Composed Error Model for Insect Damage Functions: Yield Impact of Rotation Resistant Western Corn Rootworm in Illinois

by

Paul D. Mitchell, Michael E. Gray, and Kevin L. Steffey

American Agricultural Economics Association
2002 Annual Meeting, July 28-31, 2002, Long Beach, California
Selected Paper

Composed Error Model for Insect Damage Functions: Yield Impact of Rotation Resistant Western Corn Rootworm in Illinois

Paul D. Mitchell[1]

Michael E. Gray[2]

and

Kevin L. Steffey[2]

March 2002

[1] Corresponding Author and Assistant Professor, Department of Agricultural Economics, Texas A&M University, 333 Blocker Building, College Station, TX 77843-2124, voice (979) 845-6322, fax (979) 862-1563, and e-mail p-mitchell@tamu.edu.
[2] Professors of Agricultural Entomology, Department of Crop Sciences, 1102 S. Goodwin, University of Illinois, Urbana, IL 61801, phone (217) 333-6652, and email m-gray4@uiuc.edu and ksteffey@uiuc.edu

Composed Error Model for Insect Damage Functions and Rotation Resistant Western Corn Rootworm in Illinois

Abstract

This paper describes a composed error model for estimating the conditional distribution of yield loss to serve as an insect damage function. The two-part error separates yield variability due to pest damage from other non-pest factors such as soil heterogeneity, non-uniform application of agronomic practices, and measurement errors. Various common functional forms (linear, quadratic, Cobb-Douglas, negative exponential, hyperbolic, sigmoid) for the pest damage function are presented and parameter estimation is described.

As an empirical illustration, the model is used to estimate a damage function for corn rootworm, the most damaging insect pest of corn in the United States. The estimated damage function gives expected proportional yield loss as a function of the root rating difference and is used to estimate yield loss due to rotation resistant western corn rootworm in east-central Illinois. The estimated average yield loss is 11.6%, more than enough to cover the cost of a soil insecticide application. However, tremendous variability in actual loss exists, so that the probability that actual loss is less than the cost of a soil insecticide ranges 32-45%, depending on the assumed yield and price. As a result, IPM methods potentially have great value, since they can eliminate uneconomical soil insecticide applications.

Keywords: integrated pest management, Monte Carlo integration, root rating, rotation resistance, soil insecticide.

A common problem when using experimental plot data to estimate insect damage functions is the occurrence of "negative losses." For example, a field experiment evaluating a new insecticide may find that the average crop yield for the treated plots exceeds the average yield on the untreated control plots, and that the yield difference is statistically significant. However, for some replicates within the same block, the yield for the untreated control plot exceeds the yield for the plot treated with an insecticide. If the pest is truly damaging, then the yield on a treated plot should always exceed the yield on an untreated plot under equivalent conditions. However, fields are not homogenous. Soil characteristics vary, tillage and nutrient applications are not uniform, yields are measured with error, and the experimental treatments are applied with error. Randomization and replication are used to prevent systematic biases so that valid statistical inferences can be made from the collected data. Assuming proper experimental design, the usual method of analysis is to conduct ANOVA to determine if the difference in mean yields for the treated and untreated plots is statistically significant.

For many types of economic analysis, determining whether a pest control treatment generates a statistically significant yield increase is insufficient. For example, the magnitude of the increase and how it varies with measurable factors such as pest populations is needed for determining an action threshold for integrated pest management (IPM). Similarly, the variance of the yield impact of a pest control treatment quantifies the consistency of the treatment and the risk associated with its use. In these cases, the distribution of the yield difference conditional on observable factors such as the pest population or damage measures is needed. The conditional mean of this distribution can serve as an insect damage function, while the variance of the conditional distribution can be used for analyzing the risk associated with the insect pest or its control.

Unfortunately, yield variability due to soil heterogeneity, the non-uniformity of tillage, nutrient and pesticide applications, and yield measurement error is confounded with yield variability due to the treatment. As a result, assuming all the observed yield variability is due to the pest or the treatment effect over estimates the impact of the pest or the treatment on yield variability. What is needed is a statistical technique that separately identifies the effect of the pest or the treatment on yield and the effect of these other non-pest factors on yield.

This paper presents a model to separately estimate yield variability due to these two sources. The model is for data from replicated plot experiments that use randomized complete block with split plot treatments to evaluate a pest control treatment such as an insecticide, the most common experimental design for such evaluations. The estimated conditional distribution for yield loss is an insect damage function and characterizes the yield risk due to insect damage. The model uses a composed error that separates observed yield variability into two components: (1) a mean-zero normal error to capture yield variability due to soil heterogeneity, non-uniformity of agronomic practices, measurement errors, and similar factors and (2) a strictly positive error to capture yield variability due to pest damage. Characteristics of various useful forms of the model are presented and parameter estimation is described. As an empirical illustration, the composed error model is used to estimate a damage function for corn rootworm, a group of related insect species that are the most damaging insect pest of corn in the United States. Using data from the experiments of Gray and Steffey, the estimated damage function is used to estimate expected yield loss due to rotation resistant western corn rootworm in east-central Illinois.

Composed Error Model

Let Y_c and Y_t respectively denote the measured yield on the control plot and the treated plot. Each treated plot receives a pest control treatment such as an insecticide that reduces or eliminates the pest population. The control plot paired with each treated plot receives no pest control treatment and so should suffer more pest damage than the treated plot. Define λ as proportional yield loss due to pest damage:

(1) $$\lambda = (Y_t - Y_c) / Y_t .$$

If the control yield exceeds the treated yield, λ is negative, while the opposite is true if the treated yield exceeds the control yield.

The composed error model uses two independent errors, a normal (Gaussian) error ε and a strictly positive error δ. Specifically, the model assumes

(2) $$\lambda = 1 - \exp(-(\delta + \varepsilon)) .$$

The normal error ε has a zero mean and variance σ^2, while δ has an exponential distribution with mean θ. This specification ensures that $-\infty < \lambda < 1$, which is the range consistent with the definition of λ in equation (1). Maximum loss occurs if $Y_c = 0$ and $Y_t > 0$, when equation (1) gives $\lambda = 1$. The other extreme occurs if the treated plot completely fails and $Y_c > 0$ and $Y_t = 0$, when equation (1) gives $\lambda = -\infty$.

For notation, define $y = \delta + \varepsilon$. Meeusen and van Den Broeck report the probability density function for a random variable such as y. From their specification, the appendix derives an alternative expression for $g(y)$, the probability density function of y:

(3) $$g(y) = \frac{1}{\theta} \exp\left(\frac{\sigma^2 - 2\theta y}{2\theta^2}\right)\left[1 - \Phi\left(\frac{\sigma^2 - \theta y}{\sigma\theta}\right)\right],$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function. Given $g(y)$ and using the transformation of variable technique, the density function for $\lambda = 1 - \exp(-y)$ is:

(4) $$h(\lambda) = \frac{1}{\theta}(1-\lambda)^{\frac{1-\theta}{\theta}} \exp\left(\frac{\sigma^2}{2\theta^2}\right)\left[1 - \Phi\left(\frac{\sigma^2 + \theta \ln(1-\lambda)}{\sigma\theta}\right)\right]$$

for $\lambda < 1$, and 0 otherwise (see appendix).

