An Empiricist’s Guide To Modern Coexistence Theory

1. **Introduction** 
   1. Why we care about species coexistence…
   2. Big-picture of why MCT is the best tool we have for evaluating coexistence of species.
   3. Multiple methods have been proposed for evaluating the potential for coexistence among species using empirical data. Although each of these methods is designed to evaluate Chesson’s inequality criterion, it remains unclear 1) whether these methods are functionally equivalent and 2) which methods are most appropriate for a given study system.
   4. We compare five different methods that have been proposed for evaluating coexistence and ask ….
2. Should mention it is empiricists’ guide for fluctuation independent mechanisms of the modern coexistence theory. Fluctuation independent means that species does not require fluctuations over time either in population densities or environmental factors for their coexistence. The five methods we reviewed here should be used to predict coexistence when time or space induced environment fluctuations are negligible in mediating species’ coexistence (Chesson 1994, Barabás et al. 2018). In other word, this means the methods here are calculating the first order approximation of invasion growth rate of the focal species, which is the key metric leading to mutual invasibility criteria.
3. **Describing the theoretical background for the five commonly used empirical methods**

The frequency independent part of the modern coexistence framework is funded on Chesson’s key insight toward the mutual invasibility criterium. The mutual invasibility criterium means, for any two species (e.g. *i* and *j*) to stably coexist each of them must be able to invade the carrying capacity state of other one from rare. In other word, when the mutual invasibility criteria is met, both species have positive invasion growth rate. To have positive invasion growth rate, the species must limit themselves more than they limit their competitor, i.e. *per capita* intra-specific competition coefficients (*αii* or *αjj*) must be greater than *per capita* inter-specific competition coefficient (*αij* or *αji*). Chesson showed that the mutual invasibility criteria i.e. *αii* > *αij* and *αjj* > *αji*, can be expressed in a different fashion (Chesson 1990, 2000). Note that, when making the following deductions, Chesson assumes that species’ population dynamics can be described by Lotka-Volterra model and the *per capita* competition coefficients (*α*) are density independent. First, Chesson defined the niche overlap (*ρ*) as to describe how similar the two competing species are in terms of using resources. First, Chesson defined the niche overlap (*ρ*) as to describe how similar the two competing species are in terms of using resources. The niche difference (ND) is thus 1 − *ρ*. Second, Chesson defined relative fitness difference (RDF; the *fi* is the same as the *ki* in Chesson 1990) as to describe which species should exclude the other one if they completely overlap their resource use. Accordingly, the product of *ρ* and RFD is the ratio of inter- specific to intra-specific competition coefficients, i.e. . When intra-specific competition of species *j* is greater than inter-specific competition of species *i* (), so that . By the same logic, when intra-specific competition of species *i* is greater than inter-specific competition of species *j* (), . Consequently, the mutual invasibility criteria for stable coexistence can be rewritten as the following inequality for coexistence.

(1)

We now introduce five commonly used methods to empirically measure *per capita* competition coefficients (*α*) and thus ND, RFD and finally Chesson’s inequality for coexistence (eqn. 1).

* 1. **Negative Frequency dependency (NFD) method**

The most commonly used empirical method to measure *per capita* competition coefficients (*α*) is the negative frequency method (Adler et al. 2007, Levine andHilleRisLambers 2009). One key assumption of NFD method is that a community is saturated with respect to biomass, so all the resources or niche are being occupied by either the focal species *i* or it competitor *j*. Under this assumption, decreasing the frequency of the focal species *i* frees the resources for its competitor, which increases in abundance. Therefore, decreasing the frequency of focal species *i* means the focal species *i* are competing with less individuals of its own kind than individual of the competitor. If the focal species *i* is more limited by its own than by its competitor, i.e. the intra-specific competition coefficient is greater than the inter-specific competition coefficient, increasing relative frequency of the focal species *i* should decrease its own *per capita* growth rate and frequency should thus have a negative slope. If inter-specific competition is greater than intra-specific competition, frequency dependency should be positive as the *per capita* growth rate of the focal species *i* increases with its own frequency. Therefore, as long as the frequency dependency is negative (NFD slope is negative) AND BOTH SPECIES HAVE POSITIVE GROWTH RATES WHEN RARE, species should be guaranteed to stably coexistence. Adler et al [2007] expanded upon this definition and showed that, in addition to the requirement for a negative slope of frequency dependence, both species must have positive approaching zero frequency (i.e. their intercepts are positive in figure 1). This constraint is the definition of mutual invasibility, and as long as the relationship between a species frequency and its growth rate is linear, knowing the growth rate and slope at an intermediate frequency could allow an empiricist to extrapolate to predict the growth rate approaching zero frequency and determine whether both species are mutually invasible.

We note that there are two issues associated with the NFD method. First, the magnitude of negative frequency dependency (the slope) is not equivalent to either intra- or inter-specific competition coefficients but a rather complex combination between both. As a result, the slope of NFD cannot be used to determine ND and RFD. Second, while the NFD method is consistent with Chesson’s coexistence requirement that coexisting species are mutually invasible, we argue that the assumption that NFD is constant across frequencies (constant NFD slope) is problematic in practice.

