ARTICLE

ECOLOGY
ECOLOGICAL SOCIETY OF AMERICA

Dominant species stabilize pollination services through response diversity, but not cross-scale redundancy

Mark A. Genung[1] | **Rachael Winfree**[2]

[1]Department of Biology, University of Louisiana, Lafayette, Louisiana, USA

[2]Department of Ecology, Evolution and Natural Resources, Rutgers University, New Brunswick, New Jersey, USA

Correspondence
Mark A. Genung
Email: mark.a.genung@gmail.com

Abstract

Substantial evidence suggests that biodiversity can stabilize ecosystem function, but how it does this is less clear. In very general terms, the hypothesis is that biodiversity stabilizes function because having more species increases the role of compensatory dynamics, which occur when species in a community show different responses to the environment. Here, we focus on two forms of compensatory dynamics, cross-scale redundancy (CSR) and response diversity (RD). CSR occurs when species respond to a disturbance at different scales such that scale-specific disturbances do not negatively affect all species. RD occurs when species contributing to the same function show different responses to an environmental change. We developed a new analytical approach that can compare the strength of CSR and RD in the same dataset and used it to study native bee pollination of blueberry at 16 farms that varied in surrounding agricultural land use. We then asked whether CSR and RD among bee species are associated with the stability of blueberry pollination. Although CSR and RD were both present, only RD was associated with higher stability of pollination. Furthermore, the effects of RD on stability were due to a single widespread species, *Andrena bradleyi*, that is a specialist on blueberry and, unlike other bee species, was highly abundant at farms surrounded by intensive blueberry agriculture. Thus, the stabilizing effect we observed was attributable to an "identity effect" more than to species richness per se. Our results demonstrate how CSR and RD can be empirically measured and compared and highlight how the theoretical expectations of the biodiversity–ecosystem functioning field are not always upheld when confronted with real-world data.

KEYWORDS
agriculture, biodiversity, cross-scale redundancy, disturbance, dominant species, ecosystem function, pollination, response diversity, richness, stability

INTRODUCTION

Biodiversity, in the form of higher species richness, can stabilize function across time and space, but the mechanisms behind this effect are less clear (Cardinale et al., 2012;

Ives & Carpenter, 2007; Lemanski et al., 2022). Proposed mechanisms include overyielding, or an increase in mean function that reduces the role of variability (Hector et al., 2010); statistical averaging, or the fact that community stability is higher when there are more

(rather than fewer) independently fluctuating species (Doak et al., 1998); and a broader category referred to as compensatory dynamics, which occur when species in a community show different responses to environmental conditions (Gonzalez & Loreau, 2009). One difficulty in testing for these mechanisms is a lack of analytical tools for measuring their strength, which is a key knowledge gap because richness is not itself the driver of stability (McCann, 2000); instead, richness simply increases the chances of observing stabilizing mechanisms. Here, we distinguish and quantify two forms of compensatory dynamics, namely cross-scale resilience and response diversity (RD) (defined below) and determine whether they are associated with the stability of ecosystem function.

Cross-scale resilience holds that when species contributing to the same ecosystem function respond to a disturbance at different scales, the community is better able to return to a stable state because a scale-specific disturbance will not negatively affect all species (Holling, 1988; Peterson et al., 1998; Steffan-Dewenter et al., 2002). For example, clearing natural habitat for smaller and larger farms may differentially impact bee species that forage at different radii (Benjamin et al., 2014). Cross-scale resilience has been an influential idea, but infrequently tested (Fischer et al., 2007; Nash et al., 2016; Winfree & Kremen, 2009), perhaps because it is a nuanced, multistep model that requires years of data for confirmation. A simpler approach, termed cross-scale *redundancy* (CSR), focuses on only the first condition of cross-scale resilience: that the community includes functionally redundant species that interact with their environment at different spatial scales (Allen et al., 2005; Nash et al., 2016; Wardwell et al., 2008). Recently, cross-scale redundancy has become a proxy for cross-scale resilience (Angeler et al., 2016; Martin et al., 2019; Nash et al., 2016). This is reasonable, because variability among species, especially in body size, affects the scale at which organisms perceive and use resources (Ritchie & Olff, 1999). As such, we expect different species to respond to land use change at different scales (Peterson et al., 1998; Benjamin & Winfree, 2014), and these different responses should reduce the chances of a single disturbance destabilizing the entire community. Statistical confirmation of whether CSR occurs continues to be rare (Winfree, 2013; but see Nash et al., 2016).

The second stabilizing mechanism, RD, occurs when species contributing to the same function show variation in their responses to environmental change (Elmqvist et al., 2003; Walker et al., 1999). RD is related to the idea of insurance effects (Yachi & Loreau, 1999), which holds that having many, functionally redundant species will provide insurance against decline in function resulting from species loss. However, RD goes one step further, in that not only will some species persist after others are

lost, but the persisting species may find the environment *more* favorable after environmental change. There is compelling evidence that RD exists from both experimental (Leary & Petchey, 2009; McGrady-Steed & Morin, 2000) and observational studies (Brittain et al., 2013; Burgess et al., 2021; Cariveau et al., 2013; Stavert et al., 2017). For example, Stavert et al. (2017) showed that pollination to *Brassica rapa* across an agricultural intensification gradient was stabilized by exotic pollinator species, which contrasted with native species by responding positively to agriculture. However, while differential responses to environmental change clearly *could* stabilize function, there is no guarantee that the presence of RD *will* increase stability. For example, Cariveau et al. (2013) found that RD by bee species in response to agricultural land cover was present (i.e., statistically detectable) for three crop species, but was stabilizing (i.e., correlated with a metric of stability) only when the signal of RD was generated by a species with a positive response to agricultural land cover (Cariveau et al., 2013). Thus, we should not infer that the presence of a proposed stabilizing mechanism—be it CSR, RD, or otherwise—actually stabilizes function.

Little is known about the relative strength of CSR and RD and whether they interact to affect stability (Winfree, 2013). In fact, the two mechanisms are not consistently distinguished in the literature. Some consider species operating at different scales (what we call CSR) to be another form of RD (Elmqvist et al., 2003). Others have described RD as a likely result of CSR (e.g., fig. 1 in Nash et al., 2016) or as a mechanism allowing persisting species to compensate for declines in function caused by disturbance-induced species loss, thus making the system more resilient (Sundstrom et al., 2018). We argue that CSR and RD are not completely distinct. A robust definition of RD must be informed by CSR, in that assessing variability in species' responses to disturbance should account for the different scales at which those species respond to the disturbance (Winfree & Kremen, 2009). However, we also argue that CSR and RD can be included as predictors in the same analysis because the strength of CSR and the strength of RD are not necessarily correlated. In other words, species operating at different spatial scales does not require that they also show different responses to an environmental disturbance, or vice versa (Appendix S1: Section S1). This opens the door to testing whether CSR or RD is a stronger stabilizing mechanism and the extent to which CSR and RD interact to affect the stability of ecosystem function.

Here, we present an analytical framework and novel quantitative tests for CSR and RD. Our approach is based on two analyses. First, we use statistical models that allow species identity to interact with the spatial scale at

which environmental change is measured, which would indicate CSR, and/or the extent of environmental change, which would indicate RD (Figure 1). Second, we subset our empirical data in a series of simulations, which allows us to see how variability in richness and the presence/absence of difference species affect CSR and RD. We demonstrate our approach by assessing how biodiversity, acting through CSR and/or RD, affects the spatial stability of function along a disturbance gradient, using data on wild bee pollination of highbush blueberry (*Vaccinium corymbosum*) at farms surrounded by varying amounts of agricultural land cover. In general, crop pollination by wild insects tends to decrease with increases in agricultural land cover (Kremen et al., 2007), but bee species differ in their response to land cover (Cariveau et al., 2013), and also respond to land use at different scales (Benjamin et al., 2014).

We address three hypotheses. **H1**: We predict that CSR and RD will both be statistically detectable (i.e., distinguishable from a null model; see *Methods*). We expect CSR to be distinguishable from a null model because foraging range varies across bee species, causing them to respond to agricultural land cover at different spatial scales (Benjamin et al., 2014; Kendall et al., 2022). We expect a RD to be distinguishable from a null model because a previous study of these same data using a simpler statistical approach (linear models instead of generalized additive models (GAMs), which are more flexible) has already documented such a signal (Cariveau et al., 2013). The biodiversity–stability hypothesis predicts that both signals should increase with species richness, although this is mostly unexplored empirically. **H2**: CSR and RD will increase the stability of pollination. We expect this because both CSR and RD are
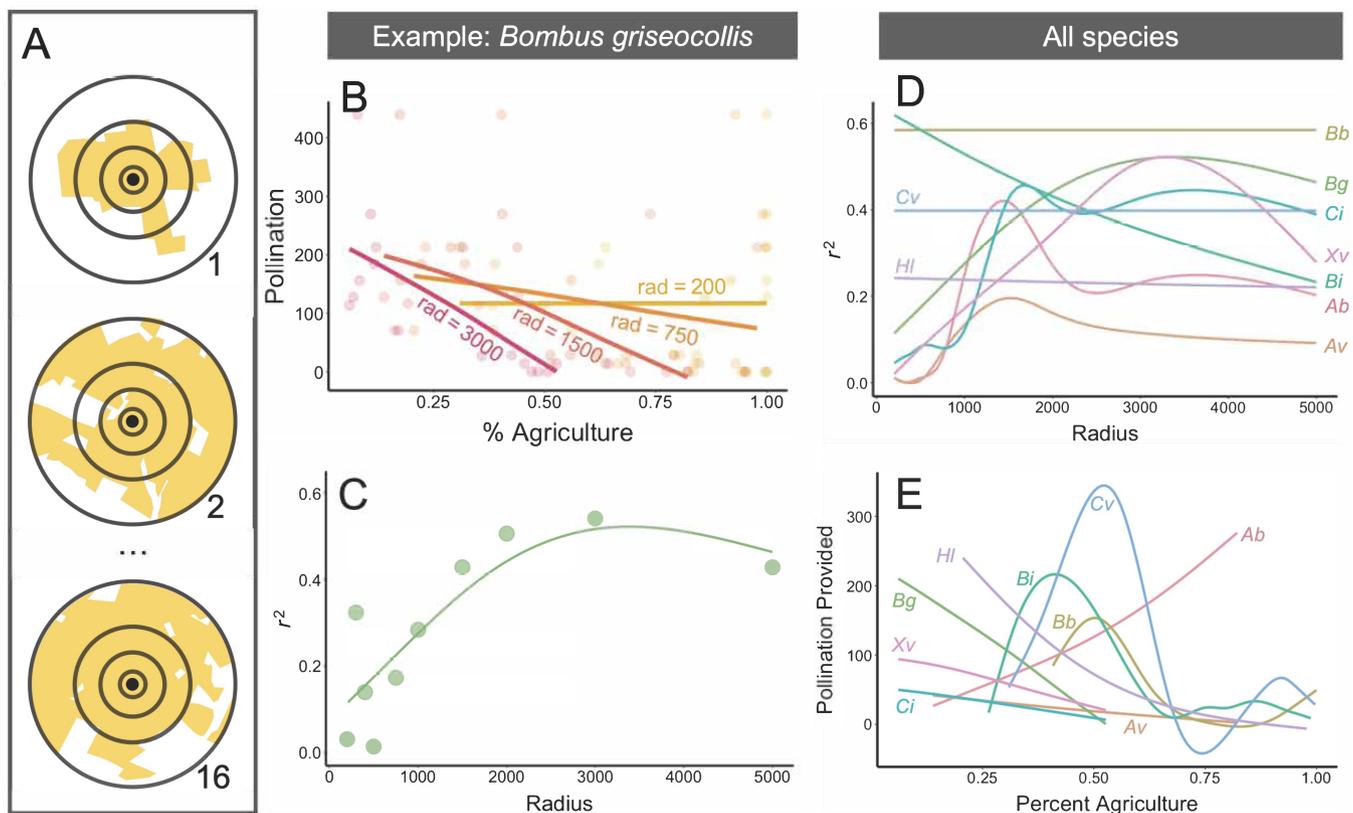

