MAIL CODE: EA, EG & ET

WORKING PAPERS

THE IMPRECISION OF TRADITIONAL WELFARE MEASURES IN EMPIRICAL APPLICATIONS

Frederick I. Johnson

WORKING PAPER NO. 86

June 1983

FTC Bureau of Economics working papers are preliminary materials circulated to stimulate discussion and critical comment. All data contained in them are in the public domain. This includes information obtained by the Commission which has become part of public record. The analyses and conclusions set forth are those of the authors and do not necessarily reflect the views of other members of the Bureau of Economics, other Commission staff, or the Commission itself. Upon request, single copies of the paper will be provided. References in publications to FTC Bureau of Economics working papers by FTC economists (other than acknowledgement by a writer that he has access to such unpublished materials) should be cleared with the author to protect the tentative character of these papers.

BUREAU OF ECONOMICS
FEDERAL TRADE COMMISSION
WASHINGTON, DC 20580

IMPRECISE WELFARE MEASURES

Frederick Ian Johnson

March 1983

I gratefully acknowledge the assistance of Dan Alger, whose many comments have been most helpful.

Introduction

We are frequently interested in measuring the welfare impacts of a change in the price of a good. Several measures, particularly consumer's surplus (CS) and, to a lesser degree, compensating variation (CV), are commonly used to estimate the change in welfare. While these measures have been exhaustively discussed and debated in the literature, the primary emphasis has been their theoretical validity. Here our principal concern is their accuracy when they are estimated empirically. We are most interested in the error in our measurements of CS and CV when these measures are based upon estimates of the relevant supply and demand estimates, estimates that in practice are often quite imprecise. In particular, how broad is the distribution of these welfare measures likely to be, and how large are the differences between these two measures relative to the error in estimating them?

In this note we explore the magnitude of empirical errors in CS and CV through a series of excise tax examples. We also compare the differences between CS and CV in relation to the measurement errors, and explore the circumstances in which deadweight loss (DWL) measured by CS differs importantly from DWL measured by CV. Because our analysis consists primarily of a series of (what we hope are typical) examples, any conclusions are intended to be suggestive rather than definitive. For the situations we consider, the following conclusions concerning CS, CV and DWL emerge:

best case: measuring the cost to the consumer of an excise tax, when supply is perfectly elastic and current equilibrium is known. For most purposes, the estimates of

CS and CV seem quite precise, and the difference between the two is small relative to their standard errors;

next best: measuring the cost to the consumer of an excise tax, when supply is upward sloping and current equilibrium is known. The estimates of CS and CV are not very precise, but the difference between the two again is small in relation to their standard errors;

worst case: measuring total benefit of a planned project, with no *a priori* constraints on demand. For most purposes, standard errors seem very large, possibly too large for any meaningful inference of benefit;

deadweight loss: the divergence between DWL_{CS} and DWL_{CV} may appear substantial, yet be swamped by errors in estimation.

Theoretical Background

Since its debut in 1841, consumer's surplus has remained a hotly disputed topic. 1/ Excoriated by theoreticians, it is nevertheless widely used by empirical economists. Two recent articles, one by Robert Willig (1976) and one by Jerry Hausman (1981), reexamine consumer's surplus and its usefulness as a welfare measure. Willig shows that, under a wide range of circumstances, consumer's surplus does not differ substantially from either compensating variation or equivalent variation, and suggests that any error will be small compared to the error in empirical estimation. 2/ Hausman points out, however, that economists are most frequently interested in measuring the net

---

1/ For a comprehensive review of the literature, see Currie, Murphy and Schmitz (1971) and Chipman and Moore (1976,1978).

2/ Robert Archibald and Robert Gillingham, in a recent study of energy demand, use Willig's technique to show a minuscule difference between CS and CV. "The Distributional Impact of Alternative Gasoline Conservation Policies," *Bell Journal of Economics*, Autumn, 1981.

change in social welfare, "dead weight loss", rather than the gross change in consumer's welfare. Errors in consumer's surplus may be small, yet large in relation to dead weight loss.