To show that NFD slope is a complex combination of intra- and inter specific competition coefficients, we attempt to derive the NFD slope based on the the Lotka-Volterra model. Conceptually, when the per capita growth rate is being plotted against the frequency of the focal species, the NFD slope is actually the ”per %” impact on the per capita growth rate. To show that negative frequency dependency (NFD) slope cannot be used directly to measure competition coefficients (*α*), we attempt to derive the NFD slope from the Lotka-Volterra model. We found that, the NFD slope cannot be readily derived from the Lotka-Volterra model without making further assumptions. First, there is only density term but no frequency term in the Lotka-Volterra model. Only when the community density is fixed, the density dependency, *αij*, is equivalent to density dependency (Adler et al. 2007). In addition, since the density dependency, *αij*, is modeled in *per capita* fashion, one-to-one conversion between the focal species *i* and its competitor also needs to be assumed. By doing so, the Lotka-Volterra competition model can be rewritten as followed.

(2)

In equation 2, *B* is the fixed community density and one unit decrease of *Ni* will lead to one unit increase of *Nj*. Note that this *B* is an arbitrarily defined constant describing a fixed community density and has nothing do to with the equilibrium of any of the species. To calculate the negative frequency dependency slope, we take derivative of equation 6 in terms of *Ni/B*.

(3)

This equation 3 describe the change of species *i*’s *per capita* growth rate with respective to the change of its own frequency in a community (Fig. 1). From equation 3 the NFD slope depends on a combination of *per capita* growth rate (*ri*) and the fixed community density (*B*) in addition to the intra- and inter-specific competition coefficients. From this equation, we first see that NFD slope is negative as long as the intra-specific competition (*αii*) is greater than the inter-specific competition (*αij*). Additionally, higher *per capita* growth rate of a species and higher community density (e.g. in the later more mature stage of the community) would lead one to estimate stronger frequency dependency (Fig. 1). Most importantly, although NFD slope has been used to estimate species coexistence empirically for annual plant communities (e.g. Godoy et al. 2014), it should be interpreted with caution as it is related but not equivalent to the competition coefficients (*αii* and *αij*) and thus should not be directly used to calculate ND and RFD in order to predict species coexistence.

To demonstrate that non-constant intra- and inter specific competition coefficients can lead to non-constant NFD slope and incorrect prediction of species coexistence, we used numerical simulation of a well-known two species consumer-resource model [Tilman 1977] to show that a species growth rate is non-linearly related to its frequency, even when the community is saturated with respect to biomass and both species are known to coexist (Figure S1). Specifically, evaluating NFD at different frequencies of the species in a saturated community can lead to the incorrect prediction regarding coexistence. When frequency dependence is evaluated when each species is approaching a frequency of 1, the method fails to predict that both species have positive growth rates when rare and, thus, fails to predict coexistence. When frequency dependence is evaluated when both species are at frequency of 0.5, the method fails to predict coexistence in most cases. We found that NFD led to correct predictions only when evaluated at frequency approaching zero for each species. While these simulations are by no means exhaustive with respect to the types of species interactions found throughout ecology, such non-linear frequency dependence patterns are likely when population dynamics are governed by non-linear relationships with either biotic (e.g. type II feeding response) or abiotic (e.g. temperature, nutrients) factors. This finding means that for an empiricist to use the NFD method, they would need to either 1) measure the growth rate of each species across the full range of frequencies to establish that the growth rate of each species is linearly related to its frequency or 2) evaluate the growth rate of each species when rare (i.e. directly demonstrate mutual invasibility).

* 1. **Fitting the classic Lotka-Volterra model**

The second empirical method is to parameterize the Lotka-Volterra model with empirical data. In the classic Lotka-Volterra model, the *per capita* growth rate of species i can be described by the following equation.

(4)

In the above equation, *Ni* and *ri* are the the density and the intrinsic growth rate of species *i* respectively. The *αii* is the *per capita* intra-specific competition coefficient, which describes the *per capita* effect of species *i* on the *per capita* growth rate of species *i*. The *αij* is the inter-specific competition coefficient, which describes the *per capita* effect of species *j* on the *per capita* growth rate of species *i*. For any two species (e.g. *i* and *j*) to stably coexist, the mutual invasibility criteria must be met, i.e. *αii* > *αij* and *αjj* > *αji*.

To use the Lotka-Volterra model to empirically predict coexistence for species *i* and *j*, one must first estimate six different parameters that are used in the Lotka-Volterra model: intrinsic growth rate of each species (*ri* and *rj*), *per capita* intra-specific competition coefficients (*αii* and *αjj*), and *per capita* inter-specific competition coefficients (*αij* and *αji*). In theory, this could be accomplished using maximum likelihood method from a single co-culture time-series dataset, where both species are introduced at low density and allowed to grow to steady-state. However, in practice, it is difficult to parameterize all six variables from a single time-series. An alternative would be to use three datasets for each species pair: each species as a monoculture and one co-culture of the two species (Fig. 2). With the empirically estimated competition coefficients, one can calculate the niche difference (ND), relative fitness difference (RFD) and access Chesson’s coexistence inequality. An important consideration is that, the Lotka-Volterra model assumes constant intra- and inter-specific competition coefficient with respect to population sizes and time, which means the first individual and the last individual have the same *per capita* effect on the growth rates.

* 1. **Sensitivity method**

The sensitivity method is another method proposed to measure niche difference (ND) and relative fitness difference (RFD) without explicitly estimating the inter-specific competition coefficients (Carroll et al. 2011). Instead, sensitivity method relies on the effect of inter-specific competition on the population dynamics. Sensitivity method is also based on the idea of mutual invasibility that species need to be able to invade its competitor from rare in order to coexist (Fig. 3). The rationale is that when the focal species *i* overlaps its niche with its competitor, the *per capita* growth rate of the focal species *i* should be lower when invading its competitor than when growing alone from rare. The more the focal species *i* overlaps its niche with its competitor, the lower invading *per capita* growth rate species *i* should have. The sensitivity metric (*Si*) is being designed to quantify such decrease and thus to quantify the effect of the competitor on the focal species *i*. Specifically, the sensitivity metric (*Si*) compares the focal species *i*’s *per capita* growth rate when invading its competitor versus the focal species *i*’s *per capita* growth rate when growing alone from rare. The difference between the two growth rates is then the proxy of the effect of competitor on the focal species *i*. If either species has a growth rate less than or equal to zero when invading, the *Si* is then no less than 1, which means that there will not be coexistence (mutual invasion criterion). According to Carroll et al. 2011, sensitivity metric (*Si*) is calculated by the following formula.

