

SI Material and Methods: Immigration Simulations

Conventional studies of coexistence mechanisms typically consider how *in situ* growth is affected by interactions within and between species, excluding immigration and emigration. Immigration itself is not seen as a mechanism of coexistence because it does not truly solve the ‘problem’ of negative local species interactions. Nonetheless, immigration is undoubtedly the reason many rare species appear persistent. While we have excluded many species from our analyses because they were clearly ephemeral, our criteria for doing so was minimal (species so incidental that parameters could not even be calculated, because they did not have more than two non-negative abundances in a row). It is certainly a possibility that other species in the 90 communities which did enter our analysis rely substantially on immigration to persist. Because immigration will affect the persistence of rare species specifically, it becomes important to consider this source of growth when discussing rare species, rather than ignore it.

Immigration raises two questions in the framework of negative frequency dependence. 1) Does immigration bias our ability to detect true *in situ* negative frequency dependence in rare species? 2) How should we incorporate the effect of immigration for rare species into the asymmetric NFD framework? To answer the first question, and begin a discussion about the second, we built on the simulations described in Appendix S3 to incorporate immigration/emigration dynamics. In this scenario, a community of 10 species experience either equal negative frequency dependence or a randomized set of parameters, so that there is no relationship between the strength of NFD and equilibrium frequency. We then add immigration and emigration with a second community: at each time step, 10% of each population is sent to the second community. The second community has dynamics based on the same model, also with 10% emigration. Its species have either opposite equilibrium frequencies to the first community (in the stabilized scenario) or randomized parameters (in the randomized scenario). The growth of each species in the focal community is:

$$u_{s,t+1} = \exp(FD_s * N_{s,t}/C_t + IGR_s) * N_{s,t}$$

$$n_{s,t+1} \sim \text{Pois}(u_{s,t+1})$$

$$N_{s,t+1} = 0.9 * n_{s,t+1} + 0.1 * n_{s2,t+1}$$

Where $u_{s,t+1}$ is the expected abundance of species s in time $t+1$ and C_t is the total community abundance in time t . IGR represents the intrinsic growth rate and FD represents the frequency dependence experienced by each species s . $n_{s,t+1}$ is the true abundance of species s in time $t+1$, after incorporating demographic stochasticity with a Poisson distribution. To this 10% immigration and emigration is incorporated using $n_{s2,t+1}$, the abundance of species s in the second community, to give $N_{s,t+1}$, the total abundance of species s in the focal community in time $t+1$. All simulations were initialized with a total community abundance of 1000 individuals.

We then incorporated measurement error by drawing observed abundances, X , from a negative binomial distribution with mean $N_{s,t+1}$ and dispersion parameter k :

$$X_{s,t+1} \sim \text{NB}(N_{s,t+1}, k)$$

As in Appendix S3, we altered the relationship between the values of the dispersion parameter and the equilibrium frequency to demonstrate the robustness of our methodology to the combined biases of sampling and immigration (Table S3, Figures S8-S9, S11-S12). There is no asymmetric NFD in this simulated community. Any significant relationship detected would be the result of a bias created by immigration, or the combination of immigration and sampling bias (Appendix S3 shows that sampling bias alone does not create a significant pattern).

Simulations show that immigration does not create a pattern of asymmetric NFD in this scenario (Figures S10, S13). The absence of asymmetric NFD is observed despite the fact that population growth due to immigration often does prevent stochastic extinction for rare species (Figures S8, S11), particularly in the randomized case, in which NFD alone would not predict persistence. (To test the effect of immigration with a more conventional model, we repeated this exercise with a Lotka Volterra model. The results were similar; immigration did not result in asymmetric NFD.) In our simulations, the emigration rate is density-dependent. Though it is possible that the population fluctuations of a rare species caused by immigration could be frequency-dependent, these simulations suggest that dependence does not ‘look’ like *in situ* negative frequency dependence.