FIGURE 1 This figure explains how our analyses detected cross-scale redundancy (CSR) and response diversity (RD). Panel (A) shows three of the 16 farms in our dataset (black dots), with the gold shaded areas representing the agricultural land cover surrounding each farm. The concentric circles show four of our 10 radii of analysis. For each of nine bee species, we used generalized additive models (GAMs) to test the relationship between that species' pollination and percent agricultural land cover at each radius. Panel (B) shows, for one bee species (*Bombus griseocollis*), how the pollination-by-agriculture relationship changes as agricultural land cover is measured at the four radii shown in panel (A). (In this example, the GAMs produced linear fits.) Each line in panel (B) translates to one point in panel (C), which shows the r^2 of the pollination-by-agriculture relationship across different radii. In panel (D), we remove the data points and show one line per bee species (Ab, *Andrena bradleyi*; Av, *Andrena vicina*; Bb, *Bombus bimaculatus*; Bg, *Bombus griseocollis*; Bi, *Bombus impatiens*; Ci, *Colletes inaequalis*; Cv, *Colletes validus*; Hl, *Habropoda laboriosa*; Xv, *Xylocopa virginica*). In panel (E), we chose the best radius for each bee species (where "best" means the radius corresponding with the highest r^2 in panel D) and used this radius to determine how the pollination of each bee species changed with percent agricultural land cover.

explicitly cast as stabilizing mechanisms and this is the prediction based on the theoretical literature, although again, it has rarely been tested empirically. **H3**: A species' effects on stability, CSR, and RD will all increase with that species' abundance. Current definitions allow both rare and common species to generate signals of CSR and RD, but we predict that rare species will have weaker effects on stability even if they generate signals of CSR and RD.

METHODS

Field methods

Site selection

We collected data from 16 commercial highbush blueberry farms in southern New Jersey from 2011 to 2012. At each farm, we assessed the percent of surrounding land cover in agriculture using spatial data from the New Jersey Department of Environmental Protection GIS (http://www.state.nj.us/dep/gis). We used agricultural land cover as a measure of disturbance because it is among the primary causes of global declines in biodiversity (Pereira et al., 2010). Land cover surrounding these farms is primarily composed of two types, agriculture and forest, which together make up 85% of land cover. Farm sites were chosen so that percent of agriculture surrounding the farm at small (300 m radius) and large (1500 m radius) spatial scales were not correlated ($r^2 = 0.11$; see Appendix S1: Section S2). This study design allows for an appropriate test of CSR, because the magnitude of the disturbance at a given farm (i.e., the proportion of agriculture surrounding the farm) will vary with the spatial scale examined. All farms were separated by at least 1000 m. For more details on site selection and the surrounding landscape, see Benjamin et al. (2014). Hereafter, "radius" refers to the scale at which agricultural land cover was measured, and "agricultural land cover" refers to the percent of agricultural land cover at a given radius.

Pollinator collections and pollination

To determine pollinator abundance, bees were hand-netted from blueberry flowers at each farm in 2011 and 2012. Collections took place during peak bloom (mid-April through mid-May in our study region), with netting taking place only on relatively warm sunny days with little wind. Within each year, we collected data at each farm on three different days; on each day, we collected bees twice, once in the morning and once in the afternoon. Bee specimens were identified to species by taxonomists (see *Acknowledgements*). We quantified per-visit pollen deposition for bee morphogroups by presenting virgin flowers to foraging bees. Morphogroups were used because not all species could be identified on the wing during these visits. After completion of a single visit, we allowed pollen tubes to germinate for 48 h. Stigmas were softened and stained with aniline blue dye, and the number of pollen tetrads with germinated pollen tubes were counted using a fluorescent microscope. We estimated the contribution of each bee species to pollination by multiplying its abundance by the mean single visit pollen deposition of its morphogroup. The use of morphogroups, and our method for estimating pollen deposition, are both common approaches for studies attempting to quantify species-level contributions to pollination (Benjamin et al., 2014; Cariveau et al., 2013; Vázquez et al., 2005). Morphogroup species lists can be found in Benjamin et al. (2014).

Analytical methods

Summary

Our analysis includes several steps, which we present in short form here before explaining the details. We begin with data collected from 16 sites, with agricultural land cover measured around each site at 10 radii (Figure 1A). For each bee species, we used GAMs and found the predictive power (r^2) of the relationship between agricultural land cover and pollination, with each site as a data point (example in Figure 1B). We use these results in two ways. First, we ask how the predictive power of agricultural land cover changes across radii (Figure 1C). Comparing these curves (the entirety of the curve, not just the peaks) across species helps us quantify CSR (Figure 1D). Second, we find the most predictive radius for each bee species (i.e., each line's peak in Figure 1D). Using that radius to define agricultural land cover, we then assess RD based on how pollination changes with agricultural land cover (Figure 1E).

We subsampled our dataset to create bee communities with different richness and species compositions, allowing us to test (1) whether CSR and RD are correlated with pollinator richness and (2) whether CSR and RD are correlated with the stability of pollination across the land use gradient. Below, we describe the subsampling process and how we define stability. Throughout our analyses, we use GAMs because we do not expect any of the relationships to be linear (Ross et al., 2023), especially not for all bee species. All analyses were done in R (R Core Team, 2023)

We fit GAMs using function gam() in package mgcv (Wood, 2015).

Bee community data

To relate pollination at each farm to surrounding land cover, bee community data were summed across both years. Land cover did not meaningfully change across years, and most bee species are so rare (i.e., fewer than five individuals collected) that one year of data tells us little about their responses to agriculture. Combining data across years allows us to analyze nine species with at least 50 individuals; these species together accounted for 90% of total pollination (Appendix S1: Section S3). No other species met even half of this abundance cutoff, making it a logical point of separation. Two species stood out as functional codominants: *Bombus griseocollis* and *Andrena bradleyi*, each of which provided ~18% of total pollination (where "total" pollination means pollination summed across sites and years). No other species provided more than 11%.

Quantifying stability

We used a common definition of stability, namely $\bar{y}/\sigma$, where $\bar{y}$ is mean pollination and σ is the SD of pollination, both taken across all sites (Lehman & Tilman, 2000; Tilman, 1999). High levels of stability result when sites have consistent, high pollination regardless of surrounding agricultural land cover. Although in the literature stability is posited as resulting from CSR and/or RD, we stress that stability is not necessarily linked with either. We could potentially find, for example, high stability without any CSR or RD, or high CSR or RD without stability. Our general method allows for other measures of stability, such as the slope of a function by disturbance relationship (Cariveau et al., 2013) or the proportion of sites meeting a threshold value of function (Carnus et al., 2015) to be substituted for $\bar{y}/\sigma$ if desired.

Simulated bee communities

To examine how stability is affected by richness, CSR, and RD, we simulated bee communities by subsampling from our data (Cariveau et al., 2013). Specifically, we simulated a community for every unique combination of three to nine species. Retained bee species kept all individuals; that is, if a bee species was retained, it had its observed abundance and pollination at every site. Each simulation produced unique values for stability, CSR,

and RD, which are used in addressing H2 and H3, as described below. To be clear, these "simulations" are just subsets of the empirical data, not new data created for any purpose.

H1: Both CSR and RD will be statistically detectable

Quantifying cross-scale redundancy

To quantify CSR, we tested whether bee species varied in the spatial scale at which they responded to agricultural land cover. Specifically, we compared two GAMs, one which did not allow for CSR (model **C1**) and one that did (**C2**). Both models had as a response variable the r^2 of the "pollination by agriculture" relationship (see Figure 1C). **C1** included the effect of species identity and radius and is analogous to a linear model with those same predictors as main effects, but without the interaction term. It fit a single smooth effect describing how the r^2 of the "pollination by agriculture" relationship varies with spatial scale (i.e., the radius at which agricultural land cover was measured), allowing species to vary only in their intercepts. In contrast, **C2** fit the smooth effect separately for all species (as in Figure 1D), making it analogous to a linear model with the main *and* interactive effects of species identity and radius. If there is a signal of CSR in the data, allowing each species to have its own smooth effect will increase predictive power enough to offset the increased number of estimated parameters. In other words, **C2** will be higher quality than **C1**, with adjusted r^2 serving as our measure of model quality. The signal of CSR is then the adjusted r^2 of **C2**, minus the adjusted r^2 of **C1**. Each simulation run has one value for the "signal of CSR," which we hereafter shorten to CSR.

Quantifying RD

To quantify RD, we tested whether bee species varied in their responses (e.g., "positive but saturating" or "negative and linear") to agricultural land cover. Specifically, we compared two GAMs, one which did not allow for RD (model **R1**) and one that did (**R2**). Both models had as a response variable the pollination provided by each bee species at each site. **R1** included the effects of species identity and agricultural land cover and is analogous to a linear model with those same predictors as main effects, but no interaction. It fit a single smooth effect describing how pollination changes with agricultural land cover, allowing species to vary only in their intercepts. In contrast, **R2** fit the smooth effect separately for each species

(as in Figure 1E), making it analogous to a linear model with the main *and* interactive effects of species identity and agricultural land cover. For both models, agricultural land cover was measured at the most predictive radius for the bee species in question (extracted from the CSR analysis, above) (Winfree & Kremen, 2009). RD is therefore linked with CSR, but we argue that this is unavoidable; either RD is linked with CSR in this way, or one is potentially measuring species' responses to environmental change at radii other than those that are most predictive. If there is a signal of RD in the data, allowing each species to have its own smooth effect will increase predictive power enough to offset the increased number of estimated parameters. In other words, **R2** will be of higher quality than **R1**, with adjusted r^2 again as our measure of model quality. The signal of RD is then the adjusted r^2 of **R2**, minus the adjusted r^2 of **R1**. Each simulation run has one value for the "signal of RD," which we hereafter shorten to RD. Consistent with recent analytical tests of RD, we assume that RD can result from any nonidentical responses of different species to environmental change, for example, RD can in principle occur even when the sign of the response is the same across species (Cariveau et al., 2013; Ross et al., 2023).

For both CSR and RD, we tested our proposed methods against simulated data with expected levels of CSR and RD and found that our proposed methods worked well (Appendix S1: Section S4).

