9-1-2011

Implications of Simultaneity in a Physical Damage Function

Kelly M. Cobourn
Boise State University

Hannah J. Burrack
North Carolina State University

Rachael E. Goodhue
University of California - Davis

Jeffrey C. Williams
University of California - Davis

Frank G. Zalom
University of California

This is an author-produced, peer-reviewed version of this article. © 2009, Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (https://creativecommons.org/licenses/by-nc-nd/4.0/). The final, definitive version of this document can be found online at Journal of Environmental Economics and Management, doi: 10.1016/j.jeem.2011.02.002
Title: Implications of Simultaneity in a Physical Damage Function

Authors and Affiliations:
(1) Kelly M. Cobourn
    Boise State University, Department of Economics
(2) Hannah J. Burrack
    North Carolina State University, Department of Entomology
(3) Rachael E. Goodhue
    University of California, Davis, Department of Agricultural and Resource Economics
(4) Jeffrey C. Williams
    University of California, Davis, Department of Agricultural and Resource Economics
(5) Frank G. Zalom
    University of California, Davis, Department of Entomology

Running Title: Implications of Simultaneity in a Damage Function

Manuscript Correspondence:
    Kelly M. Cobourn
    Boise State University, Department of Economics
    1910 University Drive
    Boise, ID  83725-1620
    kellycobourn@boisestate.edu
    phone: (208)426-3302
    fax: (208)426-2071
Title: Implications of Simultaneity in a Physical Damage Function

Abstract:
A modeler must often rely on highly simplified representations of complex physical systems when analyzing associated economic issues. Herein, we consider a management problem in which a bioeconomic system exhibits simultaneity in processes governing productivity and damage. In this case, it may benefit the producer to sacrifice productivity to reduce the costs associated with increased damage. We specify empirically a structural damage relationship that explains the biological process by which an invasive species damages a host and estimate the structural model and its reduced form with an exceptional dataset on infestation of olives by the olive fruit fly. We contrast the results of these models with the approach typically taken in the economic literature, which expresses damage as a function of pest density. The population-based approach introduces significantly greater bias into the individual grower’s choice of damage control inputs than estimates based on the structural model.

Keywords: bioeconomic modeling; simultaneous equations; pest management; damage; olive fruit fly
Because of the potential for feedback between physical and economic systems, integrated interdisciplinary models increasingly have been used in economic analyses. Even in these integrated models, the complexity of physical systems often compels the modeler to rely on highly simplified representations. Although the economic literature recognizes the potential for bias to arise from the misspecification of physical relationships, few studies examine the extent to which that bias affects conclusions about economic behavior.

We consider a case in which a biological system exhibits simultaneity in processes that contribute to the potential value of and damage to an economic output. We examine this issue specifically within the context of a pest management problem. The fundamental issue of concern is that increasing the potential value of the crop also makes it a more attractive host for a pest, increasing the costs associated with infestation. Under these circumstances, it may be optimal for a producer to sacrifice potential gains in crop value in order to reduce losses from pest damage.

Although our empirical analysis focuses on a specific pest–host relationship, it represents a broad class of environmental and resource problems. Another such example involves estimating the effect of climate change on plant growth. Climate change influences foliar growth directly via its influence on temperatures and precipitation, variables that are usually included in bioeconomic models. Climatic variables also influence ambient ozone levels, which reduce plant yields [1,13]. An analysis that fails to control for this indirect effect will produce estimates of the impact of temperature changes on yields that suffer from omitted-variable bias.

In our application, favorable weather conditions increase the size of the host, which increases both yield and output quality—for which producers receive a price premium. However, those same conditions contribute to an increase in the pest population, which damages the host, reducing yield. As in the example of the effect of climate change on plant growth, underlying biological variables
affect the level of damage both directly and indirectly. However, our case has an additional feedback loop that further complicates estimation: Larger (higher quality) hosts provide a more attractive and effective reproductive medium for the pest, contributing to an increase in the pest population and an increase in the incidence of host damage. As a result, those underlying conditions that favor growth in the host and in the pest population have a second-round impact on growth in the pest population and damage to the host.

In the economic literature on pest management, the prevalent approach is to specify damage as a proportional function of the pest population density. Because this simplified representation does not allow for simultaneity in processes governing the development of the host and the pest damage process, we argue that this approach may sometimes be misspecified. We outline a simple theoretical framework that illustrates how failure to account for such simultaneity influences economic behavior. Then we define and estimate a structural damage function that describes the pest–host damage process. We compare the fit, predictive ability, and economic conclusions yielded by the structural model with those of a population-based damage function similar to that used most often in the economic literature. We demonstrate that the latter biases the choice of pest control inputs and the estimated losses from pest damage relative to the structural approach.

**The Damage Control Literature**

A pest damage function defining yield loss as a function of a pest population, however well it fits the data, may misrepresent causality in the damage relationship. One problem that arises concerns how the pest population is defined. The population of interest in the damage function is that portion of the pest population that actually inflicts harm upon the host. Without studying the basic drivers of the damage process, it is difficult to distinguish between the portion of the pest population that is easily measured, as with traps, and that which damages the host.
For example, the modeler may link the adult pest population to yield losses, even though in many cases adult pest populations *per se* do not inflict damage. Adult males in the current population cannot cause the damage by larvae born from eggs deposited before the males became adults. Their contribution to damage inflicted by the next generation is indirect because the female adult chooses where to lay eggs. Rather, factors such as climate and weather drive reproductive activity, overall numbers of the pest, and damage.

Expressing damage as a function of underlying factors is akin to specifying a structural equation. Depending on the specifics of the structural specification, a population-based approach may be a reduced form. However, this need not be the case. Identifying the underlying factors that affect the pest–host relationship clarifies ways in which a population-based damage model departs from the structural and reduced-form damage functions, and illustrates pitfalls associated with a population-based specification. An additional advantage to formulating the structural damage model is that it can explicitly define the role of the host in the infestation process, a facet of the problem that a population-based model may obscure. A structural model can therefore incorporate management practices that reduce the susceptibility of the host to infestation.

