction (Adler et al.
96 2007). Stabilizing and equalizing forces have been identified in both fluctuation-dependent
97 mechanisms (e.g. storage effects) and fluctuation independent mechanisms of coexistence (e.g.
98 competition for a limiting resource) (Miller and Klausmeier 2017, Barabas et al. 2018). Thus,
99 Chesson's inequality provides a general framework for predicting species coexistence.

100 Since the development of Chesson's theory, much attention in ecology has turned
101 towards the empirical estimation of ND and RFD in order to determine how these two forces
102 contribute to coexistence in real communities. As a growing number of empiricists have tried to
103 quantify ND and RFD in their individual study system, the number of different empirical
104 approaches proposed for doing so has also grown. However, the various methods for
105 implementing modern coexistence theory in empirical studies have been derived from different
106 models of species interactions, make different assumptions, and use different experimental
107 designs. Therefore, it remains unclear which method(s) are best suited for a given study, whether
108 the methods give comparable estimates of ND and RFD, and whether the methods actually give
109 the same prediction regarding coexistence. If Chesson's theory is to become widely implemented

110 in empirical studies and in applied contexts, we need a ‘users guide’ to help ecologists determine
111 which empirical approach meets their needs.

112 Here we provide a summary and comparison of four methods that have been proposed to
113 measure ND and RFD empirically, and a fifth method that does not give estimates of ND and
114 RFD but has been used to predict coexistence based on Chesson’s theory. In Part 1 of our paper,
115 we describe the theoretical background of each method, illustrate how it can be implemented
116 empirically, and ask whether the methods yield the same estimates of ND and RFD and make the
117 same predictions regarding coexistence. In Part 2 of the paper we provide a list of decision steps
118 to guide empiricists in selecting the most appropriate method(s) for their study system and
119 contrast the methods in terms of the amount of effort required to implement them. In Part 3, we
120 discuss the main advantages and disadvantages of the methods and make some suggestions for
121 future empirical work on coexistence theory.

122

123 **Part 1. Summary of Five Empirical Methods for Implementing Chesson’s Theory**

124 In this part of the paper we first review the fundamentals of Chesson’s theory that are
125 most relevant to empirical work on competition and coexistence. We then summarize each of
126 five empirical methods that have been used to measure niche differences (ND) and relative
127 fitness differences (RFD) by (i) explaining how the method relates to Chesson’s theory, (ii)
128 showing how the method can be implemented empirically, and (iii) highlighting the method’s
129 key limitations and assumptions. At the end of Part 1, we use numerical simulations to compare
130 the five methods to determine if they give the same estimates of ND and RFD, and to assess
131 whether they yield the same prediction for coexistence based on the invasibility criterion.

132

133 *1.1 Brief review of Chesson's theory*

134 When Chesson first introduced his theory for coexistence, he did not prescribe a specific
135 empirical approach or experiment that should be used to estimate ND and RFD in real biological
136 communities. Instead, he used a phenomenological model of competition to show how the
137 mutual invasibility criterion, a common prerequisite for coexistence, depends on how each
138 species limits their own growth rate versus that of their competitor (Chesson 1990). Specifically,
139 he showed that the criterion for mutual invasibility can be expressed as an inequality involving
140 both ND and RFD (Equation 1).

141
$$1 - ND < RFD < \frac{1}{1-ND} \quad (1)$$

142 In this inequality, the term 1-ND represents the degree of niche overlap (ρ in Chesson (1990)),
143 which ranges from zero when species do not share any resources to one when the resource
144 requirements of species are identical. RFD represents the ratio of competition-free fitness among
145 the two species ($\frac{f_j}{f_i}$ in Chesson (1990)). If this equality is not satisfied, then one of the species is
146 unable to maintain long-term, positive growth rates and will go locally extinct.

147 Because ND and RFD are not terms that cannot be quantified directly from experiments
148 or observations, Chesson showed how these forces can be derived from the classic Lotka-
149 Volterra competition model. In this model, the *per capita* growth rate of species i is a function of
150 both intraspecific and interspecific competition as described by Equation 2:

151
$$\frac{1}{N_i} \frac{dN_i}{dt} = r_i (1 - \alpha_{ii} N_i - \sum_{j \neq i} \alpha_{ij} N_j) \quad (2)$$

152 In Equation 2, N_i is the density of species i , and r_i is the intrinsic per capita growth rate of species
153 i . The intra-specific competition coefficient α_{ii} describes the *per capita* effect of species i on the
154 *per capita* relative growth rate of species i and is equal to the inverse of the carrying capacity
155 (K_i) for species i . The inter-specific competition coefficient α_{ij} describes the *per capita* effect of

156 species j on the *per capita* relative growth rate of species i . Equations 3 and 4 relate the inter-
157 and intra-specific interaction coefficients from the Lotka-Volterra model to ND and RFD:

$$158 \quad ND = 1 - \sqrt{\frac{\alpha_{ij}\alpha_{ji}}{\alpha_{ii}\alpha_{jj}}} \quad (3)$$

$$159 \quad RFD = \sqrt{\frac{\alpha_{ij}\alpha_{ii}}{\alpha_{jj}\alpha_{ji}}} \quad (4)$$

160 ND and RFD from Equations (3) and (4) can then be used in Equation 1 to predict coexistence.
161 Because Chesson derived ND, RFD and the inequality for mutual invasibility based on the
162 Lotka-Volterra competition model, we use the same approach to explain four of the empirical
163 approaches described below and show that a fifth approach is ultimately not compatible with
164 Chesson's derivation.

165

166 *1.2 Method based on the parameterized Lotka-Volterra model*

167 1.2.1 Theoretical background. Since Chesson originally used the Lotka-Volterra model to
168 explain his criterion for coexistence, the most obvious empirical approach for estimating ND and
169 RFD is to parameterize the Lotka-Volterra competition model (Equation 2) using data collected
170 from experiments or time-series observations from natural ecosystems. The inter- and intra-
171 specific interaction coefficients (α_{ii} and α_{ij}) from the Lotka-Volterra competition model can be
172 used to estimate ND and RFD using Equations 3-4.

173 1.2.2 Empirical approaches. Empirical implementation of the Lotka-Volterra model
174 requires estimating six different parameters used in Equation 2: intrinsic per capita growth rate of
175 each species (r_i and r_j), per capita intra-specific competition coefficients (α_{ii} and α_{jj}), and per
176 capita inter-specific competition coefficients (α_{ij} and α_{ji}). The simplest way to parameterize the
177 Lotka-Volterra model from experiments would be to use three treatments for each pair of

178 species: a time-series of each species grown alone as a monoculture and one time-series
179 representing a co-culture of the two species (Figure 1). From each monoculture time series, the
180 empiricist needs to measure the population density of each species over time, from low density to
181 steady-state. From these time series, the empiricist can estimate the maximum per capita growth
182 rate of each species (r_i), which occurs as the species' density approaches zero, and the steady-
183 state population size of each species in monoculture (K_i).

184 Empiricists have two options for using the monocultures to estimate the intraspecific
185 competition coefficients (α_{ii}). The first option is to estimate these interaction coefficients using
186 the slope of each species' relative growth rate (scaled on its maximum growth rate r_i) versus its
187 population density (Figure 1A and 1B, right). This slope has the opposite sign of the intraspecific
188 interaction coefficient α_{ii} . However, in practice this slope is unlikely to be fixed across all
189 population densities. As a result, intra- and inter-specific competition coefficients measured in
190 competition experiments are likely to vary with population densities (Abrams 1980). An
191 alternative strategy is to use the assumption that intraspecific competition coefficients (α_{ii}) are
192 equal to $1/K_i$, which comes directly from the Lotka-Volterra model in Equation 2. This approach
193 yields estimates of α_{ii} that cannot, say, be used in Equation 2 to re-create transient dynamics
194 observed in time series, but can be used in Equations 3 and 4 to estimate ND and RFD and
195 predict coexistence using Chesson's inequality (Equation 1).

196 Next the empiricist would perform a competition experiment in which both of the species
197 are introduced to habitat at low density and the population density of each species is measured
198 over time (Figure 1C). From these time series, the empiricist must use non-linear regression to
199 parameterize the interspecific interaction coefficients (α_{ij} and α_{ji}) by substituting the parameter

200 estimates from the monocultures into Equation 2. Finally, the empiricist can use all four
201 interaction coefficients to compute ND and RFD using Equations 3 and 4.

202 1.2.3 Limitations. A key assumption of this approach is that the intra- and inter-specific
203 competition coefficients are fixed with respect to population sizes of either species. In other
204 words, the first individual and the last individual added to a population have the same per capita
205 effect on the growth rates of its own species or that of its competitor. This assumption is not
206 always met in real biological communities where intra- and inter-specific competition
207 coefficients can depend on species' densities (Smith-Gill and Gill 1978). The assumption also
208 does not apply when the mechanisms leading to competition are driven by non-linear
209 dependence on resources. Examples include functional responses of consumers to prey density
210 and non-linear dependence of growth rates on abiotic resource availability (e.g. the Monod
211 function). Figure A1 in the Supporting Information shows that, when applied to numerical
212 simulations based on a well-known consumer resource model parameterized with real biological
213 data, intraspecific coefficients measured in monoculture near equilibrium lead to inaccurate
214 predictions regarding coexistence. However, when the intraspecific interaction terms are
215 replaced by $1/K_i$ the method yields accurate predictions. Therefore, in those situations where
216 competition coefficients are fixed with respect to population size, or can be measured at low
217 population sizes of each species, then the empirical approaches can be used to estimate ND and
218 RFD.

219

220 *1.3 Sensitivity method*

221 The second method for estimating ND and RFD, the sensitivity method, is similar to the
222 Lotka-Volterra method in that it is phenomenological and requires information from direct
223 competition experiments.

224 1.3.1 Theoretical background. The sensitivity method quantifies the proportional
225 reduction in a species' growth rate when invading a steady-state population of its competitor
226 (Carroll et al. 2011, Narwani et al. 2013). In this method, the maximum growth rate of each
227 species in monoculture (μ_i) and when invading a steady-state population of the competitor
228 species (μ_{ij}) are used to calculate each species sensitivity to interspecific competition (S_i) using
229 Equation 5:

$$230 \quad S_i \equiv \frac{\mu_i - \mu_{ij}}{\mu_i} \quad (5)$$

231 Carrol and others (2011) and others have shown that ND is proportional to the geometric mean
232 of these sensitivity measures, whereas RFD represent variation around the mean.:

$$233 \quad ND = 1 - \sqrt{S_i S_j} \quad (6)$$

$$234 \quad RFD = \sqrt{\frac{S_i}{S_j}} \quad (7)$$

235 A species' sensitivity to competition is jointly determined by ND and RFD (Carroll et al.
236 2011, Narwani et al. 2013). Specifically, greater ND between the two species reduces the impact
237 of interspecific competition so that S_i will approach zero. Greater RFD, on the other hand, causes
238 species to be asymmetrically affected by competition such that one species' sensitivity increases
239 while the other's decreases. While Carroll and others (2011) verbally argued that this method is
240 compatible with Chesson's theory, in Appendix A we show explicitly how this method relates to
241 Chesson's theory and prove that it is identical to Equations 3 and 4 when interspecific interaction

242 coefficients of the Lotka-Volterra model are evaluated only at the steady-state population density
243 of the resident species.

244 1.3.2 Empirical approaches. The sensitivity method uses a combination of monocultures
245 and pairwise invasion experiments to quantify the reduction in each species' growth rate caused
246 by a steady-state population density of the other species at its carrying capacity (Figure 2). The
247 experiment by Narwani et al. (2013) provides an example for how to implement the sensitivity
248 method empirically. Their experimental system involved species of freshwater green algae
249 growing under controlled conditions in the laboratory. They grew each species as a monoculture
250 in fresh growth medium, starting from low biomass and allowing the populations to reach their
251 carrying capacity. From these time series, they quantified the per capita maximum growth rate of
252 each species as a monoculture, which occurs when the focal species is at low density (μ_i and μ_j).
253 After each species reaches its carrying capacity, they introduced the other species from low
254 density (e.g. 0.01% of K) and quantified the per capita growth rate of each species when
255 invading the other (μ_{ij} and μ_{ji}). Finally, for each pair of species, they used these growth rates to
256 calculate the sensitivity metrics (S_i and S_j) using Equation 5 and used those sensitivity metrics to
257 calculate ND and RFD using Equations 6 and 7.

258 1.3.3 Limitations. Using the sensitivity method requires one to perform mutual
259 invasibility experiments, which are only practical for organisms whose population growth rates
260 can be measured over tractable periods of time. Mutual invasibility experiments are harder to
261 apply to organisms that grow slowly (e.g. trees) since it would take a long time to acquire time
262 series of species densities needed to estimate per capita growth rates. Also, it is important to note
263 that the invasion growth rates (μ_{ij} and μ_{ji}) must be measured when the invader population
264 density approaches zero. Under this condition, intra-specific competition is negligible, and the

265 resident species' density is near steady-state. If the growth rate of the invader species were
266 measured at greater density of the invader species or lower density of the resident species (i.e.
267 long after invasion), then the S_i would be affected by both intra- and inter specific competition.
268 The resulting predicting regarding species coexistence would be incorrect.

269

270 *1.4 Parameterizing MacArthur's consumer resource model*

271 The third method to estimate ND and RFD from empirical data is to parameterize
272 MacArthur's consumer-resource model (MacArthur 1970) then use these parameters to calculate
273 ND and RFD using Chesson's original derivation (Chesson 1990, 2000). This method is different
274 from both the Lotka-Volterra and sensitivity methods because it does not rely on experiments
275 where the species are grown together in order to quantify how the species influence each other's
276 growth rates. Instead, this method works by parameterizing a mechanistic model that describes
277 species interactions, then reorganizing those parameters to estimate ND and RFD for assessing
278 Chesson's inequality.