The nature of the problem is illustrated in Figure 1. (The ordinary and compensated demand curves are drawn to scale for a rather extreme case, Hausman's gasoline example, which is discussed below.) The area A+B represents CS, while the area A+B+C represents CV. If C is small relative to A+B, then CS is a fair approximation of CV. Dead weight loss measured by CS (call it DWL_{CS}) will be the area B, while DWL_{CV} will be the area B+C. It follows that:

$$(1)\quad (DWL_{CV}-DWL_{CS})/DWL_{CV} = ((CV-CS)/CV)\ (1+(A/(B+C))) > (CV-CS)/CV$$

Consequently, in percentage terms CS will approximate CV more closely than DWL_{CS} approximates DWL_{CV}.

FIGURE 1

In spite of the theoretical differences, the empirical importance of these differences may be acceptably small, since typically neither CS nor CV can be measured precisely. Both welfare measures are computed from <u>estimates</u> of the ordinary and compensated demand curves. To illustrate the empirical importance of these differences, suppose that an excise tax is imposed. In an extreme case where the supply is perfectly elastic (for example, small country in a large world), then it is known that the tax is fully borne by the consumer, and imprecision in the estimates of CS and CV stems entirely from uncertainty about the change in the quantity demanded. More generally, the supply curve slopes upward. This introduces a second source of imprecision into the estimates of CS and CV, namely, the size of the net change in price to the consumer. (This point is illustrated in Figure 2: the final price to the consumer depends on the elasticities of supply and demand.) In either case, if the value of the initial equilibrium is known, the only source of error in the estimates of CS and CV is error in the estimates of the slopes of the supply and demand curves. If the initial equilibrium is not known, then estimates of the intercept term must also be factored into the computatuion of CS and CV, increasing the error in these measures.

FIGURE 2

Procedure

We consider a linear supply-demand model for which we have (hypothetical) estimates of the coefficients and their standard errors.3/ We assume that the estimated coefficients are distributed normally, and are mutually independent. Price and quantity at the initial equilibrium are each normalized to equal 1.0 . The supply and demand curves are known to pass through the point (1, 1) (this is just saying that we know the initial

---

3/ We assume linearity because it simplifies the computations. In the "small" excise tax case, an alternative choice of functional form would not substantially alter the simulations. However, when measuring total benefit, functional form is important. (See Ziemer, Musser and Hill, "Recreation Demand Equations: Functional Form and Consumer Surplus," American Journal of Agricultural Economics 62 (1) February 1980).

equilibrium), so the only uncertainty is in the price and income coefficients, of which we have only estimates.

Suppose that the (ordinary) demand curve has been estimated, and is of the following form: 4/

(2) $\Delta \hat{Q} = \hat{\beta}\, \Delta P + \hat{\gamma}\, \Delta I$

where Q, P, and I are, respectively, quantity, price, and income. From the Slutsky equation, we have a relation between the elasticities of the ordinary and compensated demand curves:

(3) $e_c = e_o + a \cdot e_i$

e_c : elasticity of the compensated demand curve
e_o : price elasticity of the ordinary demand curve
e_i : income elasticity of demand
a : proportion of income spent on the good

Since the initial equilibrium was normalized, point values of e_c and e_o equal the slopes of the corresponding demand curves, and $a \cdot e_i$ equals the income coefficient. CS and CV can then be computed as follows:

(4) $CS = \Delta P + 0.5\hat{\beta}\, \Delta P^2$

(5) $CV = CS + 0.5\hat{\gamma}\, \Delta P^2$

and dead weight loss under these measures is:

4/ The ratio is sometimes called the "coefficient of variation". However, we avoid using the term since it could too easily be confused with "compensating variation", and more importantly because "coefficient of variation" is often defined differently.

(6) $DWL_{CS} = 0.5\hat{\beta}\,\Delta P^2$

(7) $DWL_{CV} = DWL_{CS} - 0.5\hat{\gamma}\,\Delta P^2$

The moments of CS, CV, DWL_{CS} and DWL_{CV} depend on the moments of the estimated coefficients. We denote the ratio of the standard error of a random variable to its mean by δ, and assume that all the estimated coefficients have the same values of δ. 5/ The reciprocal of δ can be interpreted as a t-statistic, 6/ so $\delta = 0$ implies no error, while $\delta = 0.5$ corresponds to a t-statistic of about 2.0 . Now, through two examples, one extreme case with perfectly elastic supply and one case with upward sloping supply, we examine how the size of δ, together with with the supply elasticity, affects the precision of the estimates of the welfare measures when a "small" excise tax is imposed. Following the two excise-tax examples, we briefly examine the precision of CS and CV as measures of the total benefit of a project.