Although several studies in the pest management literature discuss the potential for factors other than the pest population to affect the observed rate of damage to a host, the majority conceptualize losses from infestation as a function of the pest population alone [10,11,23,26]. The model used by Feder [10] exemplifies the population-based approach to formulating a damage function: The model assumes damage \((D)\) is a linear function of the pest population \((N)\): \(D(N) = \delta \cdot N\). Feder treats \(\delta\) as a constant although he mentions that, in general, \(\delta\) may depend on “temperature, humidity, differences in plant susceptibility at different periods, arrival of new pest biotypes, and other factors” (p. 97).
Another vein of literature examines the productivity of pest control inputs assuming that those inputs abate damage by reducing the pest population relative to some baseline [3,19,29]. Each starts with a theoretical production function of the form \( y = f(Z, G(X)) \), where \( y \) is total output. Output is a function of a vector of inputs \( Z \), which affect production directly, and a damage abatement function \( G(\cdot) \), which depends on a vector of damage-control inputs \( X \).

Lichtenberg and Zilberman [19] mention the potential for bias in estimates of pesticide productivity when factors simultaneously affect the pest population and host productivity. However, they do not address this bias in their analysis, which is theoretical. Babcock, Lichtenberg, and Zilberman [3] specify a damage function that depends on the damage-control inputs \( X \) and other inputs \( Z \), including weather, host characteristics, and management practices. Similarly, Saha, Shumway, and Havenner [29] allow \( X \) and a subset of the vector \( Z \) to interact in the damage abatement function \( G(\cdot) \). In both of these analyses, the inclusion of \( Z \) in the damage or abatement function alters the optimal use of a damage-control input by an individual producer. These studies each share a potential weakness in that they define a constant baseline damage rate that changes only with the use of damage-control inputs.¹

Studies focused on estimating the productivity of non-pesticide damage-control inputs, particularly genetically modified (GM) crops, contribute another perspective to the damage control literature. For example, Qaim and Zilberman [24] express damage as a function of an initial pest population, insecticide use, and the adoption of GM crops. Their study emphasizes the role of GM crops in increasing productivity and in reducing the use of chemical insecticides. Smale, Zambrano, and Cartel [31] summarize the contributions of another 46 studies in the economic literature that focus on the impacts of Bt cotton adoption. These studies recognize the efficacy of control methods that reduce damage by targeting host susceptibility. However, none
explicitly breaks down the damage process into its structural components to identify underlying factors that simultaneously drive changes in the yield and quality of the output and pest damage.

In contrast, Regev, Gutierrez, and Feder [25] track dynamic and simultaneous changes in alfalfa weevil development and alfalfa growth to determine optimal treatment timing for an individual grower. The biological component of their model specifies inter- and intra-seasonal dynamics for the plant, the adult weevil, and weevil larvae (which damage the plant by feeding), all of which are related via temperature. However, their analysis does not allow for feedback between host characteristics and the pest population, and so they cannot comment on the bias introduced by failing to consider host characteristics when determining pesticide use.

Christiaans, Eichner, and Pethig [9] present a theoretical analysis that aligns most closely with this paper’s analysis. They derive a crop production function based on micro-level constrained optimizing behavior by a pest and a host. Their formulation is structural, in that it allows underlying factors to affect both the pest population and host susceptibility. By explicitly considering host susceptibility, they show that pest control may be accomplished by enhancing crop resilience to infestation, by reducing the pest population, or by altering the use of some input that accomplishes both ends. Although most closely related to our analysis, theirs is theoretical and cannot comment on the relative importance of accounting for feedback between the host and the pest population when determining optimal pest control practices. Our contribution to the literature, and our primary objective, is to address this issue.

**Theoretical Framework**

We are concerned with a vector, \( Z \), of which a single element \( z \) \((z \in Z)\) impacts output directly and indirectly, with the indirect effect operating via a damage function. We specify a production function as
(1) \[ y = f(Z, D(z, x)), \]

where output may be defined generally as a function of yield and/or output quality.

In equation (1), \( x \) is a damage-abating input such as a pesticide and \( D(.) \) represents the damage function. By specifying a damage function, as opposed to an abatement function, as in [19], we allow the damage rate to vary even if no damage-abating inputs are applied. In our case, \( z \) is a host characteristic that is influenced by exogenous biological factors and management practices. Denoting derivatives with subscripts, we assume \( f_z > 0, \ D_D < 0, \ D_z > 0, \) and \( D_x < 0 \).

We consider the case in which \( z \) affects both yield and output quality. Specifically, price per unit output, \( p \), depends on \( z \) \((p'(z) > 0)\), and a producer is paid based on quantity delivered, \( y \), which is a function of \( z, x \), and other variables. We begin with a static profit-maximization problem in which a grower chooses \( z \) and \( x \). We fix all other variables in the production function and assume that \( f \) is differentiable and concave in \( z \) and \( D \). The grower’s objective is

\[
(2) \quad \max_{z, x} p(z) \cdot y - c(z, x) \text{ subject to } \]

\[
(3) \quad y \leq f(z, D(z, x)),
\]

where \( c(·) \) is the cost of production, which depends on the use of the damage-abating input \( x \) and the input \( z \).

Assuming an interior maximum, the first-order necessary conditions require

\[
(4) \quad p(z)f_z + p'f(z, D(z, x)) = c_z - p(z)f_D D_z, \quad \text{and}
\]

\[
(5) \quad p(z)f_D D_x = c_x.
\]

In (4) and (5) the left-hand side of the expression is the marginal benefit—and the right-hand side is the marginal cost—of an additional unit of \( z \) or \( x \), respectively. We assume that the sufficient conditions for a maximum hold.
If we were to ignore the role that z plays in the damage process, (3)–(5) become

\[(3') \quad y \leq f(z, D(x)),\]

\[(4') \quad p(z)f_z + p' f(z, D(x)) = c_z, \text{ and}\]

\[(5') \quad p(z)f_DD' = c_x.\]

In (4), the profit-maximizing producer balances the benefit of an additional unit of z, which includes the value of an increase in production and the value of an increased per-unit price for higher quality output, with the marginal costs, which include the direct costs of an incremental increase in z and the indirect costs of yield losses from increased damage from pests. In (4'), the profit maximizer balances the same marginal benefit of z with only its direct marginal cost. Given the sign assumptions, optimal z per (4) is necessarily less than the optimal z obtained using the misspecified production function, as defined by (4'). Similarly, because \(p(z)\) is increasing in z, if \(D_z \leq 0\), the optimal use of x per (5') is necessarily greater than that per (5). If \(D_z > 0\), the effect of the misspecification on the use of x depends on the relative magnitude of the opposing effects.