While we feel we have created realistic scenarios to test this bias, there are also many more scenarios in the real world than we can simulate. It is possible, for example, that measurement error is inflated for immigrating individuals in a way that would further bias the sampling of rare species, beyond what we have simulated. Specific scenarios like this seem unlikely, but emphasize the need for in-depth demographic studies in specific communities, to correctly identify all sources of population buffering and their relative importance.

Though immigration does not directly alter the relationship between equilibrium frequency and NFD, it still plays an undeniable role in the persistence of rare species. This is true for anything that buffers species from stochastic extinction. That it does not leave a signature of asymmetric NFD just inspires us to consider where it does fit in the overall picture of rare species’ persistence. It certainly must play a role very complimentary to NFD. Any buffering process that disproportionately affects rare species fits well into a more general definition of asymmetric population dynamics, one that combines local species interactions with regional processes.

Table S3. Summary of the parameter values used in simulating immigration.

Species	Stabilized Scenario						Random Scenario			
	Focal Community			Second Community			Focal Community		Second Community	
	Equilibrium frequency (f)	IGR	FD	Equilibrium frequency (f)	IGR	FD	IGR	FD	IGR	FD
1	0.001	0.001	-1	0.27	0.27	-1	0.270	-1.4	0.064	-1.3
2	0.01	0.01	-1	0.18	0.18	-1	0.126	-1.6	0.018	-1.6
3	0.02	0.02	-1	0.16	0.16	-1	0.198	-1.7	0.270	-1.7
4	0.04	0.04	-1	0.149	0.149	-1	0.192	-1.9	0.126	-1.5
5	0.08	0.08	-1	0.09	0.09	-1	0.034	-1.2	0.198	-1.8
6	0.09	0.09	-1	0.08	0.08	-1	0.018	-1.0	0.0019	-1.1
7	0.149	0.149	-1	0.04	0.04	-1	0.120	-1.5	0.192	-1.9
8	0.16	0.16	-1	0.02	0.02	-1	0.1937	-1.3	0.034	-1.2
9	0.18	0.18	-1	0.01	0.01	-1	0.0019	-1.8	0.120	-1.0
10	0.27	0.27	-1	0.001	0.001	-1	0.064	-1.1	0.018	-1.4

Figure Legends

Figure S8: The frequency-dependence relationship simulated in the focal community (panel 1) and the resulting actual species abundances in the simulated community (panel 2) for the stabilized scenario.

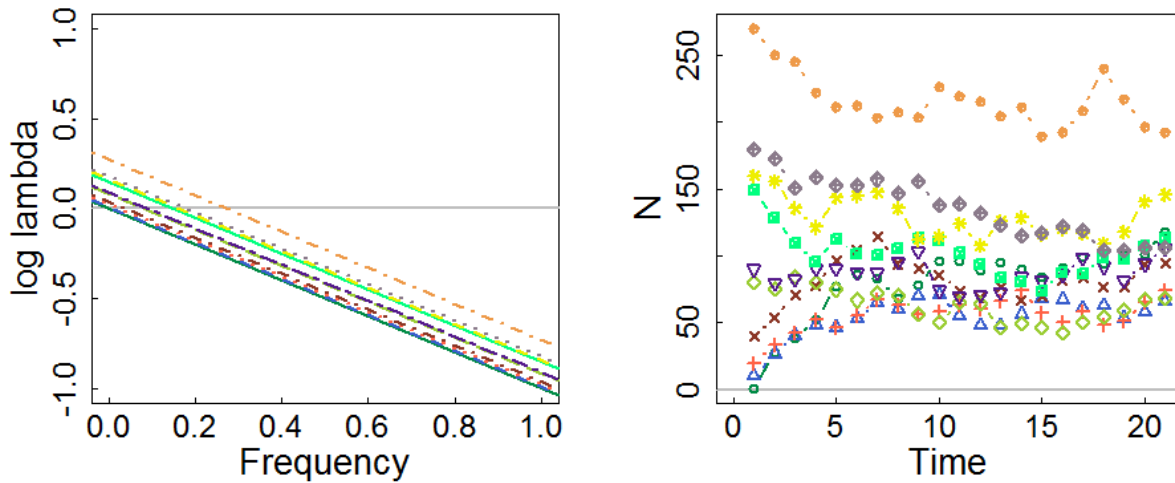


Figure S9: The resulting observed species abundances in the simulated community for the stabilized scenario, after incorporating sampling noise. Sampling noise was incorporated as a constant dispersion parameter ($k=1$, panel 1); dispersion parameter decreases with rarity ($k=10*f$, panel 2); or dispersion parameter increases with rarity ($k=0.1/f$, panel 3).