The mean of λ is $\mu_\lambda = \frac{1+\theta-\exp(0.5\sigma^2)}{1+\theta}$ and the variance is $\sigma_\lambda^2 = \frac{\exp(2\sigma^2)}{1+2\theta}$ $-\frac{\exp(\sigma^2)}{(1+\theta)^2}$ (see appendix). Figure 1 plots the probability density function $h(\lambda)$ for a variety of parameter values to illustrate the ability of the composed error model to capture a wide range of shapes for the distribution of proportional losses, from relatively symmetric to highly skewed in either direction. Maximum likelihood can be used to estimate θ and σ and their standard errors, then the results used to test whether the expected proportional yield gain due to the treatment is zero ($\mu_\lambda = 0$).

Conditional Models

Often the goal of field experiments is not to determine whether a treatment has a significant yield effect. For example, to develop an IPM action threshold, pest populations are measured in order to predict when yield loss will be sufficient to justify the expense of a pest control. Alternatively, the goal may be to assess damage *ex post* in order to determine when economic yield loss has occurred. Both cases require estimating

μ_λ as a function of an observed x, where x is some measure of the pest population or a damage assessment signal.

Maximum likelihood is useful for estimating this type of conditional model. For models presented here, assume that the treatment does not affect σ, the variance parameter of the normal error. Estimation first requires specifying a function $\theta = q(x)$ to describe how θ depends on x, then substituting this $\theta = q(x)$ into equation (4) to obtain the likelihood function in terms of x. Maximum likelihood can then be used to estimate σ and the parameters of $q(x)$. In the empirical section, several commonly desired functional forms for the damage function are derived and estimated to illustrate.

Purged Models

After estimating σ and θ, or a conditional model with $\theta = q(x)$, it is often desirable to drop the variability in yield loss due to the ε error and focus solely on variability due to the pest. For example, if the goal is to determine yield variance due to a specific pest, or the impact of pest control on yield variance. We term these "purged models," since yield loss has been purged of any dependence on non-pest factors captured by the ε error and only depends on the pest effect as captured by the δ error. To differentiate between the full and purged models, use $\tilde{\lambda}$ for proportional yield loss with the purged model, where $\tilde{\lambda} = 1 - \exp(-\delta)$.

The purged model first requires estimating the full model $\lambda = 1 - \exp(-(\delta + \varepsilon))$, then setting $\sigma = 0$ to purge the variability in yield loss of dependence on the non-pest factors associated with the ε error and using only the estimated θ or parameters of $\theta = q(x)$. The probability density function for $\tilde{\lambda}$ is

(5) $$h(\tilde{\lambda}) = \left((1-\tilde{\lambda})^{\frac{1-\theta}{\theta}} \right) / \theta$$

for $0 \le \tilde{\lambda} \le 1$, and 0 otherwise (see appendix). The cumulative distribution function is

(6) $$H(\tilde{\lambda}) = 1 - (1-\tilde{\lambda})^{1/\theta}.$$

Note that $\theta = q(x)$ in both equations if applicable. Once purged of non-pest factors, pest damage logically cannot exceed 100%, or be negative, a range consistent with the range $0 \le \tilde{\lambda} \le 1$ imposed by the purged model. The mean and variance are $\mu_{\tilde{\lambda}} = \dfrac{\theta}{1+\theta}$ and $\sigma^2_{\tilde{\lambda}} = \dfrac{\theta^2}{(1+\theta)^2(1+2\theta)}$ (see appendix). Figure 2 plots the probability density function for a variety of parameter values to illustrate the range of possible shapes.

Corn Rootworm Damage Function

To illustrate the composed error model, a conditional distribution for proportional yield loss is estimated and used as a corn rootworm pest damage function. Corn rootworm is a complex of related species that is the most damaging insect pest of corn in the United States. Yield losses and control costs have been estimated to exceed $1 billion annually (Metcalf). Generally, the most problematic species in the complex are the western corn rootworm *(Diabrotica virgifera virgifera*) and the northern corn rootworm (*Diabrotica barberi*), but in some areas the southern corn rootworm (*Diabrotica undecimpunctata howardi*) and the Mexican corn rootworm (*Diabrotica virgifera zeae*) are more damaging. Corn rootworm adult females lay eggs in the summer. These eggs hatch the next spring and the larvae feed on the roots of corn plants. These larvae pupate and emerge as adults from the soil in late summer, then mate and lay eggs.

Larval feeding damage results in direct yield loss and makes corn plants more likely to lodge and suffer additional yield loss. Because corn rootworm larvae feed almost exclusively on corn roots, females generally lay eggs in existing corn fields. As a result, crop rotations with a single year of corn are a widely used control strategy, since eggs laid in a corn field during the summer will hatch in field planted to a non-corn crop the next spring. For continuous or multi-year corn rotations, soil insecticides applied at plant are the most common control strategy in the central and eastern Corn Belt.

In recent years, yield losses and control costs have been increasing because of the development and spread of rotation resistance among western corn rootworm (Levine and Oloumi-Sadeghi). Adult females of rotation resistant western corn rootworm lay eggs not only in corn, but also in other crops. As a result, in areas where a corn-soybean rotation is common, eggs laid in soybean fields hatch in a corn field the next spring. The emerging adults mate and increase the genes responsible for this alternative egg-laying behavior among the population. Rotation resistance first appeared in east-central Illinois and has spread eastward into Indiana, Michigan and Ohio (Onstad et al.).

Because hundreds of corn rootworm larvae can infest a single plant and root feeding occurs underground, accurately measuring larval populations is difficult. As a result, corn rootworm larval damage is usually assessed by a root rating, after larvae have pupated and emerged as adults. The root rating is a measure of corn root damage based on the number of corn root nodes exhibiting feeding scars or completely destroyed by corn rootworm larval feeding. Though other root rating scales exist, the most widely used is the 1 to 6 scale of Hills and Peters, in which 1 indicates no corn rootworm feeding damage and 6 indicates three or more root nodes completely destroyed.

The composed error model is applied to estimate a corn rootworm damage function for use in estimating the annual expected yield loss due to corn rootworm in first-year corn in east-central Illinois, where rotation resistance originated. Data from experiments comparing yields and root ratings for plots treated with soil insecticide and untreated control plots are used for the estimation. The probability density function for proportional yield loss is estimated conditional on the difference in root ratings between the soil-insecticide treated and untreated plots. Field data collected in east-central Illinois concerning root ratings in untreated first-year corn are then used to determine the unconditional distribution of the root rating difference and thus the expected proportional yield loss due to rotation resistant western corn rootworm.