The theoretical framework above applies even in the case when the input z is exogenous to the grower. In this restricted problem, the relevant first-order conditions with and without z included are (5) and (5'), which implicitly define differing conditions for the optimal use of x. Whether or not z is a choice variable, the aspect of this problem fundamental to our results is that z affects the baseline damage function—defined as the level of damage observed when no damage control input (x) is used. In this case, excluding z when estimating the damage function biases the parameters of the empirical damage function. The extent and direction of bias in x is an empirical question that we examine within the context of our application to the olive fruit fly.

For the sake of simplicity, we discuss the implications of excluding z from the damage function.
when it is exogenous to the grower in order to demonstrate the potential value of estimating a structural model rather than a population-based model.

This section provides the foundation for understanding how misspecification of the damage function may affect conclusions about a grower’s incentives regarding the use of inputs $x$ and/or $z$, and any pest management recommendations that may result. However, even if a grower can do nothing to mitigate damage by a pest, incorrectly specifying the empirical damage function may still affect estimates of the losses due to the pest infestation. The remainder of this analysis focuses on specifying an empirical damage function when simultaneity between processes that drive changes in the host and pest damage exists. In order to address the potential misspecification bias discussed in this section, the modeler may have to investigate the physical or biological damage process at a greater level of detail than is typical in the economics literature and express damage as a function of its structural drivers instead of relying on a simplified population-based representation.

**Empirical Damage Function**

We investigate the structure of a damage function specifically within the case of the recent olive fruit fly infestation in California. The olive fruit fly, native to the Mediterranean, was first detected in Los Angeles County in 1998. By 2004, the fly had spread throughout all of California’s olive-producing counties, ranging from the southern border of the state to Shasta County in the north, and from the coast to the western edge of the Sierra Nevada mountain range.

This case exhibits several characteristics that simplify the analysis. Because the olive fruit fly is a single-host pest, all damage by the fly is manifest in the olive crop. Also, the olive fly is the only pest at present that cannot be controlled with pruning, so olive trees are effectively a single-pest host. Because the fly damages only the olive fruit—it does not harm a tree’s future productivity—we can
focus on damage to fruit during a single growing season. Finally, because California’s olive-growing regions account for all commercial production in the U.S. (and for the vast majority of ornamental olive trees), the fly’s further spread is not a complication [33].

An additional advantage of this specific case is that the fly has a long history in the Mediterranean region. As a result, the entomological literature is established and has extensively investigated the relationship between fruit size and the susceptibility of fruit to infestation by the fly. That literature shows a consensus that fruit size and infestation rates exhibit a strong positive correlation. There is as yet no agreement about whether flies are actually attracted to larger fruit because they are larger or whether there is some other characteristic driving infestation that is simply correlated with fruit size [5]. Factors suggested as important include shape, color, hardness of the epicarp, and the composition of olive surface waxes [16,17,28].

The entomological literature for the Mediterranean region, as well as emerging literature specific to California, indicates that weather variables and management practices influence fly populations and reproductive activity. Studies of olive phenology find that these same weather conditions and management practices contribute to fruit growth. Thus, the scientific literature identifies the factors that simultaneously affect fly populations and fruit vulnerability, and provides the foundation for our structural damage function.

The relationship between fruit size and productivity is defined by the specific structure of the olive industry in California. As previously discussed, fruit size affects both yield and output quality. The impact of fruit size on yield is straightforward. Larger fruit are heavier and processors pay for raw fruit on the basis of weight at delivery. The link between size and quality is simple in this application: Olive size is equivalent to quality, as defined and rewarded by the canning industry. Even though for many fruits, the link between size and quality is complex—Parker and Zilberman
[22] notice that because larger peaches are sweeter, a premium on peach size is actually a premium for sweetness—olives, left to soak in acid baths for as long as a year, artificially ripened with lye, and blackened with iron sulfate, are effectively uniform in terms of every characteristic except size.

Our structural damage function specifies host damage \((D)\) as a function of a host characteristic \((z)\) and the pest population \((P)\):

\[
D = f_d(z, P)
\]

The pest may be present, but if the host is not susceptible to damage, we expect a zero damage rate. The converse is also true: The host may be susceptible to infestation, but if there is no pest population, we expect zero damage.\(^3\) Any host characteristic that directly influences yield plays the role of the \(z\) in the theoretical model.

Because the fly relies on its host as a reproductive medium, the host’s characteristics may affect the level of the pest population. We specify the pest population as a function of \(z\) and a vector of other exogenous explanatory variables \(V\):

\[
P = f_p(z, V)
\]

The level of damage may also affect the host’s characteristics, as would be the case if infestation reduced the size of the host or distorted its color. The host characteristic \(z\) depends on damage and a vector of other explanatory variables \(W\):

\[
z = f_z(D, W)
\]

There may be overlap in \(V\) and \(W\). The full structural damage system consists of the three simultaneous equations described by (6)–(8).

For estimation, we condense (6)–(8) into a system of two simultaneous equations:

\[
D = f_d(z, V), \quad \text{and}
\]

\[
z = f_z(D, W)^4
\]
In general, estimation of (9) and (10) requires an instrumental variables approach. In the case of olives, we can simplify (10) based on the fact that entomologists have found no evidence that damage by the olive fly affects the size of fruit. Thus, (10) becomes

\[(10') \quad z = f_z(W),\]

and the damage system reduces to two simultaneous equations that are triangular in structure. The assumption required to drop \(D\) from (10) is our critical identifying assumption.