H2: CSR and RD will increase the stability of pollination

We used variation across simulation runs to understand how CSR and RD affected stability. Here, "CSR" and "RD" mean the delta r^2 values extracted from the GAMs described above (**C1**, **C2**, **R1**, **R2**). One CSR and RD value is produced by each simulation. Using these data, we then fit a final GAM (**S1**, because the response is stability) integrating richness, CSR, and RD as predictors of stability. **S1** had the following form: $S_p \sim s(\text{CSR}, \text{RD}) + s(\text{R})$, where S_p is the stability of pollination, $s(\text{CSR}, \text{RD})$ is a bivariate thin plate spline representing the main *and* interactive effects of CSR and RD, and $s(\text{R})$ is a univariate thin plate spline representing the main effect of richness. Using bivariate thin plate splines for the interaction of CSR and RD is appropriate as they share the same units (in this case, Δ adjusted r^2) and are measured on the same scale (Wood, 2015). Even though our primary interest is the main and interactive effects of CSR and RD on stability, we included richness as a covariate because richness likely stabilizes function in ways that are not captured by CSR and RD. However, including

richness has no substantial impact on our conclusions (Appendix S1: Section S5).

H3: Species' effects on stability, CSR, and RD will increase with abundance

For all nine bee species in the analyses, we calculated identity effects by first taking all simulation runs and dividing them into groups based on whether the bee species was present or absent in the species pool (i.e., present in at least one site). We then calculated, by a simple difference of means, how much the presence of each species affected mean values of (1) stability, (2) CSR, and (3) RD. We expect that the more abundant species will have greater identity effects on stability, CSR, and RD.

RESULTS

H1: Both CSR and RD will be statistically detectable

We found weak evidence of CSR and stronger evidence of RD. Averaging across simulations, the GAM that allowed for CSR (**C2**) explained 20.8% $\pm$ 12.5% (1 SD) more variation than the model that did not (**C1**). The model that allowed for RD (**R2**) explained 43.9% $\pm$ 22.9% (1 SD) more variation than the model that did not (**R1**). There was no convincing evidence of richness increasing either CSR or RD (Figure 2). Richness produced no detectable effect on CSR, and, although RD slightly increased with richness, this was mostly due to a sampling effect (Huston, 1997), specifically an increased chance to include *A. bradleyi*, which had a singularly positive response to agricultural land cover such that its presence generated RD (see also Cariveau et al., 2013).

H2: CSR and RD will increase the stability of pollination

We used our "stability" GAM (**S1**) to generate response surfaces that predicted stability ($\bar{y}/\sigma$) and its components (mean pollination, $\bar{y}$; and variability in pollination, σ) based on the effects of CSR and RD. There was no relationship between CSR and stability. However, RD was correlated with high stability, both because of higher mean pollination and lower variability of pollination when RD was present. As described later in the results, these effects are largely due to the presence of *A. bradleyi*. Although it is not our main question, for completeness, we explored the possibility that richness increased the
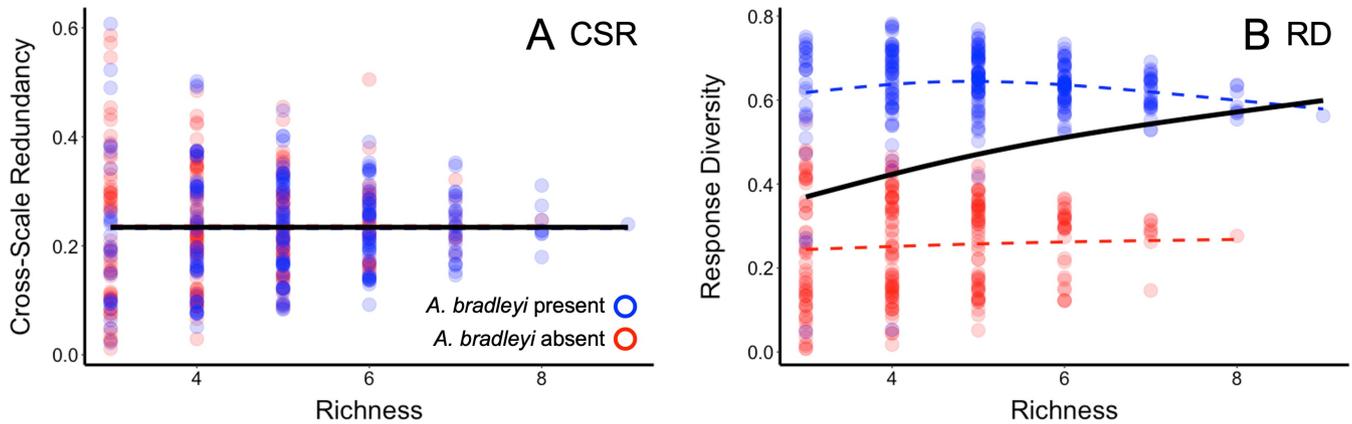
FIGURE 2 Relationship between bee species richness and cross-scale redundancy (CSR, panel A) and response diversity (RD, panel B). Each data point is one simulation, varying in richness from three to 14. Blue and red points are data points that included or did not include *Andrena bradleyi* (a common, blueberry specialist), respectively. Solid black lines are generalized additive models (GAMs) fit to all data points. Blue and red lines are GAMs fit to the blue and red data points, respectively. In panel (A), the three lines are nearly indistinguishable. In panel (B), a superficially positive effect of richness on RD was parsimoniously explained by considering the identity effects of *A. bradleyi*.

stability of pollination despite increasing neither CSR nor RD. Appendix S1: Section S5 includes this exploration, along with more details for the other results in this paragraph.

H3: Species' effects on stability, CSR, and RD will increase with abundance

A. bradleyi and *B. griseocollis*, which were among the three most abundant species and the two most important pollinators, increased stability much more than any of the other seven species (Figure 3). Both species, but especially *A. bradleyi*, provide much of the pollination at sites where total pollination is low (Appendix S1: Section S6). This increases stability ($\bar{y}/\sigma$) by substantially increasing mean pollination provided ($\bar{y}$) while lowering variability (σ). The fact that the remaining seven species contributed so little to stability provides strong support for the hypothesis that the stabilizing effects of CSR and RD are driven by the most abundant species, because *all* the bee species that we included in the analysis would be considered common (the 28 rarer species were all excluded because their low sample sizes could not yield convincing GAMs; see *Methods*).

There was no relationship between species' effects on CSR and its effect on stability (Figure 3B), likely because no species generated strong CSR in the first place. There was a stronger relationship between a species' effects on RD and its effects on stability, although this was entirely driven by *A. bradleyi* (Figure 3C); thus, it may be better to think of it as a unique identity effect rather than a

general trend attributable to richness per se. This means that although *B. griseocollis* had somewhat strong effects on stability, it was not due to CSR or RD, and other mechanisms may explain its stabilizing effects.

DISCUSSION

We present a new analytical approach that compares the strength of two often-discussed (but rarely quantified) mechanisms through which biodiversity might stabilize ecosystem function, namely RD and CSR. We build on the idea that stabilizing mechanisms, which invoke species' differential responses to the environment, can be measured as statistical interactions between species identity and a given environmental variable (Winfree & Kremen, 2009). We found that RD had stronger stabilizing effects than CSR, consistent with other studies that have identified RD as stronger than other stabilizing mechanisms (Thibaut et al., 2012; Winfree & Kremen, 2009). RD is, by definition, directly linked to ecosystem function because the strength of RD depends on whether species' function responds similarly to disturbance. CSR is one step further removed. It results from variability in the scale at which each species' function by disturbance relationship is the strongest but captures no information about the form of that relationship (e.g., negative, positive). However, the argument that RD is more directly linked to ecosystem function is only conceptual, and the studies cited above only qualitatively compared the strength of different stabilizing mechanisms. The advance here is a quantitative test comparing
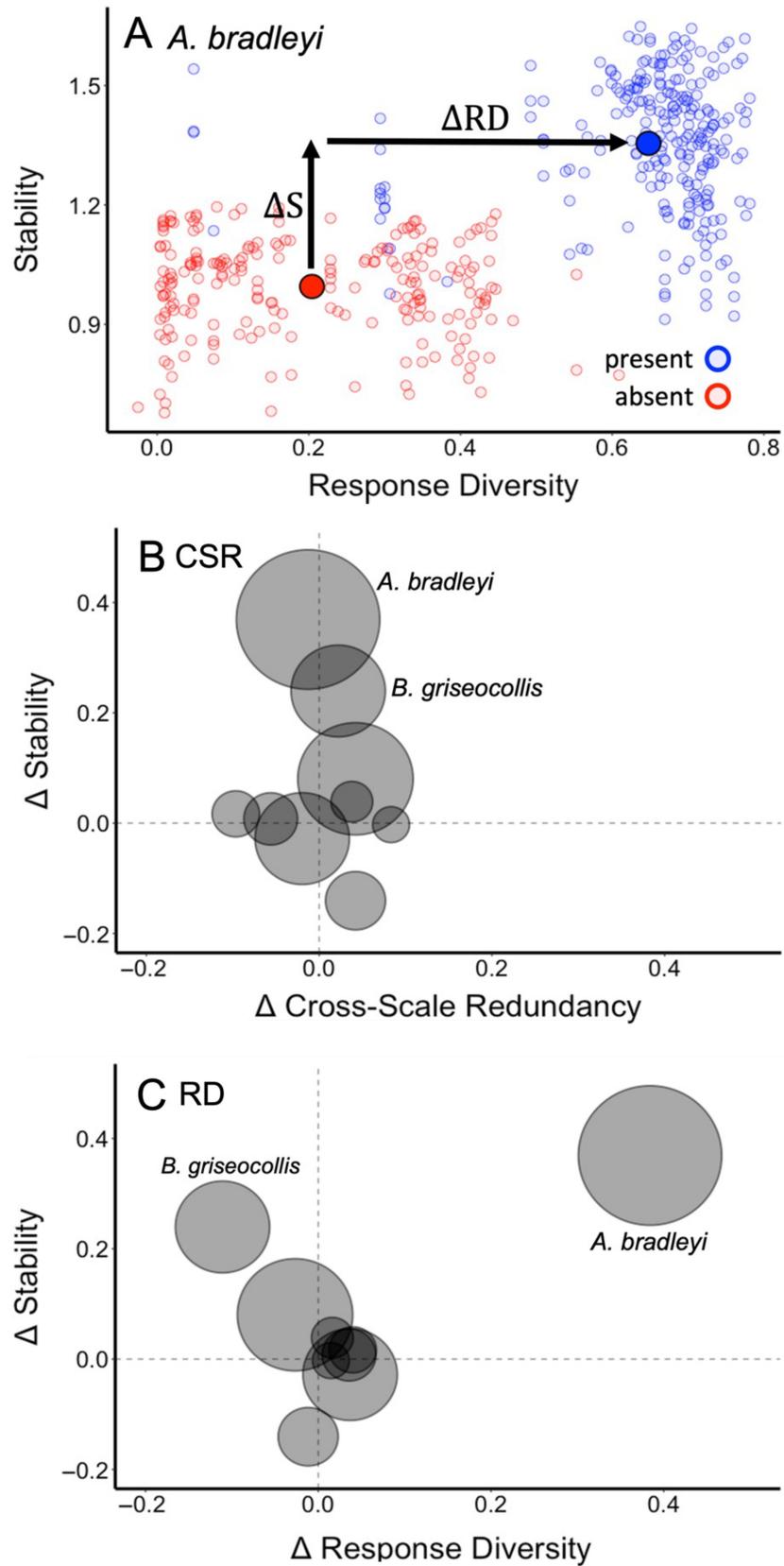
the strength of CSR and RD, applicable to any data with species-level function estimates and measurements of disturbance taken at different scales. The test we propose partially aligns with a recent framework for assessing RD (Ross et al., 2023), in that we measure traits directly linked to species' functional contributions and use GAMs to fit trait by environment relationships. However, we deviate from Ross et al. (2023) in one important way: Instead of using variation in the first derivatives of species' responses to the environment, we quantify RD based on the difference in predictive power. We consider both to be reasonable choices for translating the verbal definition of RD into statistical language.