Conditional Distribution of Proportional Yield Loss

Three years (1994-1996) of data from experiments conducted in near Urbana, Illinois were used for estimation (Gray and Steffey). Whole plot treatments were 12 commonly grown hybrids. Sub-plot treatments were 2 rows treated with the soil insecticide Counter® (terbufos) and 2 untreated rows. Depending on the year and location, 8-10 replicates for each hybrid were planted. Collected data included machine-harvested yield and the average root rating for five plants, using the 1-6 scale of Hills and Peters. Only data with treated and untreated yields and root ratings for both sub-plots were used, so that both the root rating difference and proportional yield loss could be calculated. The final result was 330 observations of the soil insecticide yield (Y_t) and average root rating (R_t) and the untreated control yield (Y_c) and average root rating (R_c). Proportional yield loss λ is calculated via equation (1) and the root rating difference x is calculated as $x = R_c - R_t$. Table 1 summarizes the data used for estimation.

Several common functional forms for expected proportional loss conditional on the root rating difference were estimated, i.e. $E[\lambda \mid x] = \mu_{\lambda}(x)$. A zero intercept was imposed so that plots with equal root ratings have the same expected yield. Note that a zero intercept may not be desired for all applications. Table 2 reports the required functions $\theta = q(x)$ for several functional forms for $\mu_{\tilde{\lambda}}(x)$. For notation, $\omega = \exp(0.5\sigma^2)$ and α and β be parameters to estimate. For the purged model, $\sigma = 0$, so that $\omega = 1$. Model names in Table 2 describe the functional form of $\mu_{\tilde{\lambda}}(x)$ for the purged model, not the conditional mean $\mu_{\lambda}(x)$ of the full model or of $q(x)$.

Table 3 reports maximum likelihood parameter estimates and standard errors, as well as goodness of fit and model selection measures, for each model. The adjusted R^2 and root mean square error (RMSE) were calculated using $\mu_{\lambda}(x)$, the conditional mean of the full model, since the data were fit to this mean. The adjusted R^2 and RMSE support the linear model, while the Likelihood Dominance Criterion (Pollack and Wales) and Akaike's Information Criterion (AIC) support the Cobb-Douglas model. Given these mixed results, we selected the linear model since it is both parsimonious in terms of the number of parameters and gives the best fit. Figure 3 illustrates the fit and indicates why the adjusted R^2 and RMSE are low for all models. The data show tremendous variation in proportional yield loss for the same root rating difference, so that no univariate model can provide a good fit.

The conditional mean of the purged model, $\mu_{\tilde{\lambda}}(x)$, is appropriate for a corn rootworm damage function, since only yield variability due to corn rootworm is pertinent. As a result, proportional yield loss follows the probability density reported in equation

(5), where $\theta = q(x)$ as reported in Table 2, with parameters as reported in Table 3. Thus mean proportional yield loss is $\mu_{\tilde{\lambda}}(x) = 0.114x$ for the linear model. The cumulative distribution given by equation (6) allows calculation of a 95% confidence interval around the purged model's predicted mean.

Empirical Application

Conditional Distribution of the Root Rating Difference

The Gray and Steffey data were used to estimate the probability density function for the root rating difference conditional on the untreated root rating. The root rating difference has upper and lower limits. When the untreated root rating is 6 and the soil insecticide treated root rating is 1, the root rating difference reaches its maximum of 5. The minimum of zero occurs when the two root ratings are equal, assuming that the untreated root rating must equal or exceed the treated root rating. The minimum and maximum in the data are 0.2 and 4.0 respectively.

Given the existence of upper and lower limits, a conditional beta distribution is assumed. Plots indicated a linear or quadratic relationship between the mean root rating difference and the untreated root rating with a constant standard deviation. A zero intercept was imposed, so that no root rating difference is expected when the untreated root rating indicates no corn rootworm damage. The lower and upper limits of the distribution were fixed at 0.0 and 5.0. Specific models for the linear and quadratic means are $\mu_x(R_c) = r_1(R_c - 1)$ and $\mu_x(R_c) = r_1(R_c - 1) + r_2(R_c - 1)^2$, with constant standard deviation σ_x.

The standard beta density with parameters ν and γ has mean $\mu_x = \nu/(\nu+\gamma)$ and variance $\sigma_x^2 = \nu\gamma/[(\nu+\gamma)^2(\nu+\gamma+1)]$. Solving these equations for ν and γ gives $\nu = [\mu_x^2(1-\mu_x)/\sigma_x^2] - \mu_x$ and $\gamma = [\mu_x(1-\mu_x^2)/\sigma_x^2] - (1-\mu_x)$. Substituting the linear or quadratic equation for μ_x into these gives the density function in terms of σ_x, r_1, and r_2 so that maximum likelihood can be used to estimate these parameters. Table 4 reports parameter estimates and standard errors for both models. Because all reported goodness of fit and model selection measures support the quadratic model, the quadratic model is used for this analysis. Figure 4 illustrates the fit.

Unconditional Distribution of the Untreated Root Rating

The experiments conducted by Gray and Steffey used late-planted corn the previous season as a trap crop to ensure high corn rootworm larval populations. As a result, their data for the untreated root rating are not indicative of the unconditional distribution of the untreated root rating in first-year corn. O'Neal et al. report monitoring data from first-year corn fields of several cooperating farmers in different counties in east-central Illinois for 1996-1999. Collected data included the average and standard deviation of the root rating in several untreated fields. These root rating data indicate the natural pressure from rotation resistant western corn rootworm laying eggs in soybeans.

Table 5, adapted from O'Neal et al. Table 1, reports the mean and standard deviation of the untreated root rating each season. Because a root rating must range 1 to 6, the beta density is an appropriate choice for the unconditional distribution for the untreated root rating in first-year corn. First, the reported means and standard deviations are rescaled to the standard beta density range of 0 to 1. For the mean, rescaling requires subtracting the minimum of 1 and dividing by the range of 6 – 1 = 5, and for the standard deviation, rescaling requires dividing by the range of 5. Table 5 reports the ν and γ for the standard beta density consistent with the rescaled means and standard deviations for each year, using the equations for ν and γ as functions of the mean μ and variance σ^2. For notation, denote the implied rescaled untreated root rating as $\widetilde{R}_c = (R_c - 1)/5$.

Assuming that the rescaled untreated root rating follows a beta density with a ν and γ equal to the average ν and γ reported in Table 5 would underestimate its actual variability. As a result, a hierarchical model is specified, in which the parameters ν and γ follow a bivariate normal distribution with means and variance-covariance matrix as reported in Table 5.

Empirical Results

For the specified model, the unconditional expected value of proportional yield loss is E[λ] = αE[x] and E[x] = 5.0(r_1(E[R_c] – 1) + r_2(E[R_c^2] – 2E[R_c] + 1)). However, calculating E[R_c] = E[$\nu/(\nu+\gamma)$] and E[R_c^2], where ν and γ follow a bivariate normal distribution, is analytically intractable. As a result, Monte Carlo integration (Greene p. 192-195) is used to estimate E[R_c] and E[R_c^2]. Similarly, the unconditional variance of proportional yield loss is Var[λ] = α^2Var[x]. However, the unconditional Var[x] is not the σ_x^2 reported in Table 4, since σ_x was estimated conditional on R_c. As a result, Monte Carlo methods are also used to estimate the unconditional Var[x] and obtain a 95% confidence interval for λ.