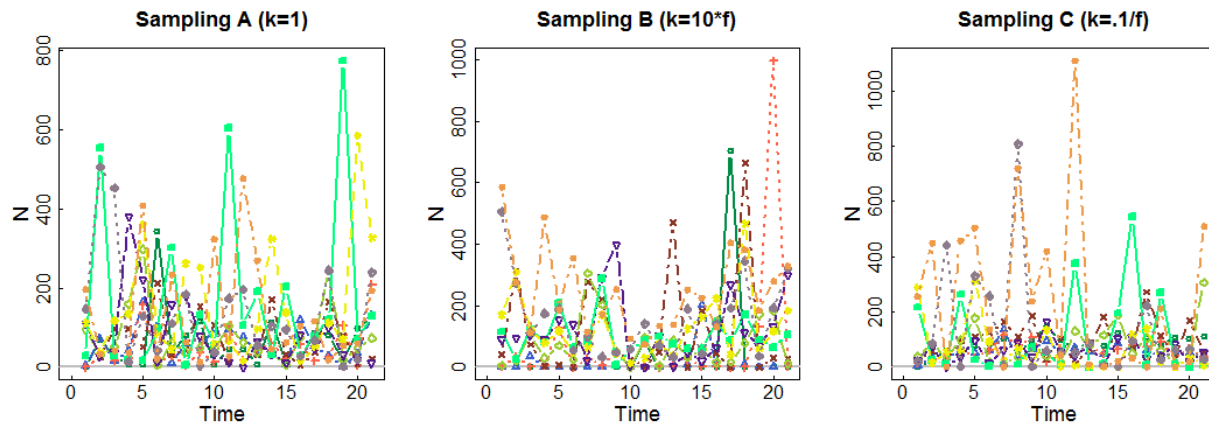


Figure S10: Results from simulated and sampled communities in the stabilized scenario, using the same methodology applied to the real community data. The first figure in each row shows the calculated relative abundances and growth rates from the simulated data for 10 species (points). Lines represent the estimated relationship for each species. The second figure shows the relationship between equilibrium frequency and the strength of NFD in each simulated community (red lines are the fitted

log-log relationship). The third figure shows histograms of the 'Null' pattern estimated from the shuffled data in 5000 randomizations. The vertical red line is the pattern estimated from the simulated data for comparison. This is how p values were calculated (for use in false discovery rate control). The dispersion parameter was constant (A, $k=1$), increased with rarity (B, $k=10 \cdot f$), or decreased with rarity (C, $k=0.1/f$). In this scenario with no relationship between equilibrium frequency and NFD, 10% immigration and emigration, and sampling uncertainty, the randomization methods correctly detect no significant relationship.