Perfectly Elastic Supply

As noted above, when supply is perfectly elastic the price to the consumer increases by the full amount of the tax, so the only source of error in estimates of CS and CV is uncertainty

---

5/ We express demand in "delta" form to avoid introducing other variables in the demand function which, for the purposes of this paper, can be treated as constants.

6/ Strictly speaking, δ is drawn from a normal distribution, while a t-statistic is drawn from a t-distribution. Therefore $1/\delta$ can be viewed as a t-statistic with an infinite number of degrees of freedom.

about the change in quantity demanded. We can compute the magnitude of this error directly. Since CS and CV are linear functions of $\hat{\beta}$ and $\hat{\gamma}$, which by assumption are normal variables, the estimates of CS and CV will also be normal. Table 1 shows the means and confidence bounds on CS and CV when a tax of 5% of the initial price is imposed. The demand elasticities in the first column are chosen to roughly correspond to Hausman's gasoline example (p.673), while those in the second column are intended to represent a more typical good, perhaps fruit. Note that for most purposes CS serves as a good proxy for CS here (the two differ by only 0.1%), and both can be estimated with a high degree of precision (the standard errors are only 0.2% to 0.3% of the mean).

TABLE 1

We noted above that the percentage difference between DWL_{CV} and DWL_{CS} will be greater -- and sometimes much greater -- than the percentage difference between CV and CS. While there may be only a slight difference between CS and CV, all of the difference will be concentrated in the measurement of DWL, as illustrated by equations (4) through (7). By the same token, the measurement error will also be concentrated in DWL. Table 1 shows mean values and standard errors for DWL_{CS} and DWL_{CV}. Note that the ratio of the standard error to the mean is exactly the same for DWL_{CS} and , as can be inferred from equation 6, and somewhat greater for DWL_{CV}. In the "fruit" example, the difference between the means for the estimates of DWL_{CS} and DWL_{CV} is small (about 0.1%), while in the gasoline example, the estimate of DWL_{CV} exceeds that for DWL_{CS} by 33%. Nevertheless, this difference is not particularly significant relative to the standard errors of both measures of DWL, as is shown in Figure 3. For δ =0.3, the two means differ by 0.81 s.e. (where s.e. is the standard error of DWL_{CV}), while for δ =0.5, they differ by only 0.25 s.e. Also, the two are not significantly different (at a 5% significance level). 7/ In short, the difference between the two means is swamped by the error in estimation.

FIGURE 3

Upward Sloping Supply

If supply is upward sloping, the net change in price to the consumer, ΔP, is also a random variable. Consequently, CS and CV will not in general follow any known distribution. Rather than provide an analytical solution for this case, we simulate the distribution of these two measures. We assume that the point elasticities of supply and demand are 2.0 and -1.0 respectively. We select a value for δ , then randomly generate values for the slope coefficients, compute the corresponding values of CV and CS, and determine the means and 95% confidence bounds. 8/

First, how do CS and CV differ? From the Slutsky equation

7/ Hausman shows a 2 percent difference between CS and CV, while our table shows only 0.1 percent. This is because the relative difference of the two terms increases with the change in price (see Figure 1): we consider a 5 percent tax against Hausman's 100 percent. Note that the ratio of DWL_{CS} to DWL_{CV} will remain largely unaffected.

8/ We simulated each 1000 times, and derived the bounds by discarding the top and bottom 25.

(equation 3), it is clear that the larger $a \cdot e_i$ relative to e_0, the greater the discrepancy between the ordinary and compensated demand curves, and hence also between CS and CV. However, our simulations did not show any significant difference between CS and CV even for large values of $a \cdot e_i$. With $a \cdot e_i = 0.5$ (this value would imply both a large fraction spent on the good and a high income elasticity) and $\delta = 0.3$, CV exceeded CS by only about 3%, while the confidence bounds on CS and CV were $\pm$ 30% of the means.