A C++ program using algorithms reported in Press et al. and Cheng drew random variables from the bivariate normal and beta distributions. First 5,000 draws of ν and γ from the bivariate normal distribution were obtained, then for each pair, 5,000 draws of the scaled untreated root rating $\widetilde{R}_c$ from the beta distribution were obtained, for a total of 25 million draws. Each $\widetilde{R}_c$ was then transformed to R_c by multiplying by 5 and adding 1. The average of these R_c and the squared R_c is the Monte Carlo integral estimate of E[R_c] and E[R_c^2] respectively.

To estimate Var[λ] and obtain a 95% confidence interval required further Monte Carlo draws of the root rating difference x and proportional yield loss λ. Each R_c was used to parameterize the beta density and draw a root rating difference x. By inverting the cumulative distribution for the purged model and using the Inverse Transform Method (Cheng), a random draw for λ is $\lambda = 1-(1-u)^{\theta}$, where u is a uniform random variable and $\theta = \alpha x/(1-\alpha x)$. The equation for θ is derived from the reported equation in Table 3 for the linear model, but for the purged form of the model with $\omega = 1$, since $\sigma^2 = 0$. The average of these λ and λ^2 is a Monte Carlo estimate of E[λ] and E[λ^2], so that the Monte Carlo estimate of Var[λ] = E[λ^2] – E[λ]^2. Similarly, the lower 2.5% and upper 97.5% quantiles are Monte Carlo estimates of the 95% confidence interval.

Table 6 reports all Monte Carlo estimates, as well as the correct values for those that can be determined analytically. The unconditional expected proportional yield loss due to rotation resistant corn rootworm in untreated first-year corn in east-central Illinois is 0.116. The standard deviation is 0.125 and the lower and upper limits of the 95%

confidence interval are 0.00149 and 0.460 respectively. These results indicate that not only is yield loss on average is quite substantial, but also quite variable.

Converting these proportional yield loss estimates into revenue loss requires using an expected yield and price and assuming that corn rootworm damage is independent of yield and price. Table 7 reports the expected revenue loss, as well as the lower and upper limits of the 95% confidence interval, using parameter estimates in Table 6. Estimates of the direct cost of purchasing and applying a soil insecticide for corn rootworm control typically range \$12-\$15/ac. Thus, the estimated revenue loss is on average more than enough to cover the direct cost of a soil insecticide.

The tremendous variability in the actual yield loss realized implies that though on average the direct cost will be covered, the probability that the cost will not be covered in a specific year on a specific field is substantial. The last column in Table 7 reports Monte Carlo estimates of these probabilities for the different yield and price assumptions. The revenue loss for each Monte Carlo draw of λ was calculated, then the losses were sorted and the cumulative probability for each loss determined empirically. Table 7 reports the probabilities that revenue loss < \$15/ac.

In general the average losses in Table 7 indicate that farmers should be concerned about corn rootworm damage on first-year corn in east-central Illinois. However, the probabilities that the loss is less than \$15/ac in Table 7 are large and indicate that applying a soil insecticide on all first-year corn acres will quite often result in a revenue loss, since the cost of the soil insecticide will not be recovered. As a result, an IPM method that measures the adult corn rootworm population or egg laying in soybean fields to be planted in corn the next season could be profitable if scouting costs are low and provide reliable information. As an example of such an IPM method, O'Neal et al. have developed an economic threshold using Pherocon AM traps to measure adult populations in soybean fields.

Conclusion

This paper describes a composed error model for use with experimental plot data to estimate a conditional distribution for yield loss to serve as an insect damage function. The model uses a two-part error to separate yield variability due to pest damage from other factors such as soil heterogeneity, non-uniform application of agronomic practices, and measurement errors. Various functional forms for the pest damage function are presented for the conditional model and parameter estimation is described.

As an empirical illustration, the composed error model is used to estimate a damage function for corn rootworm, the most damaging insect pest of corn in the United States. Using data from the experiments of Gray and Steffey, the estimated damage function is used to estimate expected yield loss due to rotation resistant western corn rootworm in east-central Illinois. The estimated average yield loss is 11.6%, which is more than enough to cover the cost of a soil insecticide application which typically ranges \$12-\$15/ac. However, tremendous variability in actual loss exists, so that the probability that actual loss is less than \$15/ac ranges 32-45%, depending on the assumed yield and price. As a result, IPM practices such as described by O'Neal et al. potentially have value, since they can eliminate uneconomical soil insecticide applications.

Various improvements or extensions of the composed error model are possible. The conditional models reported in Table 2 impose a zero intercept, so that the insect pest can only cause non-negative damage. However, some experimental evidence indicates that at low populations, some insect pests can actually increase yields by stimulating plant growth. Similarly, the zero-intercept form of the composed error model prevents estimating any negative impacts that pest control may have, such as crop damage due to herbicide application or a "yield drag" due to a transgenic gene conferring herbicide or insect resistance. As a result, some applications require models without a zero-intercept.

Additionally, the composed error model specified here uses an exponential error to capture yield variability due to the pest. The exponential error is quite restrictive in terms of the shape of the probability density function and has only one parameter. As a result, estimating models with flexible relationships for both the conditional mean and conditional variance is difficult. More flexible conditional models require a different error assumption for the pest effect, but deriving the associated composed error for the joint distribution of the errors can become difficult.

Table 1. Summary statistics for proportional yield loss and root rating difference data from Gray and Steffey used for estimation.

Variable	Year	Average	Standard Deviation	Minimum	Maximum	n	n < 0
Proportional Yield Loss	1994	0.272	0.157	-0.163	0.808	115	5
	1995	0.488	0.214	-0.363	0.850	113	2
	1996	0.197	0.110	-0.123	0.585	102	2
	Pooled	0.323	0.207	-0.363	0.850	330	9
Root Rating Difference	1994	2.76	0.53	0.8	4.0	115	0
	1995	2.68	0.68	0.2	4.0	113	0
	1996	2.03	0.52	0.6	3.2	102	0
	Pooled	2.51	0.67	0.2	4.0	330	0

Table 2. Required functions $\theta = q(x)$ for the full model that give common functional forms for the conditional mean of proportional yield loss for the purged model.

Functional Form	Purged Model $\mu_{\tilde{\lambda}}(x)$	Full Model $\mu_{\lambda}(x)$	Required $\theta = q(x)$
Linear	αx	$\alpha x\omega$	$\dfrac{\omega - 1 + \alpha x\omega}{1 - \alpha x\omega}$
Quadratic	$\alpha x + \beta x^2$	$\alpha x\omega + \beta x^2\omega$	$\dfrac{\omega - 1 + \alpha x\omega + \beta x^2\omega}{1 - \alpha x\omega - \beta x^2\omega}$
Cobb-Douglas	αx^{β}	$\alpha x^{\beta}\omega$	$\dfrac{\omega - 1 + \alpha x^{\beta}\omega}{1 - \alpha x^{\beta}\omega}$
Negative Exponential	$\alpha(1 - \exp(-\beta x))$	$\alpha(1 - \exp(-\beta x))\omega$	$\dfrac{\omega - 1 + \alpha(1 - \exp(-\beta x))\omega}{1 - \alpha(1 - \exp(-\beta x))\omega}$
Hyperbolic	$\dfrac{\alpha x}{\alpha x + 1}$	$\dfrac{\alpha x}{\alpha x + \omega}$	$\omega - 1 + \alpha x$
Sigmoid	$\dfrac{\alpha x + \beta x^2}{\alpha x + \beta x^2 + 1}$	$\dfrac{\alpha x + \beta x^2}{\alpha x + \beta x^2 + \omega}$	$\omega - 1 + \alpha x + \beta x^2$

Table 3. Estimated parameters (standard errors in parentheses) and goodness of fit measures for various corn rootworm damage functions.