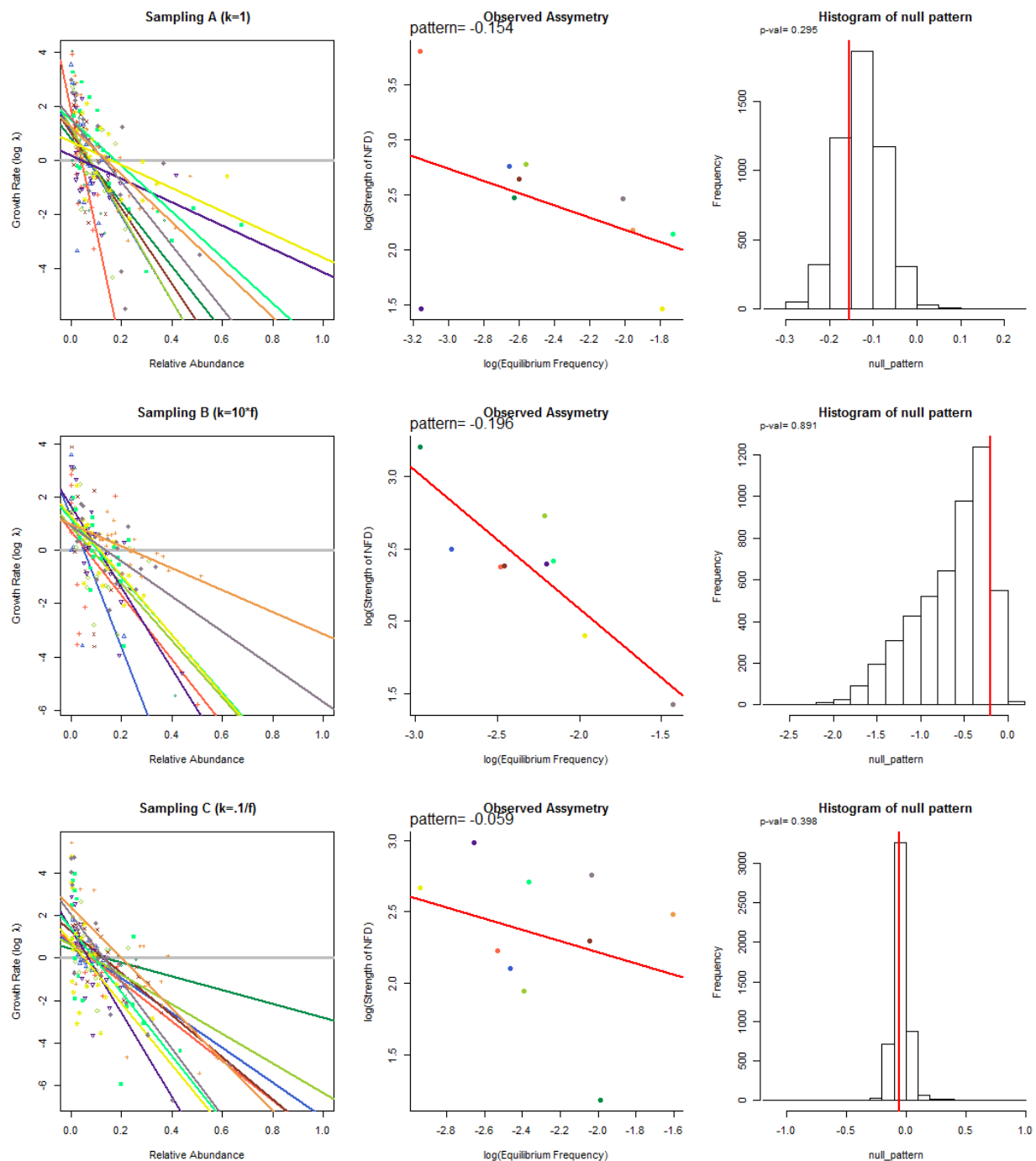


Figure S11: The frequency-dependence relationship simulated in the focal community (panel 1) and the resulting actual species abundances in the simulated community (panel 2) for the randomized scenario.