(5)

In equation 5, *μi* is the *per capita* growth rate of species *i* when growing alone from rare and *μij* is the *per capita* growth rate of species *i* when it competitor (species *j*) is at its carrying capacity. Such mutual invasion experiment not only directly tests for mutual invasibility, but also empirically estimates ND and RFD [Carroll 2011]. Therefore, the mutual invasion experiment used to estimate the sensitivity metric (*Si*) is not restricted to a specific competition model, and intuitively connects competition and population dynamics [Carroll 2011]. Here, we further this intuitive connection by showing that the sensitivity metric actually describe the impact of the entire competitor population on the *per capita* growth rate of focal species *i*. From equation 5, the reduction of species *i*’s per capita growth rate, i.e. the nominator, is actually caused by the entire population of the other species *j* because the invasion growth rate (*μij*) is measured when the other species *j* is at the carrying capacity. Accordingly, the sensitivity (*Si*) measures the ”population” impact of species *j*, but not the per capita impact of species *j* on the focal species *i*.

To show that sensitivity is actually the population level impacts, not the *per capita* level, we derive the sensitivity metric (*Si*) from the classic Lotka-Volterra competition model (equation 4). The *μi* and *μij* in equation 5 are therefore *ri* and *μij* is respectively. Accordingly,

(6)

From equation 6, we see that sensitivity (*Si*) is the equilibrium density of species *j* () times the *per capita* competition coefficient (*αij*). The sensitivity thus is a measurement of the overall population effect of species *j* on focal species i but not the ”*per capita*” effect of species *j*. Small tweak should be implemented when using the sensitivity method to estimate per capita inter-specific competition coefficients (*αij*).

Moreover, Carroll et al. 2011 verbally argued that niche difference (ND) and relative fitness difference (RFD) can be defined as the geometric mean and standard deviation of the sensitivity metric (*Si*) respectively. Here we show that geometric mean and standard deviation of sensitivity metrics are theoretical valid definitions of ND and RFD. In the Lotka-Volterra model, species’ density at the equilibrium (*N\**) is actually . Therefore, sensitivity (*Si*) can be expressed as, which represents the inter-specific competition scaled on intra-specific competition coefficient. According to Chesson (1990), niche overlap () is defined as . can then be expressed as *,* which is the geometric mean of sensitivity *Si* and *Sj*. The niche difference (ND) is therefore *.* In addition, , and is the definition of the inverse of relative fitness difference of species *j* over species *i* (RFD; in Chesson 1990).

With the above derivation, we can also derive the coexistence criteria using the sensitivity metric. From the Lotka-Volterra model model, intra-specific competition (*αii*) must be greater than the inter-specific competition (*αij*) to guarantee stable coexistence. Therefore, we can have the following deduction.

, so (7)

, so (8)

Combining equation 7 and 8, we have an inequality for coexistence expressed with sensitivity metrics, , which is in the same form as in Chesson’s coexistence framework. In brief, the sensitivity metric (*Si*) is not equivalent to the competition coefficient (*αij* in the Lotka-Volterra model), but due to its methmatic attributes, it can be used to calculate ND and RFD and to predict coexistence directly.

* 1. **Fitting the MacArthur’s consumer resource model**

The consumer resource model proposed by MacArthur in 1970s is used to describe how species compete for different prey resources (MacArthur 1969, 1970). This model is reorganized into a Lokta-Volterra form to more closely understand the rather phonological competition coefficients (*α*) between competing species (Chesson 1990, 2000). After the reorganization shown in(Chesson 1990), the following equation represent the linkage between the Lottka-Volterra model and the parameters of MacArthur’s consumer resource model.

(9)

(10)

Left-hand side of equation 9 and 10 consists of parameters in the Lotka-Volterra model, while the right-hand side consists of parameters from MacArthur’s consumer resource model. On the left-hand side, *αij* is the competition coefficient and *fi* is per capita growth rates of the species *i* in the absence of resource limitation, which determines the winner of the competition (Chesson 2000). On the right-hand side, *cil* and *cjl* are the consumption of species *i* and *j* on resource *l* respectively, *mi* is the mortality of species *i*, *wi* is the value of one unit of resource *l* to the species, and *rl* and *Kl* are the *per capita* growth rate and carrying capacity of resource *l*. Through this linkage, empirically measured parameters in MacArthur’s consumer resource model can be translated into parameters in Lotka-Volterra model and thus be used to calculate niche difference (ND) and relative fitness difference (RFD) and finally access Chesson’s coexistence inequality.

* 1. **Fitting the Tilman’s resource ratio model**

Similar to MacArthur’s consumer resource model, Tilman’s resource ratio consumer resource model (Tilman 1977) can also be translated to a Lotka-Volterra form(Letten et al. 2017). Letten et al. 2017 reorganize Tilman’s two-species consumer resource model for two essential resources to the following Lokta-Volterra form (equation 11 to 14), so that one can decipher the parameters impacting species’ *per capita* growth rate. According to Letten et al. the inter- and intra-specific competition coefficients can be expressed as following,

(11)

(12)

(13)

(14)

In the above equations, *cij* is the consumption term of consumer species *i* on resource *j*, so it contains a parameter *yij* that representt the yield of consumer species *i* per resource *j*. *D* is the dilution rate, *S* is the supply rate of resource *i*, and R\* is the minimum resource density of resource *j* that still allows the species *i* to have positive *per capita* growth rate.