Parameter	Linear	Quadratic	Cobb Douglas	Negative Exponential	Hyperbolic	Sigmoid
α	0.114	0.191	0.218	0.311	0.177	0.291
	(0.00398)	(0.0155)	(0.0238)	(0.0237)	(0.00987)	(0.0400)
β	--	-0.0297	0.286	1.0369	--	-0.0413
		(0.00541)	(0.115)	(0.350)		(0.0128)
σ	0.237	0.343	0.357	0.353	0.293	0.344
	(0.0481)	(0.0634)	(0.0674)	(0.0653)	(0.0503)	(0.0613)
Adjusted R^2*	0.123	0.060	0.060	0.053	0.109	0.062
RMSE	0.194	0.200	0.200	0.201	0.195	0.200
Log-likelihood	116.0	128.3	131.1	130.9	124.7	130.0
AIC	-228.0	-250.5	-256.2	-255.7	-245.3	-253.9

* Since a zero intercept is imposed, the adjusted R^2 is appropriate (Greene p. 255).

Table 4. Estimated parameters (standard errors in parentheses) and goodness of fit measures for the linear and quadratic mean models of the distribution of the root rating difference conditional on the untreated root rating.

Parameter	Linear	Quadratic
r_1	0.667	0.544
	(0.00491)	(0.0309)
r_2	--	0.0304
		(0.00758)
σ_x	0.348	0.336
	(0.0131)	(0.0127)
Adjusted R^2*	0.732	0.744
RMSE	0.342	0.335
Log-likelihood	-113.9	-106.0
AIC	231.7	218.1

* Since a zero intercept is imposed, the adjusted R^2 is appropriate (Greene p. 255).

Table 5. Data concerning the unconditional distribution of the untreated root rating in first-year corn in east-central Illinois.

		------- Reported* -------		----- Rescaled -----			
Year	n	Mean	Standard Deviation	Mean	Standard Deviation	ν	γ
1996	14	2.25	0.16	0.250	0.032	45.53	136.58
1997	17	3.40	0.19	0.480	0.038	82.49	89.36
1998	15	2.82	0.20	0.364	0.040	52.30	91.39
1999	28	2.26	0.15	0.252	0.030	52.53	155.91
					Average	58.21	118.31
					Variance	204.4	827.6
					Covariance	−247.3	

* Source: O'Neal et al., p. 100, Table 1.

Table 6. Monte Carlo estimates of various statistics concerning yield loss due to rotation resistant western corn rootworm in east-central Illinois.

Statistic	Monte Carlo Estimate	Correct Value
E[ν]	58.21	58.21
E[ω]	118.31	118.31
Var[ν]	272.35	272.52
Var[ω]	1103.1	1103.5
Cov[ν,ω]	-246.7	-247.3
Cor[ν,ω]	-0.450	-0.451
E[R_c]	2.694	---
E[R_c^2]	7.610	---
E[x]	1.020	---
E[λ]	0.116	---
Var[λ]	0.0156	---
Standard Deviation of λ	0.125	---
2.5% Quantile of λ	0.00149	---
97.5% Quantile of λ	0.460	---

Table 7. Monte Carlo estimated expected revenue loss due to rotation resistant western corn rootworm in first-year corn in east-central Illinois for a variety of yield and price assumptions, as well as the probability that the revenue loss is < $15/ac.

Yield (bu/ac)	Price ($/bu)	Expected Revenue Loss ($/ac)	95% Confidence Interval Lower ($/ac)	95% Confidence Interval Upper ($/ac)	Probability Revenue Loss < $15.00
120	2.00	27.84	0.36	110.30	0.448
120	2.15	29.93	0.38	118.58	0.427
120	2.30	32.02	0.41	126.85	0.408
130	2.00	30.16	0.39	119.50	0.425
130	2.15	32.42	0.41	128.46	0.405
130	2.30	34.68	0.44	137.42	0.387
140	2.00	32.48	0.42	128.69	0.404
140	2.15	34.92	0.45	138.34	0.385
140	2.30	37.35	0.48	147.99	0.368
150	2.00	34.80	0.45	137.88	0.386
150	2.15	37.41	0.48	148.22	0.367
150	2.30	40.02	0.51	158.56	0.351
160	2.00	37.12	0.47	147.07	0.369
160	2.15	39.90	0.51	158.10	0.351
160	2.30	42.69	0.55	169.13	0.335
170	2.00	39.44	0.50	156.26	0.355
170	2.15	42.40	0.54	167.98	0.336
170	2.30	45.36	0.58	179.70	0.321

