Generalized Single Index Models and Jensen Effects on Reproduction and Survival

Zi Ye, Giles Hooker and Stephen P. Ellner

Abstract

Environmental variability often has substantial impacts on natural populations and communities through its effects on the performance of individuals. Because organisms' responses to environmental conditions are often nonlinear (e.g., decreasing performance on both sides of an optimal temperature), the mean response is often different from the response in the mean environment. Ye, Hooker, and Ellner (2020) proposed testing for the presence of such variance effects on individual or population growth rates by estimating the “Jensen Effect”, the difference in average growth rates under varying versus fixed environments, in functional single index models for environmental effects on growth. In this paper, we extend this analysis to effect of environmental variance on reproduction and survival, which have count and binary outcomes. In the standard generalized linear models used to analyze such data the direction of the Jensen Effect is tacitly assumed a priori by the model’s link function. Here we extend the methods of Ye et al. (2020) using a generalized single index model to test whether this assumed direction is contradicted by the data. We show that our test has reasonable power under mild alternatives, but requires sample sizes that are larger than are often available. We demonstrate our methods on a long-term time series of plant ground cover on the Idaho steppe.

1 Introduction

Effects of environmental variability on different organisms have many important impacts on ecological systems (Vasseur and McCann, 2007). Nonlinear responses to a varying environmental factor can either increase population growth (Drake, 2005; Koons et al., 2009) or decrease it
(Lewontin and Cohen, 1969) depending on the shape of the response to the varying environmental factor, and thus can determine how changes in environmental conditions under climate change will affect species distributions (Vasseur et al., 2014). These differing effects can also play a role in maintaining coexistence among competing species (see for example Hutchinson, 1961; Chesson and Warner, 1981; Ellner, 1987; Chesson, 1994, 2000a,b). Consequently, a key challenge is to understand whether environmental variability is beneficial or detrimental to individual species’ growth rates. Classically, this analysis is obtained by examining the curvature of the reaction norm and deducing its effect by an application of Jensen’s inequality. Specifically, for an environmental variable \( Z \), a convex reaction norm function \( g(Z) \) describing growth in different environmental conditions results in \( Eg(Z) > g(EZ) \) and a positive effect of environmental variability with the reverse holding under concave responses. The difference between \( Eg(Z) \) and \( g(EZ) \) is the Jensen Effect of environmental variability on the response.

In many parametric statistical models the assumed form of \( g(\cdot) \) predetermines its curvature – for example, positive curvature in a logistic regression when the fitted probability is low, negative curvature when the fitted probability is high. Recent interest has therefore focused on nonparametric methods to assess \( g(\cdot) \). Ye and Hooker (2020) considered situations where the environment \( Z \) is a single index representation of climate summaries: \( Z = X\beta \) or as a functional single index \( Z = \int X(t)\beta(t)dt \), where high frequency climate history \( X(t) \) is available. These were then used within a nonparametric single index model (Hardle et al., 1993) \( Y = g(X\beta) + \epsilon \). The curvature of \( g \) determines the sign of the Jensen Effect. Because nonparametric estimates of curvature have high variability, Ye et al. (2020) proposed to instead estimate the effect of Jensen’s inequality directly, the Jensen Effect \( Eg_\lambda(X\beta) - g_\lambda(EX\beta) \). Here \( g_\lambda \) is obtained from a smoothing spline with smoothing parameter \( \lambda \). Ye et al. (2020) proposed to test the sign of the effect by an extension of the SiZer method of Chaudhuri and Marron (1999), representing the estimated effect as a Gaussian process over \( \lambda \) and examining its maximum deviation from zero.

In this paper, we extend the analysis of Ye et al. (2020) to generalized single index models in which a natural link function implies the sign of the Jensen Effect by default. This behavior results from the use of nonlinear link functions to impose natural boundaries in the response: positivity for counts, or constraining probabilities to \([0, 1]\) for binary outcomes. Within this framework, we write
\( EY = h(g(X\beta)) \) where \( h \) is a fixed, user-chosen function. Extending results in Ye et al. (2020) requires accounting for the curvature of \( h \) both within estimates of variability and within the bias of the resulting estimator. In the case of the identity link employed in Ye et al. (2020), a second derivative penalty tends to shrink the Jensen Effect towards zero, making the analysis conservative. In our context, a strong second derivative penalty will make \( \hat{g}_\lambda(\cdot) \) nearly linear yielding a Jensen Effect corresponding to the default behavior of \( h(\cdot) \). We thus cannot readily test whether the Jensen Effect is different from zero, but we can ask whether it has the opposite sign to what would be implied by \( h(\cdot) \), and we develop such one-sided tests below.

Our analysis can also apply to transformations used for statistical validity. Ye and Hooker (2020) examined the year-to-year growth of the sagebrush \emph{Artemesia tripartita} using a functional single index model for the log growth (where growth is the ratio of final and initial plant sizes); the log transformation was introduced as a variance stabilizing transformation. The resulting model was largely concave. Our interest, however, lies in expected growth, rather than expected log growth and we thus examine the Jensen Effect of \( \exp \left( \hat{g}_\lambda(X\hat{\beta}) \right) \). Noting that the exponential is convex, we ask whether the composite model retains the negative Jensen Effect.

While exponential transforms appear naturally in transformed data and Poisson regression models, models for the binary response of survival present further challenges. The commonly-used logistic transform may be convex or concave depending on the predicted probability, making a default sign for the Jensen effect less evident. Instead, we test whether the estimated effect differs from the value obtained by linear logistic regression. Further, our models may include covariates other than those that describe the environment; in our case study example, plant area is a significant determinant of reproduction and survival. In a partially linear model, the inclusion of further covariates \( A \) as additive terms, \( Y = A\gamma + g(Z) + \epsilon \), does not interact alter the Jensen effect; similarly when using an exponential transform, examining \( E \exp(g(Z)) - \exp(g(EZ)) \) provides an estimate of the Jensen effect at each value \( A \) up to a multiplicative scale. In a logistic model, however, the size, and even the sign, of the effect can depend on additional covariates. For such models, we therefore define the Jensen Effect to be the average effect of replacing \( Z \) with its average for each data point; when only environment covariates are present in the model, this definition reduces to the Jensen effect as previously defined.
In section 2 below we first review the procedures developed in Ye et al. (2020), and then develops extensions to the case where we are interested in transformations of the response variable. Section 3 then develops these methods for generalized single index models more generally and applies these to Poisson responses. Section 4 then develops tests specifically for logistic regression. Simulation results and example data analyses accompany each of the sections.

2 The Jensen Effect and Transformed Responses

In this section we review the Jensen Effect as developed in Ye et al. (2020) and extend its application to transformed responses. For example, Ye and Hooker (2020) examined the effect of environmental history (daily temperature and precipitation over the preceding 36 months) on the relative growth rate (RGR) of the sagebrush *Artemesia tripartita*. RGR is the change in plant size relative to its initial size, equivalent to the ratio between final size and initial size, and is a widely-used measure of plant performance (e.g., Grime and Hunt, 1975; Garnier, 1992; Cornelissen et al., 1996; Diaz et al., 2004; Agren et al., 2012). In their analyses Ye and Hooker (2020) used log-transformation as a variance-stabilizing mechanism, fitting a model

$$ \log Y = g(X\beta) + \epsilon. $$

(1)

where $Y$ is the ratio between a plant’s size at two successive annual censuses. However, it is the effect of variable environments on the original scale ($Y$ rather than $\log Y$) which is of interest,

$$ Y = \exp[g(X\beta) + \epsilon], $$

(2)

and which we pursue here.

Below, we first develop estimates of $g$ and $\beta$, corresponding exactly to those employed in Ye et al. (2020), and then extend inference methods to test the sign of $\delta = E \exp(g(X\beta)) - \exp(g(EX\beta))$.