Note that, in the generic consumer resource model, each consumption terms (*cij*) is a function of resource density, e.g. in Tilman’s 1977 deduction. However, if the consumption term is resource density dependent, competition coefficients (*αij*) becomes resource dependent. In other words, if the consumption term is resource density dependent and thus not fixed, estimate of niche difference (ND), relative fitness difference (RFD) and finally Chesson’s coexistence inequality will also vary. Therefore, Letten et al.’s derivation should be used to predict coexistence based on the mutual *per capita* effects of each species on the other at the equilibrium. To use Letten et al.’s derivation (equation 11 to 14) to calculate competition coefficients for predicting coexistence at the equilibrium, one would have to assume that the consumption of species *i* on resource *j* is evaluated at the equilibrium. For example, when at equilibrium, consumer’s consumption (*cij*) should be equal to the dilution rate (*D*) divided by the yield of consumer (*yij*). In another words, the *αij* describes impact of species *j* on the *per capita* growth rate of species *i* when the resource that limits species *j* is at the equilibrium, i.e. at the R\* level (R\*jj). This assumption is the same as the sensitivity method since both methods assume the competing species to be at the equilibrium. This assumption is also valid because the mutual invasibility criteria also assume the competing species to be at the equilibrium.

1. **Compare the five commonly used empirical methods**

After describing the theoretical background, we compare these five methods in terms of (1) whether they are identical with respective to the algebra for calculating niche difference (ND) and relative fitness difference (RFD) as well as (2) whether they make the same coexistence prediction (Table 1). In the left part of Table 1, we compare whether the five methods are identical when using their metrics to calculate niche difference (ND; ) and relative fitness difference (RFD; ). The metric of each method is listed in the parenthesis following the name of each method. We see that the negative frequency dependency is the outlier as the slope of frequency is a complex combination of *per capita* growth rate (*ri*) and the fixed community density (*B*) in addition to *per capita* the intra- and inter-specific competition coefficients. Except NFD, the Lotka-Volterra model and the sensitivity method are identical in terms of the algebra for ND and RFD calculation (eqn. 7 and 8). Both consumer-resource models are also identical with the Lotka-Volterra model and the sensitivity method because one can simply reorganize the parameters in the consumer-resource models to the Lotka-Volterra form and derive ND and RFD.

In terms of whether the five methods make qualitatively the same coexistence predictions when applying these methods in real experiment (right part of the Table 1), we see that, again, the negative frequency dependency (NFD) method differs from the other four. The sign of the NFD slope can be used to predict coexistence and the prediction is the same as the other methods. However, the value of the NFD slope cannot be used to calculate ND and RFD, and to predict species coexistence as the slope is not equivalent to the competition coefficients (*αii* and *αij*). Other then NFD, the other four methods are qualitatively the same when predicting species coexistence as they have the same algebra to calculate ND and RFD and to predict coexistence. However, the two consumer models are qualitatively the same with the Lotka-Volterra model and the sensitivity method under some assumptions. Specifically, it is assumed that the competition coefficients (*αii* and *αij*) in the Lotka-Volterra model is density independent, which means the competition coefficients (*αii* and *αij*) is constant with respect to time and population density. The constant competition coefficients assumption is also the assumed in the sensitivity method but not necessary in the two consumer-resource models.

In the two consumer-resource models, the impact of the competing species (*j*) on the focal species (*i*) one depends on competing species *j*’s consumption on the resource that limit the growth of the focal species. For example, in Tilman’s consumer-resource model, the *per capita* competition of species *j* on *i* () depends on the consumption of species *j* on resource *j* (*cjj*), which is the limiting resource of species *i*. Although we assume that the consumption of species *i* on resource *j* is evaluated at the equilibrium when applying equation 11 to 14, this might not reflect the reality in the experiment system. In addition, the resource levels under which the experiment is conducted can affect the parameter estimation in the consumer resource models and thus affect the calculation of competition coefficients (*αii* and *αij*) and ND and RFD. For example, the R\* is the Tilman’s consumer-resource model is determined by the dilution rate (*D*) so that under different dilution rate, the *per capita* competition coefficients (e.g. ) and thus ND and RFD can vary. Consequently, the prediction to species coexistence can be different from that from the Lotka-Volterra model fitting or sensitivity methods.

In conclusion, we argue that, except the NFD method, these five methods are identical in terms of the algebra to calculate *per capita* competition coefficients and thus ND and RFD. However, when applying these methods in real experiment, the two consumer-resource models can make qualitatively different coexistence predictions when the *per capita* competition coefficients is not constant with respective to time and population size.