279 1.4.1 Theoretical background. MacArthur's consumer resource model describes how
280 species consume and thus compete for two or more prey resources (MacArthur 1970). The model
281 is composed of differential equations representing the growth of each consumer species as a
282 function of resource densities (Equation 8) and a differential equation (or set) that describes the
283 population dynamics of each prey resource and their mortality due to consumption by the
284 consumers (Equation 9).

285
$$\frac{1}{X_i} \frac{dX_i}{dt} = b_i (\sum_{l=1}^m c_{il} w_{il} R_l - m_i) \quad (8)$$

286
$$\frac{1}{R_l} \frac{dR_l}{dt} = r_l \left(1 - \frac{R_l}{K_l}\right) - \sum_{i=1}^n c_{il} X_i \quad (9)$$

287 In this model X_i is population density of the consumer species i and R_l is population
288 density of the prey resource l . The term b_i represents the effect of prey consumption on the
289 growth rate of the consumer, r_l is the maximum per-capita growth rate of prey resource l , K_l is
290 the carrying capacity for the prey species l , w_{il} represents the increase in consumer population
291 density for each unit of prey resource l consumed. The term c_{il} is the resource capture rate by
292 consumer i on resource l and m_i is the density-independent mortality for consumer species i .
293 Chesson showed that, by implementing a time-scale separation technique, parameters in
294 MacArthur's consumer resource model can be used to calculate ND and RFD using Equations 10
295 and 11 (Chesson 1990, 2000):

$$296 \quad ND = 1 - \sum_l c_{ij} c_{jl} \frac{w_{il} K_l}{r_l} \quad (10)$$

$$297 \quad RFD = \frac{\sum_l c_{jl} w_{il} K_l - m_j}{\sum_l c_{il} w_{il} K_l - m_i} \quad (11)$$

298 The estimates of ND and RFD from this method can then be used to evaluate Chesson's
299 inequality and predict coexistence (Equation 1).

300 1.4.2 Empirical approaches. Using MacArthur's consumer resource model to estimate
301 ND and RFD for a pair of species requires quantifying 1) the per capita consumption rate of each
302 consumer species on each prey resource (c_{il} , units of prey consumed per unit prey density); 2) the
303 per capita maximum growth rate and carrying capacity of each prey resource when no consumers
304 are present (r_l and K_l); and 3) the yield of consumer population density or biomass relative to
305 each unity of prey consumed (w_{il}). Because we are not aware of any empirical studies that have
306 parameterized the MacArthur model for the purpose of estimating ND and RFD, we describe the
307 experimental approach that would be required (Figure 3). First, the empiricist would need to
308 identify or define the prey resources that are available to the consumer species. Each prey
309 resource would be inoculated or planted at low density into an environment free of other prey

310 resources and consumers, then the population density would be measured over time in order to
311 estimate the per capita maximum growth rate of the prey (r_l , which occurs as the prey population
312 density approaches zero) and its carrying capacity (K_l , which occurs when the prey growth rate
313 approaches zero). Next, the experimentalist would need to introduce each consumer species into
314 several different densities of each prey resource growing as a monoculture. Under those different
315 prey resource densities, the experimentalist would measure the per capita consumption rate of
316 prey resource by the consumer species (c_{il}) and the yield of consumer density or biomass per unit
317 prey resource consumed (w_{il}).

318 The precise number of parameters to be estimated depends on the number of prey
319 resources considered by the model. For example, for two consumer species and three prey
320 resources (Figure 3), the hypothetical experiment requires 18 parameters to be quantified: 3
321 different maximum per capita growth rates and 3 carrying capacities of the prey resources (r_l and
322 K_l , $l = 1$ to 3), 6 per capita consumption rates (c_{il} ; $i = 1$ and 2, $l = 1$ to 3), and 6 yields (w_{il} ; $i = 1$ and
323 2, $l = 1$ to 3). These parameters can then be used in Equations 10 and 11 to obtain ND and RFD,
324 which can subsequently be used in Equation (1) to predict coexistence.

325 1.4.3 Limitations. The MacArthur's consumer resource model gives a more mechanistic
326 understanding of competitive interactions among species, and allows one to predict coexistence
327 for pairs of species without needing to grow them together in a competition experiment.
328 However, these desirable properties come with greater number of experimental treatments
329 compared to the Lotka-Volterra and sensitivity methods. In particular, this method requires as
330 many consumption experiments as there are resources, and each of these experiments involves
331 measuring consumption rates at a range of resource species densities (Figure 3 B and C). While
332 this constraint does not impact the ability of the method to predict coexistence under defined

333 conditions, it could limit the extent to which those predictions can be applied to natural
334 environments where the number of potential prey species is large.

335

336 *1.5 Parameterizing Tilman's consumer resource model*

337 Like the method based on MacArthur's consumer-resource model, the method based on
338 Tilman's consumer resource model does not require species to be grown together in a
339 competition experiment (Letten et al. 2017). However, unlike the method based on MacArthur's
340 CRM, the method based on Tilman's CRM is specific to abiotic resources that are controlled by
341 a constant rate of supply and do not have their own intrinsic growth rate (i.e. a chemostat).

342 1.5.1 Theoretical background. Letten and others (2017) showed how a consumer resource
343 model (Tilman 1977) can be reorganized to a Lotka-Volterra form in order to estimate ND and
344 RFD and assess mutual invasibility using Chesson's inequality. This method is based on
345 Tilman's two-species consumer resource model for two essential and non-substitutable resources
346 (Tilman 1977). In this model, one set of differential equations describes the growth of each
347 consumer species as a function of resource availability (Equation 12) and another set of
348 equations describes the dynamics of abiotic resources and their depletion due to uptake by the
349 consumer and dilution (Equation 13).

$$350 \frac{1}{N_i} \frac{dN_i}{dt} = r_i \min \left[\frac{s_i}{s_i + k_{ii}}, \frac{s_j}{s_j + k_{ij}} \right] - D \quad (12)$$

$$351 \frac{1}{R_i} \frac{dS_i}{dt} = D(S_{i0} - S_i) - \sum_{i=1}^n \frac{N_i}{y_{ii}} r_i \min \left[\frac{s_i}{s_i + k_{ii}}, \frac{s_j}{s_j + k_{ij}} \right] \quad (13)$$

352 In this model, N_i is the population density of species i , r_i is the maximum per capita growth rate
353 of species i , y_{ii} is the yield of species i on resource i , and k_{ii} is the half saturation constant for
354 growth of species i on resource i . The term S_{i0} is the external supply concentration for resource i ,

355 S_i is the concentration of resource i in the environment, and D is equal to both the supply rate of
356 resources and the density-independent loss rate for both species.

357 Letten et al. (2017) showed how the parameters from Tilman's CRM can be used to
358 calculate ND and RFD. First, the empiricist must determine which species is limited by each
359 resource (e.g., using Resource-Ratio theory (Tilman 1982)). This requires comparing the supply
360 ratio for the two resources against the R^* s for each species at the pre-determined dilution rate.
361 As such, the following equations are specific to only the condition where species 1 is limited by
362 resource 2 and species 2 is limited by resource 1.

363
$$ND = 1 - \sqrt{\frac{y_{12}y_{21}}{y_{11}y_{22}}} \quad (14)$$

364
$$RFD = \sqrt{\frac{\left(\frac{Dk_{21}}{D-r_2} + S_1\right)^2 y_{11}y_{21}}{\left(\frac{Dk_{12}}{D-r_1} + S_2\right)^2 y_{12}y_{22}}} \quad (15)$$

365 Equations 14 and 15 can be used to estimate ND and RFD and then used in Chesson's inequality
366 (Equation 1) to predict species coexistence.

367 1.5.2 Empirical approaches. To illustrate how this method could be implemented
368 empirically, we describe the approach that Tilman first used to parameterize his model (Tilman
369 1977). In the chemostat environment associated with Tilman's model, the abiotic resources or
370 nutrients are delivered to the experimental system at a constant supply rate that matches the
371 density-independent mortality rate (Tilman 1977, 1982). First, he measured the growth kinetics
372 for two algae species when limited by two different essential resources (silicate and phosphate).
373 He inoculated each species as a monoculture into growth medium containing a range of
374 concentrations of the limiting resource (either silicate or phosphate) with all other resource in
375 excess. From these time series of population densities, he quantified the growth rate of each

376 species as a function of the concentration of the limiting resource following the Monod model
377 (Figure 4). For each species, this yields estimates of half saturation constants (k_{ij}) for each
378 resource and a single maximum per capita growth rate for both resources (r_i). Next, Tilman
379 quantified the yields (y_{ij}) of each species on each resource by measuring the elemental content of
380 a known number of cells. By following the same approach, an empiricist can quantify the 10
381 parameters needed to define Tilman's model.

382 As shown by Letten et al. (2017), the parameters described above can be used to predict
383 coexistence under different resource supply ratios and dilution rates in a chemostat. However,
384 the form of the equation used to get ND and RFD depends upon the empiricist knowing whether
385 or not the supply ratio is outside the ratio of $\frac{R_{ii}^*}{R_{ij}^*}$ and $\frac{R_{ji}^*}{R_{jj}^*}$ for the two species, as determined by
386 resource-ratio theory (Tilman 1982).

387 1.5.3 Limitations. The method using Tilman's consumer resource model requires an
388 empiricist to know precisely which resources the species compete for, which limits its
389 applicability to many real scenarios and ecosystems where the identity of limiting resources and
390 the supply rates may not be known. Additionally, the resource supply concentrations must be
391 fixed and the supply rates must be equal to the density-independent loss rate, which can only be
392 achieved in certain experimental settings like chemostats. Another important note is that the
393 elemental content of organisms like algae is not constant and can vary as a function of growth
394 rate, nutrient availability, and other factors (Sterner and Elser 2002). Non-constant yields due to
395 luxury uptake and variable internal stores (Grover 1991) have important consequences for
396 competition, but it has yet to be determined how this plasticity affects estimation of ND and
397 RFD.

398

399 *1.6 Negative frequency dependence (NFD) method*

400 The final method that we summarize, the negative frequency dependence method (NFD),
401 has not been proposed as a means of obtaining estimates of ND and RFD that are directly
402 compatible with Chesson's inequality (Equation 1). In fact, in Supporting Information C, we
403 explicitly show that the NFD method cannot be used to derive estimates of ND and RFD that are
404 consistent with Chesson's theory. However, the NFD method can be used to predict coexistence
405 using the criterion of mutual invasibility. Moreover, the NFD method has been proposed as a
406 way to interpret stabilizing and equalizing forces from Chesson's original theory, and the method
407 has been used to illustrate the impacts of ND and RFD in manipulative experiments (Adler et al.
408 2007, Levine and HilleRisLambers 2009).

409 1.6.1 Theoretical background. The NFD method quantifies the change in the per capita
410 growth rate of a species as a function of its frequency in a community (Adler et al. 2007, Levine
411 and HilleRisLambers 2009). Here the frequency of a species refers to the proportion of total
412 biomass or individuals in a community belonging to that species. This method makes the key
413 assumption that the community is saturated with respect to total species densities, thus a
414 frequency of 1 represents a steady-state monoculture at its carrying capacity. At all other
415 frequencies, the community composition need not be at a steady-state, but it is assumed that any
416 increase in the population density of one species will be offset by a decrease in population
417 density of another species. Using this assumption, the slope of the NFD relationship has been
418 used to reflect the difference between intra- versus inter-specific competition (Adler et al. 2007).
419 Increasing species i 's frequency means that individuals of species i will compete more with
420 individuals of its own kind than with individuals of other species, and will thus experience
421 stronger intraspecific competition than interspecific competition. Therefore, if intra-specific

422 competition is greater than inter-specific competition, the species affects its own growth rate
423 more than it affects the growth rate of other species, and the NFD slope should be negative.

424 The NFD method is most often used as a graphical approach for understanding the
425 balance of ND and RFD (Figure 5). According to Adler et al. (2007), a more negative NFD
426 slope represents a stronger stabilizing force, which they argue is proportional to the ND in
427 Chesson's inequality. Similarly, they argue that the difference between species' growth rate in
428 the absence of stabilizing forces is the equalizing force, proportional to RFD. In Supporting
429 Information C we show that this NFD slope is not equivalent to ND and that the difference in
430 elevation between species' NFD relationships is not RFD. Although the NFD approach does not
431 yield estimates of ND and RFD that are consistent with Equation 1, this method can still be used
432 to make predictions about coexistence based on Chesson's mutual invasibility criterion.

433 Specifically, if each species has a positive growth rate at a frequency approaching 0 then the
434 mutual invasibility criterion is satisfied, and coexistence is assured. Adler and others (2007)
435 showed how both the slope and elevation of the NFD plot are needed to accurately determine
436 whether this condition is met.

437 1.6.2 Empirical approaches. The NFD method could be implemented using either
438 experiments or observations from natural ecosystems (Figure 5A). Using the experimental
439 approach, an empiricist could grow both species together and measure their densities over
440 time. First, the empiricist would need to ensure that total community density or biomass was
441 saturated. One way to do this would be to use invasion experiments in which the resident species
442 is grown to steady-state and then the invader species is introduced at low density. Using the
443 invasion approach guarantees that the community is saturated since any decrease in the resident's
444 population density is met with an increase in the density of the invader. Next, the empiricist

445 could use the time series for each species' density to calculate the per capita growth rate and
446 frequency of each species at each time point. Alternatively, the NFD method could be
447 implemented using observational data from natural ecosystems (Adler et al. 2010). This
448 approach could allow an empiricist to estimate frequency dependence for species that are not
449 easily manipulated (e.g. trees). To use this approach, an empiricist would quantify the per capita
450 growth rate and the relative frequency of the species in different habitat patches or along
451 ecological gradients. Although this approach has not been applied empirically to make
452 predictions regarding coexistence (but see Yenni et al. 2017), it is one of only two methods
453 reviewed here that do not require manipulative experiments.