2.1 Model Formulation and Estimation

To ensure the identifiability of the model, we require that $\|\beta\| = 1$ and $\beta_1 > 0$. We will estimate $\beta$ and $g$ through the machinery of smoothing splines. We use a $K_1$-dimensional B-spline basis for
the link function $g$. Then for any $s$ in the domain of $\mathcal{I} = \{X_i\hat{\beta}, i = 1, \ldots, n\}$, the link function $g$ has the form

$$g(s) = \phi(s)^\top d,$$  \hspace{1cm} (3)

where $\phi(\cdot)$ and $d$ are $K_1$-dimensional column vectors, $\phi$ consisting of all basis functions evaluated at $s$. Here we note that the range of $\mathcal{I}$ depend on $\beta$, but that the $\|\beta\| = 1$ constraint allows us to specify that $|s| < \max_{ij} |X_{ij}|$.

The coefficients $\beta$ and $d$ are estimated by minimizing a penalized sum of squares

$$\text{PLS} = \sum_{i=1}^{n} [Y_i^* - g(X_i\beta)]^2 + \lambda \int_{\mathcal{I}} (g''(s))^2 \, ds \hspace{1cm} (4)$$

$$= \sum_{i=1}^{n} [Y_i^* - \phi(X_i\beta)^\top d]^2 + \lambda d^\top \Phi d.$$  

This is a nonlinear optimization problem, which we solved numerically using the BFGS algorithm as implemented in the R function optim. We implement the norm constraint on $\beta$ by standardizing within the function evaluation. We initialize the optimization at $g(s) = s$ with the linear model for $\beta$ estimated by unpenalized least squares. We ran optim twice, initializing the second run from the result of the first run to ensure convergence; see Ye et al. (2020) for further details. The resulting estimates are $\hat{\beta}_\lambda$ and $\hat{g}_\lambda(s) = \phi(s)^\top \hat{d}_\lambda$.

### 2.2 Jensen Effect and its Variance

To investigate the Jensen Effect, we need to estimate $\delta$, and conduct inference about its sign. We define $\delta$ in the exponential model as

$$\delta \doteq \frac{1}{n} \sum_{i=1}^{n} \exp[g(X_i\beta)] - \exp[g(\bar{X}\beta)],$$  \hspace{1cm} (5)

where $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$.

Inference is based on using our estimates from above in eqn. (5), for a series of different $\lambda$ values. Define a $(n + 1)$-dimensional column vector $a \doteq (\frac{1}{n}, \ldots, \frac{1}{n}, -1)^\top$ and the augmented set of evaluation points $e \doteq (E_1, \ldots, E_n, \bar{E})^\top$ with corresponding evaluation matrix $\Phi^+$, where $E_i = X_i\hat{\beta}$.
(at each environment value) and \( \bar{E} = \frac{1}{n} \sum_{i=1}^{n} X_i \hat{\beta} \) (averaged across all environment values), we express
\[
\hat{\delta}_\lambda = a^\top \exp \left( \Phi_\lambda^+ \hat{d}_\lambda \right).
\]
We use these \( \hat{\delta} \) values to perform inference on \( \delta \), as follows:

1. We develop a covariance function estimate \( \text{cov}(\hat{\delta}_{\lambda_1}, \hat{\delta}_{\lambda_2}) \) for \( \hat{\delta} \) thought of as a function of \( \lambda \).

2. We use the covariance function to evaluate critical values for \( \min_\lambda \hat{\delta}_\lambda / \sqrt{\text{var}(\hat{\delta}_\lambda)} \) which constitutes our test statistic.

Specifically, for any \( \lambda_1 \) and \( \lambda_2 \), we use a Delta method approximation \( \text{Casella and Berger (2002)} \):
\[
\Sigma_\delta (\lambda_1, \lambda_2) \doteq \text{cov} \left( \hat{\delta}_{\lambda_1}, \hat{\delta}_{\lambda_2} \right)
\]
\[
= \begin{bmatrix}
a \cdot \exp \left( \Phi_{\lambda_1}^+ \hat{d}_{\lambda_1} \right)
\end{bmatrix}^\top \Phi_{\lambda_1}^+ \text{cov} \left( \hat{d}_{\lambda_1}, \hat{d}_{\lambda_2} \right) \begin{bmatrix}
\Phi_{\lambda_2}^+ \hat{d}_{\lambda_2}
\end{bmatrix}^\top \begin{bmatrix}
a \cdot \exp \left( \Phi_{\lambda_2}^+ \hat{d}_{\lambda_2} \right)
\end{bmatrix},
\]
where \( \text{cov} \left( \hat{d}_{\lambda_1}, \hat{d}_{\lambda_2} \right) \) is calculated in Ye et al. (2020) and we review this here. Recall (4), the penalized least squares criterion to be minimized, and define the matrix of evaluations of the link function bases and derivatives \( \Phi_{ij}^{(k)} = \phi_j^{(k)} \left( X_i \hat{\beta} \right) \) and covariates matrix \( \mathbf{X} \). We derive the gradient of (4) as
\[
\begin{bmatrix}
\nabla d \\
\nabla \beta
\end{bmatrix} = \begin{bmatrix}
\Phi^\top (Y^* - \Phi^\top d) + \lambda \mathbf{P}_g d \\
\mathbf{X}^\top \text{diag} \left( \Phi^{(1)} \right) (Y^* - \Phi^\top d)
\end{bmatrix} = \begin{bmatrix}
\mathbf{Z}_g \\
\mathbf{Z}_\beta
\end{bmatrix} (Y^* - \Phi^\top d) + \begin{bmatrix}
\lambda \mathbf{P}_g d \\
\mathbf{0}_p
\end{bmatrix},
\]
and expected Hessian as
\[
\mathbb{H} \doteq \begin{bmatrix}
\mathbf{Z}_g^\top \mathbf{Z}_g + \lambda \mathbf{P}_g \\
\mathbf{Z}_\beta^\top \mathbf{Z}_g & \mathbf{Z}_\beta^\top \mathbf{Z}_\beta
\end{bmatrix}.
\]
We can now obtain the sandwich covariance estimate
\[
\text{cov} \left( \begin{bmatrix}
d \\
\beta
\end{bmatrix} \right) = \hat{\sigma}^2 \mathbb{H}^{-1} \begin{bmatrix}
\mathbf{Z}_g \\
\mathbf{Z}_\beta
\end{bmatrix} \begin{bmatrix}
\mathbf{Z}_g \\
\mathbf{Z}_\beta
\end{bmatrix}^\top \mathbb{H}^{-1}.
\]
in which the variance of the random error \( \sigma^2 \) is estimated using the model selected by GCV and
\[
\hat{\sigma}^2 = \frac{1}{df_{\text{res}}} \sum_{i=1}^{n} \left( Y_i^* - \hat{g}_{\lambda_{\text{gcv}}} \left( X_i \hat{\beta} \right) \right)^2,
\]
\[
\text{(10)}
\]
where the residual degrees of freedom is defined as

$$\text{df}_{\text{res}} = n - 2\text{tr}(S) + \text{tr}(SS^T) - p,$$

(11)
as in [Ruppert et al. (2003)] with $S = S(\lambda)$ being the approximate smoother matrix defined by the model.

This derivation follows exactly the calculations in [Ye et al. (2020)] with only the addition of a Delta method approximation to account for the exponential transform in (6). In practice, we have found that including the single index parameters $\nabla\beta$ within the covariance calculation has negligible effect on the covariance for $\hat{\delta}_\lambda$ and we have dropped these terms from our calculations in Section 3 in order to ease the notational burden.