Second, how does δ affect the estimates of CS and CV? Figure 4 shows values of CS and CV as δ increases from 0 to 0.5. As δ increases, the dispersions of CS and CV surge. With $\delta = 0.5$, the confidence bounds on CS and CV are [0.006, 0.050], that is, from 20% to 170% of the mean. These bounds contrast dramatically with confidence bounds of [0.0496, 0.0501] -- 99.5% to 100.5% of the mean -- when supply is perfectly elastic (computed from Table 1).

FIGURE 4

More surprising is the fact that, as the standard errors of the estimates increase, the <u>means</u> of CS and CV change. If the standard errors are relatively small, then the means of CS and CV approximately equal their point estimates. However, the distributions are skewed so, as the standard errors increase, the means diverge farther from the point estimates.

Measuring Total Benefit

Suppose that, instead of a <u>change</u> in welfare, we wish to measure <u>total</u> welfare. Such a situation might arise in evaluating a proposed public project: benefits are often approximated by CS. Again, how precise is CS, and how good a proxy is it for CV?

We use the same model as before, but now measure the benefits from providing 1 unit of the good. CS and CV are measured by the areas under the corresponding demand curves from Q=0 to Q=1. 9/ The first part of Table 2 provides estimates of both CS and CV when it is known that the demand curve passes through the point (1,1). In this case, the only errors entering into CS and CV result from errors in the price and income coefficients. Three things stand out. First, there is negligible difference between the estimates of CS and CV either at the means (0.1% to 7%, for the δ's chosen) or at the bounds (less than 3% for the δ's chosen). Second, the confidence intervals on each are quite broad (with δ = 0.5, the 95% confidence region is from 67% to 220% of the mean). And third, the distributions are asymmetric, as reflected by the change in the means of CS and CV as δ changes.

TABLE 2

A final variant is considered for comparison. If demand is no longer constrained to pass through (1, 1) a second source of error enters into the estimate of CS, error in the intercept

---

9/ That is, we measure total benefit by the area under the demand curve and above the horizontal axis: this area would include both revenues (if any charge) and economic rents. Whether this area represents CS, or CS plus revenues, depends on what price is charged, if any. This is, of course, a semantic rather than a substantive issue.

term. 10/ The effect of this additional error is shown in the second part of Table 2. For δ =0.3 the confidence interval on CS is [0.30, 4.00], which is from 15% to 240% of the mean. For δ=0.5, it becomes [-1.25, 10.93], which is -60% to +515% of the mean. In short, when there are no constraints on the demand curve, even reasonably good demand estimates generate a ridiculously broad interval for CS.

Concluding Remarks

In closing, two points are worth noting. First, we have tried to estimate CS and CV on their most favorable terms. In particular, we considered a "small" tax, and we emphasized situations where the initial equilibrium was known. Second, while we have examined only the cost to the consumer, estimates of producer's surplus (PS) follow the same pattern as do estimates of CS. However, note that the dispersion of the sum of CS and PS will be smaller than the dispersion of the two terms individually. The explanation is that, while we may not know the exact change in the price to the consumer or to the producer, we know that the sum of the changes is exactly equal to the amount of the tax.

We could continue to simulate other situations, but we feel our basic point stands: errors in supply and demand estimates beget yet greater errors in welfare estimates. Willig's point is well taken, and his paper has led to a wider acceptance of CS.

10/ Point estimates are chosen to generate the same demand curve.

However, because of the errors that arise in computing these measures, one should be cautious about relying heavily on these measures when evaluating policies.

REFERENCES

Chipman, John S. and Moore, James C., (1976) "The Scope of Consumers Surplus Arguments" in Anthony W. Tang, Editor, Evolution, Welfare, and Time in Economics: Essays in Honor of Nicholas Georgesco - Rosen, Lexington Books 1976 pp 69-123.

-----------, (1978) "Compensating Variation, Consumer's Surplus and Welfare" American Economic Review, December 1980, pp 933-949.