1. **When, why, and how each method should be used (narrative table 2)**
   1. Having summarized each method and its principles of operation, an empiricist is left to determine which method(s) are most appropriate for their study system, experimental approach, and goals. The upper section of Table 2 is a decision tree that divides the five methods with respect to several sequential bifurcations.
      1. The first bifurcation is whether or the empiricist knows the factors that influence population dynamics in their study system [Q1]. This question divides the five methods into two completely separate groups: phenomenological methods that are informed by quantifying species interactions but make no assumptions about mechanisms (highlighted in red), and two methods based on consumer resource models in which species are assumed to interact only through specific mechanisms (highlighted in blue).
      2. As shown in section 2, both of the consumer resource models can be used to estimate interaction coefficients and obtain estimates of ND and RFD. However, none of the phenomenological methods can be used to predict the mechanisms by which species interact in consumer-resource models.
      3. Due to the completely divergent properties of these two classes of models, several of the remaining decision steps are specific to either of the consumer-resource models or the phenomenological methods.
   2. Phenomenological Methods.
      1. The three phenomenological methods highlighted in red are similar in that an empiricist does not need to know which mechanisms regulate population dynamics, whether the species are competing for a resource, or what type of resource (biotic or abiotic) the species are competing for.
      2. One of the most consequential decisions among these methods is whether the data will come from manipulative experiments or observations from an un-manipulated system [Q2]. All three phenomenological methods will work for manipulative experiments, but only the NFD method has been applied to observational data in order to predict coexistence. This determinant is particularly important for study systems where manipulation is not feasible (e.g. long-lived species, protected habitats).
      3. Another bifurcation among the three phenomenological methods is whether the method requires data measured in monocultures [Q3]. The negative frequency dependence method is distinct because it is does not require monocultures. The Lotka-Volterra and Sensitivity methods are further distinguished by the need for each species to be grown at steady state as monocultures [Q4], either to measure steady-state abundance (i.e. carry capacity) or as a resident population for invasion experiments.
      4. The final determinant among the phenomenological methods is whether the method can be generalized to predicting coexistence among multiple species [Q5]. Theoretically, when predicting species coexistence between multiple species, the Lotka-Volterra model can consider other species individually, while the sensitivity and negative frequency dependency methods require that the other species are considered in aggregate. In other words, fitting the Lotka-Volterra model allows an empiricist to obtain the pairwise interaction coefficients for all species and thus can predict coexistence between the focal species versus the multiple species in aggregate. On the other hand, the sensitivity and negative frequency dependency methods can only be used to predict coexistence between the focal species versus the multiple species in aggregate.
      5. When using any of these phenomenological methods for more than two species at a time, an empiricist would need to assume that the multiple species consortia already stably coexist before the presence of the focal species. In addition, an empiricist would need to assume that with the impact from the multiple species aggregate on the focal species remain the same with the presence of the focal species. However, none of these three methods can deal with intransitive competition, where competition among species can be non-hierarchical.
   3. Consumer-Resource Methods.
      1. The consumer-resource models are differentiated primarily based on whether the resource is abiotic and governed by a constant rate of supply (e.g. inorganic nutrients consumed by plants) or biotic and has its own population dynamics [Q6].
      2. Another characteristic that distinguishes the consumer-resource models is the number of resources that are considered. Specifically, Letten et al [2017] demonstrated that the consumer-resouce model can be used for two abiotic resources, so it remains unclear whether the method could be expanded to consider information about additional resources. In contrast, the method based on MacArthur’s consumer-resource model works for systems where the number of relevant resources is very large.
      3. While the consumer-resource model methods have certain advantages, these methods can only be applied in a limited subset of cases where the empiricist knows all of the factors that affect the population dynamics of the species.
   4. Having considered the questions under the section ‘Decision Steps’ in Table 2, an empiricist should be able to identify the method that is most appropriate. Using the first half of the table as a guide will should result in one preferred method, or in some cases a choice between two (e.g. LV and sensitivity) that can be further informed by the inputs/outputs section of the table (see below). In the ‘Method’ section of Table 2, we direct the reader to 1) the foundational paper that describes the underlying model for population dynamics, 2) the theoretical paper that relates the model to Modern Coexistence Theory and Chesson’s Inequality, and 3) an example of an empirical study that employed the method in the context of modern coexistence theory. For the consumer resource models, we are unaware of any empirical studies that have used the MacArthur CRM to predict coexistence and the only paper to have applied the R\* CRM used previously published data.
   5. Comparison of Method Inputs
      1. These five methods differ in terms of the information that would be required as ‘inputs’ in order to estimate ND and RFD.
      2. For instance, the phenomenological methods differ in terms of the number, length, and types of time series required. As a result, the number of new experiments required for all pairwise combinations of species increase linearly or exponentially with each additional species. In contrast, the consumer-resource models require only as many additional experiments as the number of resources.
      3. While all of the the phenomenological methods require at least one co-culture of each species pair in order to quantify the strength of their interaction. The direct Lotka-Volterra method requires a minimum of one co-culture, but the sensitivity and NFD methods require two or more co-cultures. In contrast, the methods based on consumer-resource models do not require any co-culture in order to predict interaction strength.
      4. Due to the need for long time-series, some of the methods would not be tractable for long-lived species (e.g. the all-in-one LV parameterization that Oscar demonstrated). However, the NFD method can work for long-lived species using a space for time substitution.
   6. Comparison of Method Utility/Outputs
      1. We showed how each of these methods, with the notable exception of negative frequency dependence, can be used to obtain estimates of ND and RFD. While the methods differ in terms of their experimental design and assumptions about population dynamics, we expect these methods to give the same prediction regarding coexistence when applied to the same species and environmental conditions. In terms of model output then, the key differences are between phenomenological and consumer-resource methods.
      2. Only the consumer resource models are able to predict the potential for coexistence among combinations of species without growing those species together simultaneously.
      3. None of the phenomenological methods can be used to make predictions about novel combinations of species or different environmental contexts. However, consumer resource models can be used to predict ND and RFD under limited sets of different environmental conditions. For instance, Letten et al show that the Tilman R\* model can be used to predict the ND and RFD at different nutrient supply rates or dilution rates [Letten et al 2017], but if for example, temperature were changed, the model cannot be used to make predictions.
2. **Cautions and future directions**

Caution 1: Need to empirically demonstrate equivalence of the methods. To date, we are unaware of any empirical studies that have applied more than one of these methods to the same study system. According to what we have mathematically shown in the previous section, these five methods are not identical to each other in terms of how the derive ND and RFD, but there is good correspondence in terms of predicting coexistence. This means that values of ND and RFD from different experimental approaches are not comparable. As we show here, a few papers have applied empirically-derived parameter values to show that two methods are comparable (Letten et al 2017; Levine and HilleRisLambers 2009), but these ad hoc tests do not reflect the differences in experimental design, assumptions, and calculations that are outlined in Table 2.

