Expected Consumer's Surplus as an Approximate Welfare Measure[1]

Edward E. Schlee
Department of Economics
Arizona State University
Tempe, AZ 85287-3806
ed.schlee@asu.edu

Revised June 2002

Abstract

Willig (1976) argues that the change in consumer's surplus is often a good approximation to the willingness to pay for a price change: if the income elasticity of demand is small, or the price change is small, then the percentage error from using consumer's surplus is small. If the price of a good is random, then the change in *expected consumer's surplus* (ECS) equals a consumer's willingness to pay for a change in its distribution if and only if its demand is independent of income and the consumer is risk neutral. We ask how well the change in ECS approximates the willingness to pay if these conditions fail. We show that the difference between the change in ECS and willingness to pay is of higher order than the L_1 distance between the price distributions if and only if the indirect utility function is additively separable in the price and income. If additively separability fails, then the percentage error from using ECS is unbounded for small distribution changes, and is always nonzero in the limit except for knife-edge cases. Moreover, the percentage error can be large even if risk aversion, the good's income elasticity of demand and its budget share are all small. Thus, the widespread use of expected consumer's surplus as a welfare measure under uncertainty cannot be justified by approximation arguments inspired by those formulated for nonrandom prices.

JEL Classifications: D6, D8
Keywords: Expected Consumer's Surplus, Welfare under Uncertainty

[1] I thank Hector Chade, Peter Hammond, Manuel Santos and seminar participants at Arizona State University for comments.

1 Introduction

Expected consumer's surplus remains a popular measure of consumer welfare in applied microeconomic models with uncertainty. It is especially so in Industrial Organization, where incomplete information models have flourished in the last two decades.[2] Rogerson (1980) and Turnovsky, et. al. (1980) consider the validity of expected consumer's surplus as a welfare index. They show that expected consumer's surplus (ECS) represents a consumer's (expected utility) preferences over price distribution if and only if the consumer's marginal utility of income is independent of the good's price (and any other random variables entering the indirect utility function).[3] In many applications, however, expected consumer's surplus is not simply used to rank random alternatives for consumers: it is often added to the change in expected profit to evaluate the overall welfare consequences of the change. For this *sum* to be a valid welfare index (in the sense that it is positive if and only if the change is a potential Pareto improvement–the winners are willing to compensate any losers) expected consumer's surplus must not only represent each consumer's preferences, but must equal each consumer's willingness to pay for the change, a much more demanding requirement. Indeed, expected consumer's surplus equals the willingness to pay if and only if the indirect utility is quasilinear in income (Stennek, 1999), implying both that the demand for good is independent of income and that the consumer is risk neutral over income gambles.

Since these conditions are surely violated, one is naturally led to ask how well the change in expected consumer's surplus approximates the willingness to pay. For nonrandom prices, Willig (1976) argues that the change in consumer's surplus is often a good approximation to the willingness to pay for a price change: if either the income elasticity of demand is small, or the price change is small, then the percentage error from using consumer's surplus is small; in particular, the percentage error vanishes as the price change tends to zero. This result at least justifies consumer's surplus for "local" cost-benefit analysis: If the policy is indexed by a parameter and both consumer's surplus and profit are differentiable in that parameter with the sum having a positive derivative, then a *small enough* policy change will be a potential Pareto improvement. In addition, Vives (1987, Proposition 1) shows that, under some smoothness and curvature assumptions, the percentage error from using consumer's surplus to measure welfare is at most of order $1/\sqrt{\ell}$, where ℓ is the number of goods.

We show that natural analogs of these approximation arguments fail if prices are random. First, the difference between the change in ECS and willingness to

[2]To give just three examples, it is used to analyze optimal regulation (Baron and Myerson, 1982; Lewis and Sappington, 1988); vertical relationships (Rey and Tirole, 1986; Deneckere, et. al., 1997); and information sharing in oligopoly (see Vives [1999, Section 8.3] for a thorough review of the large information-sharing literature).