RD was a strong stabilizing force

We found that RD, due to species' differential responses to agricultural land cover, was a stabilizing force. As agricultural land cover increased, most bee species decreased in abundance and contributed less pollination. However, the functional codominant *A. bradleyi* increased its pollination with agricultural land use, buffering other species' decrease in pollination. *A. bradleyi* is a specialist on blueberry plants, and presumably prefers these high-agriculture sites because the agriculture consists largely of blueberry fields (Benjamin et al., 2014). Previous assessments have reported that RD varies widely — from strong to almost nonexistent — although this variation may be due to the wide range methods used to measure RD (Ross et al., 2023). Reviewing all recent RD studies is beyond the scope of this paper, but studies of bee pollination generally find evidence of RD (Cariveau et al., 2013; Fründ et al., 2013; Stavert et al., 2017; Winfree & Kremen, 2009). There is a growing body of work asking whether the presence of RD actually stabilizes function (Cariveau et al., 2013; Leary & Petchey, 2009; Sasaki et al., 2019; Schnabel et al., 2021; Thibaut et al., 2012; Winfree & Kremen, 2009), but there is no consensus on this point. Sasaki et al. (2019) and Schnabel et al. (2021) both found evidence that species in plant communities respond differently to environmental change over time, leading to RD and higher stability. Thibaut et al. (2012) found that RD was the main

contributor to the portfolio effect, which stabilized fish abundance over time. In partial contrast, Cariveau et al. (2013) found that RD stabilized pollination for only one of three crops. Given that RD is cast as a stabilizing mechanism, it is surprising that few studies have attempted to link RD with a stability of abundance or function. Further work in this area would be valuable.

Cross-scale redundancy had almost no stabilizing effects

In contrast, cross-scale redundancy, which resulted from the differential scales at which species respond to agricultural land cover, had almost no detectable effect on stability. A key part of CSR involves the contributions of trait diversity and how these traits affect the scale at which ecosystem function is provided (Massol et al., 2017; Peterson et al., 1998). Here, we expected that variation in bee body size would affect their foraging range (Kendall et al., 2022), and as such would also affect the radius at which the pollination-by-agricultural-land-use relationship was strongest. While we can visually detect this pattern (e.g., *B. griseocollis* responding at a larger scale than *A. bradleyi*; Figure 1D), the signal of CSR was much weaker than the signal of RD. Previous papers have quantified the variable scale at which different species respond to landscape characteristics (Roland & Taylor, 1997; Steffan-Dewenter et al., 2002; Wardwell et al., 2008; Winfree & Kremen, 2009). However, we are unaware of papers that have both quantified the strength of CSR and linked it to measures of stability. More work on this area would help determine whether species' scale-specific responses to disturbance are generally important to stability.

Importance of dominant species for stability mechanisms

The presence of two codominant bee species (*A. bradleyi* and *B. griseocollis*) was strongly associated with increased stability (Figure 3B,C). Because stability is defined as the

FIGURE 3 We quantified the effects of each of the nine bee species we studied on cross-scale redundancy (CSR), response diversity (RD), and stability. Panel (A) shows how species effects are defined, using *Andrena bradleyi* as an example. Points are separated by whether *A. bradleyi* is present (blue) or absent (red) from the simulated bee community. Open circles show the results of the 466 simulations, and the larger filled circles are group mean values taken across those simulations. The difference between the group means gives the change in stability (ΔS) and response diversity (ΔRD) associated with the presence of *A. bradleyi*. The same logic would apply to ΔCSR, if CSR had been on the *x*-axis. In panels (B) and (C), each data point is one bee species, with *x*- and *y*-values calculated as shown in panel A (i.e., ΔRD on the *x*-axis and ΔS on the *y*-axis). Points are sized by each species' abundance summed across sites; we stress that point sizes therefore do not reflect CIs. (B) The link between CSR and stability is weak: No species generates a strong signal of CSR, and to the extent there is any signal it is not correlated with stability. (C) In contrast, the link between RD and stability is strong, almost entirely because *A. bradleyi* strongly increases both.

ratio of mean function to variability in function (i.e., $\bar{y}/\sigma$), stabilizing species either decrease variability or increase mean function enough to offset the increase in variability (Cottingham et al., 2001; Gross et al., 2014). Early biodiversity–function research focused on how quickly variability increased with mean function, averaged across all species in a community (i.e., mean-variance scaling) (Tilman, 1999). However, more recent work has highlighted the key stabilizing effects on individual species, if they diverge from the community's overall mean-variance scaling relationship (Grman et al., 2010). We found clear evidence that *A. bradleyi* and (to a lesser extent) *B. griseocollis* did just this, by contributing much less to variability than would be expected based on their high mean function (Appendix S1: Section S7). In ecological terms, this means that *A. bradleyi* and *B. griseocollis* were important pollinators at sites that received little pollination from other species and may have traits allowing them to persist and provide pollination at sites where other species struggle. For *A. bradleyi*, the trait and associated mechanisms appear clear; it contributes to stability through RD because it is a blueberry specialist and most agricultural fields in the study region grow blueberry (Benjamin et al., 2014). Thus, the response of *A. bradleyi* to increasing agricultural land cover is positive, unlike other species which show mostly negative responses. In contrast, *B. griseocollis* had a strongly negative response to agricultural land cover such that its presence decreased RD (Figure 3B). However, as *B. griseocollis* is a key pollinator at sites with average function (Appendix S1: Section S6), it was second only to *A. bradleyi* in its stabilizing effect. We stress that abundant species will not necessarily increase stability, because they could easily add enough variability to offset even their large increase to mean function. Further exploration of whether abundant species tend to have lower variability than expected would be useful.

The outsized importance of a few species is consistent with the fact that ecological communities invariably include few common and many rare species (McGill et al., 2007). However, the important role of dominant species may also depend, at least partially, on how we chose to define stability. The across-site stability metric we chose, while very common in biodiversity research (Lehman & Tilman, 2000; Tilman, 1999), can obscure important patterns, such as whether sites are meeting key threshold values of function (Carnus et al., 2015). Rare species are more likely to be identified as important if the biodiversity–stability relationship is conceptualized as the number of species needed to meet thresholds of function at multiple sites, because such analyses capture the patchy distribution of rare species and the potential for many species to be essential for function at low-function

sites (Genung et al., 2023; Kremen et al., 2002; Simpson et al., 2022; Winfree et al., 2018).

Conclusions

When and how biodiversity stabilizes ecosystem function is an important and long-standing question in ecological research (Cardinale et al., 2012; Elton, 1958; Gonzalez & Loreau, 2009; Hector et al., 2010). Questions involving compensatory dynamics, specifically whether species respond differently to environmental changes, are central to answering this question (Gonzalez & Loreau, 2009). While studies of RD and CSR are increasingly common, few have explicitly tested whether these stabilizing mechanisms are correlated with stability, and more attention to this knowledge gap is warranted. We presented methods for detecting and quantifying the strength of two commonly cited stabilizing mechanisms: cross-scale redundancy and RD. We hope these methods contribute to a better understanding of when there is a strong signal of these stabilizing mechanisms, when this signal is linked with stability, and the relative roles of dominant and rare species.

CONFLICT OF INTEREST STATEMENT
The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT
Data (Genung & Winfree, 2024a) are available in Dryad at https://doi.org/10.5061/dryad.0rxwdbs9c. Code (Genung & Winfree, 2024b) is available in Zenodo at https://doi.org/10.5281/zenodo.13352594.

ORCID
Mark A. Genung https://orcid.org/0000-0002-9383-667X

REFERENCES
Allen, C. R., L. Gunderson, and A. Johnson. 2005. "The Use of Discontinuities and Functional Groups to Assess Relative Resilience in Complex Systems." *Ecosystems* 8: 958–966.

Angeler, D. G., C. R. Allen, C. Barichievy, T. Eason, A. S. Garmestani, N. A. Graham, D. Granholm, L. H. Gunderson, M. Knutson, and K. L. Nash. 2016. "Management Applications of Discontinuity Theory." *Journal of Applied Ecology* 53: 688–698.

Benjamin, F. E., J. R. Reilly, and R. Winfree. 2014. "Pollinator Body Size Mediates the Scale at Which Land Use Drives Crop Pollination Services." *Journal of Applied Ecology* 51: 440–49.

Benjamin, F. E., and R. Winfree. 2014. "Lack of Pollinators Limits Fruit Production in Commercial Blueberry (Vaccinium corymbosum)." *Environmental Entomology* 43(6): 1574–1583.

Brittain, C., C. Kremen, and A. M. Klein. 2013. "Biodiversity Buffers Pollination from Changes in Environmental Conditions." *Global Change Biology* 19: 540–47.

Burgess, S. C., E. C. Johnston, A. S. Wyatt, J. J. Leichter, and P. J. Edmunds. 2021. "Response Diversity in Corals: Hidden Differences in Bleaching Mortality among Cryptic *Pocillopora* Species." *Ecology* 102: e03324.

Cardinale, B. J., J. E. Duffy, A. Gonzalez, D. U. Hooper, C. Perrings, P. Venail, A. Narwani, et al. 2012. "Biodiversity Loss and Its Impact on Humanity." *Nature* 486: 59–67.

Cariveau, D. P., N. M. Williams, F. E. Benjamin, and R. Winfree. 2013. "Response Diversity to Land Use Occurs But does not Consistently Stabilise Ecosystem Services Provided by Native Pollinators." *Ecology Letters* 16: 903–911.

Carnus, T., J. A. Finn, L. Kirwan, and J. Connolly. 2015. "Assessing the Relationship between Biodiversity and Stability of Ecosystem Function–Is the Coefficient of Variation Always the Best Metric?" *Ideas in Ecology and Evolution* 7: 89–96.

Cottingham, K., B. Brown, and J. Lennon. 2001. "Biodiversity May Regulate the Temporal Variability of Ecological Systems." *Ecology Letters* 4: 72–85.

Doak, D. F., D. Bigger, E. Harding, M. Marvier, R. O'malley, and D. Thomson. 1998. "The Statistical Inevitability of Stability-Diversity Relationships in Community Ecology." *The American Naturalist* 151: 264–276.