The fact that ND and RFD are not comparable among different experimental approaches limits the possibilities for synthesis. For example, studies that relate ND and RFD to phylogenetic or functional divergence among species (Narwani et al 2013) would need to use the same empirical approach and experimental conditions for all of the species.

Caution 2: Adherence to assumptions. Although we have shown that several of the empirical approaches for predicting coexistence among species give the same qualitative outcome, this is true only under certain assumptions and conditions. When those assumptions are not recognized and justified, any of these three methods can give misleading predictions. For instance, the Lotka-Volterra method assumes that the per-capita effect of species i on species j is independent of the density of either species i or species j. However, there are cases where this assumption is clearly not met. For example, if species were limited by resources (e.g. nutrients), a positive saturating relationship between the availability of resources and per-capita growth rate means that density-dependence is weak at low population sizes and stronger at higher population densities. Thus, both the inter and intraspecific interaction coefficients would appear to be very small if measured at low population densities and very high if measured at population densities approaching the steady-state biomass. In designing and interpreting experiments, it would be unclear which value to select for the interaction coefficients. This example shows how empirically comparing two methods can reveal differences among the methods which are not readily apparent from their derivation.

Caution 3: Limits to the applicability of CRM. As shown in table 1, using either of the methods based on consumer-resource models requires that the empiricist knows the environmental factors that determine the outcome of competition, and specifically, that those factors are resources. This is more easily achieved for certain experimental systems (e.g. microbes grown under laboratory conditions) than others (e.g. ungulate herbivores). Outside of abstract experiments, it is hard to know for sure which resources or factors govern population dynamics. While these experiments are useful for isolating the *mechanism* of competition, they require detailed knowledge about natural history of the organisms, which in many cases is unknown.

Caution 4: Chesson’s inequality for predicting coexistence is only applicable to two-species system. Chesson’s coexistence framework, and the methods derived from it, are designed to predict coexistence among pairs of species. This ND/RFD framework has not been generalized to multi-species communities (but see Carroll et al 2011). For example, the ND between three species is not as straightforward as the ND between two species. In terms of experimentation, the sensitivity and the negative frequency dependency method can be used in one-to-many species contexts, provided some assumptions discussed previously. Importantly, this emphasis on pairwise interactions and experimentation means that intransitive competitive interactions, if present, are unaccounted for.