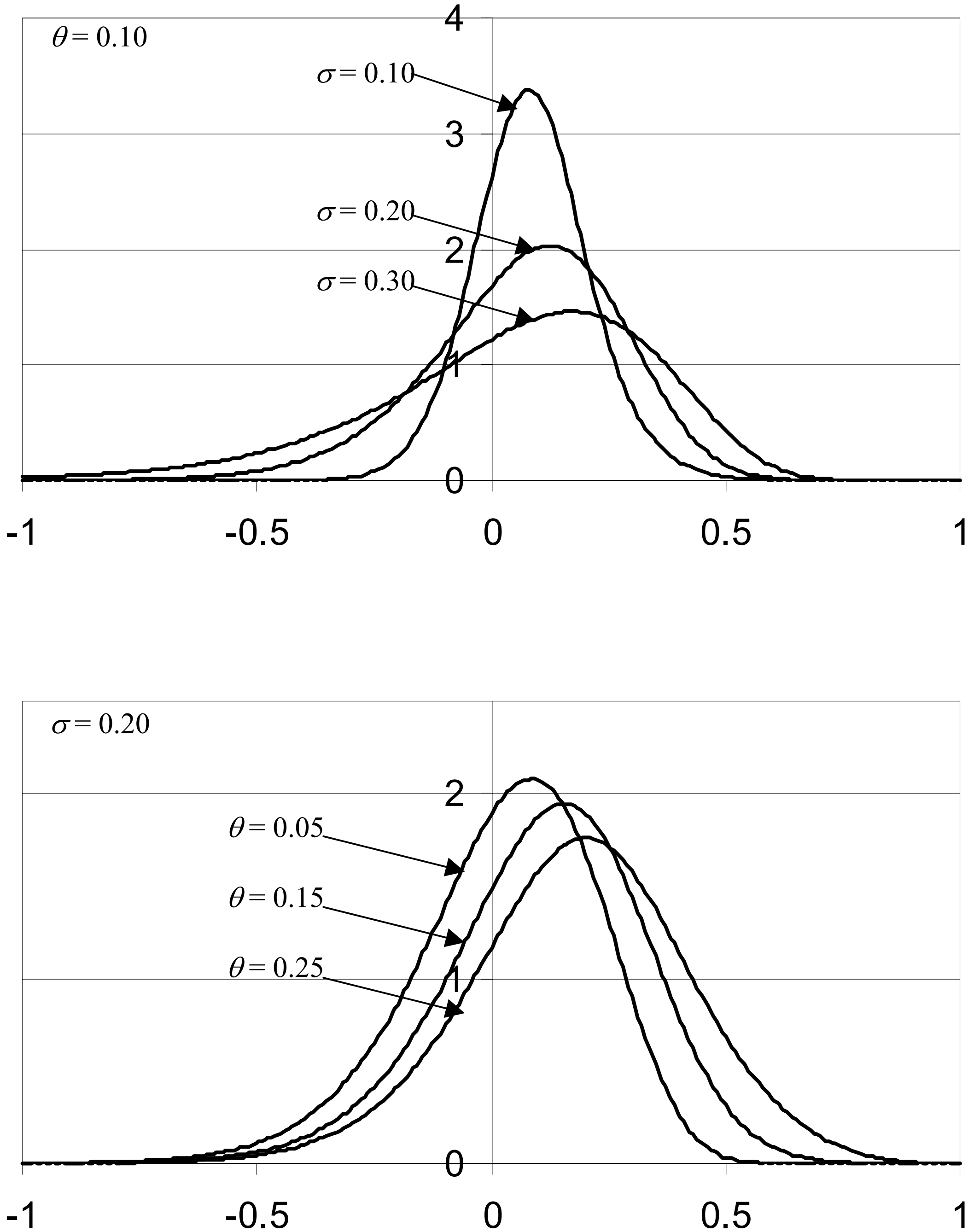

Figure 1. Probability density function $h(\lambda)$ with $\theta = 0.10$ and $\sigma = 0.10$, 0.20, and 0.30 (top) and $\sigma = 0.20$ and $\theta = 0.05$, 0.15, and 0.25 (bottom).

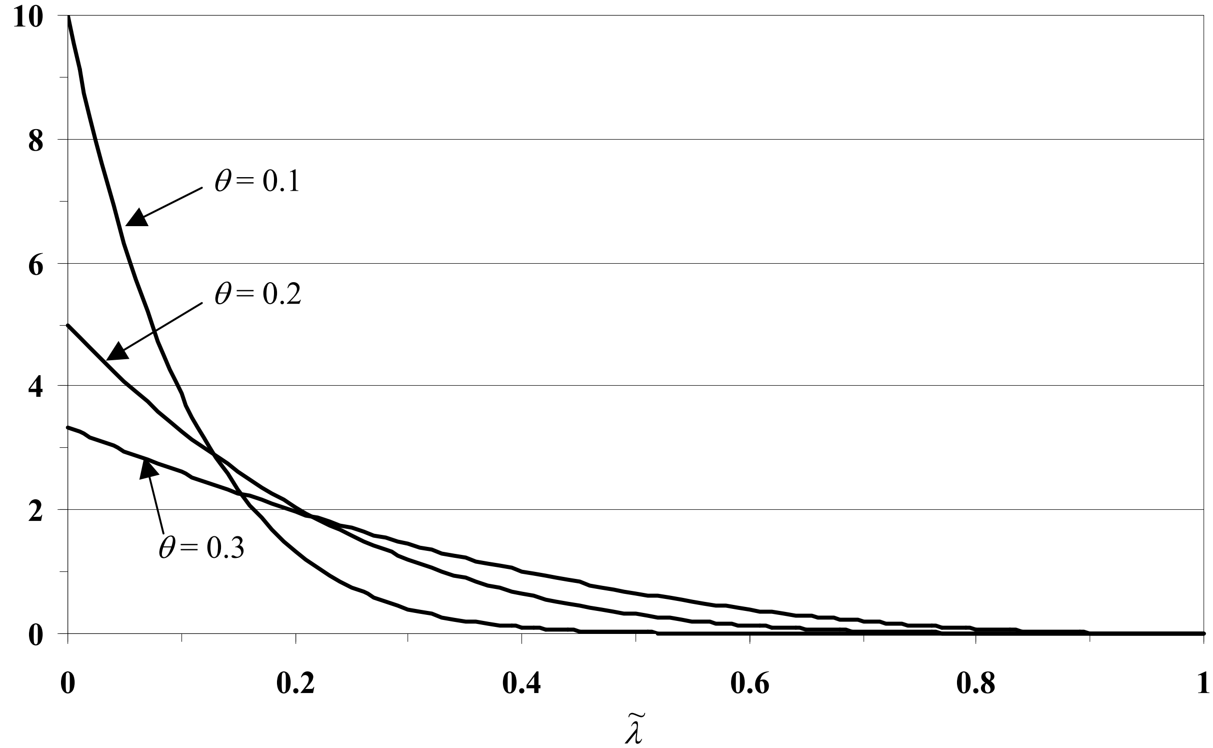

Figure 2. Probability density function $h(\tilde{\lambda})$ with $\theta = 0.10$, 0.20, and 0.30.

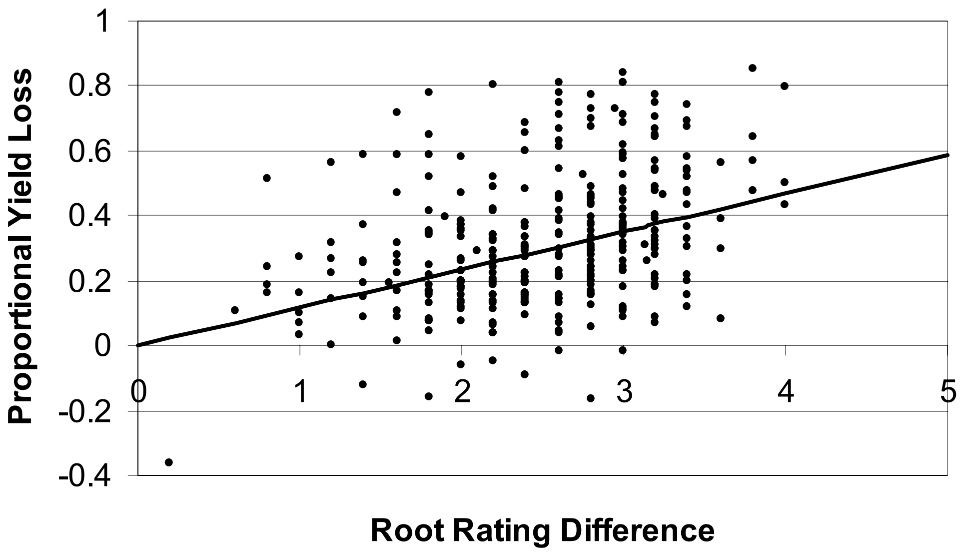

Figure 3. Observed proportional yield loss and predicted mean as a function of the root rating difference for the linear model.