We can now form the normalized $\hat{\delta}_\lambda$ process, $t_\lambda = \hat{\delta}_\lambda / \sqrt{\text{var}(\hat{\delta}_\lambda)}$. This process has covariance approximated by

$$\text{cov}(t_{\lambda_1}, t_{\lambda_2}) = \frac{\text{cov}(\hat{\delta}_{\lambda_1}, \hat{\delta}_{\lambda_2})}{\sqrt{\text{var}(\hat{\delta}_{\lambda_1})} \sqrt{\text{var}(\hat{\delta}_{\lambda_2})}},$$

(12)

These combine to form the variance-covariance matrix for $t_\lambda$ over a grid of $\lambda$ values:

$$\Sigma_t = [\text{diag}(\Sigma_\delta)]^{-\frac{1}{2}} \Sigma_\delta [\text{diag}(\Sigma_\delta)]^{-\frac{1}{2}}.$$

(13)

Even when it is true that $\delta = 0$, we will get $Et_\lambda > 0$ for large values of $\lambda$, because the penalty forces $\hat{g}_\lambda(\cdot)$ to be close to linear, and the convexity of the exponential transform produces a positive Jensen Effect. We thus consider a one-sided alternative, and test for $\delta < 0$. To carry this out, we use $\min_\lambda t_\lambda$ as our test statistic. In practice, the minimum is evaluated at a grid of $m \approx 20$ values of $\lambda$, and we obtain critical values of the minimum by simulating from the multivariate normal distribution $t_{\text{null}} \sim N(0_m, \Sigma_t)$ and taking the $\alpha$ quantile of $\min t_{\text{null}}$.

### 2.3 Simulation Results

We evaluated the hypothesis test introduced in the previous section in a series of simulation studies. Testing for a nonzero Jensen Effect requires estimating the quantity $\delta$ with a range of smoothing parameters $\lambda$. We expect that our estimate $\delta_\lambda$ will exhibit high variance at small $\lambda$ because little smoothing is done on the link function estimate, and large positive bias at high $\lambda$ when the estimated
link function is strongly shrunk towards linear, but that if the underlying relationship is generally concave, we will be able to detect a negative Jensen Effect.

Our simulation study starts with \( p = 5 \) covariates generated uniformly on \([0, 0.5]\). Since we will eventually need to take a log-transformation, these covariates must be restricted to the positive real line. The true coefficients are \( \beta = \frac{1}{\sqrt{p}} 1_p \) so that \( \|\beta\| = 1 \).

We considered 4 different link functions, (1) \( g^*(s) = \exp (s) \), (2) \( g^*(s) = \sqrt{s} \), (3) \( g^*(s) = \sin (s) \), (4) \( g^*(s) = s \), where the domain is \( s \in [0, 0.5] \). For the estimate of the link function \( g \) in the single index model we used a 25-dimensional quintic B-spline basis. For each link function, we simulated 1000 data sets of size 1000, for error standard deviation 0.01. We obtained critical values for our test by simulating 5000 samples from the multivariate Gaussian null distribution. The method’s performance at such a low level of noise can be viewed as a “sanity check” on our method – if it failed in this setting, we need to look for a different approach. Fortunately, it did not.

Figure 1 presents a sample of \( \delta_\lambda \) and \( t_\lambda \) functions versus \( \log(\lambda) \) for link function \( g^*(s) = \sqrt{s} \); plots for the other link functions are in Appendix A. The rejection rates for the four link functions were 0%, 100%, 98.3%, 0.1% respectively. Note that the 0% rejection rate for the first case is expected as \( \delta > 0 \) in this case, while we tested for the opposite sign. In the case of concave composite link functions, we correctly reject with high power, because the noise level in this case was very low. In the final case with a linear link function, the result is conservative relative to the nominal 5% rejection rate when the null hypothesis holds.

### 2.3.1 Power Analysis

To investigate the statistical power of our hypothesis test, we considered a series of values of \( \sigma \), the standard deviation of the random errors, along with four sample sizes. We considered \( \sigma \) values of \{0.03, 0.05, 0.07, 0.09, 0.1\}, and sample sizes of 100, 200, 500 and 1000. For each combination we generated 400 data sets from which to calculate the rejection frequency.

Figure 2 presents the rejection rates plotted against \( \sigma \) value for different sample sizes. We observe a sharp decrease as \( \sigma \) increases unless the sample size is large. As we expected, decreasing noise and increasing sample size increase the rejection rate.
Figure 1: Illustration of the Jensen Effect for a concave function in the exponential single index model, where the link function is $g^*(s) = \sqrt{s}$. Left: a sample of the Jensen Effect values $\delta_\lambda$ as a function of smoothing parameters $\lambda$. Right: the corresponding t-statistics $t_\lambda$ functions.

Figure 2: Statistical power of the hypothesis test for the Jensen Effect being negative in the exponential single index model. The rejection rates of the Jensen Effect test are plotted against $\sigma$ values as the standard deviation of the random error. Left: concave function $g^*(s) = \sqrt{s}$. Right: concave function $g^*(s) = \sin(s)$. 
2.3.2 Empirical Application

Returning to our motivating ecological study, we again analyzed the growth of the sagebrush *Artemisia tripartita* (ARTR) observed at the U.S. Sheep Experiment Station (USSES) between 1926 and 1957. USSES is located 9.6 km north of Dubois, Idaho (44.2°N, 112.1°W), 1500m above sea level. Ecologists at the USSES established 26 1 m² quadrats between 1926 and 1932. Among them, 18 quadrats were distributed among 4 ungrazed exclosures, and 8 were distributed in 2 pastures grazed at medium intensity spring through fall. All quadrats were located on similar topography and soils.

During the period of data collection, mean annual precipitation was 270 mm and mean temperatures ranged from −8°C (January) to 21°C (July). The vegetation is dominated by the shrub, *Artemisia tripartita*, and the C3 perennial bunchgrasses *Pseudoroegneria spicata*, *Hesperostipa comata*, and *Poa secunda*. These 4 species, the focus of our analyses here, comprised over 70% of basal cover (grasses) and 60% of canopy cover (shrubs and forbs).

Each individual *A. tripartita* plant was mapped in each of 26 1m² ungrazed quadrats. Mapping was not done every year, however plant growth over one year (the increase in ground area covered by the plant canopy) could measured for 22 year-to-year transitions. The response of interest is the relative change in area between two successive annual censuses. Figure 3 plots residuals versus fitted values for a principal components regression using either the original measurements or logged values where we see that there is some heteroscedasticity in the response variable that is corrected after taking the log.

We have a total of 1003 year-to-year transitions across all individual plants (genets), quadrats, and years. Because not all plants survive each year, and new plants establish naturally each year, not all plants are observed for all 22 years. For simplicity, the model we use here does not include individual plant effects, which tend to be small relative to year-to-year variability. For each observed annual transition, we record the following variables:

1. $\Delta \logarea$ — the change in the logarithm of total area from one year to the next.

2. $W$ — a measure of competition given by a weighted sum of the area occupied by conspecific individuals in a series of concentric annuli extending from the genet’s center.
Figure 3: Fitted values versus residuals for a principal components regression for the shrub *Artemisia tripartita*. Left: untransformed size data. Right: log-transformed size data. We can see that log-transformation greatly reduces the variance heteroscedasticity. Note that clumping in the fitted values is due to multiple observations from the same year having the same climate history.

3. \( t(s) \) – daily temperature history (average of daily minimum and maximum temperatures) over the 36 months prior to the second census (\( s \) giving historical time). A small number of missing values in the temperature record were filled in by spline interpolation.

4. \( p(s) \) – daily precipitation over the previous 36 months.