**Tables**

Table 1

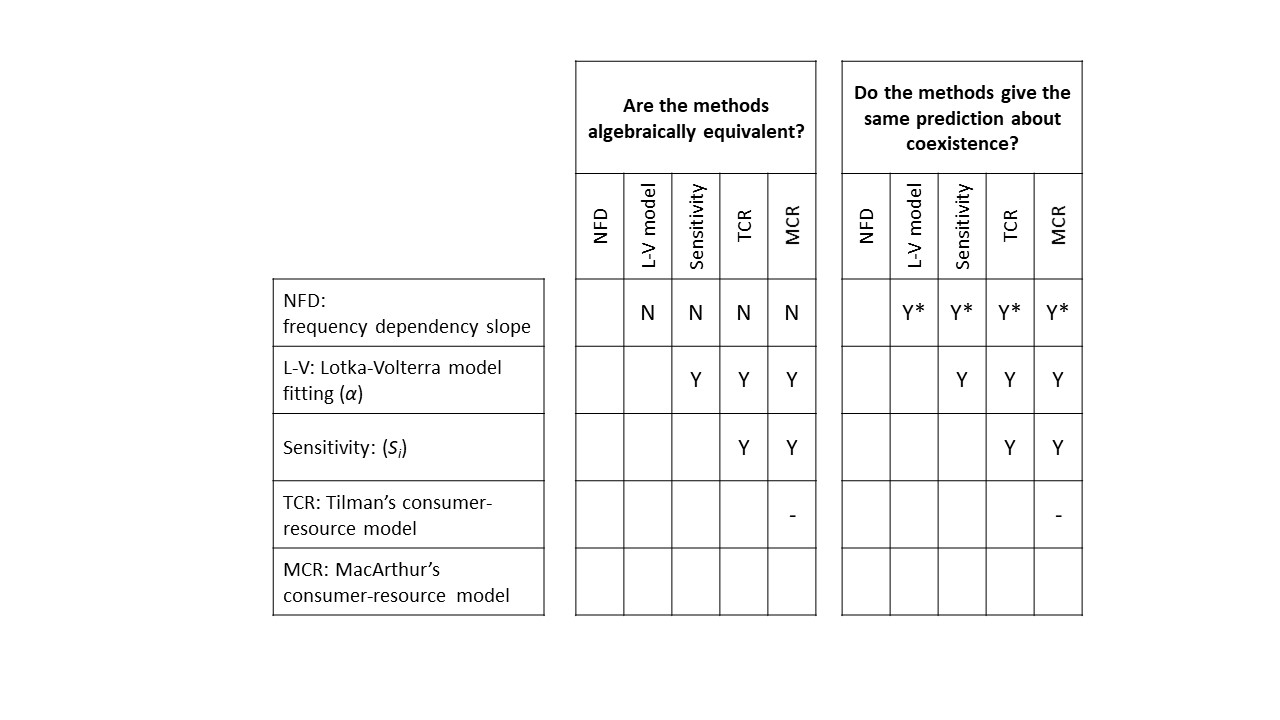


Table 2



**Figures**

Figure 1

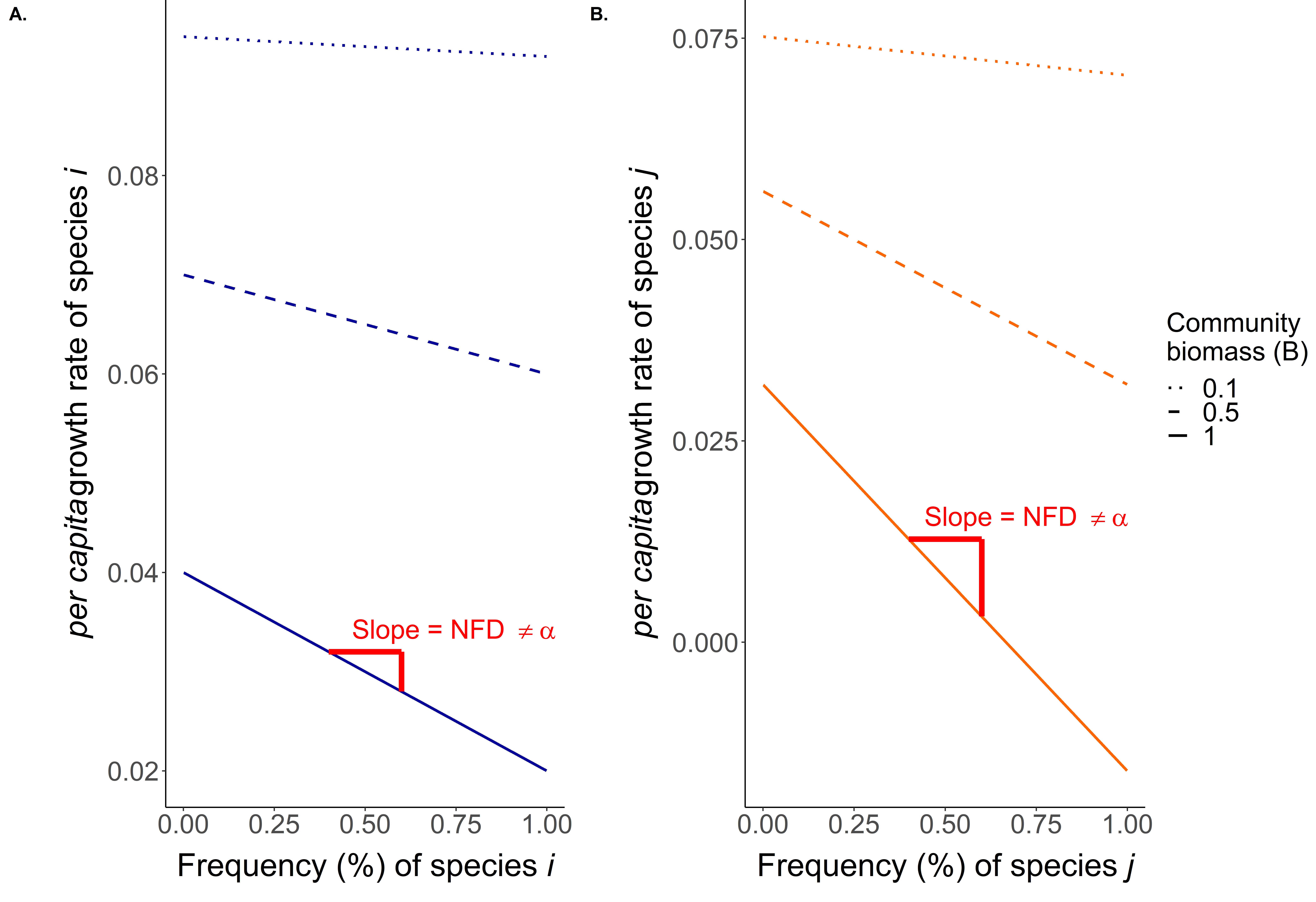


Figure 1. An example plot showing the negative frequency dependency (NFD) of species *i* (panel a.) and species *j* (panel b.). To calculate NFD, we first determined an arbitrary community biomass (B) and gradually increased the frequency of the focal species (species *i* in panel a. and species *j* in panel b.) to calculate the *per capita* growth rate of the focal species. Note that we directly calculated the *per capita* growth rate from equations of the Lotka-Volterra model not from numerical simulations. As the figure shows, the NFD depends on the arbitrarily community biomass (*B*). The resulting NFDs (slopes) match that are expected by equation 7. For example, when community biomass is 1 (dashed line in the middle), NFD of species i is -0.02 and NFD of species j is -0.045.

Figure 2



Figure 2. An example plot showing the estimation of αij from fitting Lotka-Volterra model to time series. The points represent the density of species i (open circles) and j (solid dots) respectively. The dashed lines are the two fitted growth curve. The species densities were generated by a Lotka- Volterra model and added some random noise.

Figure 3

Figure 3. An example plot showing the estimation of sensitivity (*Sij*) of species *i* invading species *j* (panel a.) or the reverse (panel b.). In both panels, solid dots are the growth curve species when growing alone and the open circles are the growth curve of species when invading the carrying capacity of the other species. To estimate the sensitivity of the focal species (e.g. species *i*), the competing species (species *j*) is fixed at the equilibrium and invade the focus species to estimate the invading growth rate of the focal species.