Currie, John M.; Murphy, John A.; and Schmitz, Andrew (1971) "The Concept of Economic Surplus and its Use in Economic Analysis" The Economic Journal (December 1971)

Hausman, Jerry A. (1981) "Exact Consumer's Surplus and Deadweight Loss", American Economic Review, September 1981 pp. 662-676

Willig, Robert D. (1976) "Consumer's Surplus Without Apology" American Economic Review 66(4), (September 1976) pp.

TABLE 1

Means and Standard Errors of CS, CV, DWL_{cs} and DWL_{cv} where supply is perfectly elastic

	"gasoline"			"fruit"		
	$\varepsilon_o = -0.2$, $a \cdot \varepsilon_i = 0.05$			$\varepsilon_o = -1.0$, $a \cdot \varepsilon_i = 0.01$		
	mean	standard error/mean		mean	standard error/mean	
		$\delta = 0.3$	$\delta = 0.5$		$\delta = 0.3$	$\delta = 0.5$
CS	0.04975	0.00151	0.00251	0.04875	0.0077	0.0128
CV	0.04981	0.00155	0.00259	0.04876	0.0077	0.0128
DWL_{cs}	0.00025	0.3000	0.5000	0.00125	0.300	0.500
DWL_{cv}	0.0001875	0.4123	0.6880	0.00124	0.302	0.504
(CV-CS)/CV		0.1%			< 0.1%	
$(DWL_{cv}-DWL_{cs})/DWL$		33.3%			0.1%	

Price and quantity are each initially equal to 1.0

tax = 0.05 (5%)

TABLE 2: BOUNDS ON CS AND CV, TOTAL BENEFIT

CONSTRAINED TO PASS THROUGH (1,1)

δ	CS			CV[a]		
	LOWER	UPPER	MEAN	LOWER	UPPER	MEAN
0.0			1.500			1.502
0.3	1.311	2.156	1.564	1.311	2.163	1.567
0.5	1.227	3.941	1.705	1.227	4.042	1.827

UNCONSTRAINED: POINT ESTIMATE PASSES THROUGH (1,1)

δ	CS		
	LOWER	UPPER	MEAN
0.0			1.500
0.3	0.298	4.000	1.706
0.5	-1.245	10.927	2.123

(For CV to be well-defined we must know with certainty one point on the curve; i.e. the initial (price, quantity) combination. In the unconstrained case we presume no such knowledge.)

a. Income elasticity is 0.5, expenditures on good are 1% of income.