Note that all records from a given year have exactly the same temperature and precipitation history; see Teller et al. (2016) for further details. We also include the competition variable \( W \) as a further scalar covariate with its own coefficient. Note that no intercept is included as this is confounded with the estimate of \( g \). The resulting model is

\[
\log \left( \frac{A_{t_1}}{A_{t_0}} \right) = g \left( \alpha W + \int p \beta_1 + \int t \beta_2 \right),
\]

(14)

\[
\frac{A_{t_1}}{A_{t_0}} = \exp \left[ g \left( \alpha W + \int p \beta_1 + \int t \beta_2 \right) \right],
\]

(15)

where \( \log \left( \frac{A_{t_1}}{A_{t_0}} \right) \) is the difference of plant’s logarithm of area between time \( t_0 \) and \( t_1 \). (15) is an exponential functional single index model, and the intercept \( \alpha \), the functions \( g, \beta_1 \) and
\( \beta_2 \) need to be estimated. Note that the left-hand side of (15) differs by 1 from the plant's relative growth rate \((A_{t_1} - A_{t_0}) / A_{t_0}\). So to investigate the Jensen Effect on relative growth rate, we need to get an estimate of the composite link function \( g^* = \exp(g) \).

We use a 15-dimensional Fourier basis for the two coefficient functions \( \beta_1 \) and \( \beta_2 \). We generated initial values to solve the nonlinear optimization problem by assuming that the composite link function is linear, such that \( g^*(s) = s \). Our analysis rejects the null hypothesis \( \delta > 0 \), thus concluding that environmental variability reduces relative growth rate in Artemesia. This confirms the findings in Ye and Hooker (2020), even after accounting for the exponential transform.

We observe that most of the Jensen Effect \( \delta \) values are negative in Figure 4, when they are estimated by different \( \lambda \) values. This lends credence to the conclusion that \( g^* \) is concave in its effects over the range of observed environments, and environmental variability is harmful for growth. The remaining plots in Figure 4 are estimates of functions \( g^*, \beta_1 \) and \( \beta_2 \), where the smoothing parameter \( \lambda \) is selected by generalized cross-validation.

### 3 Generalized Single Index Model

We now extend the methods of the previous section to generalized linear single index models, to allow for Poisson distributed responses (such as annual number of offspring), or binary outcomes such as survival.

We define a Generalized Single Index model as

\[
\begin{align*}
\eta &= g(X\beta), \\
\mu &= h(\eta) = h[g(X\beta)], \\
Y &\sim \text{Distribution} \left( \mu \right).
\end{align*}
\]

Here, \( \eta \) is the response from a single index model, \( h \) is the GLM link function, and we are interested in a possible Jensen Effect on the composite link function defined as \( h \circ g \).

In particular, we examine the two most common generalized linear model applications. In Section 3.1, we examine a Poisson regression framework to model the number of offspring produced by Artemesia in a given year. Since the Poisson model shares the exponential link function \( h \) with
Figure 4: Jensen Effect for relative growth rate of the shrub *Artemisia tripartita*, using the exponential single index model formulation. Top-left: plot of the Jensen Effect estimates $\delta$ against the logarithm of smoothing parameter, $\log(\lambda)$. Top-right: plot of the estimate of the link function $g^*$ corresponding to the $\lambda$ selected by GCV. Bottom: plots of estimates of the coefficient functions $\beta_1$ and $\beta_2$ selected by GCV.
the exponential model, we can borrow some ideas from the exponential model to assess the Jensen Effect: whether environmental variability will increase or decrease the number of offspring of a species.

3.1 Poisson Model

In a Poisson single index model, we observe a count response $Y$ drawn from a Poisson($\mu$) distribution with $\mu = \exp(\eta)$ where $\eta$ is generated by a single index model:

$$\eta := g(X\beta), \quad \mu = \exp(\eta) = \exp\left[g(X\beta)\right].$$  \hfill (19)

Recall that $\mu$ is the mean and variance of a Poisson distribution and the canonical GLM link function $h$ is exponential.

We assume $n$ observations $(X_i, Y_i)$ with $Y_i$ generated from the Poisson distribution above. The Jensen Effect $\delta$ is defined as

$$\delta = \frac{1}{n} \sum_{i=1}^{n} \exp\left[g(X_i\beta)\right] - \exp\left[g(\bar{X}\beta)\right],$$  \hfill (21)

where $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$.

In a generalized linear model, we replace squared error by the log-likelihood, given in the Poisson model by

$$\ell = \sum_{i=1}^{n} \left[ Y_i \eta_i - \exp(\eta_i) \right] = \sum_{i=1}^{n} \left[ Y_i g(X_i\beta) - \exp\left[g(X_i\beta)\right]\right].$$  \hfill (22)

As above, we use B-spline basis to represent the link function $g$, such that $g(s) = \phi^T(s) d$ and $P_g$ as the second derivative penalty matrix for $g$. The penalized log-likelihood is

$$PLL = -\ell + \lambda d^T P_g d$$  \hfill (23)

$$= -\sum_{i=1}^{n} \left\{ Y_i \phi^T(X_i\beta) d - \exp\left[\phi^T(X_i\beta) d\right]\right\} + \lambda d^T P_g d.$$

For a fixed smoothing parameter $\lambda$, we estimate the coefficients $\beta$ and $d$ by minimizing the penalized log-likelihood \hfill (23). Denote the resulting estimates by $\hat{\beta}_\lambda$ and $\hat{d}_\lambda$, we have $\hat{g}_\lambda = \phi^T \left(X\hat{\beta}_\lambda\right) \hat{d}_\lambda = \ldots$
**Φ**_λ_**d**_λ and **µ**_λ = exp (**g**_λ), where **g**_λ and **µ**_λ are _n_-dimensional column vectors evaluated at each observation using the estimated _β_.

Using the notation from Section 2.2, we can write

\[
\hat{\delta}_λ = \frac{1}{n} \sum_{i=1}^{n} \exp [\hat{g}_λ (E_i)] - \exp [\hat{g}_λ (\bar{E})] = a^\top \exp \left( \Phi_λ^+ \hat{d}_λ \right).
\] (24)

By the Delta method (Casella and Berger, 2002), for smoothing parameters λ₁ and λ₂, the covariance between δ_λ₁ and δ_λ₂ is

\[
\Sigma_δ (λ₁, λ₂) = \text{cov} (\hat{\delta}_λ) = \begin{bmatrix} a \cdot \exp (\Phi_{λ₁}^+ \hat{d}_λ_1) & a \cdot \exp (\Phi_{λ₂}^+ \hat{d}_λ_2) \end{bmatrix}^\top \Phi_{λ₁}^+ \text{cov} \left( \hat{d}_λ_1, \hat{d}_λ_2 \right) \Phi_{λ₂}^+ a^\top \exp \left( \Phi_{λ₂}^+ \hat{d}_λ_2 \right).
\] (25)

Applying standard arguments, the covariance between **d**_λ₁ and **d**_λ₂ is

\[
\text{cov} (\hat{d}_λ_1, \hat{d}_λ_2) \approx \left( \frac{\partial^2 (PLL)}{\partial d_{λ₁}^2} \right)^{-1} \text{cov} \left( \frac{\partial (PLL)}{\partial d_{λ₁}}, \frac{\partial (PLL)}{\partial d_{λ₂}} \right) \left( \frac{\partial^2 (PLL)}{\partial d_{λ₂}^2} \right)^{-1}
\] (26)

\[
\approx \left( \Phi_{λ₁}^T \mathbb{W}_λ \Phi_{λ₁} + λ₁ \mathbb{P}_g \right)^{-1} \Phi_{λ₁}^T \text{cov} (Y) \Phi_{λ₂} \left( \Phi_{λ₂}^T \mathbb{W}_λ \Phi_{λ₂} + λ₂ \mathbb{P}_g \right)^{-1}.
\] (27)

where \( \mathbb{W}_λ = \text{diag} (\hat{µ}_λ) \) is an adjusted weight matrix. Finally, \( \text{cov} (Y) = µ \) and this does not change with λ.