454 Having obtained pairs of growth rate and frequency from either experiments or
455 observational studies, the empiricist can construct plots of the NFD relationship (Figure 5B).
456 This relationship can be used to estimate the growth rate when each species approaches
457 frequency of zero. If either of the two species does not have a positive growth rate when rare,
458 then the pair will not coexist. Since the NFD method does not yield estimates of ND and RFD
459 that can be used in Equation 1, the utility of this method is completely dependent on its ability to
460 accurately predict mutual invasibility (i.e. growth rates at frequency of 0). In many cases, this
461 prediction will be based on observations made at species frequencies greater than zero. As long
462 as the relationship between a species' frequency and its growth rate is linear, the NFD slope and
463 elevation can theoretically be used to predict whether both species will have positive growth
464 rates when rare, thus meeting the mutual invasibility criterion. If the NFD relationship is not
465 linear, then the NFD method can give inaccurate predictions (Figure 5B, see Limitations below).
466 Indeed, even when applied to numerical simulations from the simple Lotka Volterra model, the

467 NFD slope is not constant across species frequencies in a saturated community (Supporting
468 Information Figure C1).

469 Levine et al. (2009) demonstrated how the NFD method can be implemented
470 experimentally. In their study with 10 species of grassland plants, they manipulated the relative
471 frequency of each focal species by varying the proportion of seeds belonging to the focal species
472 versus all other species. At the end of the growing season, they quantified the growth rate of each
473 species in their study plots by multiplying the number of seeds belonging to each species by the
474 proportion of those seeds that were viable the next year. They then quantified the slope of NFD
475 by plotting the growth rate of each species against its frequency in the initial community.

476 Although the slope of NFD is not equal to Chesson's ND and the difference in intercepts is not
477 equal to RFD, the authors showed that the effect of niche differences on growth rates can be
478 removed by experimentally maintaining each species' density at a constant, non-equilibrium
479 level that is not subject to competition from other species. Their experiment showed that
480 effectively removing niche differences among species (even without measuring them) led to
481 dominance by the species with the highest per capita growth rates. In other words, in the absence
482 of ND the outcome of competition was determined by RFD. It is important to note that while this
483 approach is based on Chesson's inequality, it does not require measuring ND and RFD.

484 Similarly, other studies have measured the slope of NFD as evidence for the importance of
485 stabilizing forces, but did not directly interpret the slope as ND or the intercepts as RFD (Yenni
486 et al. 2017).

487 1.6.3 Limitations. Despite some of the desirable aspects of the NFD method in terms of
488 empirical approaches (above), it has three key limitations. First, unlike the other four methods
489 summarized in this paper, the NFD method does not yield estimates of ND and RFD. This may

490 not be a concern if the purpose of the study is simply to predict species coexistence and does not
491 focus on explaining why certain pairs coexist while other pairs do not. Second, the NFD method
492 assumes that the community density is saturated across the range of species' frequencies
493 observed. Meeting this assumption in experiments requires sufficiently long time series to show
494 that total biomass of a community is fixed. In observational studies based on natural ecosystems,
495 it might not be possible to ensure that total biomass is saturated.

496 The third limitation of the NFD method is that the relationship between a species'
497 frequency and growth rate is often non-linear (Figure 5). In Appendix C, we show the NFD
498 method can lead to incorrect predictions about species coexistence when applied to systems with
499 non-linear relationships between species' growth rates and densities. If the slope and elevation of
500 the NFD plot are evaluated over a narrow range of species frequencies, and those data were used
501 to extrapolate to predict growth rate as frequency approaches zero, then the method could make
502 inaccurate predictions about mutual invasibility and coexistence. If the relationship between each
503 species' frequency and its growth rate is not linear, then an empiricist must adequately describe
504 the relationship to account for the non-linearity. This means that for an empiricist to use the NFD
505 method, they would need to either 1) measure the growth rate of each species across the full
506 range of frequencies to establish that the growth rate of each species is linearly related to its
507 frequency or 2) evaluate the growth rate of each species when rare (i.e. directly demonstrate
508 mutual invasibility). Both of these options would dramatically increase the effort required but
509 may be necessary in systems where only observational studies are possible.

510

511 *1.7 Do the methods give the same prediction regarding coexistence?*

512 Although each of the five methods can be used to predict coexistence, the experimental
513 approaches required for those methods are different, and it is not clear that the methods would
514 yield the same predictions (or values of ND and RFD) if applied to the same study system. Here
515 we use numerical simulations to investigate whether four of those methods, when implemented
516 as shown in Figures 1, 2, 3, and 5, lead to the same prediction regarding coexistence and give the
517 same estimates of ND and RFD. We could not include both the method based on MacArthur's
518 CRM and the method based on Tilman's CRM since these mechanistic models have
519 incompatible assumptions – the resources in MacArthur's CRM have their own population
520 dynamics while the resources in Tilman's CRM are abiotic and governed by a constant rate fo
521 supply. We chose to use numerical simulation for this demonstration since we are unaware of
522 any empirical dataset that has been, or could be, analyzed using more than two of the methods.
523 The numerical simulations were based on Tilman's consumer-resource model (Tilman 1977)
524 with two species of phytoplankton competing for two essential resources (See Supporting
525 Information A). For each set of resource conditions, we performed numerical simulations that
526 represent four distinct methods: 1) fitting the Lotka-Volterra model to monocultures and a co-
527 culture (Figure 1), 2) the sensitivity method (Figure 2), 3) the method using Tilman's CRM
528 (Figure 4), and 4) the NFD method (Figure 5).

529 Figure 6 shows that under specific limiting assumptions, all four methods made the same
530 prediction about coexistence and that these predictions matched the outcome based on the
531 equilibrium condition from simulations. Across the different resource conditions that we
532 explored, the two species were predicted to coexist when the resource supply conditions caused
533 each species to be limited by a different resource, consistent with resource-ratio theory (Tilman
534 1977). However, this agreement among the methods was conditional on how the Lotka-Volterra

535 and NFD methods were parameterized. The Lotka-Volterra method only matched the predictions
536 for coexistence from the other methods when we assumed that intraspecific competition
537 coefficients were equal to the inverse of the carrying capacity (Supporting Information Figure
538 A1; Section 1.2). When we estimated the intraspecific coefficients directly from the monoculture
539 time series as they approached their carrying capacity, the method produced incorrect predictions
540 and overestimated the range of parameter space allowing for coexistence. Similarly, the NFD
541 method only matched the predictions for coexistence from the other methods when we 1)
542 evaluated the slope of NFD when species' frequencies were approaching zero and 2) used both
543 the slope and the intercept to predict the growth rate when frequency approaches zero. Unless
544 these conditions were met, the NFD method tended to over- or under-estimate the region of
545 resource conditions that allows for coexistence (Supporting Information Figure A1).

546

547 *1.8 Do the methods yield the same values of ND and RFD?*

548 Although the methods gave the same predictions regarding coexistence, Figure 6 (F and
549 G) shows that the methods do not yield the same values of ND and RFD, even when applied to
550 the same simulated study system. The Lotka-Volterra method (using the simplification that
551 $\alpha_{ii}=1/K_i$) and the sensitivity method gave identical estimates of ND and RFD across the range of
552 resource conditions used, but these estimates differed from the method based on Tilman's
553 consumer resource model. This disparity can be explained by the fact that the Lotka-Volterra and
554 sensitivity methods assume that per capita inter- and intraspecific interaction coefficients are
555 independent of species' densities. Although this assumption is likely to be violated when species'
556 population dynamics are affected by mechanisms that produce non-linearity between population
557 densities and growth rates, using the assumption that $\alpha_{ii} = 1/N_i^*$ allows for accurate predictions

558 regarding coexistence. In contrast, the method based on Tilman's consumer resource model does
559 not assume that interaction coefficients are independent of species densities, but instead
560 quantifies both inter- and intraspecific interaction coefficients only at the steady-state densities
561 predicted for monocultures that undergo invasion. This means that the interaction coefficients,
562 and thus ND and RFD, measured according to either of the phenomenological methods (Figures
563 1 and 2) are unlikely to match the values predicted from a mechanistic method, even though both
564 can correctly predict mutual invasibility.

565 This comparison of methods highlights an important caution, namely that estimates of
566 ND and RFD obtained by different methods are not always comparable. Therefore, future
567 syntheses or meta-analyses should not combine studies that measured ND and RFD by different
568 methods. Even within a single method (e.g. the Lotka Volterra method) there can be substantial
569 differences in the estimates of ND and RFD depending on the experimental design and how the
570 interaction coefficients are parameterized. Similarly, estimates of ND and RFD from the
571 mechanistic methods are dependent on the non-biological parameters used in those models (e.g.
572 dilution rates). If a future study were to compile these values from different studies without
573 ensuring that the same assumptions were used throughout, the results and interpretation of the
574 synthesis would be meaningless.

575

576 **Part 2. An Empiricist's Guide to Selecting a Method To Estimate ND and RFD**

577 Having described and compared the foundation of each empirical method, here in Part 2
578 of the paper we offer practical guidance to help empiricists determine 1) which method(s) are
579 most appropriate for their study system and 2) how much experimental effort is required for a
580 given method. To aid our discussion, we have summarized the methods in Table 1, which is

581 organized into three sections. The section of the table labeled ‘Decision Steps’ is a decision tree
582 that allows an empiricist to identify the most appropriate method for their study system. The
583 section labeled ‘Method’ directs the empiricist to the key literature needed to implement the
584 approach. Last, the section of the table labeled ‘Experimental Requirements’ outlines key aspects
585 of the experiments that are required to use the method.

586

587 *2.1 Decision Steps - deciding which method to use*

588 The first section of Table 1, ‘Decision Steps’, uses a sequence of questions about the
589 study system to help an empiricist identify the most appropriate method for their work. The first
590 question asks whether the method must yield estimates of ND and RFD that can be directly
591 related back to Chesson’s inequality (Equation 1). Thus, Question 1 isolates the NFD method
592 from all others. This distinction is important since the outputs from this method cannot be
593 compared to the other four methods. However, the NFD method can accurately predict
594 coexistence based on mutual invasibility and, depending on the answers to the remaining
595 questions, it could be the most appropriate method for certain study systems. In particular, the
596 NFD method is the only option that does not require an empiricist to grow each species alone as
597 a monoculture (Question 2). This could be an advantage for study systems where experimental
598 manipulations are not feasible (e.g. long-lived species, protected habitats). Several empirical
599 studies have successfully implemented the NFD method in experiments (Levine and
600 HilleRisLambers 2009, Chung and Rudgers 2016), and other similar studies have used NFD in
601 observational studies (Adler et al. 2010).

602 The next question in the ‘Decision Steps’ is whether the empiricist knows which specific
603 resources the species are competing for and can quantify the dependence of each species’

604 population dynamics on those resources (Question 3). This question separates the four methods
605 for estimating ND and RFD into two separate groups. The phenomenological methods (Lotka-
606 Volterra and sensitivity method) are those that are informed by directly quantifying species
607 interactions, but which make no assumptions about the resources that species are competing for
608 (highlighted in green). The mechanistic methods based on MacArthur's CRM or Tilman's CRM
609 assume that species interact only by competing for shared resources (highlighted in blue). To use
610 either of the mechanistic methods, an empiricist needs to know which resources define niche
611 differences. In certain cases, it will not be possible for an empiricist to answer 'yes' to Question
612 3, because the resources required for species to grow are either not known or cannot be readily
613 quantified (e.g. non-essential resources). When one cannot answer yes to Question 3, then the
614 Lotka-Volterra and sensitivity methods may be appropriate because they can still quantify ND
615 and RFD even if the empiricist does not have a good understanding of which resource(s) species
616 are competing for, and thus, which resources define their niche. Because Question 3 is so
617 consequential, the remaining steps are particular to either the phenomenological or mechanistic
618 methods.

619 Deciding between the Lotka-Volterra method and the sensitivity method
620 (phenomenological methods, highlighted in green), depends on the answers to whether the
621 method must work for observational datasets (Question 4), whether it is necessary to
622 experimentally grow each species as at steady-state near its carrying capacity (Question 5), and
623 whether the method can be used to predict coexistence among 3 or more species simultaneously
624 (Question 7). An empiricist working with long-lived species or in protected habitats would likely
625 answer 'observational' to Question 4, eliminating the sensitivity method. In this case, the
626 empiricist would need to decide whether it is essential to obtain values of ND and RFD

627 compatible with the other four methods (requiring the Lotka Volterra method) or whether the
628 NFD method could be employed to predict mutual invasibility and thus coexistence. Similarly, if
629 an empiricist is unable to answer ‘yes’ to Question 5, she/he would be forced to use either the
630 Lotka Volterra method or the NFD method. Question 5 could be particularly important for
631 studies performed using slow-growing species where it is possible to parameterize the carrying
632 capacity term from a time series of species densities, but it would take too long for the species to
633 approach the carrying capacity to justify beginning an invasion by the other species. Lastly, the
634 two phenomenological methods differ in terms of whether they can predict species coexistence
635 among three or more species simultaneously (Question 7). While the Lotka-Volterra model can
636 be parameterized to obtain all pairwise competition coefficients for a pool of species, it has not
637 been applied to predicting coexistence of more than two species simultaneously. The sensitivity
638 method can be used beyond pairwise species interactions (Carroll et al. 2011); however, doing so
639 is limited to situations where all non-focal species can be considered in aggregate (e.g. species i
640 invading a community of $j + k + l$).