Elmqvist, T., C. Folke, M. Nyström, G. Peterson, J. Bengtsson, B. Walker, and J. Norberg. 2003. "Response Diversity, Ecosystem Change, and Resilience." *Frontiers in Ecology and the Environment* 1: 488–494.

Elton, C. 1958. *Ecology of Invasions by Animals and Plants.* New York: Springer.

Fischer, J., D. Lindenmayer, S. Blomberg, R. Montague-Drake, A. Felton, and J. Stein. 2007. "Functional Richness and Relative Resilience of Bird Communities in Regions with Different Land Use Intensities." *Ecosystems* 10: 964–974.

Fründ, J., S. L. Zieger, and T. Tscharntke. 2013. "Response Diversity of Wild Bees to Overwintering Temperatures." *Oecologia* 173: 1639–48.

Genung, M., and R. Winfree. 2024a. "Data From: Dominant Species Stabilize Pollination Services through Response Diversity, but not Cross-Scale Redundancy [Dataset]." Dryad. https://doi.org/10.5061/dryad.0rxwdbs9c.

Genung, M., and R. Winfree. 2024b. "Data From: Dominant Species Stabilize Pollination Services through Response Diversity, but not Cross-Scale Redundancy." Zenodo. https://doi.org/10.5281/zenodo.13352594.

Genung, M. A., J. Reilly, N. M. Williams, A. Buderi, J. Gardner, and R. Winfree. 2023. "Rare and Declining Bee Species Are Key to Consistent Pollination of Wildflowers and Crops across Large Spatial Scales." *Ecology* 104: e3899.

Gonzalez, A., and M. Loreau. 2009. "The Causes and Consequences of Compensatory Dynamics in Ecological Communities." *Annual Review of Ecology, Evolution, and Systematics* 40: 393–414.

Grman, E., J. A. Lau, D. R. Schoolmaster, Jr., and K. L. Gross. 2010. "Mechanisms Contributing to Stability in Ecosystem Function Depend on the Environmental Context." *Ecology Letters* 13: 1400–1410.

Gross, K., B. J. Cardinale, J. W. Fox, A. Gonzalez, M. Loreau, H. W. Polley, P. B. Reich, and J. Van Ruijven. 2014. "Species Richness and the Temporal Stability of Biomass Production: A New Analysis of Recent Biodiversity Experiments." *The American Naturalist* 183: 1–12.

Hector, A., Y. Hautier, P. Saner, L. Wacker, R. Bagchi, J. Joshi, M. Scherer-Lorenzen, E. Spehn, E. Bazeley-White, and M. Weilenmann. 2010. "General Stabilizing Effects of Plant Diversity on Grassland Productivity through Population Asynchrony and Overyielding." *Ecology* 91: 2213–20.

Holling, C. 1988. "Temperate Forest Insect Outbreaks, Tropical Deforestation and Migratory Birds." *The Memoirs of the Entomological Society of Canada* 120: 21–32.

Huston, M. A. 1997. "Hidden Treatments in Ecological Experiments: Re-Evaluating the Ecosystem Function of Biodiversity." *Oecologia* 110: 449–460.

Ives, A. R., and S. R. Carpenter. 2007. "Stability and Diversity of Ecosystems." *Science* 317: 58–62.

Kendall, L. K., J. M. Mola, Z. M. Portman, D. P. Cariveau, H. G. Smith, and I. Bartomeus. 2022. *"The Potential and Realized Foraging Movements of Bees are Differentially Determined by Body Size and Sociality." Ecology* 103(11): e3809.

Kremen, C., N. M. Williams, M. A. Aizen, B. Gemmill-Herren, G. LeBuhn, R. Minckley, L. Packer, S. G. Potts, T. A. Roulston, and I. Steffan-Dewenter. 2007. "Pollination and Other Ecosystem Services Produced by Mobile Organisms: A Conceptual Framework for the Effects of Land-Use Change." *Ecology Letters* 10: 299–314.

Kremen, C., N. M. Williams, and R. W. Thorp. 2002. "Crop Pollination from Native Bees at Risk from Agricultural Intensification." *Proceedings of the National Academy of Sciences of the United States of America* 99: 16812–16.

Leary, D. J., and O. L. Petchey. 2009. "Testing a Biological Mechanism of the Insurance Hypothesis in Experimental Aquatic Communities." *Journal of Animal Ecology* 78: 1143–51.

Lehman, C. L., and D. Tilman. 2000. "Biodiversity, Stability, and Productivity in Competitive Communities." *The American Naturalist* 156: 534–552.

Lemanski, N. J., N. M. Williams, and R. Winfree. 2022. "Greater Bee Diversity Is Needed to Maintain Crop Pollination Over Time." *Nature Ecology & Evolution* 6: 1516–23.

Martin, E. A., B. Feit, F. Requier, H. Friberg, and M. Jonsson. 2019. "Assessing the Resilience of Biodiversity-Driven Functions in Agroecosystems under Environmental Change." *Advances in Ecological Research* 60: 59–123.

Massol, F., F. Altermatt, I. Gounand, D. Gravel, M. A. Leibold, and N. Mouquet. 2017. "How Life-History Traits Affect Ecosystem Properties: Effects of Dispersal in Meta-Ecosystems." *Oikos* 126: 532–546.

McCann, K. S. 2000. "The Diversity–Stability Debate." *Nature* 405: 228–233.

McGrady-Steed, J., and P. J. Morin. 2000. "Biodiversity, Density Compensation, and the Dynamics of Populations and Functional Groups." *Ecology* 81: 361–373.

McGill, B. J., R. S. Etienne, J. S. Gray, D. Alonso, M. J. Anderson, H. K. Benecha, M. Dornelas, et al. 2007. "Species Abundance Distributions: Moving Beyond Single Prediction Theories to Integration within an Ecological Framework." *Ecology letters* 10(10): 995–1015.

Nash, K. L., N. A. Graham, S. Jennings, S. K. Wilson, and D. R. Bellwood. 2016. "Herbivore Cross-Scale Redundancy Supports Response Diversity and Promotes Coral Reef Resilience." *Journal of Applied Ecology* 53: 646–655.

Pereira, H. M., P. W. Leadley, V. Proenca, R. Alkemade, J. P. Scharlemann, J. F. Fernandez-Manjarres, M. B. Araujo, et al. 2010. "Scenarios for Global Biodiversity in the 21st Century." *Science* 330: 1496–1501.

Peterson, G., C. R. Allen, and C. S. Holling. 1998. "Ecological Resilience, Biodiversity, and Scale." *Ecosystems* 1: 6–18.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Ritchie, M. E., and H. Olff. 1999. "Spatial Saling Laws Yield a Synthetic Theory of Biodiversity." *Nature* 400(6744): 557–560.

Roland, J., and P. D. Taylor. 1997. "Insect Parasitoid Species Respond to Forest Structure at Different Spatial Scales." *Nature* 386: 710–13.

Ross, S. R. J., O. L. Petchey, T. Sasaki, and D. W. Armitage. 2023. "How to Measure Response Diversity." *Methods in Ecology and Evolution* 14, no. 5: 1150–1167.

Sasaki, T., X. Lu, M. Hirota, and Y. Bai. 2019. "Species Asynchrony and Response Diversity Determine Multifunctional Stability of Natural Grasslands." *Journal of Ecology* 107: 1862–75.

Schnabel, F., X. Liu, M. Kunz, K. E. Barry, F. J. Bongers, H. Bruelheide, A. Fichtner, W. Härdtle, S. Li, and C.-T. Pfaff. 2021. "Species Richness Stabilizes Productivity via Asynchrony and Drought-Tolerance Diversity in a Large-Scale Tree Biodiversity Experiment." *Science Advances* 7: eabk1643.

Simpson, D. T., L. R. Weinman, M. A. Genung, M. Roswell, M. MacLeod, and R. Winfree. 2022. "Many Bee Species, Including Rare Species, Are Important for Function of Entire Plant–Pollinator Networks." *Proceedings of the Royal Society B* 289: 20212689.

Stavert, J. R., D. E. Pattemore, A. C. Gaskett, J. R. Beggs, and I. Bartomeus. 2017. "Exotic Species Enhance Response Diversity to Land-Use Change but Modify Functional Composition." *Proceedings of the Royal Society B: Biological Sciences* 284: 20170788.

Steffan-Dewenter, I., U. Münzenberg, C. Bürger, C. Thies, and T. Tscharntke. 2002. "Scale-Dependent Effects of Landscape Context on Three Pollinator Guilds." *Ecology* 83: 1421–32.

Sundstrom, S. M., D. G. Angeler, C. Barichievy, T. Eason, A. Garmestani, L. Gunderson, M. Knutson, K. L. Nash, T. Spanbauer, and C. Stow. 2018. "The Distribution and Role of Functional Abundance in Cross-Scale Resilience." *Ecology* 99: 2421–32.

Thibaut, L. M., S. R. Connolly, and H. P. Sweatman. 2012. "Diversity and Stability of Herbivorous Fishes on Coral Reefs." *Ecology* 93: 891–901.

Tilman, D. 1999. "The Ecological Consequences of Changes in Biodiversity: A Search for General Principles." *Ecology* 80: 1455–74.

Vázquez, D. P., W. F. Morris, and P. Jordano. 2005. "Interaction Frequency as a Surrogate for the Total Effect of Animal Mutualists on Plants." *Ecology Letters* 8: 1088–94.

Walker, B., A. Kinzig, and J. Langridge. 1999. "Plant Attribute Diversity, Resilience, and Ecosystem Function: The Nature and Significance of Dominant and Minor Species." *Ecosystems* 2: 95–113.

Wardwell, D. A., C. R. Allen, G. D. Peterson, and A. J. Tyre. 2008. "A Test of the Cross-Scale Resilience Model: Functional Richness in Mediterranean-Climate Ecosystems." *Ecological Complexity* 5: 165–182.

Winfree, R. 2013. "Global Change, Biodiversity, and Ecosystem Services: What Can We Learn from Studies of Pollination?" *Basic and Applied Ecology* 14: 453–460.

Winfree, R., and C. Kremen. 2009. "Are Ecosystem Services Stabilized by Differences among Species? A Test Using Crop Pollination." *Proceedings of the Royal Society of London B: Biological Sciences* 276: 229–237.

Winfree, R., J. R. Reilly, I. Bartomeus, D. P. Cariveau, N. M. Williams, and J. Gibbs. 2018. "Species Turnover Promotes the Importance of Bee Diversity for Crop Pollination at Regional Scales." *Science* 359: 791–93.

Wood, S. 2015. "Package 'mgcv'." *R Package Version* 1: 29.