FIGURE 1

AREAS TO BE MEASURED FOR CS, CV AND DWL

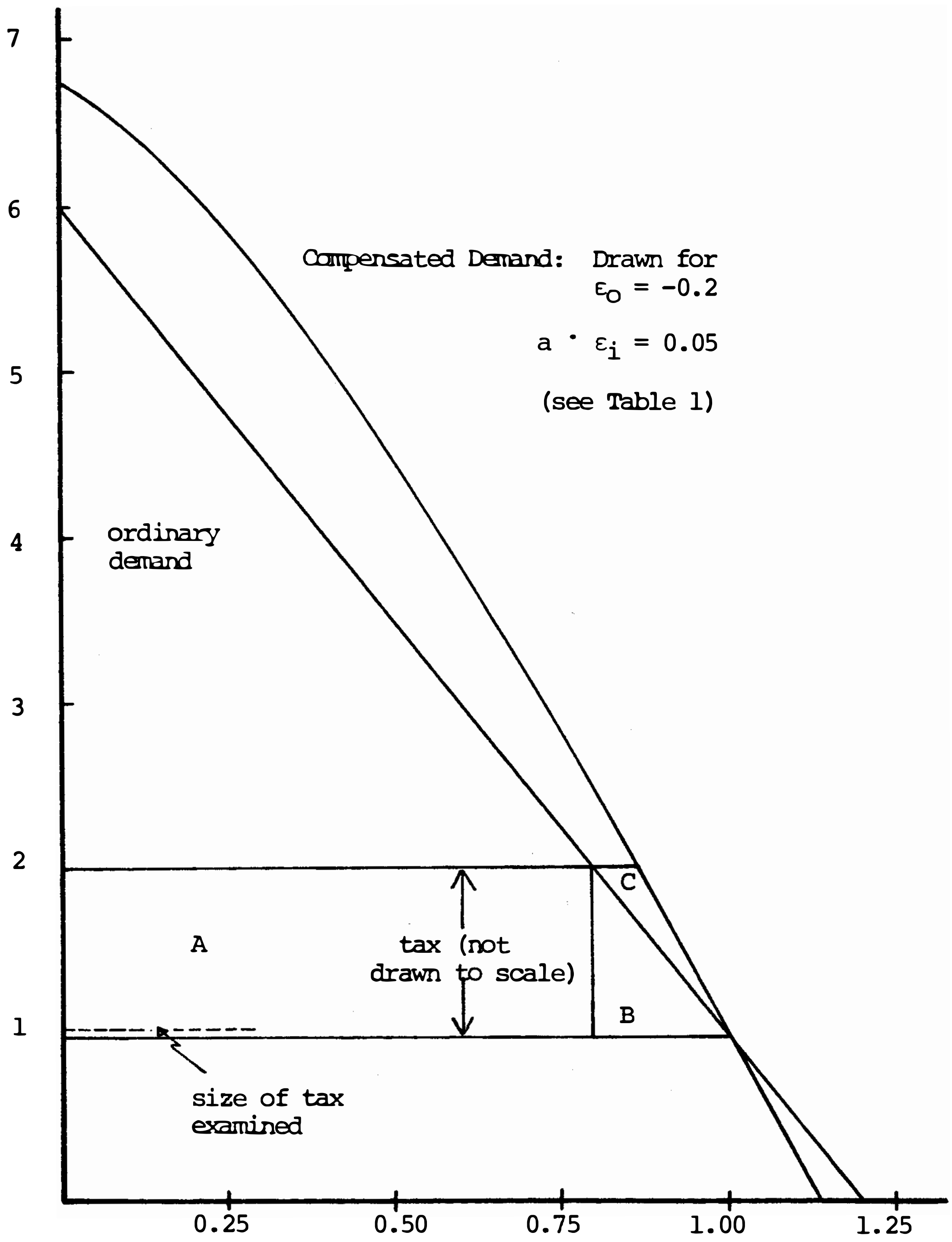


FIGURE 2

With upward sloping supply, the change in the price to the consumer depends on the demand curve. As pictured, a tax of \$0.75 could raise consumer price by \$0.25 (D_A) or by \$0.50 (D_B).