The reduced form of the damage system may be obtained by substituting (10') into (9) to obtain

\[(11) \quad D = g(W, V).\]

The reduced form may be of greater practical use as a management tool for producers than the structural damage function in cases where \(z\) is difficult to observe or measure and/or when \(W\) and \(V\) contain variables for which data are easily accessible (such as temperature and precipitation). If the variables in \(W\) predict \(z\) with perfect accuracy, the true reduced form (11) contains no additional information about the damage process than does the structural model. In practice, however, differences may arise between the two if unobservable factors affect \(z\), in which case the structural model has an advantage in that it captures the full variation in \(z\).

The reduced form in (11) is not the same as the population-based damage function approach typically used in the economic literature:

\[(12) \quad D = g(P).\]

The population-based specification (12) expresses damage as a function of a supposedly exogenous pest population. In using (12), the modeler omits \(z\), which affects damage directly. As shown in the theoretical example, the modeler who relies on (12) risks misspecifying (1) and introducing bias into economic decisions.
The system given by (9) and (10') explicitly incorporates the host characteristic $z$ when estimating the parameters associated with the damage process. In our particular application, fruit size ($z$) is an object of direct interest to growers. Because fruit size is observable and measurable and determines yield and fruit quality, growers must consider the simultaneous effects of management decisions on fruit size and pest damage. When $z$ is not an object of interest or a choice for the grower, the reduced-form (11) may be a simpler approach to estimating an empirical damage function. In this case, constructing the structural model identifies (11) as the correct reduced form. Our objective in this paper is to demonstrate that a population-based approach, which is not a reduced-form version of the structural damage model and in some cases may constitute an ad hoc specification of the damage process, potentially introduces greater bias into both individual behavior and policy recommendations than does the structural model.

Data

To estimate the structural damage function, we follow the preponderance of the scientific evidence and define $z = S$, where $S$ denotes fruit size. Literature from the fields of entomology and pomology identifies the relevant variables for inclusion in $V$ and $W$. In particular, temperatures influence the level of reproductive activity exhibited by the olive fly and overall population levels [12,18]. Orchard management practices, such as irrigation and post-harvest sanitization, also affect fly activity [38]. The same temperature and management practices affect fruit size. Thus, $V$ and $W$ overlap in that both contain information about temperatures and management practices. However, the scientific literature highlights the role that humidity and precipitation play in $W$ and supports their exclusion from $V$ [30].

To estimate the structural damage function, we use a data set developed by H.J. Burrack and F.G. Zalom in the Department of Entomology at the University of California, Davis.6
Burrack and Zalom chose orchards for their study to represent a variety of California microclimates in northern and southern olive-growing regions, and at inland and coastal locations. Collaborating researchers collected field infestation data weekly from May to December of 2005. Trees at each site were untreated with pesticides and had sufficient fruit to provide a random sample of up to 100 olives each week.

The entomologists and collaborators recorded information on fruit measurements, infestation damage, and the number of adult flies trapped in the vicinity of each tree. Also recorded were categorical data about management practices at each site, including the use of irrigation and ground cover. Data for other variables that might impact the fly’s preference for certain cultivars (such as surface wax composition) were not collected as part of the experiment because, as previously discussed, entomologists believe that fruit size may aptly capture the difference in observed infestation rates between cultivars. To measure infestation damage, each olive in the weekly sample from each tree was dissected and examined under a microscope. Infestation indicators include the number of stings on the outside of each olive, eggs inside each olive, live and dead larvae, and larval exit holes. In total, the dataset contains information on 81,267 olives.

The condensed dataset that we use in this analysis is a panel, where each sampled tree constitutes an “individual” and we average data across sampled olives on one tree. The dataset contains 1,140 observations for 48 trees over 29 weeks. There are two primary reasons that we choose to examine intra-seasonal damage rates. The first is that changes in damage correspond to gains in fruit size over the course of the growing season. Thus, growers may alter their use of inputs that cause fruit to grow (i.e. reduce yield and fruit quality) in order to avert fly damage. The second is that temporal fluctuations in damage rates drive the optimal timing and intensity of
pesticide applications within a growing season. Table I describes the variables used in estimation and their sources.

The dataset contains several measures of infestation that we could use to define damage. We use one that coincides with practice in the processing industry. Processors of canning olives in California define damage as at least one visible sting on the surface of a raw olive. Therefore, we use the proportion of fruit with at least one sting to measure damage. To estimate the population-based damage specification, we use the number of adult female flies trapped in the vicinity of each tree (denoted TR) as a proxy for the pest population (P). Because we define damage as stings to the surface of fruit, and female flies inflict those stings, this is the best available measure of the damaging pest population. As noted in Table I, the number of female flies trapped per week at each site is included in Burrack and Zalom’s dataset. The use of trapping numbers as a proxy for the pest population also aligns with current pest management practices in the industry. Specifically, integrated pest management (IPM) guidelines recommend beginning treatment for the olive fly when trapping numbers increase in early summer [34].

We consider variations on (12) that include either contemporaneous trapping numbers only or both contemporaneous and lagged trapping numbers (with a lag of up to four weeks). We refer to the specification with contemporaneous trapping numbers only as (12a) and that with both contemporaneous and lagged trapping numbers as (12b). We include specification (12b) to test the entomological hypothesis that population peaks are often followed by a lagged increase in infestation rates.

A final population-based specification, which we call an “augmented population-based” specification, includes the variables in W in addition to a measure of the pest population. The augmented population-based specification
\[ D = g(P, W) \]

constitutes an intermediate case between (12) and the detailed structural model. Specification (13) includes a direct measure of the pest population and the exogenous factors that affect the damage rate via host susceptibility, but excludes a direct measure of host susceptibility.\(^\text{10}\) Based on the previous discussion, by excluding a direct measure of fruit size, (13) potentially omits unobserved variables that may simultaneously affect fruit size and the pest population. The consequence is that (13) may produce inconsistent estimates of the damage function parameters. However, specification (13) is not intended to be the best available model of the system. Rather, the purpose of estimating (13) is to examine whether easily-obtainable data on weather and management practices can be combined with trapping data—which growers already rely upon—for use as a practical management tool.\(^\text{11}\)

**Estimation Methods**

Several unusual aspects of the empirical specification and the dataset influence the choice of estimation methodology. First, the empirical damage system is triangular in its coefficient structure. Because of possible omitted factors that simultaneously affect the dependent variables in (9) and (10'), the equations’ error terms could be correlated. Second, for the variable describing damage, a large probability mass exists at zero, which creates an estimation problem technically identical to that of censoring. Finally, the dataset includes observations from different trees across space and time, among which there could be unobserved but systematic differences in damage and fruit size.

