University of New England

Graduate School of Agricultural and Resource Economics

A Unified Approach to Sensitivity Analysis in Equilibrium Displacement Models: Comment

by

William Griffiths and Xueyan Zhao

No. 99-2 – March 1999

Working Paper Series in

Agricultural and Resource Economics

ISSN 1442 1909

 ISBN 1 86389 567 1

A Unified Approach to Sensitivity Analysis in Equilibrium Displacement Models: Comment

William Griffiths and Xueyan Zhao*

Abstract

It is pointed out that the Chebychev confidence intervals and maximum *p*-values advocated by Davis and Espinoza for sensitivity analysis of equilibrium displacement models are unnecessary. Desired probability intervals and probabilities can be accurately estimated without resorting to gross approximations.

Key Words: simulation, probability distributions, empirical quantiles.

* William Griffiths is Professor of Econometrics and Member of the Graduate School of Agricultural and Resource Economics, University of New England. Xueyan Zhao is a PhD student in the Graduate School of Agricultural and Resource Economics, University of New England. They thank Roley Piggott, Garry Griffith, John Mullen and George Davis for inspiration. Contact information: Graduate School of Agricultural and Resource Economics, University of New England, Armidale, NSW 2351, Australia. Email: wgriffit@metz.une.edu.au.

A Unified Approach to Sensitivity Analysis in

Equilibrium Displacement Models: Comment

In a recent paper in this *Journal*, Davis and Espinoza (DE) argue for a distributional approach to sensitivity analyses in equilibrium displacement models (EDMs). They advocate placing a subjective probability distribution on a parameter vector θ, and using simulation to estimate the corresponding probability distribution of an EDM outcome vector $Y^* = f(X^*, \theta)$, conditional on values of some exogenous shifters X^*. Our interest stems from a paper of ours (Zhao *et al*) which also argues for such an approach.

There are two aspects of the DE paper that we think are worthy of comment. The first is a conceptual matter. They use the term "posterior" distribution to describe the subjective probability distribution on Y^*, and refer to their procedure as a special case of Bayesian or mixed estimation. However, the essential ingredients in Bayesian estimation are a prior distribution, a likelihood function, and a posterior distribution. From Bayes' theorem, the posterior distribution is proportional to the product of the prior and the likelihood. In DE's analysis there is no likelihood function which is combined with the prior distribution on θ via Bayes' theorem to obtain a posterior distribution on θ. The distribution on Y^* is simply that which is induced by the distribution on θ, through the transformation $Y^* = f(X^*, \theta)$. Thus, in this sense, the distribution on Y^* is still a prior distribution not a posterior distribution.

Our second point is more substantive and has considerable relevance for applications. DE go on to say: "Because the actual distributions of the P^*_{xa}'s (Y^* in our notation) are not known, the distribution free Chebychev inequality is used to construct a minimum 95% confidence interval around each mean and develop maximum *p*-values." (p.875) We believe that the distribution for Y^* is known (or at least accurately estimated), and hence the Chebychev inequality is unnecessarily imprecise. We suggest alternatives to the "confidence intervals" and "*p*-values" calculated by DE.

Draws from the joint probability distribution for the elements in θ are used to compute draws or observations from the joint probability distribution for the elements in Y^*. The distribution of θ could be a posterior or a prior distribution, depending on whether sample data are available or not. In DE's example it is a prior distribution. A marginal distribution for one of the elements in Y^* is estimated by constructing a histogram of the draws of that element. The preciseness of the histogram as an estimate of the marginal probability distribution is controlled by the number of draws. We can make the histogram as accurate as we please by increasing the number of draws. In other words, the distribution of Y^* *is* known, although not analytically. The distribution of Y^* is implicitly specified by specifying the subjective distribution for θ, and the relationship between Y^* and θ. Let Y_i^* be the element of Y^* in which we are interested. DE are interested in finding an interval (a, b) such that $P(a < Y_i^* < b) = 0.95$. We prefer to use the term "probability interval", rather than "confidence interval", to describe intervals like (a, b). The term confidence interval has traditionally been used for inference about an unknown parameter in a repeated-sampling theory framework, not for intervals from subjective probability distributions. Irrespective of what we call it, an accurate estimate of the required interval can be found without using Chebychev's inequality. Sort the simulated observations on Y_i^* according to magnitude. Discard the first 2.5% and the last 2.5%.

The first and last values of the remaining ordered observations are the 0.025 and 0.975 empirical quantiles; they define the required interval. This is the approach taken by Zhao et al.