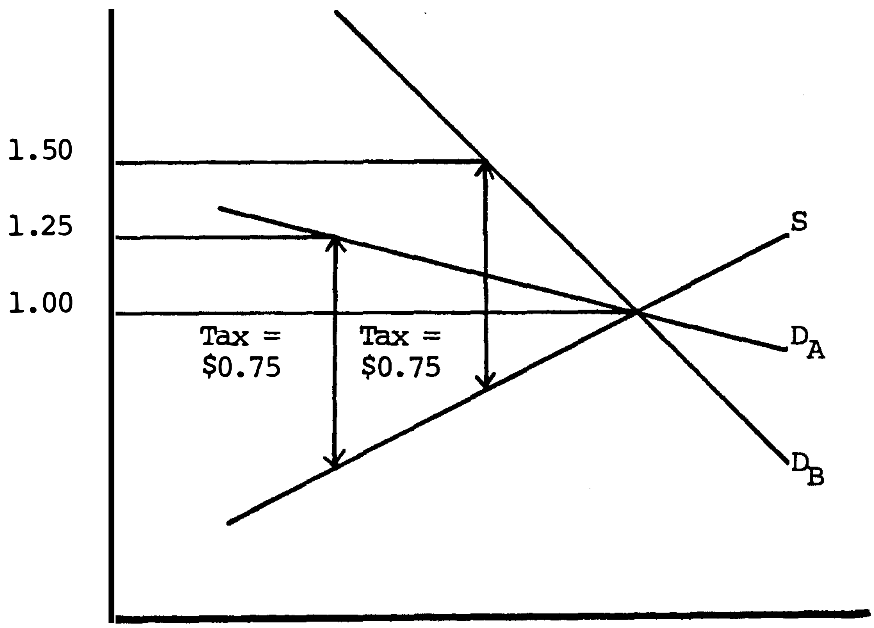
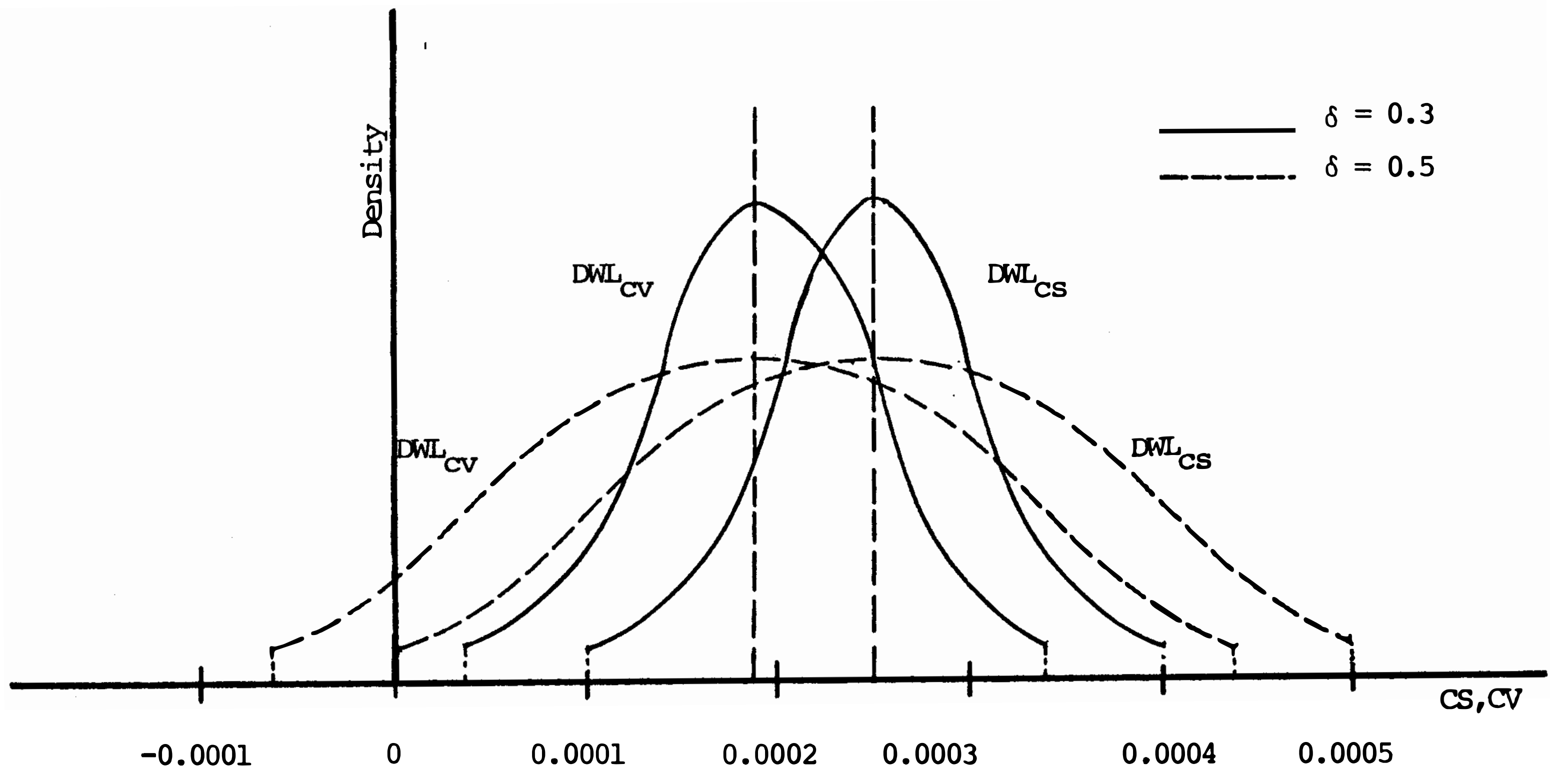


FIGURE 3

Distributions of DWL_{cs} and DWL_{cv}
(Tails Truncated at $\pm$ 2.5% Significance)



Tax = 0.05 (5%)