641 Deciding between the MacArthur and Tilman CRM methods (mechanistic methods,
642 highlighted in blue), is straightforward and depends on whether the resources that the species
643 compete for are abiotic and governed by a constant rate of supply (e.g. inorganic nutrients
644 consumed by plants) or biotic with their own population dynamics (Question 6). It is worth
645 noting that Tilman’s R^* concept has been extended to include competition mediated by predators
646 (e.g. P^* , (Tilman 1982)). However, to date, models including both predation and abiotic resource
647 competition have not been related to Chesson’s ND and RFD. Additionally, MacArthur’s
648 consumer resource model can theoretically work for more than two species at a time, but this has
649 not been demonstrated for the method based on Tilman’s consumer-resource model (Question 7).

650 Using this decision tree, an empiricist can determine which method(s) are appropriate for
651 their study system. Depending upon the study system or experimental constraints, an empiricist
652 may have multiple options for which method to use. In these cases, it can be useful to consider
653 the experimental requirements of each method (Table 1, *Experimental Requirements*) and the
654 tradeoffs among the methods in terms of their utility as discussed in Part 3.

655

656 *2.2 Experimental requirements*

657 In addition to the ‘Decision Steps’ outlined in Table 1, there are important practical
658 differences for the experimental or observational studies needed to quantify ND and RFD for
659 each method. The most important difference in study design among these methods is whether or
660 not they require the species to be grown together in order to make a prediction about coexistence.
661 The NFD method and the two phenomenological methods require each pair of species to be
662 grown together in at least one co-culture, but the mechanistic methods do not require these co-
663 cultures. This distinction means that only the mechanistic methods can be used to make
664 predictions about coexistence of species without performing pairwise competition experiments or
665 analyzing time series from co-cultures. For example, consider a typical competition experiment
666 involving a pool of three species (A, B, and C). The mechanistic methods can make predictions
667 about species coexistence for all pairwise combinations of the species (A+B, A+C, and B+C)
668 based solely on information about each species when grown individually. The phenomenological
669 methods, however, require at least one co-culture for each pairwise combination of species,
670 which means that information from monocultures and pairs A+B and A+C cannot be used to
671 make any prediction about coexistence for the pair B+C.

672 The need for species to be grown together in co-culture has important implications for the
673 total number of experimental treatments that would be required to quantify ND and RFD.
674 Depending on the study design, experiments using the phenomenological methods can require
675 more experimental treatments to predict pairwise coexistence among a pool of species than the
676 mechanistic methods do. For the phenomenological methods, the number of experimental
677 treatments required for all pairwise combinations of species increases exponentially with each
678 additional species being considered. In contrast, for the mechanistic methods the total number of
679 experimental treatments required increases linearly with the number of species being considered.
680 This is because the methods based on consumer-resource models do not require any direct
681 competition experiments in order to estimate competition coefficients (α_{ij}), while all of the
682 phenomenological methods require at least one co-culture for each species pair (and often more
683 than one) in order to quantify the competition coefficients. As a result, the relative efficiency of
684 the phenomenological versus mechanistic methods depends upon both the number of species
685 being considered and also the number of resources. When the number of species being
686 considered is small and the number of limiting resources is few, the difference in experimental
687 effort can be modest. For example, to predict pairwise coexistence among a pool of four species,
688 using the sensitivity method requires 16 experimental treatments (time series): 4 monocultures to
689 parameterize both maximum growth rate and carrying capacity and 12 invasions to parameterize
690 sensitivity (A invading B, B invading A, etc.). In contrast, using either of the consumer resource
691 models (two limiting resources) would require two experiments per species for a total of 8
692 experiments. If the mechanistic methods require parameterizing four or more limiting resources,
693 then the phenomenological methods may be more efficient for a pool of four species. However,
694 for larger pools of species the difference can be substantial. Obtaining pairwise estimates of ND

695 and RFD for a pool of 10 species requires between 55 and 180 treatments for the
696 phenomenological methods but as few as 20 treatments for mechanistic methods.

697 In addition to the number of experimental treatments required for each method, it is
698 important to consider the amount of effort and time required for each experimental treatment.
699 Specifically, the NFD, Lotka Volterra, and sensitivity methods require time series of species
700 densities in the experimental or observation study. In the case of the NFD and sensitivity
701 methods, these time series may be short in duration (i.e. at least one generation) and focused only
702 on population dynamics when species densities are very low or near the steady-state density of
703 monocultures. However, the Lotka-Volterra method requires longer time series in order to
704 parameterize both the interaction coefficients and carrying capacities. Longer time series in
705 monoculture and co-culture are more easily attainable for quickly-growing species like microbes
706 and invertebrates, but even short time series could be prohibitively arduous for slowly growing
707 species like trees.

708 Ultimately, the total effort and resources required for a study is jointly determined by the
709 method, number of species, number of limiting resources (if applicable), length of time series,
710 level of replication, and any other design elements. Using Table 1 as a guide, an empiricist
711 should be able to select a method and begin to design a study that satisfies their aims.

712

713 **Part 3. Tradeoffs Among Methods and Suggested Future Directions**

714 Having explained how to select and implement the five methods, we end the paper by
715 offering some advice for empiricists about how to navigate tradeoffs among the methods, how to
716 compare and synthesize measurements of ND and RFD from different methods, and lastly, key
717 future directions for implementing modern coexistence theory empirically.

718

719 *3.1 Tradeoffs between phenomenological and mechanistic methods*

720 Given the substantial differences in experimental design requirements and effort that are
721 required to execute the five methods described in Part 1 of the paper, it is highly likely that
722 empiricists will face choices that require tradeoffs when selecting a particular method for their
723 study system. The most obvious and important tradeoffs occur between the phenomenological
724 methods and the mechanistic methods, which differ in three important ways. First, the
725 phenomenological methods (i.e. the NFD, Lotka-Volterra, and sensitivity methods) make no
726 assumptions about the resources that species compete for. This could be beneficial for
727 empiricists who can still measure ND and RFD even if they lack detailed information about the
728 biological resources that species compete for. But the trade-off for this lack of knowledge is the
729 need for pairwise experiments to directly quantify ND and RFD. Second, a key disadvantage of
730 all three phenomenological methods is that they require each pair of species to be grown together
731 in competition, which causes the total effort to increase exponentially as more species are
732 considered. Third, the results of phenomenological experiments are specific to each pair of
733 species tested and cannot be generalized to interactions beyond that pair. Furthermore, the
734 predictions from the phenomenological methods are specific to the exact environmental
735 conditions, like resource density or resource supply rates, used in that experiment and cannot be
736 generalized outside of those same conditions.

737 But for those empiricists who can identify the resources that species compete for, use of
738 the mechanistic methods allows for potentially fewer experiments that are more easily
739 generalized to predict coexistence among all species in the focal species pool. Indeed, an
740 empiricist who is able to answer ‘yes’ to Question 3 in Table 1 could use a mechanistic method

741 to predict coexistence (or not) for not only the species pair of interest, but any and all species
742 pairs of interest based solely on experiments that are performed with each species grown alone in
743 monoculture. Importantly, the mechanistic methods also offer the ability to make predictions
744 about species coexistence under different environmental conditions. For example, Letten et al.
745 showed that the Tilman consumer resource model can be used to predict the ND and RFD at
746 different nutrient supply concentrations or dilution rates (Letten et al. 2017). The ability of the
747 mechanistic methods to handle some changes to environmental context, while limited, could be
748 useful for predicting how anthropogenic stressors (e.g. nutrient pollution) are likely to affect
749 species coexistence. The ability to make predictions about combinations of species without the
750 need to perform all pairwise competition experiments has already been touted as a benefit of the
751 mechanistic models (Tilman 1982), and it could be useful for addressing certain ecological
752 questions that do not always lend themselves well to manipulative experiments (e.g. invasions by
753 introduced species, coexistence of rare or endangered species).

754

755 *3.2 Comparing and synthesizing measurements of ND and RFD*

756 To date, only three of the four methods proposed for measuring niche and relative fitness
757 differences have been used empirically. No one, to our knowledge has used the MacArthur
758 consumer-resource model to quantify ND and RFD in any real system, despite publications
759 showing that it is possible. That means that most of our inferences about ND and RFD that have
760 been measured empirically stem from the phenomenological methods. Furthermore, we are
761 unaware of any study that has applied more than one method to the same empirical study system.
762 As such, we have no way to compare the performance of the methods empirically. Therefore, we
763 believe an important avenue for future research is to focus on measuring ND and RFD using

764 more mechanistic models, and for studies that measure ND and RFD using different methods in
765 the same study system so that we can compare results and attempt to demonstrate equivalence or
766 non-equivalence of these methods.

767 Even as we call for more mechanistic experiments and comparative studies, we caution
768 against the inevitable urge to synthesize ND and RFD in an informal data synthesis or more
769 formal meta-analysis. Although we have shown that all five existing methods should correctly
770 predict the qualitative outcome of coexistence, the methods are by no means mathematically or
771 practically equivalent. In some cases the methods will not yield the same ND and RFD, even
772 when applied to the same species and environmental conditions. Indeed, given the differences in
773 how the methods are implemented (Figures 1-5), there is no reason to expect, *a priori*, that the
774 quantitative values of ND or RFD measured for a particular group of organisms using one
775 method will produce quantitatively similar values of ND (or RFD) for that same group of
776 organisms using a different method. As such, the methods are not directly comparable, and the
777 measurements they produce should not be mixed-and-matched to produce some synthesized
778 estimate of the niche or fitness difference for, say, grassland plants.

779

780 *3.3 Future directions for implementing modern coexistence theory*

781 In our view, there are at least two important new directions that work on species
782 coexistence must go if Chesson's modern coexistence theory is to become widely implemented
783 and more practical. First, each of the empirical methods described in this review are focused on
784 fluctuation-independent mechanisms. To also include fluctuation-dependent mechanisms of
785 coexistence in Chesson's framework, we need to expand the scope of the five methods reviewed
786 here or even develop new empirical methods. To our knowledge, there have been limited

787 empirical studies that explicitly quantify the fluctuation dependent mechanisms, i.e. relative
788 nonlinearities and storage effects (but see (Angert et al. 2009, Letten et al. 2018)). Even so, it is
789 well-known that environmental fluctuations mediate species coexistence in some empirical
790 systems (Caceres 1997, Jiang and Morin 2007) and any modern theory of coexistence is
791 incomplete without them. It is also important to note that all of the methods developed to date are
792 only applicable to competitive communities, and cannot be applied to cases where species
793 facilitate the growth of each other (but see (Bimler et al. 2018)).

794 Second, empirical studies on coexistence need to move beyond prediction of pairwise
795 species interactions. Several authors have recently emphasized that modern coexistence theory is
796 under-developed for multi-species systems (Carroll et al. 2011, Levine et al. 2017). In theory, the
797 pairwise competitive hierarchy between species i versus j and j versus k might not directly
798 translate to species i and k , particularly when these species are engaged in intransitive
799 competition or higher-order interactions (Levine et al. 2017). In fact, none of the three
800 phenomenological methods (the NFD, Lotka-Volterra and sensitivity methods) can deal with
801 intransitive competition or higher-order interactions. Importantly, the emphasis to date on
802 pairwise interactions and experimentation means that intransitive competitive interactions and
803 higher-order interactions, if present, are unaccounted for in our understanding. Chesson's
804 coexistence framework has been a major advance for understanding coexistence among pairs of
805 species, and how to expand this framework to multi-species systems should be a priority for the
806 field.

807

808 **Author Contributions**

809 All three authors designed the synthesis and wrote the manuscript, FHC performed the analytical
810 derivations, CMG wrote the numerical simulation code and drafted the conceptual figures.

811

812 **Supporting Information**

813 In the supporting information section we provide: (A) Numerical Simulation of Experiments To
814 Measure ND and RFD and Predict Coexistence, (B) Relating the Sensitivity Method to
815 Chesson's Definition of ND and RFD Using the Lotka Volterra Model, and (C) Relating the
816 Negative Frequency Dependence Method to Chesson's ND and RFD. Annotated computer code
817 is provided as a separate file.

818

819 **References**

- 820 Abrams, P. 1980. Are Competition Coefficients Constant? Inductive Versus Deductive
821 Approaches. *The American Naturalist* 116:730-735.
- 822 Adler, P. B., S. P. Ellner, and J. M. Levine. 2010. Coexistence of perennial plants: an
823 embarrassment of niches. *Ecol Lett* 13:1019-1029.
- 824 Adler, P. B., J. HilleRisLambers, and J. M. Levine. 2007. A niche for neutrality. *Ecol Lett* 10:95-
825 104.
- 826 Angert, A. L., T. E. Huxman, P. Chesson, and D. L. Venable. 2009. Functional tradeoffs
827 determine species coexistence via the storage effect. *Proc Natl Acad Sci U S A*
828 106:11641-11645.
- 829 Barabas, G., R. D'Andrea, and S. M. Stump. 2018. Chesson's coexistence theory. *Ecological*
830 *Monographs* 88:277-303.
- 831 Bimler, M. D., D. B. Stouffer, H. R. Lai, M. M. Mayfield, and O. Godoy. 2018. Accurate
832 predictions of coexistence in natural systems require the inclusion of facilitative
833 interactions and environmental dependency. *Journal of Ecology* 106:1839-1852.
- 834 Caceres, C. E. 1997. Temporal variation, dormancy, and coexistence: A field test of the storage
835 effect. *Proceedings of the National Academy of Sciences* 94:9171-9175.