Yachi, S., and M. Loreau. 1999. "Biodiversity and Ecosystem Productivity in a Fluctuating Environment: the Insurance Hypothesis." *Proceedings of the National Academy of Sciences of the United States of America* 96: 1463–68.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Appendix S1

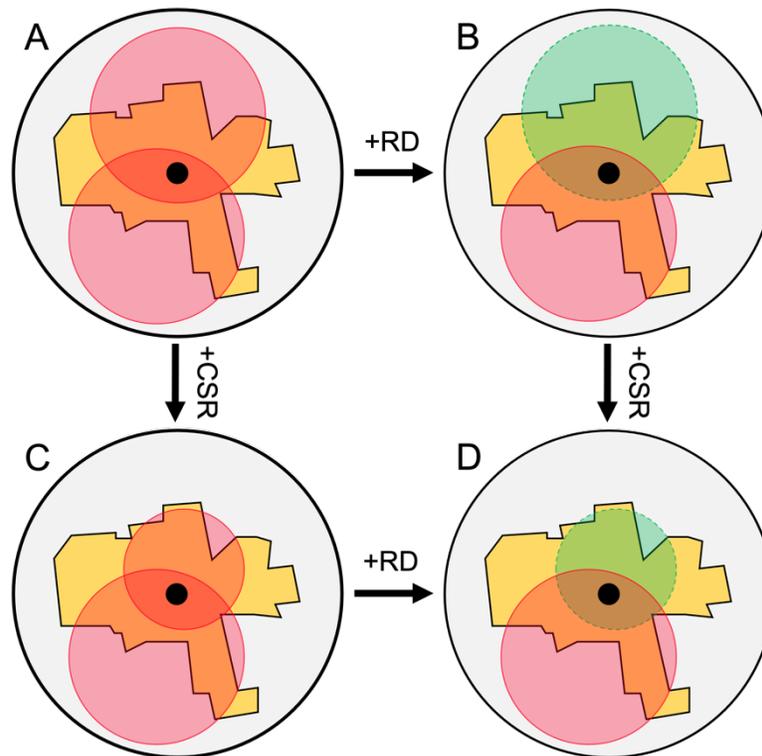**Section S1: Illustration of cross-scale redundancy and response diversity.**

Figure S1. A cartoon illustrating response diversity (RD) and cross-scale redundancy (CSR). Each panel shows the area around a farm (represented by a black point), with gold shading representing agricultural land cover, which is the same in each case. Red (solid borders) and green (dashed borders) circles are the foraging ranges of two bee species with negative and positive, respectively, responses to agricultural land cover (RD). Circle sizes correspond to bee species with different foraging range sizes (CSR). In (A), neither RD nor CSR is present. If we imagine the gold area being converted from natural habitat to agriculture, pollination provided by both bee species would decline with no stabilizing mechanism to buffer that decline. Moving from (A) to (B), the presence of RD means that pollination would be stabilized by the green species' positive response to agriculture. Moving from (A) to (C), adding CSR would not stabilize function because the species whose foraging range shrank the most now experiences an even higher proportion of agricultural land cover. However, if the same contraction had occurred for a species with positive response to agricultural land cover (i.e., moving from (B) to (D)),

adding CSR would stabilize function. These examples are illustrative, not exhaustive; CSR and RD can have a wide range of positive and negative effects on the stability of function.

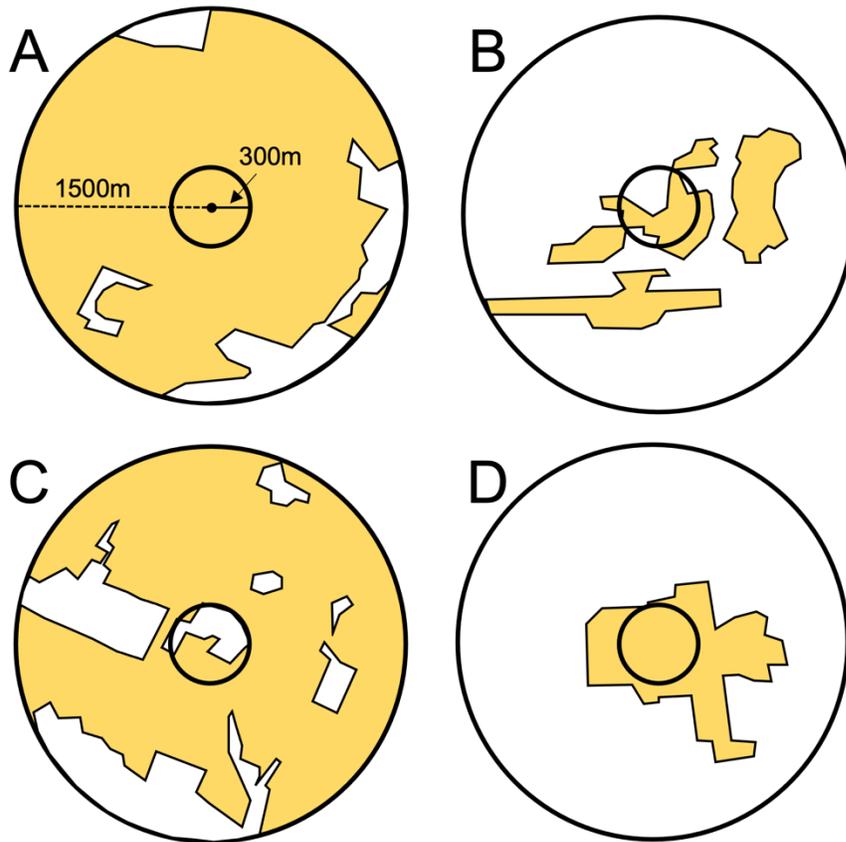**Section S2: Land use variation across scales at studied farms.**

Figure S2. Farms were chosen such that land use at smaller scales (300m radius, smaller inner circles) was not correlated with land use at larger scales (1500m radius, larger outer circles). Data were collected at the center point; yellow areas represent agricultural land cover surrounding four of our 16 study farms. While it is easy to find study sites corresponding to A and B, for which agricultural land cover is similar across scales, the key contribution of our study design was to include sites such as shown in panel C (low agricultural land cover at smaller scales, but high agricultural land cover at larger scales) and panel D (high agricultural land cover at smaller scales, but low agricultural land cover at larger scales). Adapted from Benjamin et al. (2014).

Section S3: Bee species abundance and function.

Table S1. This table shows the abundance, per-visit pollen deposition rate, and total pollination (a product of the first two values) for each bee species. Our analyses focused on the nine bee species with the highest contribution to total pollination (in **bold**), which were also the nine most abundant bee species.

Species	Abundance	Pollen per visit	Total Pollination
Bombus griseocollis	132.0	14.2	1872.2
Andrena bradleyi	202.0	9.2	1864.1
Colletes validus	132.0	8.6	1138.7
Habropoda laboriosa	83.0	13.4	1110.0
Xylocopa virginica	162.0	6.1	996.2
Bombus bimaculatus	65.0	14.2	921.9
Bombus impatiens	50.0	14.2	709.2
Colletes inaequalis	57.0	8.6	491.7
Andrena vicina	74.0	5.2	385.9
Bombus perplexus	23.0	14.2	326.2
Colletes thoracicus	20.0	8.6	172.5
Colletes cressoni	9.0	8.6	77.6
Ceratina calcarata	9.0	8.3	74.9
Andrena carlini	10.0	5.2	52.2
Lasioglossum pilosum	6.0	8.3	49.9
Augochlora pura	5.0	8.3	41.6
Augochlorella aurata	4.0	8.3	33.3
Lasioglossum oblongum	4.0	8.3	33.3
Andrena fenningeri	3.0	9.2	27.7
Andrena imitatrix	3.0	9.2	27.7
Andrena morrisonella	3.0	9.2	27.7
Lasioglossum weemsi	3.0	8.3	25.0
Lasioglossum zephyrum	3.0	8.3	25.0
Andrena banksi	1.0	9.2	9.2
Andrena barbara	1.0	9.2	9.2
Andrena carolina	1.0	9.2	9.2
Andrena cressonii	1.0	9.2	9.2
Andrena hilaris	1.0	9.2	9.2
Andrena ilicis	1.0	9.2	9.2
Andrena mandibularis	1.0	9.2	9.2
Andrena screpteropsis	1.0	9.2	9.2
Lasioglossum acuminatum	1.0	8.3	8.3
Lasioglossum coeruleum	1.0	8.3	8.3

Lasioglossum fuscipenne	1.0	8.3	8.3
Lasioglossum leucocomum	1.0	8.3	8.3
Lasioglossum subviridatum	1.0	8.3	8.3
Lasioglossum versatum	1.0	8.3	8.3

Section S4: Simulations for methods validation.

Part 1. Observed vs. expected CSR and RD.
Because we are proposing new methods for measuring the strength of CSR and RD, it is important to test whether these methods are performing as expected. To do this, we simulated a range of datasets in which we factorially manipulated the strength of CSR and RD in the data. This simulated data gives us <u>expected</u> values of CSR and RD, based on the verbal definitions of each. (Below, we detail how we simulated weak and strong values of CSR and RD.) We then run the simulated data through our proposed method to check whether it returns <u>observed</u> values of CSR and RD that match the expected values.

CSR and RD as the difference in explanatory power between two models.
We stress that we are using the difference in r2 values between two models. For example, the strength of RD is the difference between a (1) model predicting species' pollination based on two main effects: percent agriculture at the site, and species identity; and (2) the same model but allowing the main effects to interact, i.e., for each species to have its own unique "pollination by percent agriculture" relationship. If all species have the same "pollination by percent agriculture" relationship, the second model would provide no extra explanatory power and we would conclude that RD was very weak or nonexistent. The same argument applies to CSR.

Generating weak and strong values of CSR and RD.
Our simulated datasets had ten levels of CSR, ranging from what should be detected as "very weak" to "very strong". Weak CSR means species responded at the same radii to agriculture, while strong CSR means species responded at different radii. Because there were ten radii, we were able to set up ten levels of CSR. Our simulated datasets also had 40 levels of RD, ranging from what should be detected as "very weak" to "very strong". Weak RD means species showed the same response to agriculture, while strong RD means species showed different response. How we manipulated the simulated data to achieve this variation in CSR and RD strength is described below. We factorially combined the ten levels of CSR and 40 levels of RD for 400 total simulations.

Data needed for simulations.
In each simulation, we assigned species to sites based on their preference for agricultural land use at a given radius. To do this, we needed three things.

First, we generated a matrix with the (simulated) percent agriculture surrounding ten sites at ten different radii (the same ten radius values as in the main text). Percent agriculture at the smallest radius was randomly chosen between 10 and 100%. We set up the data such that percent agriculture tended to (but was not required to) decline at larger radii, matching our empirical data. This matrix was re-generated for all 400 simulations.

Second, we needed realistic variation in species abundances. We pulled 50 values from a log-normal distribution with a log(mean) of 3.5 and a log(SD) of 1.5. We kept the ten most abundant species. These choices were designed to mimic our empirical data, and our decision to analyze only abundant species. These abundance values are <u>regional</u>, i.e. summed across all sites. The regional species abundance values were re-generated for all 400 simulations.

Third, we needed a vector of each species' preference for agriculture. Our simulations included ten bee species. The first nine always preferred 0% agriculture. The last species gradually shifted from preferring 0% to 100% agriculture across the 40 levels of RD strength. We randomly chose which species was last; in other words, we didn't require it to be the most or least abundant.