An appropriate estimation methodology is one that allows for all of these characteristics. We follow Smith and Blundell [32] and combine Heckman’s method for efficient estimation of a triangular system with the Tobit estimation methodology.\(^\text{12}\) Conceptually, combining the Tobit
and Heckman methodologies involves adding a third equation to the empirical damage system. This combination produces a two-stage conditional maximum likelihood estimator (CMLE). Blundell and Smith [4] show that this CMLE performs well compared to a number of alternative estimators.\textsuperscript{13} We assume homoskedastic and normally-distributed errors, a necessity for the consistency of the Tobit estimator [20]. We test the validity of this assumption using a Hausman test based on the Tobit and Censored Least Absolute Deviation (CLAD) estimators [36].

Finally, we control for unobserved time-invariant heterogeneity among individual trees using fixed effects. We anticipate potential correlation between the unobserved individual-specific effects and the explanatory variables. For example, unobserved heterogeneity in damage due to sun exposure and growing degree-days may be correlated. In this case, using a fixed-effects specification ensures consistent parameter estimates.\textsuperscript{14} To capitalize on this advantage, we sacrifice inclusion of any time-invariant variables, such as dummies for the use of irrigation, ground cover, site location, and cultivar. The fixed-effects estimates capture the effect of these variables on damage levels in a single intercept shifter.

To facilitate the comparison of the population-based specifications with the structural damage function, we use the same estimation methodology: To estimate (12) and (13), we use the Tobit methodology and fixed effects but presume that the pest population is exogenous.\textsuperscript{15} We test these specifications against the structural damage function of (9) and (10') based on both goodness-of-fit and predictive ability. For fit, we employ a variety of test statistics to compare the nested population-based specifications against one another. To test the population-based against the structural damage specification, we use Vuong’s test for non-nested specifications [35]. To evaluate the models’ predictive ability, we use data from four trees that we exclude from the econometric analysis. We withhold observations on an oil olive tree in Amador County and a
table olive tree in each of Butte, Solano, and Ventura Counties. We select these trees and locations to represent oil and table olive cultivars at inland and coastal locations.

**Results**

Table II reports the estimation results for the structural damage system of (9) and (10'). The figures reported are the marginal effects for the censored dependent variable. The signs of the coefficient estimates for the fruit size equation (10') correspond with our expectations based on the scientific literature. This consistency lends credibility to the empirical analysis.

The coefficient estimates for the damage equation (9) in Table II indicate that olive size is indeed positively correlated with damage. However, the size effect diminishes late in the growing season. This diminution is likely due to the fact that each time a female fly stings an olive, she leaves behind a pheromone that deters another fly from ovipositing in the same olive [12]. Despite increased size, an olive is less attractive for oviposition when it is more saturated with stings late in the season.

As expected, an increase in the number of days that flies are inactive leads to a decrease in the observed rate of damage. As the number of inactive days increases or the growing season progresses, this effect diminishes. In contrast, the presence of irrigation intensifies the negative relationship between inactivity and damage: All else constant, an increase in pest activity translates into a greater increase in damage in the presence of irrigation.

As indicated by the coefficient for the inactivity-oil interaction term, an increase in activity leads to less of an increase in damage for an oil olive cultivar than for a canning cultivar, holding size constant. This result suggests that, even accounting for differences in size, oil cultivars experience less of an increase in damage from increased pest activity than do canning
cultivars. Therefore, there may be important host characteristics that differ between the two types of cultivars that are not captured by size.

The Hausman test rejects the null hypothesis that the error terms on equations (9) and (10') are uncorrelated, as indicated by the significance of the marginal effect for the fruit size residual in Table II. This evidence suggests that common unobserved factors drive simultaneous changes in fruit size and damage.16

We also test whether the time-invariant fixed factors that affect damage differ uniformly across trees within a site, or by trees of a single cultivar within a site. For the three sites from which more than one cultivar are sampled, we reject the null hypothesis that the fixed effects do not vary within the site. For a single cultivar within each site, we cannot reject the null hypothesis of identical fixed effects in 11 of 15 cases. This evidence suggests that time-invariant unobserved factors operate at the cultivar, rather than the site, level.

Table III presents results for the population-based specifications (12a), (12b), and (13). Estimates from all specifications indicate that contemporaneous trap numbers are significant and positively correlated with observed the level of damage. Estimates for (12b) indicate that the strength of the correlation between the number of trapped females and damage decreases over time. However, there is a small increase in the magnitude of the coefficient for trapped females three weeks prior to the current period. This result lends some support to the hypothesis that lagged pest numbers influence current damage rates.

For specification (13), the majority of the signs of the coefficient estimates for W coincide with those of the structural model. The use of trapping as a proxy for the pest population may circumvent the simultaneity described in (6)–(8). If trapping numbers are not a good measure of the population, use of the proxy may render (7) irrelevant, in which case
reducing the system to a single equation by combining (6) and (8) could yield consistent estimates of the damage function parameters.

Tables II and III provide quantitative descriptions of model fit. The quantitative measures include the maximized value of the log likelihood function, McFadden’s pseudo-$R^2$, Akaike’s information criterion (AIC), and the correlation coefficient for in-sample predicted and observed damage rates. These quantitative measures suggest that the augmented population-based model outperforms the population-based specifications (12a) and (12b). Vuong’s test rejects (12a) and (12b) in favor of the structural model, but cannot reject the null hypothesis that the structural model is no better than the augmented population-based model (13). Out-of-sample predictions show that the performance of the reduced-form (11) closely parallels that of (13). This suggests that the reduced-form may be useful as a practical management tool for olive producers. The advantage of (11) comes from its use of easily-accessible data on weather and management practices instead of trapping data, which are time-consuming to collect.