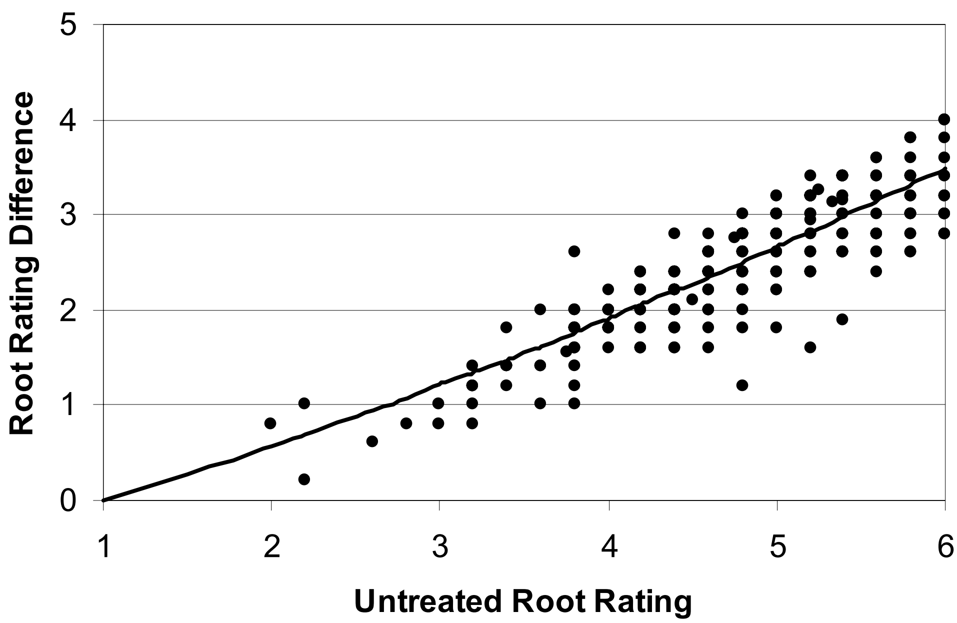

Figure 4. Observed root rating difference and predicted mean as a function of the untreated root rating for the quadratic model.

References

Cheng, R. C. H. "Random Variate Generation." *Handbook of Simulation: Principles, Methodology, Advances, Applications and Practice*. J. Banks, ed. New York: John Wiley, 1998.

Evans, M., N. Hastings, and B. Peacock. *Statistical Distributions*, 2nd ed. New York: John Wiley, 1993.

Gray, M. E., and K. L. Steffey. "Corn Rootworm (Coleoptera: Chrysomelidae) Larval Injury and Root Compensation of 12 Maize Hybrids: An Assessment of the Economic Injury Index." *J. Econ. Entomol*. 91(1998): 723-740

Greene, W. H. *Econometric Analysis*, 3rd ed. Upper Saddle River, NJ: Prentice Hall. 1997.

Hills, T. M., and D. C. Peters. "A Method of Evaluating Postplanting Insecticide Treatments for Control of Western Corn Rootworm Larvae." *J. Econ. Entomol*. 64(1971): 764-765.

Meeusen, W., and J. van Den Broeck. "Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error." *Internatl. Econ. Rev.* 18(1977): 435-444.

Metcalf, R. L. "Forward." *Methods for the Study of Pest Diabrotica*. J. L. Krysan and T. A. Miller, eds. New York: Springer-Verlag, 1986.

Levine, E., and H. Oloumi-Sadeghi. "Western Corn Rootworm (Coleoptera: Chrysomelidae) Larval Injury to Corn Grown for Seed Production Following Soybeans Grown for Seed Production." *J. Econ. Entomol*. 89(1996): 1010-1016.

O'Neal, M.E., M.E. Gray, S. Ratcliffe, and K.L. Steffey. "Predicting Western Corn Rootworm (Coleoptera: Chrysomelidae) Larval Injury to Rotated Corn with Pherocon AM Traps in Soybeans." *J. Econ. Entomol.* 94(2001): 98-105.

Onstad, D.W., M. Joselyn, S. Isard, E. Levine, J. Spencer, L. Bledsoe, C. Edwards, C. Di Fonzo, and H. Wilson. "Modeling the Spread of Western Corn Rootworm (Coeloptera: Chrysomelidae) Populations Adapting to Soybean-Corn Rotation." *Environ. Entomol.* 28(1999): 188-194.

Pollak, R. A., and T. J. Wales. "The Likelihood Dominance Criterion: A New Approach for Model Selection." *J. Econometrics*. 47(1991):227-242.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C++: The Art of Scientific Computing*, 2nd ed. Cambridge: Cambridge University Press, 1992.

Appendix

Derivation of equation (3)

Meeusen and van Den Broeck report the probability density function for $w = z + v$, where z has an exponential distribution with mean $1/\lambda$ and v is normal with zero mean and variance σ^2. Converting their notation to the notation used in this paper, $w = y$, $z = \delta$, $v = \varepsilon$ and the parameters $\theta = 1/\lambda$ and $\sigma^2 = \sigma^2$. Making these conversions, their equation (4) is the probability density function of y:

(A1) $$g(y) = \frac{1}{2\theta}\exp\left(\frac{\sigma^2 - 2\theta y}{2\theta^2}\right)\operatorname{erfc}\left(\frac{\sigma^2 - \theta y}{\sigma\theta\sqrt{2}}\right),$$

where erfc(·) = 1 – erf(·) is the complementary error function and $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x \exp(-s^2)ds$ is the error function (Press et al., p. 220). Greene (p.187) reports that $\Phi(x) = 0.5 + 0.5\operatorname{erf}(x/\sqrt{2})$, where $\Phi(\cdot)$ is the standard normal cumulative distribution function. Rearrange this expression to obtain $1 - \operatorname{erf}(x/\sqrt{2}) = 2(1 - \Phi(x))$, and then use this result to give $\operatorname{erfc}\left(\frac{\sigma^2 - \theta y}{\sigma\theta\sqrt{2}}\right) = 2\left(1 - \Phi\left(\frac{\sigma^2 - \theta y}{\sigma\theta}\right)\right)$. Substitute this into equation (A1) and simplify to obtain equation (3).

Derivation of equation (4)

Given probability density function $g(y)$ for y, the transformation of variable technique gives the probability density function $h(\lambda)$ for $\lambda = 1 - \exp(-y)$. Since $y = -\ln(1-\lambda)$ and $\partial y / \partial\lambda = \frac{1}{1-\lambda}$:

$$h(\lambda) = g(y(\lambda))\left|\partial y \big/ \partial \lambda\right|$$

$$h(\lambda) = \frac{1}{\theta}\exp\left(\frac{\sigma^2 + 2\theta\ln(1-\lambda)}{2\theta^2}\right)\left[1-\Phi\left(\frac{\sigma^2+\theta\ln(1-\lambda)}{\sigma\theta}\right)\right]\left|\frac{1}{1-\lambda}\right|$$

(A2) $$h(\lambda) = \frac{1}{\theta}\exp\left(\frac{\sigma^2}{2\theta^2}\right)\exp\left(\frac{\ln(1-\lambda)}{\theta}\right)\left(\frac{1}{1-\lambda}\right)\left[1-\Phi\left(\frac{\sigma^2+\theta\ln(1-\lambda)}{\sigma\theta}\right)\right].$$

Because $\exp\left(\frac{\ln(1-\lambda)}{\theta}\right) = \exp(\ln(1-\lambda))^{\frac{1}{\theta}} = (1-\lambda)^{\frac{1}{\theta}}$, $\exp\left(\frac{\ln(1-\lambda)}{\theta}\right)\left(\frac{1}{1-\lambda}\right)$ simplifies to

$\frac{(1-\lambda)^{\frac{1}{\theta}}}{1-\lambda} = (1-\lambda)^{\frac{1}{\theta}-1} = (1-\lambda)^{\frac{1-\theta}{\theta}}$. Substitute this simplification into equation (A2) to obtain equation (4).