How individuals were assigned to sites.
We distributed each species' regional abundance across sites based on how well their <u>preference for agriculture</u> matched the <u>simulated agriculture</u> at each site, at a given radius. Details can be seen in the publicly available R code, but we include some key points here. Specifically, we used:

```
sample(1:nsites, sum(sim.abuns[,j]), rep=T, prob=ag.match^2)
```

Meaning we sampled all available sites with replacement, a number of times based on the abundance of species `j`, with each site having a probability based on how close its percent agriculture (at a given radius) was to species `j`'s preference. `ag.match` was defined as:

```
1 - abs(ag.pref[j] - ag[,chosen.ag[j]])
```

Where `ag.pref[j]` is species `j`'s preference for agriculture, and `ag[,chosen.ag[j]]` is a column giving the percent agriculture for all sites, at the radius preferred by species `j`, which is `chosen.ag[j]`. `ag.match` is squared to strengthen the signal of species' preference for a given level of agriculture in the simulated data.

The first nine species always used the smallest radius (200m). The last species shifted from using the smallest radius to the largest radius (5000m) across the ten levels of CSR strength. As above, the "last" species was randomly chosen.

Comparing observed and expected values.
Finally, we ran all 400 simulated datasets through the same code used for our empirical analysis. If our proposed methods are working well, we should see <u>observed</u> values for CSR and RD that match the <u>expected</u> values (based the ten levels of CSR and 40 levels of RD described above). In

short, we found that our proposed methods worked well (see Figures and Figure Legends, below).
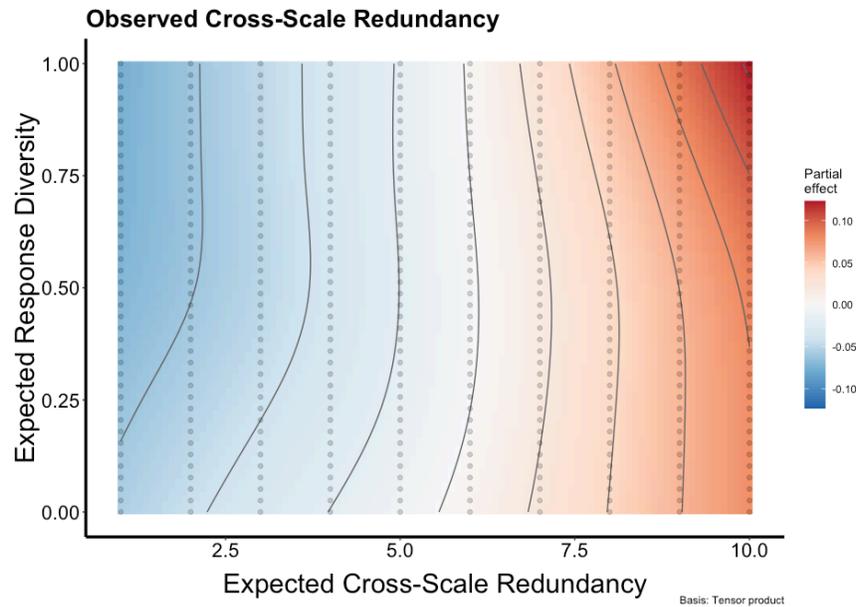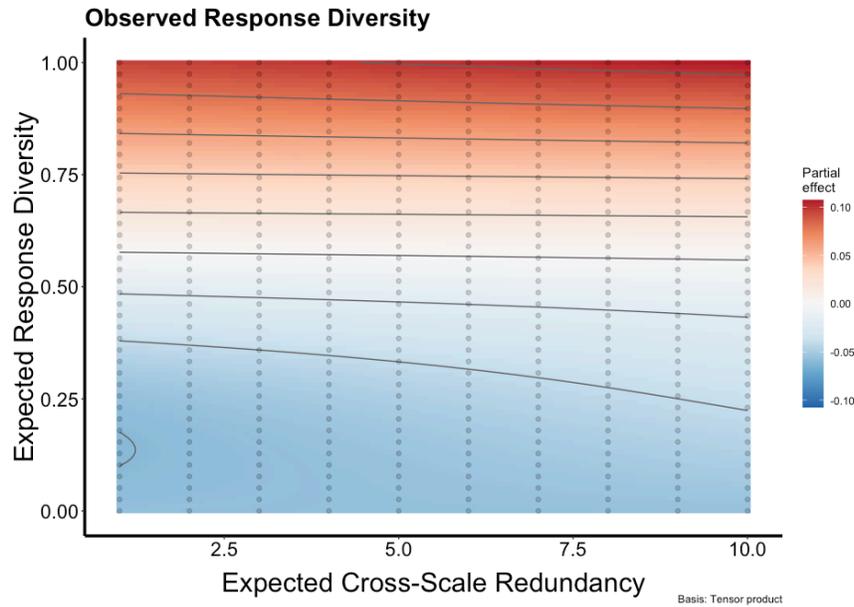

Figure S3. In the CSR surface plot above, the x-axis shows expected CSR (the ten levels of CSR strength) and the y-axis shows expected RD (the 40 levels of RD strength). Observed CSR increases (more red, less blue) with expected CSR. The nature of the relationship changes slightly with expected RD, but it is always positive. We were not necessarily expecting a linear relationship, but anything other than positive monotonic would have been concerning.

Observed Response Diversity

Figure S4. In the RD surface plot above, the x-axis shows expected CSR (the ten levels of CSR strength) and the y-axis shows expected RD (the 40 levels of RD strength). Observed RD increase (more red, less blue) with expected RD. The nature of the relationship changes slightly with expected RD, but it is always positive, with a near-trivial exception at very low values of CSR. Again, we were not necessarily expecting a linear relationship, but anything other than positive monotonic would have been concerning.

Alternate versions of simulations.
We experimented with alternate versions of the simulation, but nothing substantially changed our conclusions. For example, we tried using one fixed regional species abundance distribution, or one fixed agriculture by radius matrix, instead of re-generating these two things for all 400 combinations of expected CSR and RD. We also changed the parameters of the regional species abundance distribution. Finally, we experimented with requiring the "different" species (the "last" species, in the simulation descriptions) to be the most abundant species. Again, none of these changes had a major effect.

Part 2: Relative importance of variation in slopes vs. error estimates around slopes.
Our method uses r^2 values (specifically, the change in r^2 between two models; see Methods) to calculate the strength of CSR and RD. As such, our method is at least somewhat accounting for the amount of error (i.e., the tightness of points) around a smoother. For example, when calculating RD, it is accounting for the amount of error around the "abundance by percent agriculture" relationship for a given bee species. This is only a problem if our method is over-emphasizing the amount of error at the expense of variability among species in the steepness and signs of each smoother.

We created synthetic datasets in which we separately manipulated (1) how much bee species varied in their preference for agriculture, and (2) how strong these preferences were. The distinction is easier to clarify with a visual aid, so please see Figure S5 below.
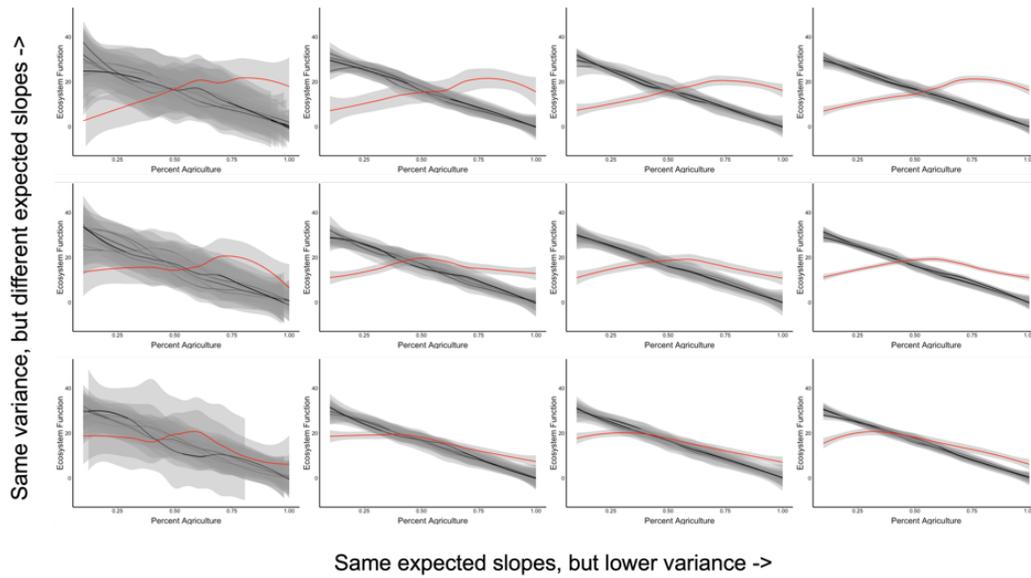

Figure S5. As we move from the bottom to the top panels of the figure, one species increasingly has a different preference for agriculture than the other nine. As we move from the left to the right panels in the figure, the degree of species' preferences for a given level of agriculture became stronger (i.e., error around the smoother decreases). The point of this figure is to illustrate the range of possibilities that we simulated. The more important issue is how these possibilities correlated with measured levels of RD, which we explore in Figure S6.
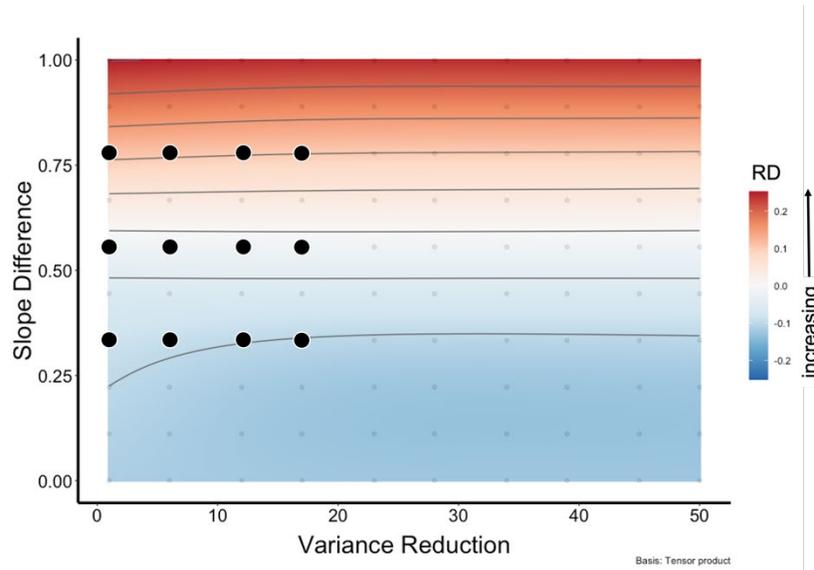
Figure S6. Observed strength of RD, with the 12 black points corresponding (i.e., matched row and column position) with the 12 panels in Figure S5. The response surface shows that our measured strength of RD is almost entirely due to how much bee species varied in their preference for agriculture (i.e., moving up and down rows in Figure S5). We consider this a desirable property because RD has traditionally focused on the sign and steepness of species' responses to environmental change, rather than the estimated error around these responses.