- 836 Carroll, I. T., B. J. Cardinale, and R. M. Nisbet. 2011. Niche and fitness differences relate the
837 maintenance of diversity to ecosystem function. *Ecology* 92:1157-1165.
- 838 Chase, J. M., and M. A. Leibold. 2003. Ecological niches: linking classical and contemporary
839 approaches. University of Chicago Press, Chicago.
- 840 Chesson, P. 1990. MacArthur's Consumer-Resource Model. *Theoretical Population Biology*
841 37:26-38.
- 842 Chesson, P. 2000. Mechanisms of Maintenance of Species Diversity. *Annual Review of Ecology*
843 and Systematics
- 844 31:343-366.
- 845 Chung, Y. A., and J. A. Rudgers. 2016. Plant-soil feedbacks promote negative frequency
846 dependence in the coexistence of two aridland grasses. *Proc Biol Sci* 283.
- 847 Gause, G. F. 1934. Experimental analysis of Vito Volterra's mathematical theory of the struggle
848 for existence. *Science* 79:16-17.
- 849 Godoy, O., and J. M. Levine. 2014. Phenology effects on invasion success: insights from
850 coupling field experiments to coexistence theory. *Ecology* 95:726-736.
- 851 Grover, J. P. 1991. Resource competition in a variable environment: phytoplankton growing
852 according to the variable-internal-stores model. *American Naturalist* 138:811-835.
- 853 HilleRisLambers, J., P. B. Adler, W. S. Harpole, J. M. Levine, and M. M. Mayfield. 2012.
854 Rethinking Community Assembly through the Lens of Coexistence Theory. *Annual
Review of Ecology, Evolution, and Systematics* 43:227-248.
- 855 Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. Princeton
856 University Press, Princeton, N.J.
- 857 Huisman, J., and F. J. Weissing. 1999. Biodiversity of plankton by species oscillations and
858 chaos. *Nature* 402:407-410.
- 859 Hutchinson, G. E. 1961. The Paradox of the Plankton. *The American Naturalist* 95:137.
- 860 Jiang, L., and P. J. Morin. 2007. Temperature fluctuation facilitates coexistence of competing
861 species in experimental microbial communities. *J Anim Ecol* 76:660-668.
- 862 Leibold, M. A. 1995. The Niche Concept Revisited: Mechanistic Models and Community
863 Context. *Ecology* 76:1371-1382.
- 864 Letten, A. D., M. K. Dhami, P. J. Ke, and T. Fukami. 2018. Species coexistence through
865 simultaneous fluctuation-dependent mechanisms. *Proc Natl Acad Sci U S A* 115:6745-
866 6750.

- 867 Letten, A. D., P. J. Ke, and T. Fukami. 2017. Linking modern coexistence theory and
868 contemporary niche theory. *Ecological Monographs* 87:161-177.
- 869 Levine, J. M., J. Bascompte, P. B. Adler, and S. Allesina. 2017. Beyond pairwise mechanisms of
870 species coexistence in complex communities. *Nature* 546:56-64.
- 871 Levine, J. M., and J. HilleRisLambers. 2009. The importance of niches for the maintenance of
872 species diversity. *Nature* 461:254-257.
- 873 MacArthur, R. 1970. Species packing and competitive equilibrium for many species. *Theoretical
874 Population Biology* 1:1-11.
- 875 MacArthur, R., and R. Levins. 1967. The limiting similarity, convergence, and divergence of
876 coexisting species. *The American Naturalist* 101:377-385.
- 877 Miller, E. T., and C. A. Klausmeier. 2017. Evolutionary stability of coexistence due to the
878 storage effect in a two-season model. *Theoretical Ecology* 10:91-103.
- 879 Narwani, A., M. A. Alexandrou, T. H. Oakley, I. T. Carroll, and B. J. Cardinale. 2013.
880 Experimental evidence that evolutionary relatedness does not affect the ecological
881 mechanisms of coexistence in freshwater green algae. *Ecol Lett* 16:1373-1381.
- 882 Oksanen, L., S. D. Fretwell, J. Arruda, and P. Niemela. 1981. EXPLOITATION ECOSYSTEMS
883 IN GRADIENTS OF PRIMARY PRODUCTIVITY. *The American Naturalist* 118:240-
884 261.
- 885 Rescigno, A., and I. W. Richardson. 1965. On the competitive exclusion principle. *Bull Math
886 Biol* 27:85-89.
- 887 Smith-Gill, S. J., and D. E. Gill. 1978. Curvilinearities in the Competition Equations: An
888 Experiment with Ranid Tadpoles. *The American Naturalist* 112:557-570.
- 889 Sterner, R. W., and J. J. Elser. 2002. Ecological stoichiometry: the biology of elements from
890 molecules to the biosphere. Princeton University Press, Princeton, NJ.
- 891 Tilman, D. 1977. Resource competition between planktonic algae: an experimental and
892 theoretical approach. *Ecology* 58:338-348.
- 893 Tilman, D. 1982. Resource competition and community structure. Princeton University Press,
894 Princeton.
- 895 Volterra, V. 1931. Variations and fluctuations of the number of individuals in animal species
896 living together. Pages 409-448 in R. N. Chapman, editor. *Animal Ecology*. McGraw Hill.

- 897 Yenni, G., P. B. Adler, and S. K. M. Ernest. 2012. Strong self-limitation promotes the
898 persistence of rare species. *Ecology* 93:456-461.
- 899 Yenni, G., P. B. Adler, and S. K. M. Ernest. 2017. Do persistent rare species experience stronger
900 negative frequency dependence than common species? *Global Ecology and*
901 *Biogeography* 26:513-523.
- 902

903

Decision Steps:	1: Must the method yield estimates of ND and RFD (i.e. Chesson's inequality)?	No	Yes	Yes	Yes	Yes
	2: Can you grow each species as a monoculture?	Yes or No	Yes	Yes	Yes	Yes
	3: Do you know the resources the species are competing for?	No	No	No	Yes	Yes
	4: What type of study will be used to parameterize the model?	Experimental or Observational	Experimental or Observational	Experimental	Experimental	Experimental
	5: Does the method require ≥ 1 species at steady-state near its carrying capacity?	No	No	Yes	No	No
	6: Is the resource abiotic (non-dynamic), biotic (dynamic), or undefined?	Undefined	Undefined or biotic (dynamic)	Undefined	Biotic (dynamic)	Abiotic* (non-dynamic)
	7: Can the method predict coexistence among 3 or more species simultaneously?	Yes	No	Yes	Yes	No
Method	Negative frequency dependence					
	Foundational paper for model		Lotka-Volterra	Sensitivity	MacArthur's CRM	Tilman's CRM
	Theoretical paper linking model to MCT		Volterra 1928	MacArthur 1970	MacArthur 1970	Tilman 1977
Experimental Requirements	Empirical paper using model for MCT		Chesson 2000	Carrol et al 2011	Chesson 1990	Letten 2017
	Does the method require the species to be grown together?		Godoy and Levine 2014	Narwani et al 2013	none	Letten 2017
	How many experimental treatments required to measure ND and RFD for all pairs of n species?		$n+[n(n-1)/2]$	n^2	$m*n$, where m= number of resources	$m*n$, where m= number of resources
Are time-series of species densities required?		Yes, enough generations to approach k	Yes, ≥ 1 generation	No	No	No

904

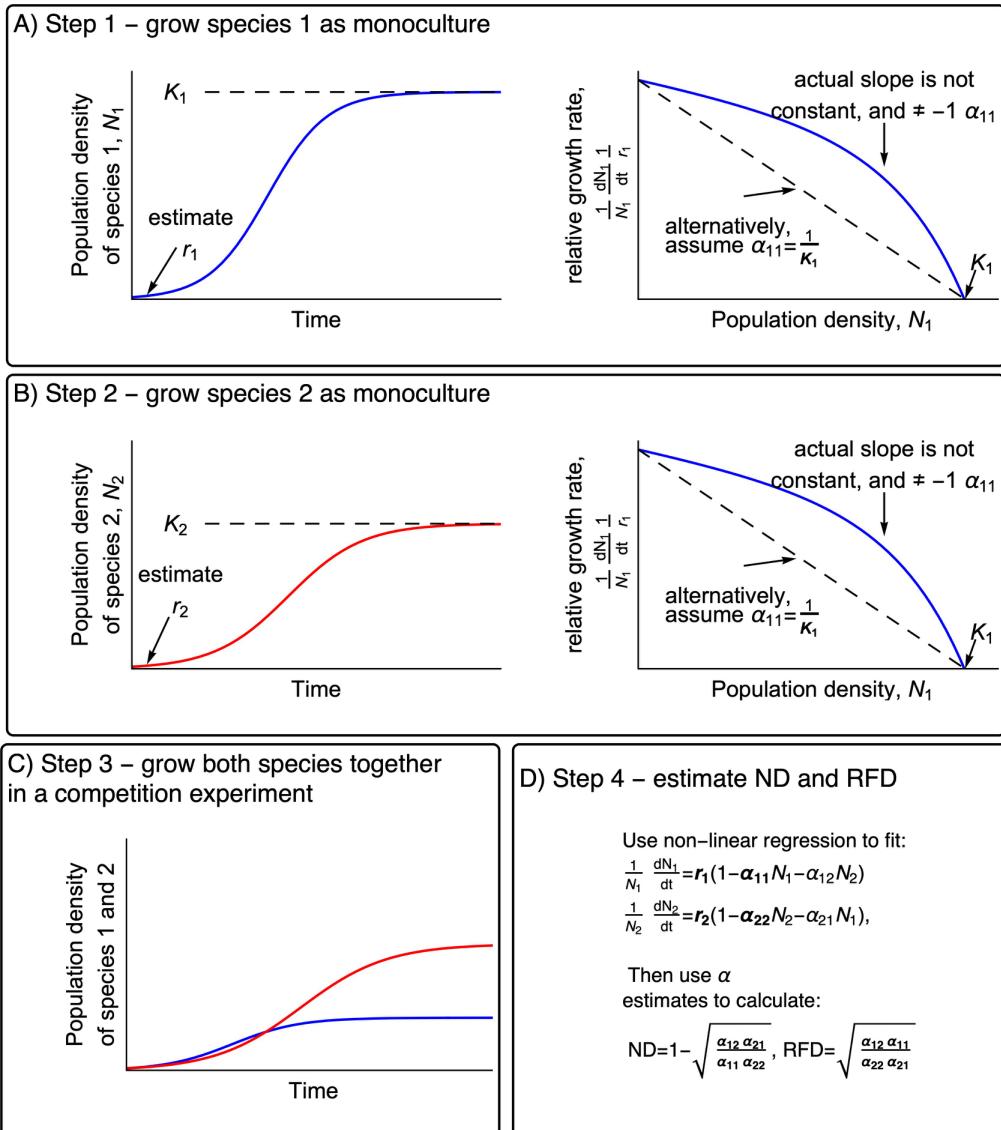
905 Table 1. A practical guide to help empiricists determine which method(s) are most appropriate for a given study system and question.

906 See Part 2 for a step-by-step explanation of this guide. * While consumer-resource models that include a second trophic level (e.g.

907 predators, P*) have been developed and used empirically, these approaches have not been linked to ND and RFD.

908

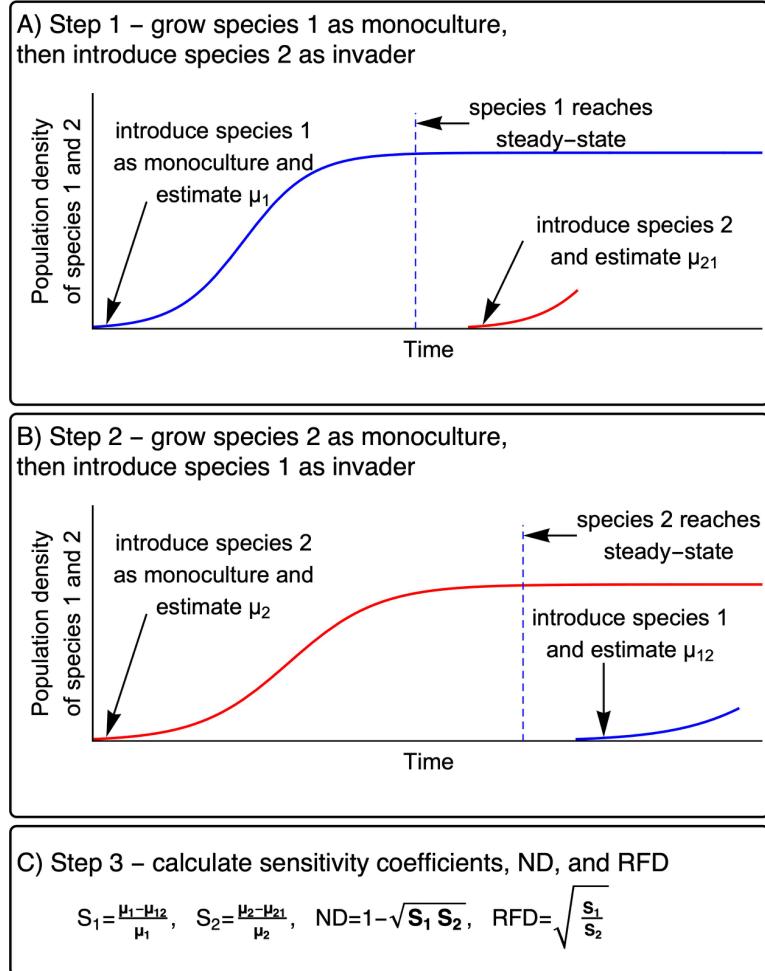
Figures



909

910 Figure 1. Conceptual plots illustrating how to use the Lotka-Volterra method to estimate ND and
911 RFD for use in Chesson's inequality (Equation 1). In each panel, unknown parameters are
912 displayed in regular typeface and previously-estimated parameters are listed in bold typeface. In
913 panels a and b, the left-hand plots show the time course of the experiment. In panels a and b the
914 right-hand plots show the relative growth rate as a function of population density – the slope of
915 this relationship is equal to the intraspecific competition coefficient (sign reversed).