To estimate 26 we use \( \mathbb{W}_λ \) with λ selected by the GCV criterion adapted to a penalized generalized linear model as in Wood (2006):

\[
\text{GCV} (λ) = \frac{n \| \sqrt{\mathbb{W}_λ} (z_λ - \Phi_{λ}^\top \hat{d}_λ) \|^2}{[n - \text{tr} (S_λ)]^2},
\] (28)

where the smoother matrix is

\[
S_λ = \Phi_{λ}^T \left( \Phi_{λ}^T \mathbb{W}_λ \Phi_{λ} + λ \mathbb{P}_g \right)^{-1} \Phi_{λ}^T \mathbb{W}_λ.
\] (29)
Putting (25) and (29) together, we have an estimate of the covariance matrix of $\hat{\delta}$. The covariance estimate for the t-statistic vector $t = \text{diag}\Sigma^{-\frac{1}{2}}\hat{\delta}$ is

$$\text{cov} (t_{\lambda_1}, t_{\lambda_2}) = \frac{\text{cov} \left( \hat{\delta}_{\lambda_1}, \hat{\delta}_{\lambda_2} \right)}{\sqrt{\text{var} \left( \hat{\delta}_{\lambda_1} \right)} \sqrt{\text{var} \left( \hat{\delta}_{\lambda_2} \right)}},$$

(30)

$$\Sigma_t = [\text{diag} (\Sigma_{\delta})]^{-\frac{1}{2}} \Sigma_{\delta} [\text{diag} (\Sigma_{\delta})]^{-\frac{1}{2}}.$$

The corresponding null distribution is simulated as a multivariate normal distribution $t_{\text{null}} \sim \mathcal{N} (0, \Sigma_t)$, with critical value obtained from the quantiles of $\min_{\lambda_t} t$.

### 3.1.1 Simulation Results

As with the exponential model, we test for the Jensen Effect by calculating the quantity $\delta$ over a range of smoothing parameters $\lambda$. In the Poisson single index model, if the composite link function $g^* = \exp (g)$ is convex, the $\delta$ estimate will be positive for most $\lambda$ values; in other cases we expect a positive bias as $\lambda$ increases. For each simulation, we simulate the null distribution from a Gaussian process, and conduct a one-sided hypothesis test to see if $\delta < 0$.

Our simulation study starts with $p = 5$ covariates generated uniformly on $[0, 20]$. The coefficients $\beta = \frac{1}{\sqrt{p}} \mathbf{1}_p$ so that $\|\beta\| = 1$. To illustrate the Jensen Effect, we choose 3 different link functions,

1. $g^* (s) = \exp \left( \frac{s}{5} \right)$,
2. $g^* (s) = \frac{30}{1 + \exp (\frac{-s}{5}) - 15}$,
3. $g^* (s) = s$. In each case we represent $g$ by a 25-dimensional quintic B-spline basis. For each link function, we simulated 200 data sets of size 1000. We obtained critical values for our test by simulating 5000 normal samples from the null distribution. Figure 5 presents a sample of $\hat{\delta}_{\lambda}$ and $t_{\lambda}$ functions versus $\log (\lambda)$ for $g^* (s) = \frac{30}{1 + \exp (\frac{-s}{5}) - 15}$; plots for the other link functions are in Appendix B. The rejection rates for these functions are: 0%, 82.5%, 1.5% respectively. The convex function has positive Jensen Effect value $\delta$ so that we observe a zero rejection rate as expected.
Figure 5: Illustration of the Jensen Effect for a concave composite link function in the Poisson single index model, where the link function is $g^*(s) = \frac{30}{1 + \exp(-\frac{s}{a})}$. Left: a sample of the Jensen Effect values $\delta_\lambda$ as a function of smoothing parameters $\lambda$. Right: the corresponding t-statistics $t_\lambda$ functions.

3.1.2 Power Analysis

To investigate the power of the hypothesis test in a Poisson single index model, we use the concave link function

$$(h \circ g)(s) = \exp(g)(s) = \frac{30}{1 + \exp(-\frac{s}{a})} - 15. \quad (31)$$

and employed values of $a$ in $\{2, 4, 6, 8, 10, 12, 14, 16\}$ along with data sets of size 100, 200, 500 and 1000. From the plots of $g^*$ in Figure 6, we can observe that $g^*$ becomes more concave for the first half of the $a$ values, and becomes less concave for the second half. The rejection rates and the true Jensen Effect $\delta$ values are in Figure 7. We see that as the link function becomes more concave, the rejection rate increases to 1, and when it becomes less concave, the rejection rate decreases.

3.1.3 Empirical application

We return to the USSES data to examine different endpoints. The vegetation in the exclosures is dominated by the shrub, *Artemisia tripartita* (ARTR), and the C3 perennial bunchgrasses *Pseudoroegneria spicata*, *Hesperostipa comata*, and *Poa secunda* (PSSP, HECO, and POSE, respec-
Figure 6: Concave link functions $g^*$ in a Poisson single index model, with $(h \circ g)(s) = \exp(g)(s) = \frac{30}{1 + \exp(-\frac{s}{a})} - 15$. Left: the values of $a$ are $\{2, 4, 6, 8\}$. The link function becomes more concave as the value of $a$ increases. Right: the values of $a$ are $\{10, 12, 14, 16\}$. The link function becomes less concave as the value of $a$ increases.

Figure 7: Power analysis based on 400 simulations of a test for concavity in a Poisson single index model with link function $\exp(g)(s) = \frac{30}{1 + \exp(-\frac{s}{a})} - 15$, where the values of $a$ determine concavity. Left: the rejection rates plotted against the values of $a$. Right: the true Jensen Effect $\delta$ values plotted against the values of $a$. 
tively). These four species, the focus of our analyses here, comprised over 70% of basal cover (grasses) and 60% of canopy cover (shrubs and forbs).

For each species, we use the number of seedlings in each quadrat as the response variable, and employ five covariates: \( \text{totParea} \) (total area currently covered by the species), \( \text{ppt1, ppt2} \) (total precipitation in each of the past two years), \( \text{TmeanSpr1, TmeanSpr2} \) (mean Spring temperature in each of the past two years). Unlike our growth analysis, we have restricted to a relatively simple model without functional effects. This is because recruitment data are only available at the quadrat level (i.e., it is impossible to know which parent plant produced a particular new recruit), which significantly reduces the number of data points for each year in the study.

Because the response variable is the count of new seedlings, we use a Poisson single index model for the data. Define \( \eta \) as

\[
\eta = \gamma A + g(X\beta),
\]

where \( A \) denotes the total area, and \( X \) is the covariate vector for temperature and precipitation. Here we continue to measure

\[
\delta_\lambda = \frac{1}{n} \sum \left( \exp{(g_\lambda X\beta)} - \exp{(g_\lambda \bar{X}\beta)} \right)
\]

noting that the prediction model is \( \exp(\gamma A) \exp(g(X\beta)) \) and hence \( \delta_\lambda \) provides a scaled measure of the Jensen effect at each \( A \) and its sign is unaffected.

Our hypothesis test rejects the null hypothesis for \( \text{Hesperostipa comata (HECO)} \) and \( \text{Pseudoroegneria spicata (PSSP)} \): environmental variability has a detrimental effect on the average number of HECO and PSSP seedlings. In Figures 8 and 17, we can observe that nearly all \( \delta_\lambda \) values are negative, confirming that a less variable environment would benefit their reproduction. The right-hand panels in these figures illustrate an example of \( \exp(g) \) estimate, corresponding to the maximum \( \lambda \) for which \( \hat{\delta}_\lambda \) was significantly different from zero. For PSSP the estimated response \( \exp(\hat{g}) \) says that seedling production is maximized at intermediate values of the single index – a “Goldilocks” zone of intermediate conditions favorable for reproduction – and environmental variability results in many years being outside that zone. For HECO there are several peaks in the estimated response, which are hard to explain biologically, but they may be a result of the
GCV-selected estimate being somewhat under-smoothed. Nonetheless, the overall pattern is for reproduction to be highest at intermediate values of the environment index.