What are the consequences of using Chebychev's inequality instead? The minimum 95% probability intervals constructed by DE are likely to be much wider than necessary, and the probability content of their intervals is likely to be much greater than 0.95. Consider, for example, a normal random variable Z with a mean of zero and a variance of one. The precise 0.95 probability interval is (–1.96, 1.96). An estimate of this interval that we obtained by drawing 10,000 simulated observations is (–1.964, 1.957). The "minimum interval" obtained using Chebychev's inequality is (–4.472, 4.472). The probability content of the minimum interval is $P(-4.472 < Z < 4.472) = 0.999992$. As an example of a distribution that is not symmetric, consider a gamma random variable G with scale parameter equal to 1 and shape parameter equal to 2. This distribution has a mean of 2, a mode of 1 and a standard deviation of $\sqrt{2}$. A precise 0.95 probability interval, obtained by having a probability of 0.025 in each of the tails, is (0.242, 5.572). An estimate of this interval that we obtained by drawing 10,000 simulated observations is (0.244, 5.536). The minimum interval obtained using Chebychev's inequality is (–4.324, 8.324), which is not very helpful given the range of a gamma random variable is $(0, \infty)$. Also, $P(0 < G < 8.324) = 0.9977$.

DE calculate maximum p-values to test whether each of Gardner's three exogenous shifters will result in significant changes in the retail- farm price ratio, and they claim that these p-values illustrate the usefulness of the proposed approach. We believe that (i) the term "p-value" is an unfortunate one, (ii) probability statements about possible values for Y_i^* are, indeed, very important, and (iii) probability statements can be accurately estimated without the imprecision associated with

Chebychev's inequality. Consider point (ii) first. In empirical applications that utilize the EDM for aiding research evaluation and priority setting, it is often important to know whether a model outcome exceeds some break-even point for making a policy recommendation. For example, it may be important whether the total benefits from a research program exceed the cost of the research investment, whether the farmers' share of total benefits is smaller than their share of the levy for funding the research, or whether farmers have a bigger share of benefits from farm research than they do from processing research. Assuming certainty about the magnitudes of the research-related exogenous variables in X^*, the answers to these questions depend on the θ values used in the model. With conventional use of the EDM, fixed parameter values are often chosen based on the researcher's belief on the most likely values for the parameters, and conclusions are made regarding the above questions. However, what if the parameter values used are incorrect? Since we are often uncertain about the parameter values, there is a chance that we will make wrong decisions. The proposed stochastic approach in DE and Zhao *et al* quantifies the uncertainty in the parameters with a subjective distribution for θ; based on the consequent derived distribution for Y_i^*, it allows us to compute the probability of making a wrong decision. A natural way for presenting information on policy-related statements such as those above is to give the probability that a statement is true (or false). By doing so, we are placing a level of confidence on the policy-related conclusions from the EDM. The required probabilities are easily obtained by calculating the proportion of simulated observations for Y_i^* that fall in the relevant pre-defined region.

The *p*-values considered by DE are, we believe, a special case of the more general problem of finding the probability that Y_i^* falls in a given interval. The term

p-value is an unfortunate one because a repeated-sampling concept is being forced into a subjective probability framework. The marriage is not ideal, but we have a suggestion which we believe is a precise estimate of the p-value sought by DE. Suppose interest centers on the point hypothesis H_0: $Y_i^* = C$. One approach to testing this hypothesis against a two-tail alternative is to not reject H_0 if C lies within a 95% probability interval (constructed as described above), and to reject H_0 if C falls outside the interval. This approach is similar to what Zellner (Ch.10) describes as Lindley's procedure. The idea is that, if C falls within a 0.95 probability interval, then it is a reasonably likely value, and hence cannot be rejected. It also seems to be the idea DE is conveying. It is equivalent to rejecting H_0 if $P(Y_i^* > C) < 0.025$ or $P(Y_i^* < C) < 0.025$, and not rejecting H_0 otherwise. Now, as mentioned above, the probability that Y_i^* lies in any pre-specified interval is accurately estimated by the proportion of simulated draws that falls in that interval. Thus, probabilities such as $P(Y_i^* > C)$ and $P(Y_i^* < C)$ are easy to obtain. Indeed, we think that conveying information to policy makers in the form of probability statements such as these is preferable to conveying a relatively arbitrary decision to reject or not reject H_0. Both are related of course. The "p-value" that eluded DE and led them to consider a "maximum p-value" instead is given by $P(Y_i^* > C)$ when $E(Y_i^*) < C$ and by $P(Y_i^* < C)$ when $E(Y_i^*) > C$. (Strictly speaking, this comparison should be between C and the median of Y_i^*, not the mean. DE give the conditions as $E(Y_i^*) < 0$ and $E(Y_i^*) > 0$, respectively; they are implicitly assuming $C = 0$.) Suppose that $E(Y_i^*) < C$ and denote the mean and standard deviation of Y_i^* by $\mu = E(Y_i^*)$ and σ, respectively. To demonstrate that DE's maximum p-value is an upper bound for $P(Y_i^* > C)$, we can write:

$$p\text{-value} = P(Y_i^* > C) = P[(Y_i^* - \mu) > (C - \mu)].$$

From Chebychev's inequality,

$$P[\,|Y_i^* - \mu| > k\sigma] \leq \frac{1}{k^2},$$

and hence

$$P[(Y_i^* - \mu) > k\sigma] \leq \frac{1}{k^2}.$$

Setting $k = (C - \mu)/\sigma$ yields:

$$p\text{-value} = P[(Y_i^* - \mu) > (C - \mu)] \leq \frac{\sigma^2}{(C-\mu)^2}.$$