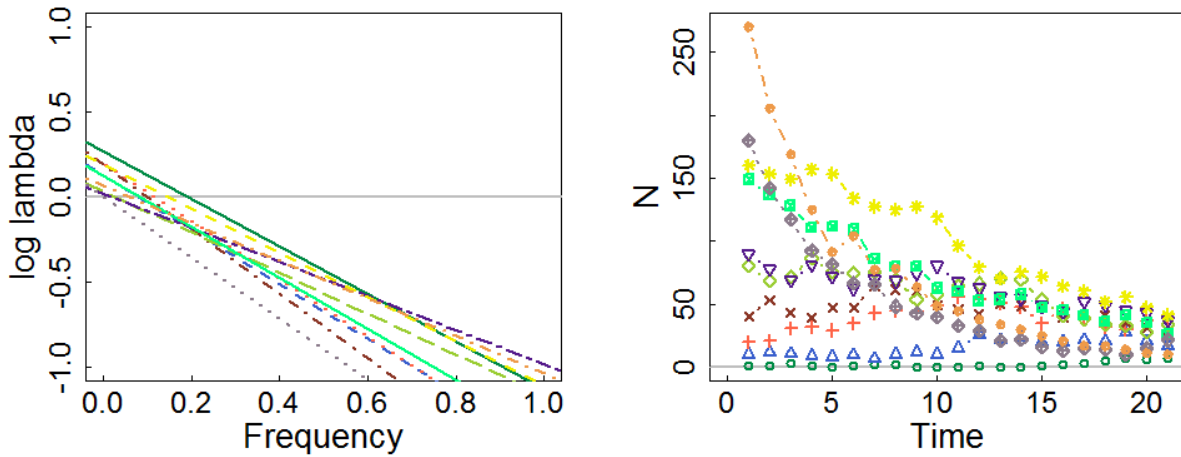


Figure S12: The resulting observed species abundances in the simulated community for the randomized scenario, after incorporating sampling noise. Sampling noise was incorporated as a constant dispersion parameter ($k=1$, panel 1); dispersion parameter decreases with rarity ($k=10 \cdot f$, panel 2); or dispersion parameter increases with rarity ($k=0.1/f$, panel 3).

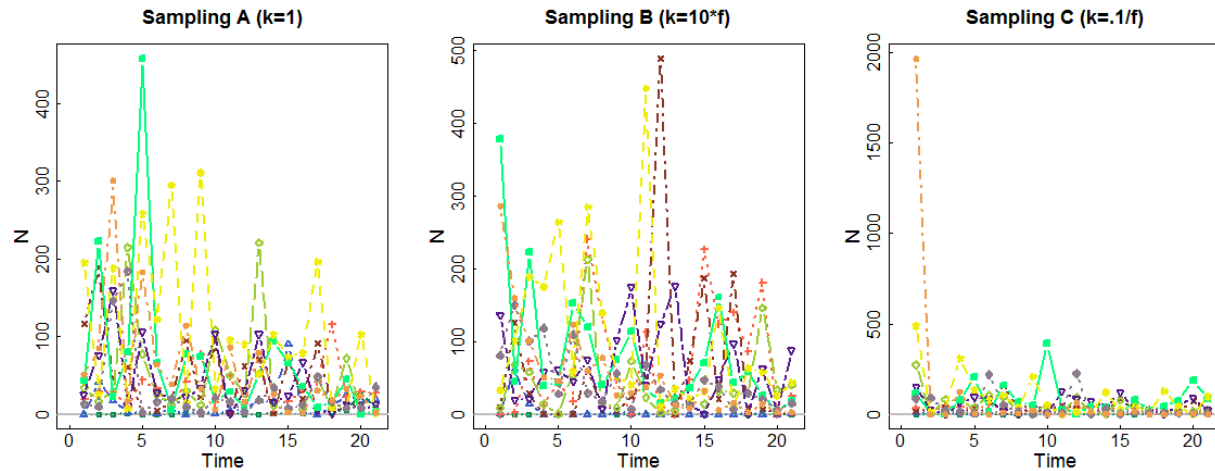


Figure S13: Results from simulated and sampled communities in the randomized scenario, using the same methodology applied to the real community data. The first figure in each row shows the calculated relative abundances and growth rates from the simulated data for 10 species (points). Lines represent the estimated relationship for each species. The second figure shows the relationship between equilibrium frequency and the strength of NFD in each simulated community (red lines are the fitted log-log relationship). The third figure shows histograms of the 'Null' pattern estimated from the shuffled data in 5000 randomizations. The vertical red line is the pattern estimated from the simulated data for comparison. This is how p values were calculated (for use in false discovery rate control). The dispersion

parameter was constant (A, $k=1$), increased with rarity (B, $k=10*f$), or decreased with rarity (C, $k=0.1/f$). In this scenario with no relationship between equilibrium frequency and NFD, 10% immigration and emigration, and sampling uncertainty, the randomization methods correctly detect no significant relationship.