Figure 8: Jensen Effect for reproduction by the grass *Pseudoroegneria spicata* (PSSP) in USSES data set, where we model the seedling count data using Poisson single index model. Left: the Jensen Effect values $\delta$ plotted against the values of $\lambda$. Right: the curve of the exponential link function $\exp(\hat{g})$, where $\lambda$ is corresponding to the minimum of $\delta$. 
Figure 9: Jensen Effect for reproduction by the shrub *Artemisia tripartita* (ARTR) in USSES data set, where we model the data using a Poisson single index model. Left: the Jensen Effect values $\delta$ plotted against the values of $\lambda$. Right: the curve of the exponential link function $\exp(\hat{g})$, where $\lambda$ is selected by GCV.

For the two other species, the null hypothesis is not rejected, and most of the $\delta_\lambda$ estimates appear to be positive: environmental variability will increase the average number of seedlings. Figures 9 and 16 in Appendix C show the estimated response functions $\exp(\hat{g})$ with $\lambda$ selected by generalized cross-validation. In both cases, the estimated response functions say that reproduction is highest at one extreme end of the range of the environment index $X\beta$, and much lower under any other conditions. Any substantial reduction in environmental variability would eliminate the rare years that are favorable for reproduction. Thus, temporal niche partitioning between HECO and PSSP (reproducing best under intermediate, near-average conditions) versus ARTR and POSE (reproducing best under extreme, atypical conditions) may be a factor contribute to coexistence of these species in the community.

4 Logistic Model

A final common endpoint in demographic studies is survival: did this individual survive from one census to the next? This is usually modeled by a logistic regression in which the curvature of the
response is positive when $P(Y = 1)$ lies below 0.5 and negative above it. The default sign of the Jensen Effect, that we wish to challenge through our hypothesis test, therefore depends on the distribution of the environmental index. This makes it difficult to obtain large power in a test against the default: if the environmental index varies enough for accurate estimation of $g$, it is likely to include regions with different curvature in the logistic. We therefore develop a test that the Jensen Effect is different from the expected value assuming a linear-logistic response, as well as a test of difference in sign when that can be determined.

Unlike the case of exponential models, in this model we cannot separate the effect of variability in environmental covariates from variability in other covariates. However, as noted above, for this case we define the Jensen effect as the average effect of replacing the environmental covariates with their average value, while keeping the remaining features at their observed values. That is, given observed values $(A_1, X_1), \ldots, (A_n, X_n)$, with $A_i$ being covariates that on included within environmental variation, we define the Jensen Effect to be

$$
\delta = \frac{1}{n} \sum_{i=1}^{n} h\left(g(A_i \gamma + X_i \beta)\right) - h\left(g(A_i \gamma + \bar{X} \beta)\right)
$$

(33)

where $s(\cdot)$ is the logistic function. This change is straightforward to implement within our framework and we note that when there are no $A_i$ present, this more general definition reduces to the definition of the Jensen Effect in the models above.

### 4.1 Model Formulation and Estimation

The survival status (binary response) of a species, $Y \in \{1, 0\}$, is generated by a Bernoulli($\pi$) distribution, where $\pi$ is the probability of survival, given by

$$
\eta = g(A \gamma + X \beta),
$$

(34)

$$
\pi = h(\eta) = \frac{1}{1 + \exp\left[-g(X \beta)\right]}.
$$

(35)

with Bernoulli responses $Y_1, \ldots, Y_n$ corresponding to the covariates above.

As above, we represent $g$ by a basis expansion, $g(s) = \phi^\top (s) d$ and estimate parameters $(\gamma, \beta)$
and \( d \) by maximizing the penalized logistic log-likelihood

\[
PLL = -\ell + \lambda d^T P_g d = -\sum_{i=1}^{n} \left\{ Y_i \phi^T (A_i \gamma + X_i \beta) d - \log \left[ 1 + \exp \left( \phi^T (A_i \gamma + X_i \beta) d \right) \right] \right\} + \lambda d^T P_g d.
\]

At fixed \( \lambda \), our estimate of the Jensen Effect is

\[
\hat{\delta}_\lambda = \frac{1}{n} \sum_{i=1}^{n} \left[ \hat{g}_\lambda (A_i \gamma + E_i) - \hat{g}_\lambda (A_i \gamma + \bar{E}) \right] = a^T h \left( \Phi_\lambda^+ \hat{d}_\lambda \right).
\]

where we redefine \( \Phi_\lambda^+ \) to be the evaluation matrix for \( A_1 \gamma + X_1 \beta, A_2 \gamma + \bar{X} \beta, \ldots, A_n \gamma + X_n \beta, A_n \gamma + \bar{X} \beta \) and \( a = (1, -1, \ldots, 1, -1) \).

We can now obtain an estimate of \( \text{cov} \left( \hat{\delta}_1, \hat{\delta}_2 \right) \) as in (25), (26) with the substitute that for \( \hat{\pi}_\lambda = \text{logistic}(\hat{g}_\lambda) \) we define the weight matrix \( W_\lambda = \text{diag} \left[ \hat{\pi}_\lambda \left( 1 - \hat{\pi}_\lambda \right) \right] \) which also gives the variance of \( Y \).

In the logistic model, we have found that \( \Sigma_\delta \) can have non-positive eigenvalues, likely associated with numerical error. When this is the case, we form the SVD

\[
\Sigma_\delta = U_\delta V_\delta U_\delta^T.
\]

and truncate negative entries of \( V_\delta \) at 0, yielding \( V_\delta^+ \) and reformulate the covariance of \( \delta \) to be

\[
\Sigma_\delta = U_\delta V_\delta^+ U_\delta^T.
\]

The t-statistics are again obtained from \( t = \text{diag} \Sigma_\delta^{-\frac{1}{2}} \hat{\delta} \). The null distribution is simulated as a multivariate normal distribution \( t_{null} \sim N(0, \Sigma_t) \) with a critical value obtained from the quantiles of \( \max_\lambda t \).

### 4.2 Testing an Alternative Null

Because the curvature of the logistic model changes sign, there will not necessarily be a natural sign to test against. Here we develop a test for the quantitative size of the Jensen effect being different from that obtained from standard logistic regression on the environment index (i.e., different from the value that would obtain if \( g \) is linear). We can frame this as a test of the null hypothesis that

\[
\delta_\lambda^+ = \delta_\lambda - \delta_\infty = 0
\]
because, for the second derivative smoothing penalty that we use, \( \delta_\infty \) corresponds to the Jensen effect when \( g \) is linear. In fact, we implement the calculation of \( \delta_\infty \) by carrying out standard logistic regression with linear \( g \), and calculating

\[
\hat{\delta}^+_{\lambda} = a^T \left[ \text{logistic}(\Phi_\lambda^T \hat{d}_\lambda) - \text{logistic}(X\hat{\beta}_\infty) \right]
\]

The procedures for the test now carry through as before, except that we require an estimate of \( \text{cov} \left( \hat{\delta}^+_{\lambda_1}, \hat{\delta}^+_{\lambda_2} \right) \). For this, we employ a Taylor expansion for the difference in fitted values between the single index and logistic regression models as

\[
\tilde{Y}_\lambda - \tilde{Y}_\infty \approx \left[ W_\lambda (\Phi_\lambda^T \Phi_\lambda + \lambda g)^{-1} \Phi_\lambda^T - W_\infty (X^T W_\infty X)^{-1} X^T \right] Y = H_\lambda^T Y
\]

from which

\[
\text{cov} \left( \hat{\delta}^+_{\lambda_1}, \hat{\delta}^+_{\lambda_2} \right) \approx a^T H_\lambda^T \hat{W}_{\lambda g \text{CV}} H_{\lambda_2}^T a
\]

and our test proceeds as above.