Figure 4



Figure 4. An example plot demonstrating the idea of niche difference (ND; *ρ*) in MacArthur’s con- sumer resource model. Consumption of species *j* on resource *l* (*cjl*) are plotted against consumption of species *i* on resource *l* (*cil*). The closeness between *cjl* and *cjl* are the closeness between these points to the 1:1 line and is expressed as *ρ*. For example, if all *cjl* are equal to *cil*, which means all points are on the 1:1 line, *ρ* is 1 and ND between species *i* and species *j* is 1−*ρ* = 0. The inset plot is an example plot showing the data required to measure consumption of species *j* on resource *l* (*cjl*).

Figure 5

Figure 5. Example plots showing the empirical data required to estimate R\* in Tilman’s consumer resource model for species *i* and *j* on resource *i* and *j*. To obtain data one would need to grow the focal species (*i* and *j*) under different resource level and measure the corresponding *per capita* growth rate. *Rij* means the minimum level of resource *j* that still allows species *i* to have positive per capita growth rate. In this example scenario, species *i* (*j*)is limited by resource *j* (*i*), so that the () is greater than (). The R\* is the one being used to calculate *αii* and *αij* (*αjj* and *αji*) because they determine how sensitivity species *i* will be affected by the resource (resoruce *j*) that limits the growth rate of species *i*.

Figure S1. Results of numerical simulation using Tilman’s consumer-resource model with two species of phytoplankton. The purpose of these simulations was to determine whether the NFD method produces a linear relationship between growth rate and frequency, under the assumptions that the community biomass is saturated and that the two species are capble of mutual invasion. The equations and parameter values are taken exactly from Tilman 1977. Simulations were performed using the function NDSolve in Mathematica 11.2 (Wolfram Research), employing a variable step size. For each set of resource supply concentrations, we performed four simulations: (1) species 1 as a monoculture growing from rare to near its equilibrium density; (2) species 2 as a monoculture growing from rare to near its equilibrium density; (3) species 1 at its equilibrium density with species 2 invading from rare; (4) and species 2 at its equilibrium density with species 1 invading from rare. From simulations 3 and 4, we used the time series to obtain pairs of each species frequency and its growth rate. These values were used to construct the relationships depicted in panels a and b. At any frequency of the two species, the NFD method requires that we use the slope to extrapolate and estimate the growth rate at a frequency approaching zero. For frequencies where this method predicts mutual invasibility for both species, the lines are blue. For frequencies where the method leads to the incorrect prediction, the lines are red. Both species have positive growth rates when their frequency approaches zero, indicating that they are mutually invasible. The coexistence equilibrium occurs at a frequency of 0.88 for species 1 and 0.12 for species 2. At point A, corresponding to frequency of 0.5 for each species, this method does not predict coexistence because while species 1 would be expected to have a positive growth rate when rare (also has negative NFD slope), species 2 is predicted to have a negative growth rate when rare (positive slope of NFD). If an empiricist did their work between species 1 frequency of ~0.05 to ~0.85, and used the slope of NFD, they would incorrectly predict that the species will not coexist. At point B, which is close to their equilibrium frequencies, both species have negative NFD slope and should have positive growth rates when rare based on this method.

Panel c shows the fidelity of prediction from the NFD method as a function of Si:P ratio and the frequency of the two species. The vertical dashed line represents the slice depicted in panels a and b. For all of the Si:P ratios shown in panel c, the species are mutually invasible and will coexist. This plot indicates that using NFD will often predict that the species will not coexist, when in fact they do coexist. This is important because without doing the full range of species frequencies in an experiment, one would not know whether and where the relationship between frequency and growth rate is non-linear.

Figure S2. Results of applying four empirical methods for evaluating MCT to numerical simulation using Tilman’s consumer-resource model with two species of phytoplankton. The purpose of these simulations was to determine whether the different methods give the same prediction under a limited range of resource supply concentrations. For each simulation, the parameter values were taken directly from Tilman 1977 and only the supply concentration of phosphate and silicate were varied. For each set of resource conditions, we performed four different numerical simulations: : (1) species 1 as a monoculture growing from rare to near its equilibrium density; (2) species 2 as a monoculture growing from rare to near its equilibrium density; (3) species 1 at its equilibrium density with species 2 invading from rare; (4) and species 2 at its equilibrium density with species 1 invading from rare. We performed an additional numerical simulation (“competition experiment”, panel a) to determine whether the species can coexist when both are introduced at low densities and allowed to reach equilibrium biomass.

Panel b shows the predictions from Letten’s method, which is not based on numerical simulation, but uses the parameter values and resource supply concentrations to predict the inter- and intra-specific interaction coefficients as specified by the original authors.

Panel c shows the predictions based on the sensitivity method, where growth rates were determined from simulations 1-4 and used to compute sensitivities as directed by Carroll et al XXXX.

Panel d shows the prediction based on computing Lotka-Volterra interaction coefficients. In this method, the per - capita impact of an invader on the resident' s per - capita growth rate was calculated using the time series from simulations 3 and 4. Additionally, the intraspecific interaction coefficient was determined as the per-capita impact of each species on its own growth rate when the population size approaches 99% of its equilibrium value (simulations 1 and 2).

Panels e-j show the predictions based on the NFD method, evaluated at different frequencies of the two species and both using only the slope of NFD (e-g) and the extrapolation to estimate growth rates when rare (h-j).