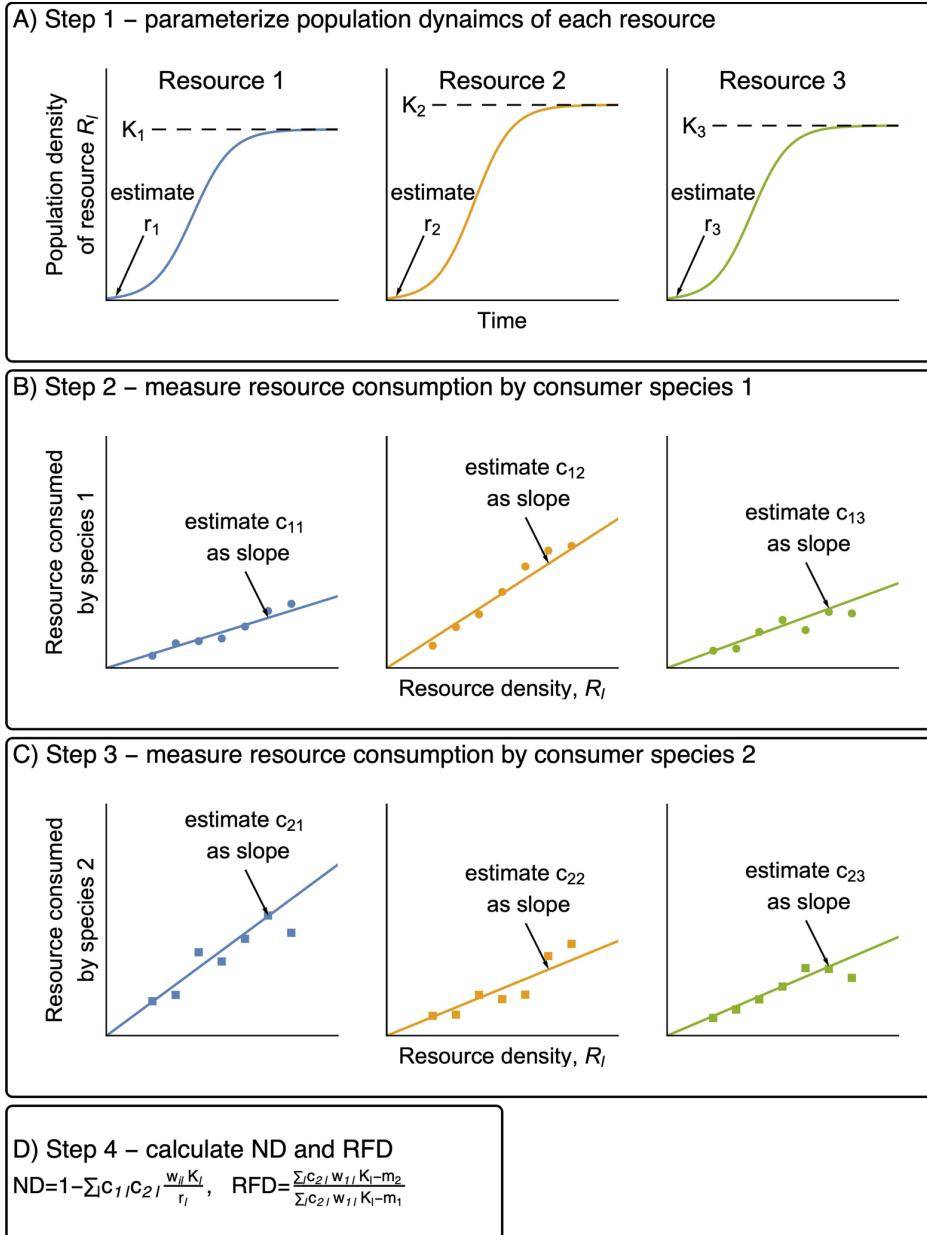
916



917

918 Figure 2. Conceptual plot depicting how to implement the sensitivity method in an experiment.

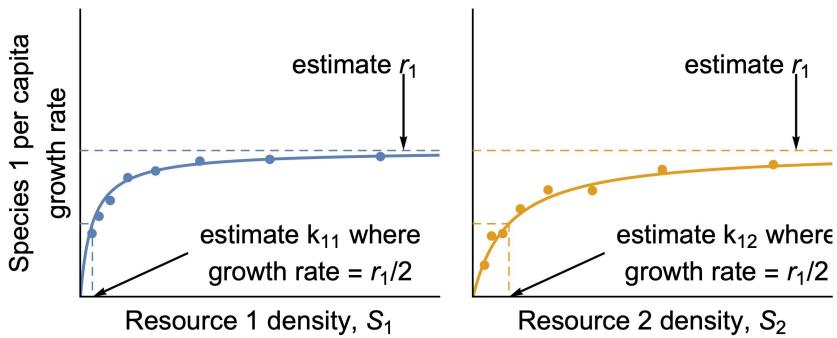
919



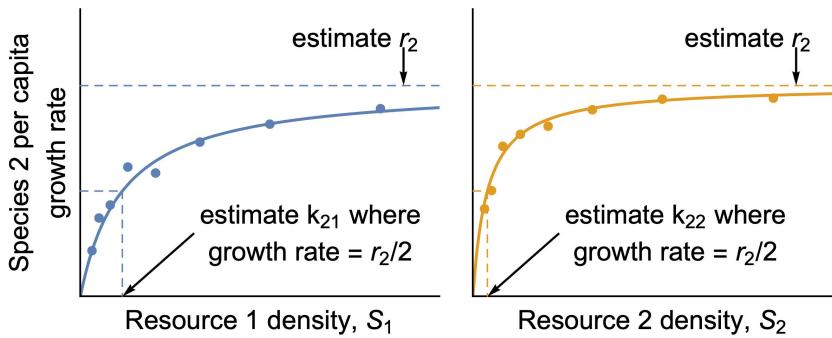
920

921 Figure 3. Conceptual plots depicting how the method based on Mac Arthur's CRM could be
 922 implemented. The yield term (w_{il} , increase in consumer units per unit prey resource consumed),
 923 can be estimated by measuring these changes for each combination of consumer and resource.

A) Step 1 – measure resource-use kinetics for consumer species 1



B) Step 2 – measure resource-use kinetics for consumer species 2



C) Step 3 – calculate ND and RFD

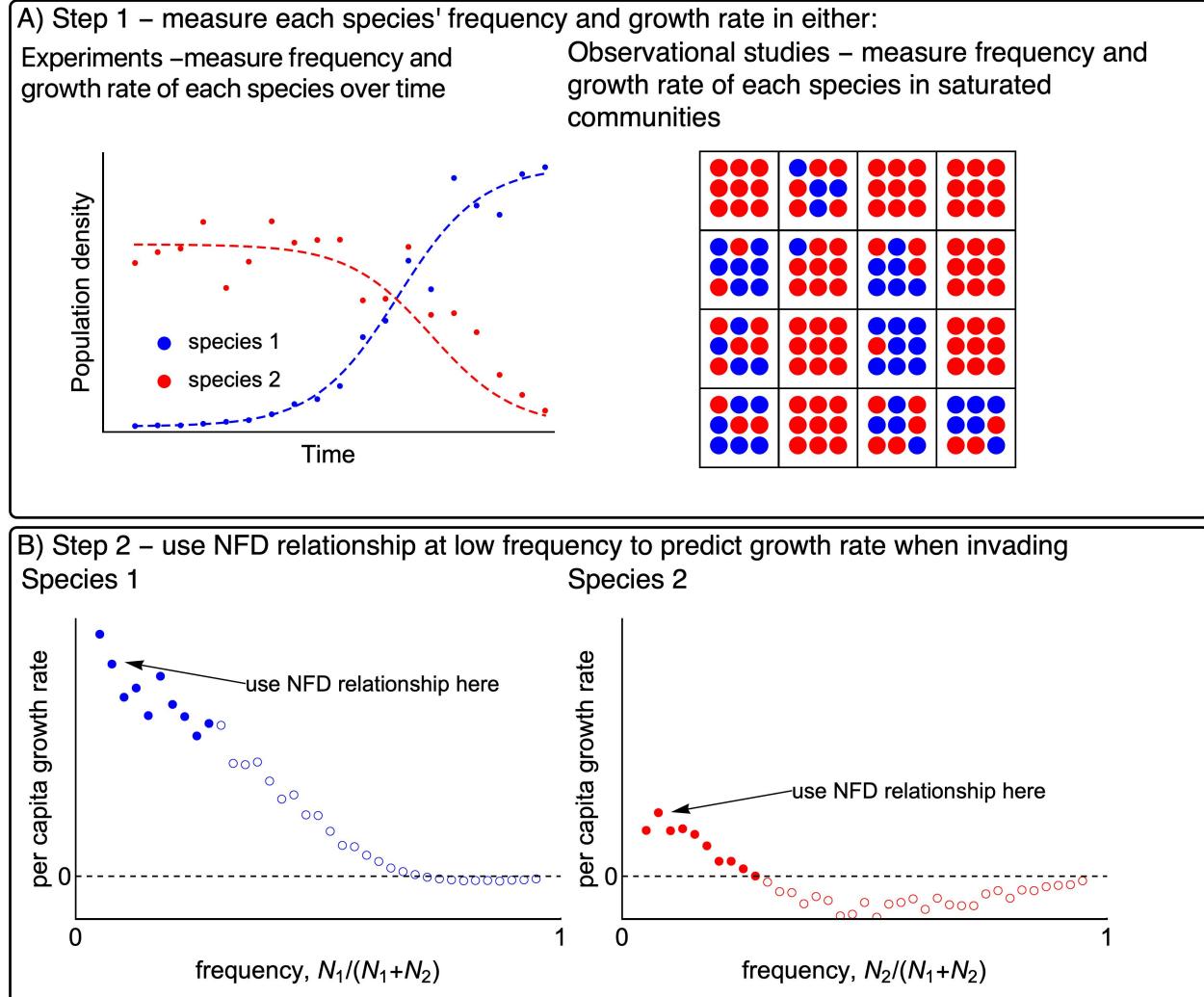
For the condition where species 1 is limited by resource 2 and species 2 is limited by resource 1:

$$ND = 1 - \sqrt{\frac{y_{12} y_{21}}{y_{11} y_{22}}}, \quad RFD = \sqrt{\frac{\left(\frac{Dk_{21}}{D-r_2} + S_1\right)^2 y_{11} y_{21}}{\left(\frac{Dk_{12}}{D-r_1} + S_2\right)^2 y_{12} y_{22}}}$$

924

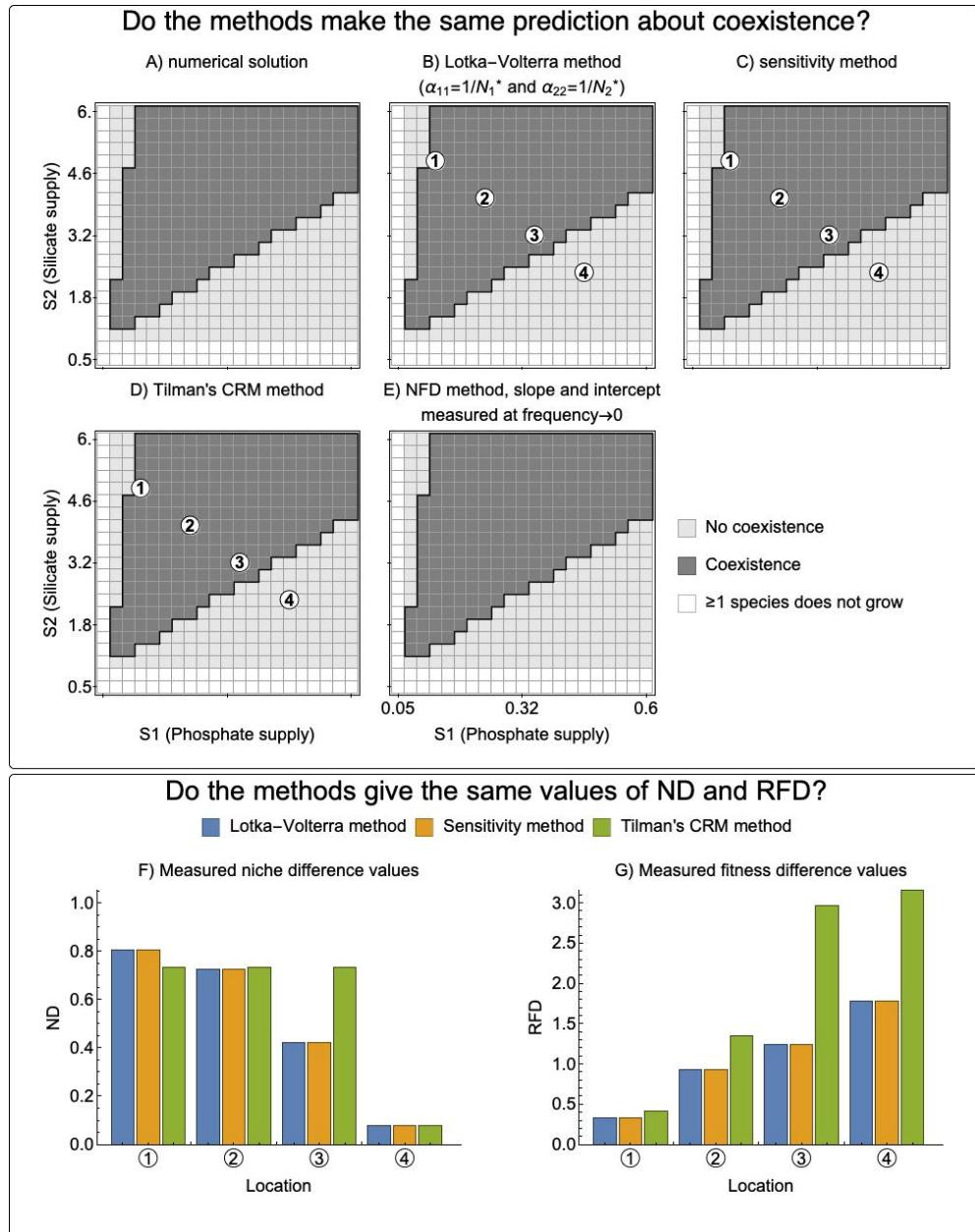
925 Figure 4. Conceptual plot depicting how to parameterize the method based on Tilman's
 926 consumer resource model. Panels A and B show the experiments needed to parameterize the
 927 maximum growth rates and Monod half-saturation constants for growth on each resource,
 928 separately for each species. The yield of each species on reach resource (y_{ii}) can be estimated by
 929 measuring the amount of resource consumed by a known number of individuals.