Note that this same framework can be developed to test for a departure from the Jensen effect that would result under particular parametric models in both the exponential and Poisson cases considered above. We did not pursue such extensions, because each of those models includes a natural sign for the Jensen effect, and so a difference in sign is most likely to be the quantity of interest.

4.3 Power Analysis

As in the Poisson model, we test for the Jensen Effect by calculating the quantity \( \delta \) over a range of smoothing parameters \( \lambda \). For comparison with previous models, our simulation assumes all covariates are associated with environmental conditions, resulting in the same Jensen effect definition. We generate \( p = 5 \) covariates uniformly on \([0, 0.5]\). The coefficient \( \beta = \frac{1}{\sqrt{p}} 1_p \) so that \( \|\beta\| = 1 \) and note that \( x\beta > 0 \) in all cases so that the logistic model is concave across the range of the single index and we test whether \( \delta > 0 \). To estimate \( g \) and \( \beta \), we represent \( g \) by a 25-dimensional quintic B-spline basis.
To illustrate the statistical power of the hypothesis test, we choose the single index to produce the convex composite link

\[(h \circ g)(s) \doteq \text{logistic}(g)(s) = \frac{\exp(-as) - \exp(-a)}{2(1 - \exp(-a))} + \frac{1}{2}, \tag{40}\]

which, for \( s > 0 \) always lies in \([0, 1]\) on which the logistic function is naturally concave. Our first test is therefore of the positivity of the Jensen Effect. We employed values of \( a \in \{1, 2, 3, 5, 8, 11, 15\} \) and sample sizes 100, 200, 500 and 100, conducting 400 simulations at each value combination. There is relatively little power for this test (fig. 10, left) due to the restricted range of the probabilities and the corresponding limited range of \( \delta \) values that can be achieved.

Using the same simulated data sets, we investigated our second hypothesis, that the numeric value of the Jensen Effect differs which would obtain under a standard logistic regression model (linear \( g \)). Here we find the power of the test unexpectedly decreasing with \( a \) (fig. 10, right). This may be due to the increasing steepness of the logistic model that best approximates the true relationship, moving the observed \( X\beta \) further from the origin, where the curvature is smaller, as is the corresponding Jensen effect.
4.4 Empirical Application

We again employ the USSES data for models of plant survival. We employ the same species and climate covariates as in Section 3.1.3 and also record the area covered by each plant, and a measure of competition given by a weighted sum of the areas covered by conspecifics in concentric annuli around the plant (Teller et al., 2016). We then measure survival from one year to the next as our outcome. We elected to use the four-variable climate summary employed in Section 3.1.3 rather than as functional covariates due to the smoothness of the effects for the growth model and the lowered power associated with logistic regression. After removing two observations with extreme measures of local competition, we had 5410 observations of *Pseudoregneria speciata*, 2394 of *Hesperostipa comata*, 3648 for *Poa secunda* and 1154 of *Artemesia tripartita*. Within our models, we regarded both climate and competition as environmental covariates, but kept the logarithm of plant area as an additional covariate. This corresponds to our choices in Sections 2.3.2 and 3.1.3; from the perspective of any one plant the location and size of its competitors is an exogenous factor like rainfall, but we recognize that one can argue for or against including local competition as an “environmental” variable.

In all species, a linear logistic model resulted in most survival probabilities being estimated above 50% and a correspondingly negative Jensen Effect. The Jensen effect from the single index model was also estimated to be negative at every level of smoothness for all species, which also failed to be statistically distinguishable from the value found from a linear logistic model. Plots the estimated adjustment to the linear logistic model and corresponding Jensen Effect over smoothing parameters have been reserved for Appendix D.

5 Discussion

The analysis of demographic data has tended to follow classical statistical paradigms, employing generalized linear models, and using data transformations to ensure reasonable correspondence with statistical assumptions. However, when those models are used to explore the effects of environmental variability, the conclusion is often predetermined by the link function used in the fitted model. Effects of environmental variability are central to many classic and current questions in ecology.
and evolution (Teller et al., 2016), so it is important to assess those effects using models where any outcome is possible, depending on the data.

In this paper, we have extended our prior analysis of nonparametric estimation of Jensen Effects, making it possible to assess the impact of environmental variability on reproduction and survival, as well as accounting for common data transforms in the analysis of growth. Specifically, the methods presented here allow us to test whether the curvature (and thus the sign of Jensen Effects) tacitly assumed by a standard GLM model is contradicted by the data. Because our methods are based on searching over a wide range of smoothing parameters, we are also able to detect relationships that might not be picked up at the one smoothing parameter value chosen by a model selection criterion.

In our empirical applications, we found that environmental variability is harmful for growth of the shrub *Artemisia tripartita* in western US sagebrush steppe, but essential for reproduction which only occurs under rare, extreme values of the environmental index. Variability is also essential for reproduction by the grass *Poa seconda*, but harmful for reproduction by the grasses *Pseudoroegneria spicata* and *Hesperostipa comata*. However, variability is harmful for the survival of individuals of all species.

Due to the bias towards linearity in smoothed function estimates, we can only conduct a one-sided test to determine if the Jensen Effect is opposite to that implied by the natural link function. While that is often the relevant question, it may also be useful to test whether the Jensen Effect is different from zero. Smoothing bias prevents us from conducting that test in our framework. The alternative approach developed here is to test whether the Jensen Effect is different from that implied by the link function. In the USSES data it appears that survival in two species is, in fact, harmed by environmental variability more than a logistic model would suggest, although this is not statistically significant.

We note that while our analysis has focused on single index models as the most parsimonious nonparametric representation of the reaction to multiple environmental factors affecting individual performance, a very similar analysis can be applied to more general smoothing models for effects of environmental covariates. Employing a smoothing penalty for the composite link $h \circ g$ may also be possible, but its statistical properties are not well understood.
In our own analysis of the USSES data, we have included plant area either as an additional linear term outside of the single index, in situations where structure of the link function meant that we could separate environmental effects from other covarites, and otherwise as an element of the single index. Applying our generalized definition of the Jensen Effect in Section 4, we could have made either choice for any model. We could also have employed separate smoothing models for temperature and precipitation if we felt that the data would support that level of model complexity. Finally, we could partition the Jensen Effect into components due to temperature, precipitation and competition. These can all be accomplished within the same framework as our current tests, albeit requiring the implementation of new models and the corresponding estimates of $\delta$.

References

Agren, G. I., J. A. M. Wetterstedt, and M. F. K. Billberger (2012, JUN). Nutrient limitation on terrestrial plant growth - modeling the interaction between nitrogen and phosphorus. New Phytologist 194(4), 953–960.

Casella, G. and R. L. Berger (2002). Statistical inference, Volume 2. Duxbury Pacific Grove, CA.

Chaudhuri, P. and J. S. Marron (1999). SiZer for exploration of structures in curves. Journal of the American Statistical Association 94(447), 807–823.

Chesson, P. (1994). Multispecies competition in variable environments. Theoretical population biology 45(3), 227–276.

Chesson, P. (2000a). General theory of competitive coexistence in spatially-varying environments. Theoretical population biology 58(3), 211–237.

Chesson, P. (2000b). Mechanisms of maintenance of species diversity. Annual review of Ecology and Systematics 31(1), 343–366.

Chesson, P. L. and R. R. Warner (1981). Environmental variability promotes coexistence in lottery competitive systems. The American Naturalist 117(6), 923–943.
Cornelissen, J., P. Diez, and R. Hunt (1996, OCT). Seedling growth, allocation and leaf attributes in a wide range of woody plant species and types. *Journal of Ecology* 84(5), 755–765.