In Figure 1, we include predictions using specifications (12a), (13), and the structural damage specification. Across the four predictions, the structural model approximates infestation rates and their intra-seasonal fluctuations better than the population-based specifications. In particular, the structural model predicts variable damage over the course of a season, while the population-based specification (12a) projects a relatively constant level of damage. The augmented population-based specification of (13) does not exhibit variations in damage to the same degree as the structural model, although it does reflect increasing damage over the course of the growing season.

In all but one case among the four, the structural model mutes the degree of damage variability slightly. For Sevillano olives in Solano County, the structural model over-predicts
variation in infestation substantially. This difference illustrates that even the structural model
does not characterize the damage process perfectly. For example, although Sevillano olives are
the largest cultivar, they also have a high pit-to-pulp ratio that may hinder larval development.
This is only one example of a host characteristic that likely varies over time and affects the
female fly’s preference across cultivars. Future data collection that considers host characteristics
other than size may increase the accuracy of the structural model relative to observed infestation
rates. Even so, refinements to the structural model would only exaggerate the contrast in both fit
and predictive ability between the structural and population-based specifications.

Our earlier theoretical framework illustrates the way in which using the misspecified
population-based damage function introduces bias into estimates of the optimal use of the
damage-abating input x. To estimate the size of this bias for our application, we assume a simple
functional form for the grower’s production function, and solve analytically for x* as a function
of the baseline damage rate in each time period.18 We compare the predictions of each model
with the level of x* implied by observed damage rates over the course of the 2005 growing
season. In all cases, the structural and reduced-form models, which produce statistically identical
values for mean squared prediction error (MSPE), substantially outperform the population-based
specification. For canning olive cultivars, the structural and reduced forms produce an MSPE of
13.18 and 12.21, respectively; the population-based model has an MSPE of 31.43. The same
holds for oil olives, with a structural MSPE of 3.4, a reduced-form MSPE of 2.1, and a
population-based MSPE of 11.5. The direction of bias differs between cultivars, growing
regions, and weeks during the growing season.

Based on this simple illustration, even when no insecticide is used, the population-based
specification overstates the losses associated with the infestation by 2.6 percent while the
structural and reduced-form estimates are on par with observed losses. Though rough and sensitive to the limitations of the simple model including the choice of functional forms, this estimate supports our hypothesis that the models yield differing results in the case when a grower can do nothing to abate pest damage. This estimate constitutes a lower bound on the amount of bias introduced into the problem by the misspecified damage function. When the grower can adjust $x$, the bias increases. Logically, when the grower can adjust inputs on more than one margin ($x$ and $z$), the bias in estimated losses from the infestation increases further.\textsuperscript{19}

**Conclusion**

Overall, the estimation results highlight the importance of host size, weather, and management factors in explaining variable damage from olive fruit fly infestation over a growing season. The results across the structural and population-based models indicate that presuming a damage process to be driven solely by the presence of a pest population may miss the contribution that a host plays in determining its own susceptibility to infestation. A mischaracterization of the damage process may bias actions undertaken by economic agents to control pests.

By exploring the underlying factors that influence the pest population and the host, we uncover key variables that introduce simultaneity into the damage function, thereby identifying the structural damage model and the appropriate reduced form. We demonstrate empirically that these models outperform a population-based approach in terms of fit and intra-seasonal predictive ability. Furthermore, the structural and reduced-form models reduce bias in the use of productive inputs and the estimated losses from infestation. Whenever a population-based model is used to set pest management guidelines or policy, the associated bias may prove problematic. For example, insecticide use guidelines dependent only on pest population thresholds obtained through sampling will not consider the role of fruit size and the time of season in the optimal pest
management strategy, thereby encouraging inefficient treatment patterns. An example of a potentially problematic policy is a pesticide use regulation that restricts insecticide applications to a uniform amount per week, based on predictions of a relatively constant damage rate. The consequence of this type of policy would be to reduce the flexibility that growers have to address variable infestation rates.

Economists have long recognized that simultaneity complicates the estimation of causal relationships using non-experimental data [15,37]. This application draws on econometric methodologies to address simultaneity in processes governing the productivity of the host and damage by a pest. Entomologists and pomologists report the essential variables for inclusion in the structural damage function. They also provide the backing for several key assumptions, foremost of which is the identifying assumption of a lack of feedback from damage to host size. The benefits to addressing simultaneity in a damage function are not without cost in terms of added methodological complexity and data requirements, however. The value of the population-based specification lies in its use of trapping numbers, with which producers and researchers are familiar. Each case is probably different; the modeler of an integrated system needs to know enough about the physical processes involved to determine whether a likely misspecification alters conclusions about economic behavior. Whenever underlying factors create simultaneity in the damage function, the concerns raised in this analysis are relevant. Indeed, without careful scientific backing, the concerns should be conceded to be relevant.
References


[31] M. Smale, P. Zambrano, and M. Cartel, Bales and Balance: A Review of the Methods Used to Assess the Economic Impact of Bt Cotton on Farmers in Developing Economies,