930



932 Figure 5. Conceptual diagram depicting how the NFD method could be implemented, either
933 through and experiment or using observational data from different communities. Panel A shows
934 two options for obtaining pairs of each species' frequency and its growth rate in a saturated
935 community. The first option is to track population densities over time in one or more competition
936 experiments. The Second option is to obtain pairs of frequency and growth rate from different
937 Communities or habitat patches in a natural ecosystem. Panel B depicts how the data from either
938 experiments or observational studies would be used to estimate the growth rate when frequency
939 approaches zero.

940



941

942 Figure 6. Comparison of four methods for predicting coexistence and estimating ND and RFD.

943 The methods were compared using simulations based on Tilman's parameterized CRM

944 (Supporting Information A). In panels A-E, the predictions for coexistence are compared against

945 the coexistence outcome based on numerical simulation. White shading means that at least one

946 species does not grow under that combination of resource supply concentrations, light gray

947 shading indicates that the method predicts that the species will not coexist, and dark shading

948 indicate indicates that the model predicts that the species will coexist. All four of the methods
949 give the correct predictions regarding coexistence across this region, but this is conditional on
950 specific limiting assumptions for the Lotka-Volterra and NFD methods (see Supplement C). The
951 methods did not give the same values for ND and RFD (f and g). The labeled locations in panels
952 F and G correspond to marked locations in panels b-d and show that the disagreement among the
953 methods is smaller toward the center of the parameter space that allows for coexistence. The raw
954 RFD values from the sensitivity method were converted to the same ordering as used in the other
955 methods (species i in the denominator rather than the species with the greater sensitivity).
956 Because the NFD method cannot be used to produce values of ND and RFD that are comparable
957 with the other four methods, only the predictions regarding coexistence are plotted.

958

959

960 **Supporting Information A: Simulation of Experiments To Measure ND and RFD and**
961 **Predict Coexistence**

962 In this supplement, we performed numerical simulations to compare the outcomes from
963 three methods for measuring ND and RFD and also the NFD method for predicting coexistence.
964 We used Tilman's parameterized consumer-resource model for two species of phytoplankton
965 competing for essential and non-substitutable resources (Tilman 1977). Annotated code for the
966 simulations is provided in a supplemental file. Simulations were performed using the function
967 NDSolve in Mathematica 11.2 (Wolfram Research), employing a variable step size. For each set
968 of resource supply concentrations, we performed four simulations: (1) species 1 as a
969 monoculture, growing from rare to near its equilibrium density; (2) species 2 as a monoculture,
970 growing from rare to near its equilibrium density; (3) species 1 at its equilibrium density, with
971 species 2 invading from rare; (4) and species 2 at its equilibrium density, with species 1 invading
972 from rare. Additionally, we performed numerical simulation where both species are introduced at
973 low densities and asked whether they coexist at the equilibrium. For each set of simulations, we
974 manipulated the supply concentration of the two resources in order to determine whether the
975 methods consistently agree.

976 We implemented the Lotka-Volterra method using information from all four simulations
977 described above. Simulations 1 and 2 were used to estimate r_i , K_i , and intraspecific interaction
978 coefficients α_{ii} . We estimated the intraspecific interaction terms using two different approaches
979 (Section 1.2). First, we estimated α_{ii} as the slope of the relative growth rate versus population
980 density (sign reversed) as the monoculture simulations approach equilibrium (Figure 1).
981 Alternatively, we used the assumption that $\alpha_{ii}=1/K_i$. We then used the parameter values from the
982 monocultures , along with simulations 3 and 4, to solve Equation 2 when each species is at low

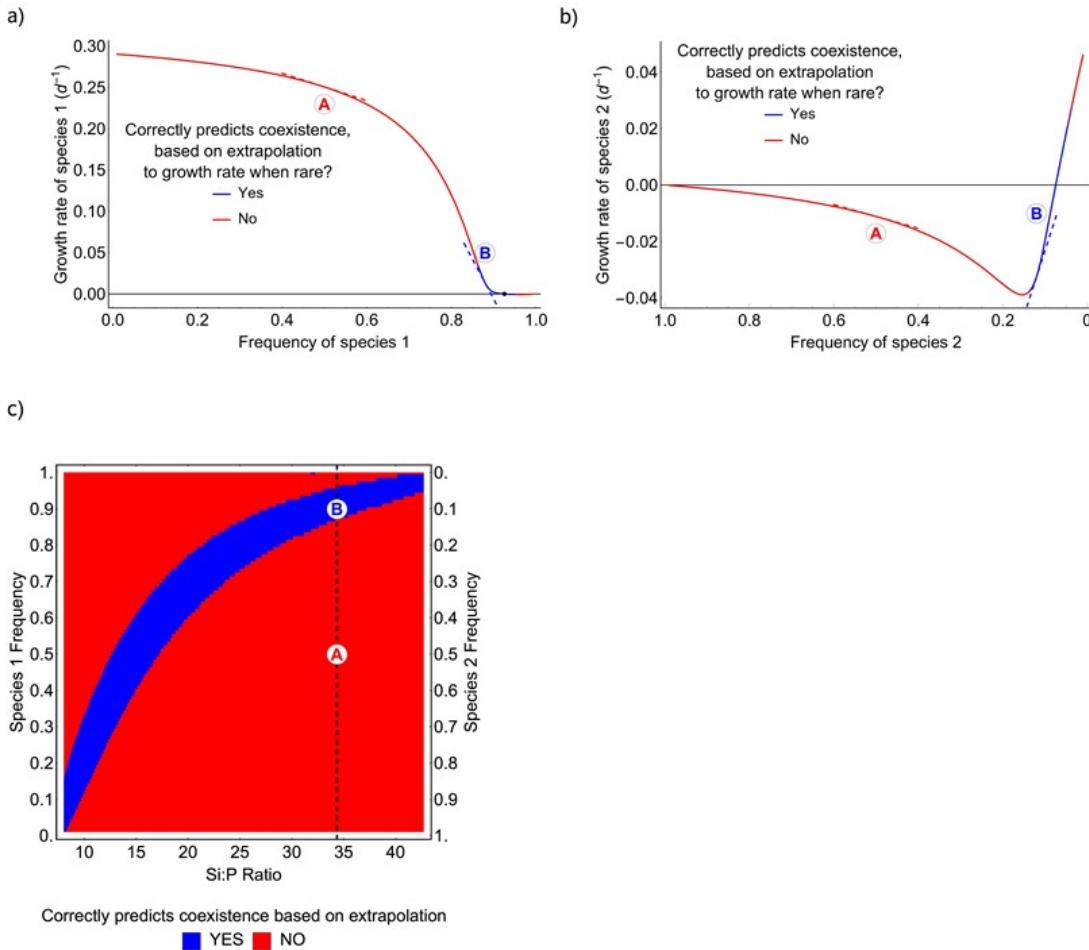
983 density and the other is near equilibrium. We used all four interaction coefficients to calculate
984 ND and RFD using Equations 3 and 4. We implemented the sensitivity method following
985 Equations 5 through 7, using output from all four simulations. The raw RFD values from the
986 sensitivity method were converted to the same ordering as used in the other methods (species i in
987 the denominator rather than the species with the greater sensitivity).

988 Next, we compared the four methods including the sensitivity method, the method based
989 on Tilman's CRM, the Lokta-Volterra method, and the NFD method using the numerical
990 simulations described above. Under specific assumptions, the methods gave the same prediction
991 regarding coexistence (Figure A2), though the methods did not produce consistent estimates of
992 ND and RFD (Figure 6).

993 As described in Appendix C, the NFD method cannot be used to get ND and RFD
994 estimates that are consistent with the other methods, but nonetheless this method can be used to
995 predict coexistence based on the same criterion. However, as shown in Figure 5 (using the
996 Lotka-Volterra model), accuracy of the NFD method depends on the range of frequencies used to
997 get the slope and elevation. To illustrate how the non-constant NFD slope is problematic in
998 predicting species coexistence, we used the simulations of mutual invasion (simulations 3 and 4),
999 described above, to construct pairs of each species' frequency and their growth rate in a saturated
1000 community. For all of these simulations, we used only supply concentrations of the resources
1001 that are known to allow for coexistence. For each value of a species frequency between 0 and 1,
1002 we calculated the slope of growth rate versus frequency. Figure A1 shows that this slope is not
1003 constant and actually changes sign depending on the species' frequencies used. Thus, using only
1004 the slope of the NFD relationship is inadequate to predict coexistence.

1005 Next, we used both the slope and elevation from the NFD method to extrapolate to
1006 frequency of 0 and predict whether the species is capable of invasion from rare (Figure A2
1007 panels A and B). Figure A2 shows that for supply conditions known to allow coexistence, the
1008 accuracy of the predictions from the NFD method depends on the range of frequencies over
1009 which the slope of NFD was measured. We discuss two points (A and B in Figure A2) to explain
1010 this effect. At the point Labeled “A”, the slope of NFD for species 1 predicts a positive growth
1011 rate as frequency approaches 0, but at the complementary frequency of species 2, the slope of
1012 species 2’s NFD predicts a negative growth rate when rare. However, based on the other 3
1013 methods, numerical simulation, and Tilman’s resource ratio theory, the species are predicted to
1014 coexist. Thus, measuring NFD under the red regions in Figure A2 will incorrectly predict
1015 exclusion even though the species will coexist. At the point labeled “B” in Figure A2, the slope
1016 and elevation of NFD for both species predicts a positive growth rate when rare. This region,
1017 depicted in blue, includes the equilibrium frequency for the two species. If an empiricist made
1018 their measurements between frequency of ~0.05 to ~0.85 for species 1, and used the slope of
1019 NFD, they would incorrectly predict that the species will not coexist. Since the frequency at
1020 which the species reach equilibrium depends on the resource supply ratio, there is no single
1021 frequency of the species that consistently leads to the correct predictions (Figure A2). While
1022 certain intermediate frequencies of the two species can be used to make accurate predictions, an
1023 empiricist would not know these frequencies without performing the competition experiments or
1024 examining frequency dependence across the entire range of frequencies. As a result, the only
1025 reliable way of implementing the NFD is to measure the slope and elevation for each species
1026 where its frequency approaches zero.