The maximum p-values computed by DE are given by $\sigma^2/(C - \mu)^2$. Because these maximum values can be very much greater than the actual values, they are of limited usefulness. Consider the first of the two earlier examples where $Z \sim N(0, 1)$, and let $C = 2$. In this case $P(Z > 2) = 0.023$. An estimate of this quantity that we obtained using 10,000 simulated draws is 0.022. The corresponding maximum p-value is 0.25. Consider the second example where G is gamma random variable with scale parameter 1 and shape parameter 2. Let $C = 3$. In this case $P(G > 3) = 0.199$, an estimate from 10,000 simulated draws is 0.200, and the maximum p-value is 2. The impreciseness of Chebychev's inequality also led DE to obtain several maximum p-values that were greater than one.

It is useful to quantify the accuracy with which a probability can be estimated. The estimates we obtained from 10,000 draws were very close to the true probabilities. However, can we be sure that such estimates will always be this accurate? The standard deviation of an estimated proportion or probability is $[p(1-p)/n]^{1/2}$ where p is the unknown probability and n is the number of draws. The maximum value of this standard deviation over all possible values for p is

$0.5/n^{1/2}$. Furthermore, for large n, an estimated proportion $\hat{p}$ is normally distributed. Thus, for 10,000 simulated draws, there is at least a 0.95 probability that the estimation error will be less than $1.96 \times 0.5/(10000)^{1/2} = 0.0098$. With 50,625 draws, the number used by DE, there is at least a 0.95 probability that the estimation error will be less than 0.0044.

To further demonstrate (1) the effect of using empirical quantiles and estimated proportions instead of Chebychev intervals and maximum p-values, and (2) the accuracy with which characteristics of a distribution can be estimated when a large number of simulated observations are available, we attempted to replicate DE's illustration from Gardner's model. The results appear in Table 1. In the last 2 columns we compare the 95% probability intervals obtained using empirical quantiles with DE's "95% confidence intervals", and the estimated proportions with DE's "maximum p-values". Standard errors for the estimated characteristics are provided. (See Mood, Graybill and Boes, p.257 for the standard error of an empirical quantile.) The simulation was programmed in SHAZAM code, using algorithms found in Rubinstein for generating beta and gamma random variables.

We successfully reproduced DE's results for their scenario I, but were unable to do so for their scenario II. After extensive checking and experimentation, we were led to conjecture that the scenario II discrepancies are caused by at least one unlikely parameter value obtained by DE. The distribution for the input substitution elasticity σ in scenario II is truncated normal defined over (0, 1) with location $\mu = 0.74$ and scale $\psi = 0.12247$. The smallest draw from this distribution reported by DE was 0.18. The probability of obtaining a draw this small or smaller is 0.245×10^{-5}. The expected number of draws this small or smaller in 50,625 draws is 0.12; the probability of getting one or more draws less than or equal to 0.18 in 50,625 draws is 0.13. Thus, a value of σ = 0.18 is possible, but it is unlikely, and it did not happen in our simulation. (Our minimum was 0.27.) Furthermore, our experimentation suggested that smaller values of σ tend to lead to values of $P^*_{xa,N}$, $P^*_{xa,W}$ and $P^*_{xa,T}$ which are larger in absolute value, a result which is consistent with the observed discrepancies. The larger variances obtained by DE are also consistent with their larger range for σ.

Although our different results warrant an explanation, they do not alter the points we are making. In Table 1 we have reported our results to more decimal places than those reported by DE so that the standard errors are meaningful. The small standard errors show how accurately the characteristics of a simulated distribution can be estimated. In particular, the estimated end points of the 95% probability intervals are very accurate indeed, and the intervals are much narrower than those based on Chebychev bounds. The estimated proportions are also much more useful than the maximum *p*-values. For $P^*_{xa,N}$ the maximum *p*-values for both scenarios are one when estimates of the desired probabilities are both 0.279, with standard errors of 0.002. That is, for a retail demand shift, there is a 28% chance that the retail-farm price ratio will increase, and a 72% chance it will decrease. Furthermore, there is a 95% chance that the percentage change in the retail-farm price ratio will be between – 47% to 23%. In the sense that zero is a reasonably likely value of $P^*_{xa,N}$, the hypothesis $P^*_{xa,N} = 0$ can not be rejected. For $P^*_{xa,W}$ and $P^*_{xa,T}$, the prior distributions on the parameters are such that zero is not a feasible value. We know, *a priori*, that $P(P^*_{xa,W} \geq 0) = P(P^*_{xa,T} \leq 0) = 0$. That is, the retail-farm price ratio will always decrease from a farm supply shift and it will always increase from a marketing supply shift. Thus, *p*-values greater than zero are not meaningful, and the required probabilities are known without estimation error. As zero is not even a possible value, the hypotheses $P^*_{xa,W} = 0$ and $P^*_{xa,T} = 0$ will be rejected.