In Part 2, CSR is analogous.

Throughout Part 2 of this appendix, we focus on RD because we consider it conceptually easier. But everything we explore relates to mechanistic concerns about how GAMs assign r^2 values. As such, everything presented here also applies to our measure of CSR. All this is necessary is that we replace the x- and y-axis labels with "radius at which agriculture is measured" and "r^2 of pollination by agriculture relationship".

Section S5: Effects of including richness as a covariate when generating response surfaces.
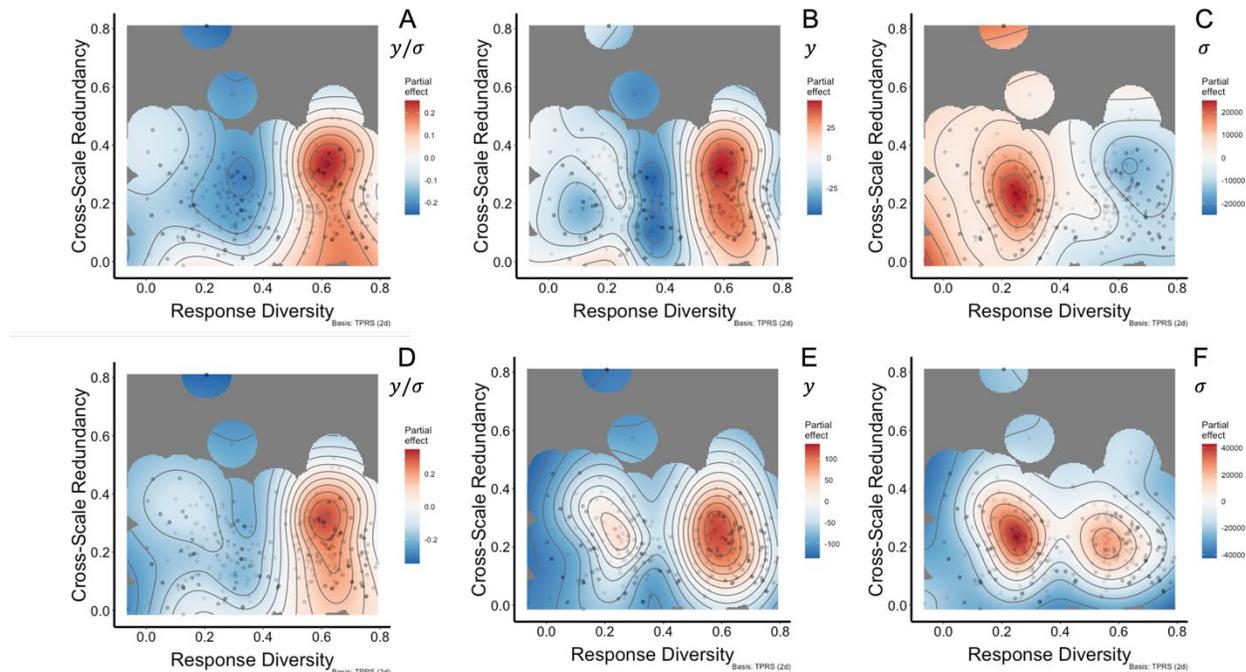

Figure S7. Surface plots showing the interactive effects of cross-scale redundancy (CSR) and response diversity (RD) on stability and its components. Stability is defined as $\bar{y}/\sigma$. Because stability is a ratio, we also examine its numerator and denominator to develop mechanistic insight. Legend colors show the changes relative to the average value across all simulations. Gray areas reflect uncertainty in the response surface, because no simulations had those combinations of CSR and RD. In the top row (A-C), richness is included as a covariate; in the bottom row (D-F) richness is omitted. In both cases, conclusions for stability are similar. RD is mostly positively correlated with stability, CSR is not, and peak stability occurs at above-average levels of CSR and RD. The inclusion of richness does change how CSR and RD affect the components of stability (i.e., $\bar{y}$ and σ), but not in a way that affects our conclusions about how CSR and RD affect stability.
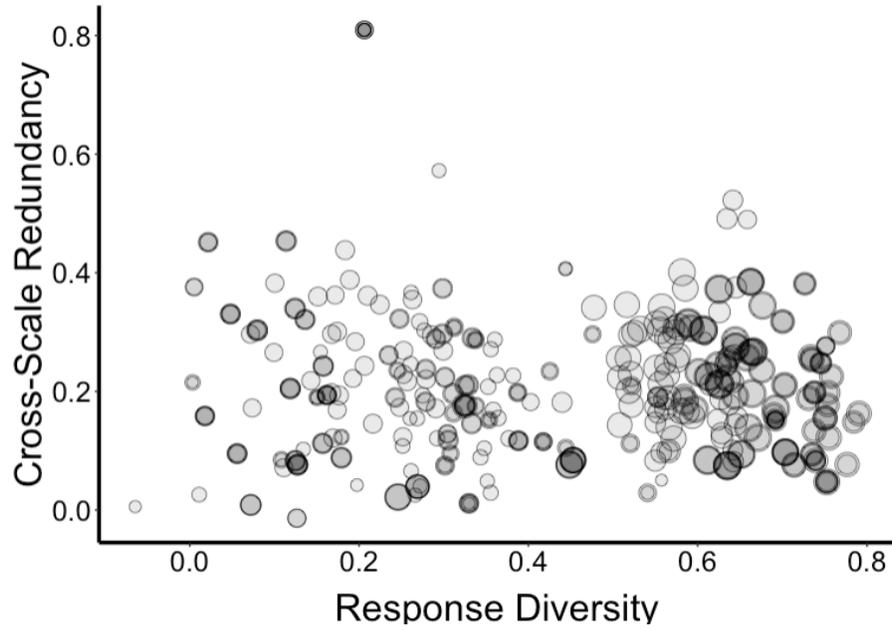
Figure S8. Each point represents one of the 466 simulations. The figure is designed to show two things. First, it shows a raw-data view (in contrast with the GAM-produced surface plot) of how CSR and RD affect stability, as the size of each point is proportional to that simulations' stability. Second, it shows that cross-scale redundancy (CSR) and response diversity (RD) are not strongly correlated in our data, increasing our ability to test for interactive effects of CSR and RD.

Section S6: Effects of removing *A. bradleyi* and *B. griseocollis*.
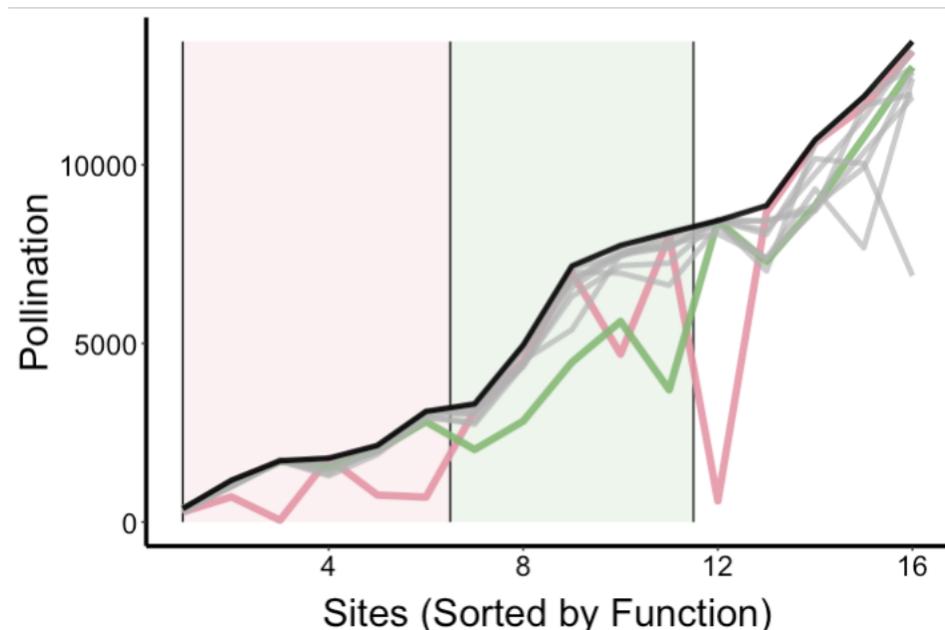

Figure S9. The x-axis is sites (blueberry farms), sorted from lowest to highest function. The black line shows pollination at each site when all bee species are present. The pink and green lines show the effects of removing pollination provided by the functional co-dominants (*Andrena bradleyi* and *Bombus griseocollis*, respectively). Using the same color scheme, the pink and green shaded areas highlight groups of farms where *A. bradleyi* or *B. griseocollis* tend to be the most important pollinator. Each of the gray lines shows the effect of removing pollination provided by one of the other bee species; they are included for comparison to clarify the importance of *A. bradleyi* and *B. griseocollis*.

Section S7: Species identity effects on mean pollination and the variability of pollination.
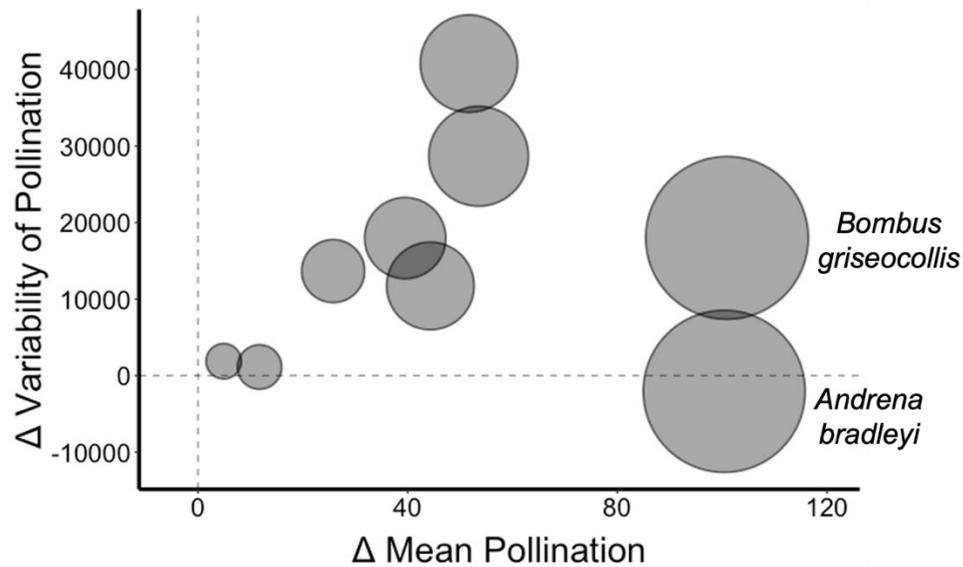

Figure S10. We calculated each species' identity effects (using simulations that included all possible subsets of three to nine species, see Main Text Figure 3A) on the change in mean pollination ($\bar{y}$) and the change in the variability of pollination (σ), both across sites. Note that most species show a large increase in $\Delta\sigma$ as $\Delta\bar{y}$ increases. The functional co-dominants, *Andrena bradleyi* and *Bombus griseocollis*, are exceptions. Both contribute far less to $\Delta\sigma$ than would be expected based on their $\Delta\bar{y}$.