1027 For the NFD method, accurate predictions required that the slope of NFD was evaluated
1028 approaching frequency of zero for each species (i.e. invasion conditions). In Figure S2 d-f, we
1029 show that evaluating the NFD slope at other frequencies leads to the wrong predictions. We used
1030 the NFD plot to evaluate coexistence at three frequencies, including near 0% (panel e and h of
1031 Figure S2), 50% (panel f and i of Figure C2) and near 100% (panel g and j of Figure C2), and
1032 either with (panel e-g of Figure C2) or without considering the elevation in addition to the slope
1033 (panel h-j of Figure C2). We see that using the NFD slope evaluated at near 0% frequency will
1034 consistently yield accurate predictions of species coexistence that match the those of the other
1035 methods.
1036



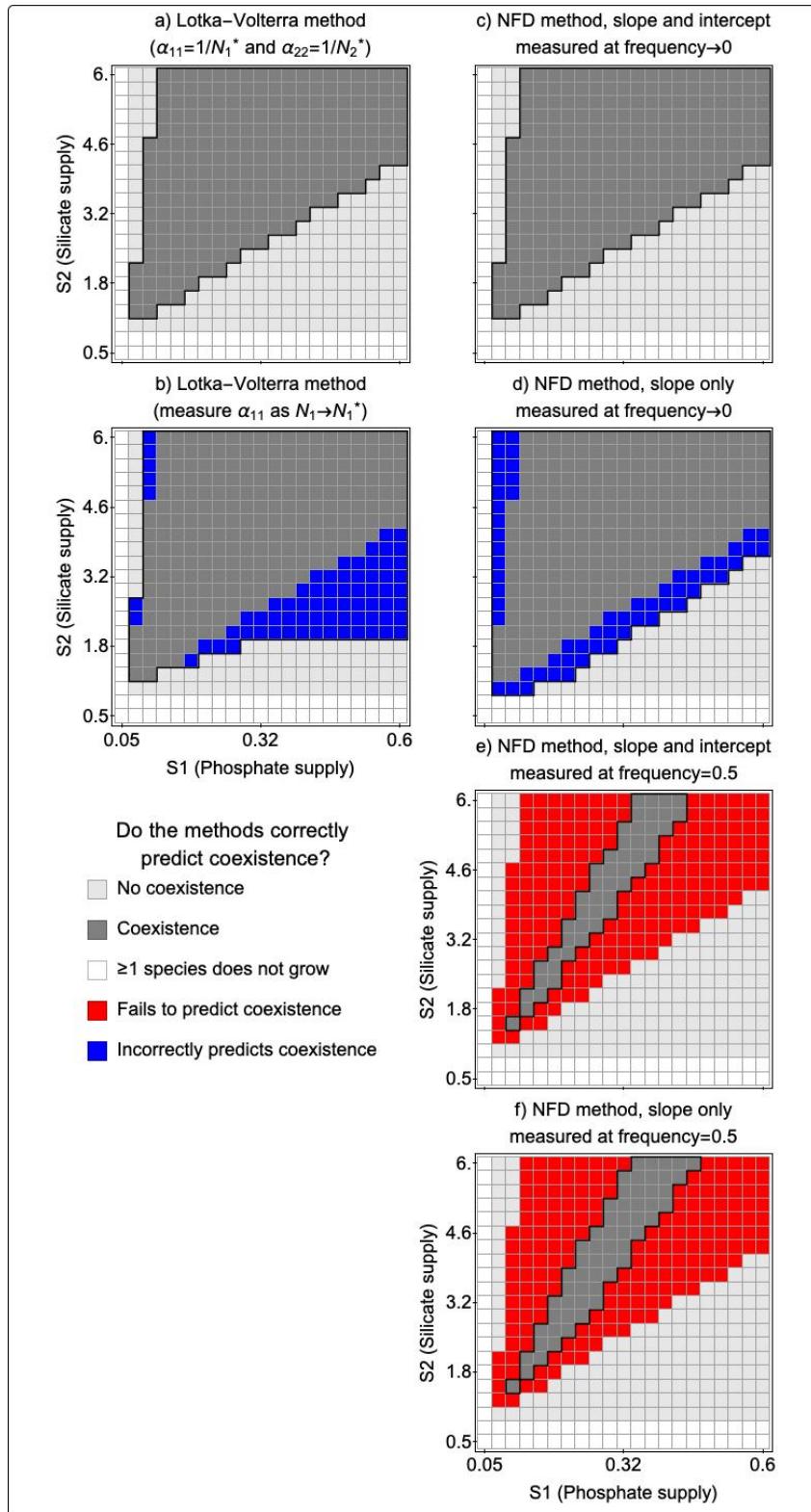
1037

1038 Supporting Information Figure A1. Results of simulation experiments using the NFD method.

1039 Panels a and b show per capita growth rate versus the frequency of species 1 and 2. At any
 1040 frequency of the two species, the NFD method requires that we use the slope to extrapolate and
 1041 estimate the growth rate when frequency approaches zero (the extrapolated vertical intercept).
 1042 For frequencies where this method predicts mutual invasibility for both species, i.e. species can
 1043 coexist, the lines are blue. For frequencies of the two species where the method leads to the
 1044 incorrect prediction, the lines are red. Both species have positive growth rates when their
 1045 frequency approaches zero, indicating that they are mutually invasible. Panel c shows the
 1046 accuracy of the NFD method as a function of the supply Si:P ratio and the frequency of the two
 1047 species at which the method was applied. The vertical dashed line represents the slice depicted in

1048 panels a and b. For all of the Si:P ratios shown in panel c, the species are mutually invasible and
1049 will coexist. This plot indicates that using NFD will often predict that the species will not
1050 coexist, when in fact they do coexist. This is important because without examining the full range
1051 of species frequencies in an experiment, one would not know whether and where the relationship
1052 between frequency and growth rate is non-linear.

1053



1054

1055 Supporting Information Figure A2. Results of simulation experiments comparing predictions
1056 from the Lotka-Volterra and NFD methods.

1057 **Supporting Information B: Relating the Sensitivity Method to Chesson's Definition of ND**
1058 **and RFD Using the Lotka Volterra Model**

1059 Here we show that sensitivity method is identical to the Lotka Volterra method given the
1060 specific limiting assumptions of the sensitivity method. To do this, we derive the sensitivity
1061 metric (S_i) from the Lotka-Volterra competition model (Equation 2). The μ_i in Equation 5 is the
1062 maximum per capita growth rate in monoculture, equal to r_i in Equation 2. The μ_{ij} is the invasion
1063 growth rate, so that we can replace N_j with species j's carrying capacity, K_j , and replace N_i with
1064 0, so that $\mu_{ij}=r_i(1-\alpha_{ij}K_j)$. Using this substitution, we show in Equation B1 that the sensitivity
1065 metric (S_i) is the equilibrium density of species j (K_j) multiplied by the *per capita* competition
1066 coefficient (α_{ij}).

1067
$$S_i \equiv \frac{\mu_i - \mu_{ij}}{\mu_i} = \frac{r_i - r_i(1 - \alpha_{ij}N_j^*)}{r_i} = \alpha_{ij}K_j \quad (\text{B1})$$

1068 Since the intraspecific competition coefficients in the Lotka Volterra model are equal to
1069 the inverse of the equilibrium population density for the monoculture ($\alpha_{ii} = \frac{1}{K_i}$), the sensitivity
1070 metric can be shown to be equivalent to the ratio of interspecific to intraspecific interaction
1071 coefficients (Equation B2).

1072
$$S_i = \alpha_{ij}K_j = \alpha_{ij} \frac{1}{\alpha_{ii}} = \frac{\alpha_{ij}}{\alpha_{jj}} \quad (\text{B2})$$

1073 From this substitution, we can relate the sensitivity metric to Chesson's ND (Equation B3), RFD
1074 (Equation B4), and use these estimates to assess the conditions for coexistence (Equation 1)

1075
$$\text{ND} = 1 - \sqrt{\frac{\alpha_{ij}\alpha_{ji}}{\alpha_{ii}\alpha_{jj}}} = 1 - \sqrt{S_i S_j} \quad (\text{B3})$$

1076
$$\text{RFD} = \sqrt{\frac{\alpha_{ij}\alpha_{ii}}{\alpha_{jj}\alpha_{ji}}} = \sqrt{\frac{S_i}{S_j}}. \quad (\text{B4})$$

1077 **Supporting Information C: Relating the Negative Frequency Dependence Method to
1078 Chesson's ND and RFD**

1079 Here we show that in order for the NFD slope to be constant, the community density
1080 must be both saturated and fixed across all frequencies of the species. To do so, we attempt to
1081 derive the NFD slope and intercept from the two species Lotka-Volterra competition model
1082 (Equation 1). Since there is no variable representing each species' frequency in the Lotka-
1083 Volterra model, we have to assume a fixed community density, B . This assumption also satisfies
1084 the assumption of the NFD method that the community density is always saturated. Fixing the
1085 community density makes the interspecific density dependence, α_{ij} , equivalent to frequency
1086 dependence (Adler et al. 2007), and allows species' frequency to be represented by N_i/B . The
1087 two-species Lotka-Volterra competition model can then be rewritten as follows

1088
$$\frac{1}{N_i} \frac{dN_i}{dt} = r_i \left(1 - \frac{\alpha_{ii}N_i + \alpha_{ij}(B - N_i)}{B} \right) \quad (\text{C1})$$

1089 , where B is the fixed community density and one unit decrease of N_i will lead to one unit
1090 increase of N_j . From Equation C1, we derive the NFD slope and intercept in the following
1091 equations.

1092
$$\text{NFD slope} \equiv \frac{\partial \frac{1}{N_i} \frac{dN_i}{dt}}{\partial \frac{N_i}{B}} = \frac{\partial \frac{1}{N_i} \frac{dN_i}{dt}}{\frac{1}{B} \partial N_i} = B \frac{\partial \frac{1}{N_i} \frac{dN_i}{dt}}{\partial N_i} = r_i(\alpha_{ij} - \alpha_{ii}) \quad (\text{C2})$$

1093
$$\text{NFD intercept} \equiv r_i(1 - \alpha_{ij}) \quad (\text{C3})$$

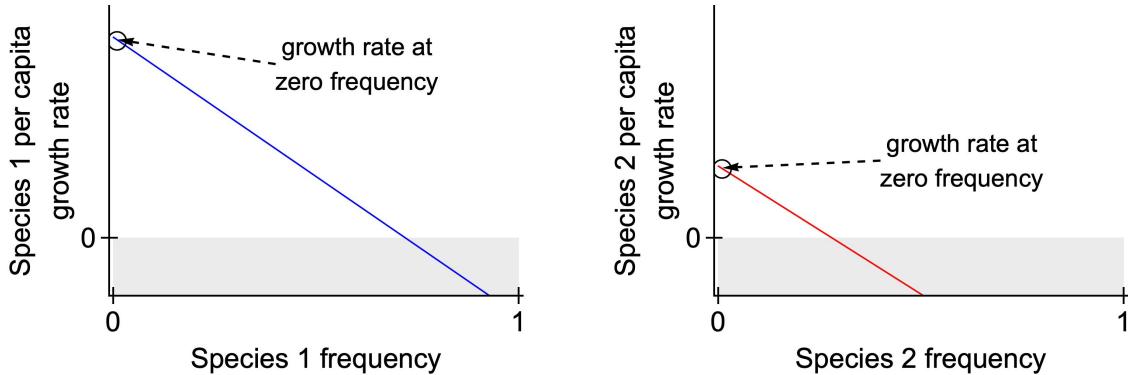
1094 In Equation C2 the NFD slope becomes constant, which means that species' per capita growth
1095 rate linearly depends on its frequency.

1096 In addition, from Equations C2 and C3, we argue that both NFD intercept and slope
1097 should be used with caution in evaluating Chesson's inequality. First, the NFD intercept
1098 represents whether species can successfully invade a steady-state population of its competitor at

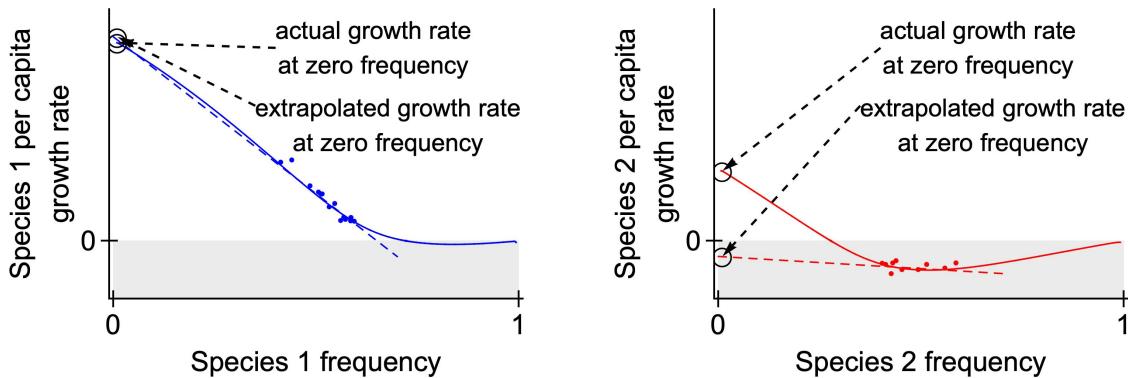
1099 its carrying capacity, so it can be used to accurately assess mutual invasibility. However, neither
1100 the difference nor the ratio of two species' NFD intercepts (Equations C2 and C3) take an
1101 analogous form to Chesson's definition of ND and RFD. It is worth noting, however, that the
1102 slope of NFD relationship has been used to represent ND for annual plant communities (Yenni et
1103 al. 2012, Yenni et al. 2017). Thus, while the NFD method can correctly predict mutual
1104 invasibility, the NFD intercept and slope should not be interpreted as RFD and ND in order to
1105 evaluate Chesson's inequality.

1106 The utility of the NFD method depends on its ability to correctly predict whether species
1107 have positive growth rates when their frequencies approach zero. If the relationship between a
1108 species frequency and its growth rate is non-linear, however, then the accuracy of the NFD
1109 method is critically dependent on the range of species frequencies used by an empiricist. In
1110 Figure C1 we show that the NFD relationship is non-linear even when the underlying population
1111 dynamics are governed by the Lotka-Volterra model. The result of this non-linearity is that,
1112 depending on the range of species' frequencies used to estimate the NFD slope and intercept, this
1113 method can give inaccurate predictions.

A) Ideal scenario – growth rate is linearly related to frequency



B) Common scenario – growth rate is non-linearly related to frequency



1114

1115 Figure C1: Panel A shows a hypothetical situation where species frequencies are linearly related
1116 to their growth rate and an empiricist can extrapolate to predict growth rates at frequency of zero
1117 and diagnose mutual invasibility. Panel B shows the more likely scenario in which growth rates
1118 are non-linearly dependent upon species frequencies. These plots were made using numerical
1119 simulation of the two-species Lotka Volterra Model, using parameter values that should allow
1120 for coexistence (at frequency of 0.72 for species 1). The points in Panel B represent empirical
1121 measurements collected at intermediate frequency of both species. Using those measurements

1122 and extrapolating to zero frequency yields the incorrect prediction that the species will not
1123 coexist.