Figure 1. Observed Damage and Out-of-Sample Predictions for Four Trees
<table>
<thead>
<tr>
<th>Variable (Abbr.)</th>
<th>Variable Description (Units)</th>
<th>Range</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Damage ($D$)</td>
<td>Proportion of olives with at least one sting</td>
<td>$[0,1]$</td>
<td>H.J. Burrack and F.G. Zalom, 2005 Fruit Development &amp; Infestation Data</td>
</tr>
<tr>
<td>Olive Size ($S$)</td>
<td>Mean olive volume (mm$^3$)</td>
<td>$[1.5,9006.9]$</td>
<td></td>
</tr>
<tr>
<td>Irrigation ($IR$)</td>
<td>Not irrigated/Irrigated</td>
<td>${0,1}$</td>
<td></td>
</tr>
<tr>
<td>Ground Cover ($GC$)</td>
<td>None/Grass</td>
<td>${0,1}$</td>
<td>Development &amp; Infestation Data</td>
</tr>
<tr>
<td>Cutivar ($OIL$)</td>
<td>Table/Oil</td>
<td>${0,1}$</td>
<td></td>
</tr>
<tr>
<td>Pests Trapped ($TR$)</td>
<td>Female flies trapped per week</td>
<td>$[0,129.6]$</td>
<td></td>
</tr>
<tr>
<td>Degree-Days ($CD$)</td>
<td>Accumulated growing degree-days for olive fruit</td>
<td>$[940.9,5220.7]$</td>
<td>California Irrigation Management Information System (CIMIS) [7], UC IPM Online Weather Database [34]</td>
</tr>
<tr>
<td>Adult Fly Inactivity Days ($AD$)</td>
<td>Accumulated days with temperatures outside 19-38 degrees Celsius</td>
<td>$[79,276]$</td>
<td></td>
</tr>
<tr>
<td>Late Season ($LT$)</td>
<td>Before/After August 1</td>
<td>${0,1}$</td>
<td>California Climate Data Archive [6]</td>
</tr>
<tr>
<td>Humidity ($HD$)</td>
<td>Relative humidity (percent)</td>
<td>$[23.2,90.4]$</td>
<td></td>
</tr>
<tr>
<td>Precipitation ($PR$)</td>
<td>Total precipitation (inches)</td>
<td>$[0,1.8]$</td>
<td></td>
</tr>
</tbody>
</table>
Table II. Marginal Effect Estimates for the Structural Damage Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Damage (9)</th>
<th>Fruit Size (10')</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growing Degree-Days (CD)</td>
<td>–</td>
<td>0.61 ***</td>
</tr>
<tr>
<td>Growing Degree-Days Squared (CD²)</td>
<td>–</td>
<td>–1.47 x 10⁻⁴ ***</td>
</tr>
<tr>
<td>Adult Inactivity Days (AD)</td>
<td>–2.17 x 10⁻² ***</td>
<td>–</td>
</tr>
<tr>
<td>Adult Inactivity Days Squared (AD²)</td>
<td>4.15 x 10⁻⁵ ***</td>
<td>–</td>
</tr>
<tr>
<td>Adult Inactivity Days (AD) x Late Season (LT)</td>
<td>2.03 x 10⁻³ ***</td>
<td>–</td>
</tr>
<tr>
<td>Humidity (HD)</td>
<td>–</td>
<td>–22.47 ***</td>
</tr>
<tr>
<td>Precipitation (PR)</td>
<td>–</td>
<td>48.13</td>
</tr>
<tr>
<td>Growing Degree-Days (CD) x Humidity (HD)</td>
<td>–</td>
<td>8.19 x 10⁻³ ***</td>
</tr>
<tr>
<td>CD/AD x Irrigation (IR)</td>
<td>–3.27 x 10⁻³ ***</td>
<td>0.58 ***</td>
</tr>
<tr>
<td>CD/AD x Ground Cover (GC)</td>
<td>–1.07 x 10⁻³</td>
<td>0.34 ***</td>
</tr>
<tr>
<td>CD/AD x Oil (OIL)</td>
<td>8.99 x 10⁻³ ***</td>
<td>–0.65 ***</td>
</tr>
<tr>
<td>Fruit Size (S)</td>
<td>1.06 x 10⁻³ ***</td>
<td>–</td>
</tr>
<tr>
<td>Fruit Size (S) x Late Season Indicator (LT)</td>
<td>–2.45 x 10⁻⁴ ***</td>
<td>–</td>
</tr>
<tr>
<td>Fruit Size Residuals</td>
<td>–5.51 x 10⁻⁴ ***</td>
<td>–</td>
</tr>
<tr>
<td>Constant</td>
<td>–</td>
<td>877.21 ***</td>
</tr>
</tbody>
</table>

Log Likelihood Value                  | –1038.11 | –                |
Adjusted or Pseudo-R²                 | 0.39      | 0.86             |
AIC                                      | 2174.21 | –                |
Correlation(ŷ,y)                        | 0.81      | –                |
N (Tree-Week)                           | 1033      | 1132             |
Censoring Rate (percent)                | 40.65     | –                |

*** denotes significance at the one percent level. All other coefficients are not significant at the ten percent level.
### Table III. Marginal Effect Estimates for the Population-Based Models

<table>
<thead>
<tr>
<th>Variable</th>
<th>Current (12a)</th>
<th>Lagged (12b)</th>
<th>Augmented (13)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trapped Females, t</td>
<td>8.53 x 10^{-3} ***</td>
<td>6.86 x 10^{-3} ***</td>
<td>6.87 x 10^{-3} ***</td>
</tr>
<tr>
<td>Trapped Females, t-1</td>
<td>–</td>
<td>6.46 x 10^{-4} ***</td>
<td>6.73 x 10^{-5}</td>
</tr>
<tr>
<td>Trapped Females, t-2</td>
<td>–</td>
<td>6.23 x 10^{-4} ***</td>
<td>2.24 x 10^{-4}</td>
</tr>
<tr>
<td>Trapped Females, t-3</td>
<td>–</td>
<td>6.57 x 10^{-4} ***</td>
<td>1.94 x 10^{-4}</td>
</tr>
<tr>
<td>Trapped Females, t-4</td>
<td>–</td>
<td>5.64 x 10^{-4} ***</td>
<td>4.06 x 10^{-5}</td>
</tr>
<tr>
<td>Growing Degree-Days (CD)</td>
<td>–</td>
<td>–</td>
<td>– 1.85 x 10^{-4}</td>
</tr>
<tr>
<td>Growing Degree-Days Squared (CD^2)</td>
<td>–</td>
<td>–</td>
<td>– 4.13 x 10^{-8} *</td>
</tr>
<tr>
<td>Humidity (HD)</td>
<td>–</td>
<td>–</td>
<td>– 0.03 ***</td>
</tr>
<tr>
<td>Precipitation (PR)</td>
<td>–</td>
<td>–</td>
<td>0.12 **</td>
</tr>
<tr>
<td>Growing Degree-Days (CD) x Humidity (HD)</td>
<td>–</td>
<td>–</td>
<td>8.11 x 10^{-6} ***</td>
</tr>
<tr>
<td>Growing Degree-Days (CD) x Irrigation (IR)</td>
<td>–</td>
<td>–</td>
<td>3.56 x 10^{-4} ***</td>
</tr>
<tr>
<td>Growing Degree-Days (CD) x Ground Cover (GC)</td>
<td>–</td>
<td>–</td>
<td>5.78 x 10^{-5}</td>
</tr>
<tr>
<td>Growing Degree-Days (CD) x Oil (OIL)</td>
<td>–</td>
<td>–</td>
<td>2.07 x 10^{-4} ***</td>
</tr>
<tr>
<td>Log Likelihood Value</td>
<td>–1254.28</td>
<td>–1229.46</td>
<td>–1039.28</td>
</tr>
<tr>
<td>Pseudo-R^2</td>
<td>0.26</td>
<td>0.27</td>
<td>0.38</td>
</tr>
<tr>
<td>AIC</td>
<td>2596.56</td>
<td>2554.91</td>
<td>2190.55</td>
</tr>
<tr>
<td>Correlation(ŷ, y)</td>
<td>0.58</td>
<td>0.60</td>
<td>0.79</td>
</tr>
<tr>
<td>N (Tree-Week)</td>
<td>1018</td>
<td>1014</td>
<td>1014</td>
</tr>
<tr>
<td>Censoring Rate (percent)</td>
<td>40.47</td>
<td>40.24</td>
<td>40.24</td>
</tr>
</tbody>
</table>