Diaz, S., J. Hodgson, K. Thompson, M. Cabido, J. Cornelissen, A. Jalili, G. Montserrat-Martí, J. Grime, F. Zarrinkamar, Y. Asri, S. Band, S. Basconcelo, P. Castro-Diez, G. Funes, B. Hamzeh-hee, M. Khoshnevi, N. Perez-Harguindeguy, M. Perez-Rontome, F. Shirvany, F. Vendramini, S. Yazdani, R. Abbas-Azimi, A. Bogaard, S. Boustani, M. Charles, M. Dehghan, L. de Torres-Espuny, V. Falczuk, J. Guerrero-Campo, A. Hynd, G. Jones, E. Kowsary, F. Kazemi-Saeed, M. Maestro-Martínez, A. Romo-Diez, S. Shaw, B. Siavash, P. Villar-Salvador, and M. Zak (2004, JUN). The plant traits that drive ecosystems: Evidence from three continents. *Journal of Vegetation Science* 15(3), 295–304.

Drake, J. M. (2005). Population effects of increased climate variation. *Proceedings of the Royal Society of London B: Biological Sciences* 272(1574), 1823–1827.

Ellner, S. (1987). Alternate plant life history strategies and coexistence in randomly varying environments. In *Theory and models in vegetation science*, pp. 199–208. Springer.

Garnier, E. (1992). Growth analysis of congeneric annual and perennial grass species. *Journal of Ecology* 80(4), 665–675.

Grime, J. and R. Hunt (1975). Relative growth rate - its range and adaptive significance in a local flora. *Journal of Ecology* 63(2), 393–422.

Hardle, W., P. Hall, H. Ichimura, et al. (1993). Optimal smoothing in Single-Index Models. *The annals of Statistics* 21(1), 157–178.

Hutchinson, G. E. (1961). The paradox of the plankton. *The American Naturalist* 95(882), 137–145.

Koons, D. N., S. Pavard, A. Baudisch, J. E. Metcalf, et al. (2009). Is life-history buffering or lability adaptive in stochastic environments? *Oikos* 118(7), 972–980.

Lewontin, R. C. and D. Cohen (1969). On population growth in a randomly varying environment. *Proceedings of the National Academy of Sciences* 62(4), 1056–1060.
Ruppert, D., M. P. Wand, and R. J. Carroll (2003). *Semiparametric regression*. Cambridge University Press.

Teller, B. J., P. B. Adler, C. B. Edwards, G. Hooker, and S. P. Ellner (2016). Linking demography with drivers: climate and competition. *Methods in Ecology and Evolution* 7(2), 171–183.

Vasseur, D. A., J. P. DeLong, B. Gilbert, H. S. Greig, C. D. G. Harley, K. S. McCann, V. Savage, T. D. Tunney, and M. I. O'Connor (2014). Increased temperature variation poses a greater risk to species than climate warming. *Proceedings of the Royal Society B - Biological Sciences* 281, 20132612.

Vasseur, D. A. and K. S. McCann (2007). *The Impact of Environmental Variability on Ecological Systems*. Springer Netherlands.

Wood, S. N. (2006). *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC.

Ye, Z. and G. Hooker (2020). Local quadratic estimation of the curvature in a Functional Single Index Model. *Scandinavian Journal of Statistics* 47(4), 1307–1338.

Ye, Z., G. Hooker, and S. Ellner (2020). The Jensen Effect and Functional Single Index Models: Estimating the ecological implications of nonlinear reaction norms. *Annals of Applied Statistics* 14(3), 1326–12341.
A Diagnostic Plots for the Jensen Effect: Exponential Model

Figure 11: Jensen Effect in a exponential single index model with composite link function $g^*(s) = \exp(s)$. Left: a sample of the Jensen Effect values $\delta_\lambda$ as a function of smoothing parameters $\lambda$. Right: the corresponding t-statistics $t_\lambda$ functions.

Figure 12: Illustration of the Jensen Effect for a concave composite link function in the exponential single index model, where the link function is $g^*(s) = \sin(s)$. Left: a sample of the Jensen Effect values $\delta_\lambda$ as a function of smoothing parameters $\lambda$. Right: the corresponding t-statistics $t_\lambda$ functions.
Figure 13: Illustration of the Jensen Effect for a linear composite link function in the exponential single index model, where the link function is $g^*(s) = s$. Left: a sample of the Jensen Effect values $\delta_\lambda$ as a function of smoothing parameters $\lambda$. Right: the corresponding t-statistics $t_\lambda$ functions.

B Diagnostic Plots for the Jensen Effect: Poisson Model

Figure 14: Jensen Effect in Poisson regression with composite link $g^*(s) = \exp \left( \frac{s}{8} \right)$. Left: a sample of the Jensen Effect values $\delta_\lambda$ as a function of smoothing parameters $\lambda$. Right: the corresponding t-statistics $t_\lambda$ functions.
Figure 15: Jensen Effect in Poisson regression with composite link $g^*(s) = s$. Left: a sample of the Jensen Effect values $\delta_\lambda$ as a function of smoothing parameters $\lambda$. Right: the corresponding t-statistics $t_\lambda$ functions.

**C  Plots for USSES Plants data in Poisson Single Index Model**

Figure 16: Jensen Effect for reproduction by the grass *Poa secunda* (POSE) in USSES data set, where we model the seedling count data using a Poisson single index model. Left: the Jensen Effect values $\delta$ plotted against the values of $\lambda$. Right: the curve of the exponential link function $\exp(\hat{g})$, where $\lambda$ is selected by GCV.
Figure 17: Jensen Effect for reproduction by the grass *Hesperostipa comata* (HECO) in USSES data set, where we model the seedling count data a Poisson single index model. Left: the Jensen Effect values $\delta$ plotted against the values of $\lambda$. Right: the curve of the exponential link function $\exp(\hat{g})$, where $\lambda$ is corresponding to the minimum of $\delta$. 

![Diagram of Jensen Effect for HECO](image-url)
D  Plots for USSES Plants data in Logistic Single Index Model

Figure 18: Tests for the Jensen effect on survival for *Artemisia tripartita*. Left plots provide the value of $\delta$ estimated for each $\lambda$ (solid red lines) along with uniform confidence bands (dashed) as well as the value obtained by logistic regression (straight horizontal line) with uniform bands for the difference between that and the single index model for each $\lambda$ (blue dashed). Right plots estimate $g$ for the single index model with thick lines giving the estimate that minimizes GCV and the right plot giving the corresponding single index values.
Figure 19: Tests for the Jensen effect on survival for *Poa secunda*. Left plots provide the value of $\delta$ estimated for each $\lambda$ (solid red lines) along with uniform confidence bands (dashed) as well as the value obtained by logistic regression (straight horizontal line) with uniform bands for the difference between that and the single index model for each $\lambda$ (blue dashed). Right plots estimate $g$ for the single index model with thick lines giving the estimate that minimizes GCV and the right plot giving the corresponding single index values.
Figure 20: Tests for the Jensen effect on survival for *Hesperostipa comata*. Left plots provide the value of $\delta$ estimated for each $\lambda$ (solid red lines) along with uniform confidence bands (dashed) as well as the value obtained by logistic regression (straight horizontal line) with uniform bands for the difference between that and the single index model for each $\lambda$ (blue dashed). Right plots estimate $g$ for the single index model with thick lines giving the estimate that minimizes GCV and the right plot giving the corresponding single index values.
Figure 21: Tests for the Jensen effect on survival for *Psuedoregneria spicata*. Left plots provide the value of $\delta$ estimated for each $\lambda$ (solid red lines) along with uniform confidence bands (dashed) as well as the value obtained by logistic regression (straight horizontal line) with uniform bands for the difference between that and the single index model for each $\lambda$ (blue dashed). Right plots estimate $g$ for the single index model with thick lines giving the estimate that minimizes GCV and the right plot giving the corresponding single index values.