Derivation of mean and variance of λ

By equation (2), $\lambda = 1-\exp(-\delta)\exp(-\varepsilon)$. Define two random variables $a = \exp(-\delta)$ and $b = \exp(-\varepsilon)$, so that $\lambda = 1-ab$. As Evans, Hastings and Peacock report, since $\varepsilon \sim \mathrm{N}(0,\sigma^2)$, *b* has a lognormal distribution with mean and variance

(A3) $$\mu_b = \exp(0.5\sigma^2)$$

(A4) $$\sigma_b^2 = \exp(2\sigma^2) - \exp(\sigma^2).$$

Since δ has an exponential distribution with mean θ, it has probability density function $w(\delta) = \exp(-\delta/\theta)/\theta$. The transformation of variable technique gives *f*(*a*), the probability density function of *a*:

(A5) $$f(a) = \frac{1}{\theta}a^{\frac{1}{\theta}-1},$$

for $0 \le a \le 1$ and 0 otherwise. The mean of *a* is $\mu_a = \int_0^1 af(a)da = \frac{1}{\theta}\int_0^1 a^{\frac{1}{\theta}}da$, which is

(A6) $$\mu_a = \frac{1}{1+\theta}.$$

The variance of *a* is $\sigma_a^2 = \int_0^1 a^2 f(a)da - \left(\int_0^1 af(a)da\right)^2$. The first term is $\int_0^1 a^2 f(a)da =$

$\frac{1}{\theta}\int_0^1 a^2 a^{\frac{1}{\theta}-1}da = \frac{1}{\theta}\int_0^1 a^{\frac{1}{\theta}+1}da = \frac{1}{1+2\theta}$. Using (A6), the second term is $\frac{1}{(1+\theta)^2}$. Thus

$\sigma_a^2 = \int_0^1 a^2 f(a)da - \left(\int_0^1 af(a)da\right)^2 = \frac{1}{1+2\theta} - \frac{1}{(1+\theta)^2}$, which can be simplified:

(A7) $$\sigma_a^2 = \frac{\theta^2}{(1+2\theta)(1+\theta)^2}.$$

The mean of $\lambda = 1 - ab$ is $\mu_\lambda = 1 - \mu_a\mu_b$ because *a* and *b* are independent, since δ and ε are independent. Using (A3) and (A5),

(A8) $$\mu_\lambda = 1 - \frac{\exp(0.5\sigma^2)}{1+\theta}.$$

The variance of $\lambda = 1 - ab$ is $\sigma_\lambda^2 = \mathrm{Var}[ab]$. Because *a* and *b* are independent,

$\mathrm{Var}[ab] = \sigma_a^2\sigma_b^2 + \sigma_a^2\mu_b^2 + \sigma_b^2\mu_a^2$. Substitute (A3)-(A5) into this equation and simplify:

$$\sigma_\lambda^2 = \frac{\theta^2\left(\exp(2\sigma^2) - \exp(\sigma^2)\right)}{(1+2\theta)(1+\theta)^2} + \frac{\theta^2\exp(\sigma^2)}{(1+2\theta)(1+\theta)^2} + \frac{\exp(2\sigma^2) - \exp(\sigma^2)}{(1+\theta)^2}$$

$$\sigma_\lambda^2 = \frac{\theta^2\exp(2\sigma^2) - \theta^2\exp(\sigma^2) + \theta^2\exp(\sigma^2)}{(1+2\theta)(1+\theta)^2} + \frac{\exp(2\sigma^2) - \exp(\sigma^2)}{(1+\theta)^2}$$

$$\sigma_\lambda^2 = \frac{\theta^2\exp(2\sigma^2) + (1+2\theta)\left(\exp(2\sigma^2) - \exp(\sigma^2)\right)}{(1+2\theta)(1+\theta)^2}$$

$$\sigma_{\lambda}^{2} = \frac{\theta^2 \exp(2\sigma^2) + (1+2\theta)\exp(2\sigma^2) - (1+2\theta)\exp(\sigma^2)}{(1+2\theta)(1+\theta)^2}$$

$$\sigma_{\lambda}^{2} = \frac{(1+2\theta+\theta^2)\exp(2\sigma^2) - (1+2\theta)\exp(\sigma^2)}{(1+2\theta)(1+\theta)^2}$$

$$\sigma_{\lambda}^{2} = \frac{(1+\theta)^2 \exp(2\sigma^2) - (1+2\theta)\exp(\sigma^2)}{(1+2\theta)(1+\theta)^2}$$

(A9) $$\sigma_{\lambda}^{2} = \frac{\exp(2\sigma^2)}{(1+2\theta)} - \frac{\exp(\sigma^2)}{(1+\theta)^2}.$$

Derivation of equation (5)

By definition, $\tilde{\lambda} = 1 - \exp(-\delta) = 1 - a$, where $a = \exp(-\delta)$. From (A8), a has probability density function $f(a) = \frac{1}{\theta} a^{\frac{1}{\theta}-1}$, so that the transformation of variable technique can be used to find $h(\tilde{\lambda})$, the probability distribution function of $\tilde{\lambda}$. Because $\tilde{\lambda} = 1 - a$, $a = 1 - \tilde{\lambda}$ and $\left| \partial a \big/ \partial \tilde{\lambda} \right| = |-1| = 1$. Thus

(A10) $$h(\tilde{\lambda}) = f(a(\tilde{\lambda}))\left| \partial a \big/ \partial \tilde{\lambda} \right| = \frac{1}{\theta}(1-\tilde{\lambda})^{\frac{1}{\theta}-1}.$$

Derivation of mean and variance of $\tilde{\lambda}$

Because $\tilde{\lambda} = 1 - a$, $\mu_{\tilde{\lambda}} = \mathrm{E}[\tilde{\lambda}] = 1 - \mathrm{E}[a]$. By equation (A6), $\mathrm{E}[a] = \frac{1}{1+\theta}$, so that

(A11) $$\mu_{\tilde{\lambda}} = 1 - \frac{1}{1+\theta} = \frac{1+\theta-1}{1+\theta} = \frac{\theta}{1+\theta}.$$

Similarly, $\sigma_{\tilde{\lambda}}^{2} = \mathrm{Var}[\tilde{\lambda}] = \mathrm{Var}[1-a] = \mathrm{Var}[a] = \sigma_{a}^{2}$, which is reported in equation (A7).