Like DE, we believe that using subjective distributions greatly enhances applications of the equilibrium displacement model. The contribution of this comment is to point out that, in such analyses, the very rough probability approximations obtained from Chebychev's inequality are unnecessary. Precise probability estimates can be found easily from the relevant proportions of simulated draws. A number of extensions to DE's analysis are made in Zhao *et al*. Hierarchical distributions are suggested as a way of representing potentially diverse views about parameters. Also, an average sensitivity elasticity is defined and used as a measure of the relative importance of the individual parameters in θ.

References

Davis, G. and M.C. Espinoza. "A Unified Approach to Sensitivity Analysis in Equilibrium Displacement Models." *Amer.J.Agr.Econ*. 80 (November 1998): 868-79.

Gardner, B.L. “The Farm-Retail Price Spread in a Competitive Food Industry.” *Amer.J.Agr.Econ*. 57 (August 1975): 399-409.

Mood, A.M., F.A. Graybill and D.C. Boes. *Introduction to the Theory of Statistics*, 3rd ed. Tokyo: McGraw-Hill, 1974.

Rubinstein, R.Y. *Simulation and the Monte Carlo Method*. New York: John Wiley and Sons, 1981.

SHAZAM. *User’s Reference Manual Version 8.0*. New York: McGrow-Hill, 1997.

Zellner, A. *An Introduction to Bayesian Inference in Econometrics*. New York: John Wiley and Sons, 1971.

Zhao, X., W.E. Griffiths, G.R. Griffith and J.D. Mullen. "Probability Distributions for Economic Surplus Changes." Unpublished manuscript, University of New England, Armidale, Australia, May 1998.

Table 1. A Comparison of Summary Statistics, Probability Intervals and *p*-values

Price Change	Scenario[a]	Variance	Mean	Median	Mode	95% Probability Interval[b]		$P(P_j \neq 0)$ [c]
$P^*_{xa,N}$	I(DE)	.10	–.14	–.10	–.05	–1.58	1.30	1.0
	I(GZ)	.102	–.138	–.096	–.05	–0.887	0.394	0.279
		(.002)[d]	(.001)	(.001)		(0.008)	(0.006)	(0.002)
	II(DE)	.05	–.10	–.08	–.05	–1.07	0.87	1.0
	II(GZ)	.0296	–.088	–.0742	–.045	–0.468	0.243	0.279
		(.0003)	(.001)	(.0008)		(0.003)	(0.003)	(0.002)
$P^*_{xa,W}$	I(DE)	.006	–.46	–.48	–.49	–0.80	–0.13	0.03
	I(GZ)	.00541	–.4625	–.4756	–.497	–0.589	–0.273	0.0
		(.00006)	(.0003)	(.0002)		(0.001)	(0.002)	(0.0)
	II(DE)	.008	–.37	–.39	–.41	–.077	0.03	0.06
	II(GZ)	.00623	–.3353	–.3525	–.380	–0.4432	–0.138	0.0
		(.00005)	(.0004)	(.0004)		(0.0004)	(0.001)	(0.0)
$P^*_{xa,T}$	I(DE)	.005	.53	.52	.50	0.20	0.85	0.02
	I(GZ)	.00513	.5256	.5155	.502	0.395	0.702	0.0
		(.00006)	(.0003)	(.0002)		(0.001)	(0.002)	(0.0)
	II(DE)	.007	.42	.43	.44	0.06	0.78	0.04
	II(GZ)	.00399	.3764	.3890	.408	0.217	0.4645	0.0
		(.00004)	(.0003)	(.0003)		(0.001)	(0.0003)	(0.0)

Notes:

[a] DE refers to the values obtained by Davis and Espinoza. GZ refers to the values that we calculated.

[b] These values are the Chebychev Bounds in the DE rows and the .025 and .975 empirical quantiles in the GZ rows.

[c] These values are upper bounds in the DE rows and sample proportions in the GZ rows. For P_N and P_W the relevant probability is $P(P_j > 0)$. For P_T it is $P(P_j < 0)$.

[d] The numbers in parentheses are asymptotic standard errors.