***, **, * denote significance at the one, five, and ten percent levels.
One potential issue identified by Norwood and Marra [21] is that excluding a measure of pest pressure from estimation of the damage abatement function may result in biased estimates of pesticide productivity.

The California olive industry consists of a canning and an oil sector. The canning sector accounts for the vast majority of the olives processed (96 percent). Quality categories are defined by U.S. Department of Agriculture regulation and depend on the mean and distribution of fruit sizes in a delivery. In 2008, Manzanillo and Mission olives received the following prices per ton by size category from canning processors: $350 for Sub-petite, $400 for Petite, $650 for Small, and $1210 for Medium/Large/Extra Large; Sevillano olives received $300 per ton for Extra Large ‘L’, $350 for Extra Large ‘C’, and $1050 for Jumbo/Colossal/Super Colossal olives. At present, California olive oil processors do not pay a premium based on fruit quality. In 2008, oil olives earned a flat price of $450 per ton [8].

We consider only pest-related damages, not yield damage from adverse growing conditions.

This specification excludes a direct measure of the pest population. This is an additional benefit whenever measuring pest populations is problematic, as is typical with mobile pests.


The data are unpublished, but are available upon request from the contact author.

The data were collected from seven orchards located in Amador, Butte, Solano, Sonoma, Tulare, Ventura, and Yolo counties. Across the seven sites, researchers collected oil and table olive cultivars, though effort centered on Manzanillo, Mission, and Sevillano fruit, which accounted for 95 percent of the olives produced in California in 2005 [33]. Sites from Glenn and
Tehama Counties (the second and third top olive producing counties) could not be included. During 2005, these two counties operated pest management districts which enforced treatment for the fly. Thus there were no untreated trees from which to sample.

8 We average by tree instead of using data on individual olives because the tree constitutes the smallest unit for which it is possible to construct a time series.

9 The panel dataset is unbalanced. Even so, we do not anticipate a sample selection problem. For each tree, the weekly sampling began prior to infestation and ran until the growing season ended.

10 Specification (13) excludes the vector V because of likely correlation between the elements of V and error arising from using trapping numbers (TR) as a proxy for the pest population (P). V includes variables based on a count of days during which temperatures fell outside of adult fly activity thresholds (AD), which is a measure of the incidence of extreme temperatures. There is evidence that extremely hot temperatures lead to a decline in the response of the fly to available trapping lures, which differs systematically from the decline in fly activity in response to the same stimulus [27]. We therefore expect including V in (13) to result in inconsistent estimates of the damage function parameters.

11 The reduced form in (11) may be more useful in practice as it depends only on observable and exogenous weather and management factors. Even so, growers may be reluctant to give up trapping data as an indicator of damage, in which case (13) may be a more palatable alternative.

12 The inclusion of humidity and precipitation in W but not in V prevents perfect collinearity when using the Heckman methodology.

13 We obtain estimates of standard errors with bootstrapping, as in [2].
We do not anticipate complications due to incorporating fixed effects into the CMLE. The dataset’s time series dimension is $T = 29$ and the number of nuisance parameters to be estimated is relatively small, so we anticipate negligible incidental-parameters bias \[14\].

Our population-based specifications are far more detailed than most in the literature. We keep the specifications as closely aligned as possible to facilitate their comparison. \[15\]

Hypothesis testing of a Tobit against a CLAD specification for the damage equation supports the null hypotheses of homoskedastic, normally-distributed errors in the damage system. \[16\]

Because they so closely parallel the results of the augmented population-based model in terms of both fit and predictive power, we refrain from reporting the results of the reduced form. \[17\]

The production function is given by $y = y_{\text{max}} [1 - D]$ where $D = D_0 [1 - G(x)]$. $D_0$ is the baseline level of damage as predicted by the empirical damage model under consideration and $G(x) = 1 - K^\lambda x^{-\lambda}$. The abatement function $G$ expresses the proportional reduction in damage due to an application of the insecticide $x$, and is assumed to be distributed Pareto, a functional form used in the literature [see 19], with parameter $\lambda = 3.5$. To simplify the profit function, we assume a constant price, $p$, of $350 per ton, a constant marginal cost of insecticide applications, $c_x$, of $6 per acre, and a maximum potential yield, $y_{\text{max}}$, of 5 tons per acre (parameters are consistent with raw olive prices, cost of production studies, and the estimated efficacy rate for GF-120 Naturalyte Bait, the predominant insecticide used against the olive fruit fly). \[18\]

This is found to be the case by K.M. Cobourn in “Incentives for Individual and Cooperative Management of a Mobile Pest: An Application to the Olive Fruit Fly in California,” unpublished Ph.D. Dissertation, University of California, Davis (2009). \[19\]