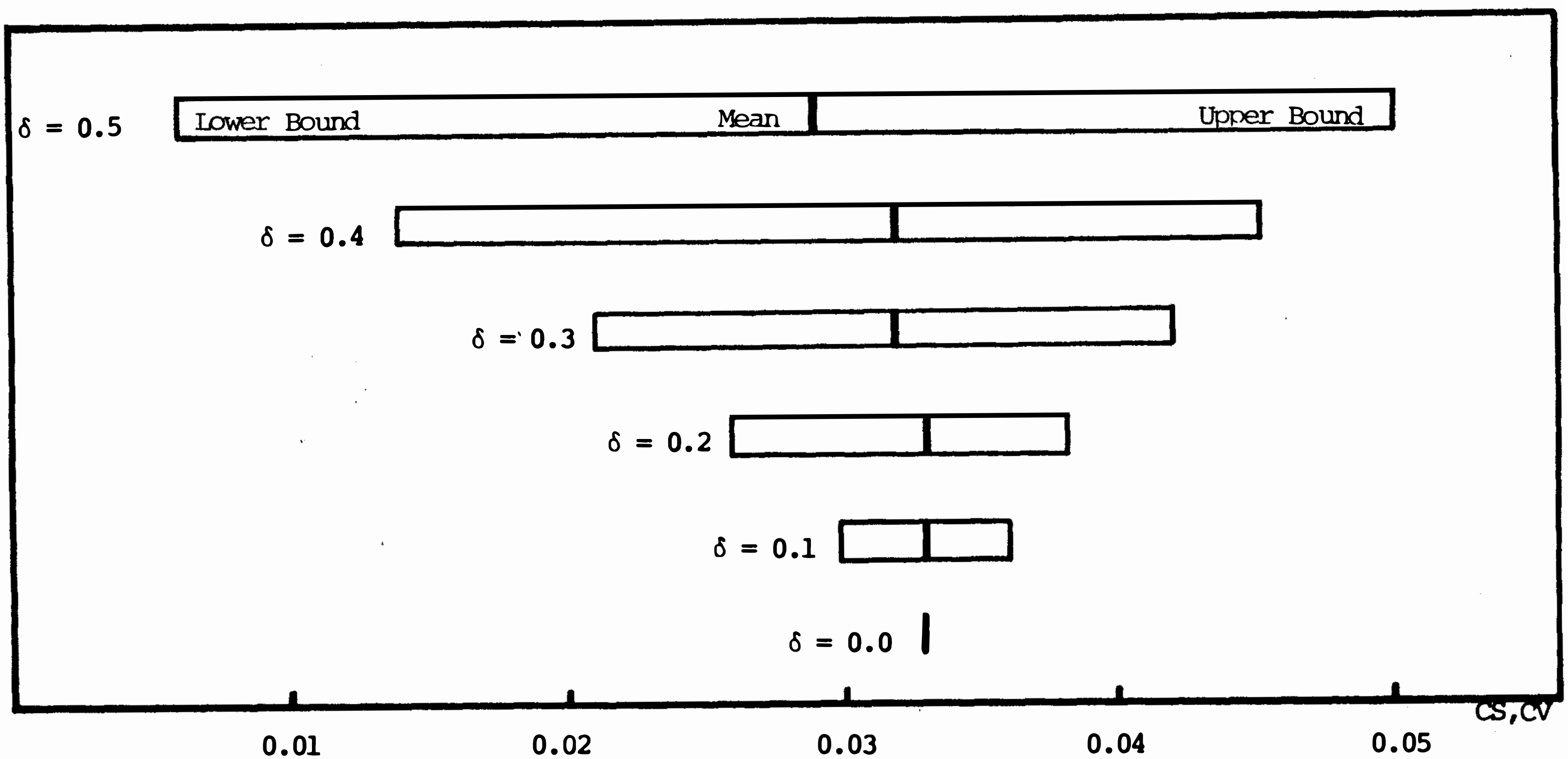
Supply: Perfectly Elastic

$a \cdot e_i = 0.05$

Demand Elasticity = -0.2

FIGURE 4

Bounds on CS and CV, for Different δ, at a 5% Significance Level[a]



a. Moments of CS and CV are essentially identical

Tax = 0.05 (5%) Supply Elasticity = 2.0

$a \cdot e_i = 0.005$ Demand Elasticity = -1.0