[3]Turnovsky et. al. (1980, Section 2 and p. 145) do not phrase their result exactly this way. They show that the change in consumer's surplus is proportional to the change in the utility precisely when the marginal utility of income is independent of any prices that change. Since a von Neumann-Morgenstern utility is unique up to an increasing affine transformation, this is equivalent to showing that ECS represents the consumer's preferences.

pay is of higher order than the L_1 distance between the price distributions if and only if the indirect utility function is additively separable in the price and income (Theorem 1). In this case, ECS is a good approximation for small changes in the price distribution: as the distribution change tends to zero, the percentage error from using ECS tends to zero (Proposition 1). If, however, additive separability fails, then the percentage error from using ECS is unbounded for small distribution changes; moreover, if the percentage error does tend to zero for some path of changes, a small perturbation of the path leads to a nonzero limit (Theorem 2)–a zero limit for the percentage error is thus a knife-edge phenomenon. Additively separability of indirect utility implies, among other things, that the price elasticity of demand is independent of income, an implication at odds with empirical studies of demand (e.g. Blundell et. al., 1993). Hence expected consumer's surplus is apt to be a poor guide even for local cost-benefit analysis, a point we illustrate with two models of investment under uncertainty (Section 4).

In addition, we show that the percentage error from using ECS can be large even if a consumer's income elasticity and risk aversion are small (Theorem 3(b) and Examples 1 and 2), in marked contrast to consumer's surplus under certainty. In sum, the widespread use of expected consumer's surplus[4] in applied microeconomics cannot be justified by approximation arguments inspired by those formulated for nonrandom prices.

2 Preliminaries

Let there be $\ell \geq 2$ goods, with prices of all goods but the first fixed and strictly positive. We assume that the preference relation over nonnegative commodity bundles is complete, transitive, continuous, nonsatiated and strictly convex. Strict convexity ensures that demands are single-valued; with nonsatiation it ensures that the budget constraint always binds. The indirect utility function is $(p, m) \longmapsto V(p, m)$, where $p > 0$ is the (non-stochastic) price of good 1 and $m > 0$ is income. Recall that V is continuous, increasing in $(-p, m)$, and quasiconvex in (p, m). The price of good 1 is random, taking on values in some compact interval $P \equiv [\underline{p}, \overline{p}]$, where $0 < \underline{p} < \overline{p}$. Let D be the space of cumulative distribution functions (c.d.f.'s) on P, endowed with the topology of weak convergence. We denote the c.d.f. that assigns probability one to the point p by δ_p.

We assume that preferences over elements of D and income levels satisfy the expected utility hypothesis: for $m', m \in \mathbb{R}_{++}$ and $G, F \in D$, a consumer weakly prefers (m', G) to (m, F) if and only if

$$\int V(p, m')dG(p) \geq \int V(p, m)dF(p),$$

[4]Since the change in expected profit can easily be zero, these results also extend to the use of expected total surplus as a measure of aggregate welfare, a point to which we return at the end of Section 3.

where V is a representation of the consumer's indirect utility function. The indirect utility function embodies two distinct aspects of the consumer's preferences over price distributions and incomes: the consumer's preferences over non-random commodity vectors; and risk-preferences, which depend on the particular representation of the indirect utility function.

Suppose that the consumer initially faces a price distribution F. Let $\pi(G, F, m)$ denote the income the consumer is willing to pay to replace F with G :

$$\int V(p, m)dF = \int V(p, m - \pi(G, F, m))dG.^5 \tag{1}$$

The number $\pi(G, F, m)$ is sometimes called the *ex ante* compensating variation for the change. If G is a degenerate c.d.f., $\pi(G, F, m)$ is sometimes called the risk premium. Clearly, $\pi(G, F, m) \geq 0$ if and only if the consumer weakly prefers G to F at income m. Moreover, for any $G \in D$, $-\pi(G, \cdot, m)$ represents the consumer's preferences on $D \times \{m\}$. (Note, however, that $\pi(\cdot, F, m)$ need not represent preferences on $D \times \{m\}$.)

Let $(p, m) \longmapsto d(p, m)$ denote the consumer's demand function for good 1. Since we are only interested in changes in consumer's surplus and prices lie in $P \equiv [\underline{p}, \overline{p}]$, there is no loss of generality in setting consumer's surplus at p equal to $\int_p^{\overline{p}} d(\omega, m)d\omega \equiv cs(p, m)$ for any $p \in P$.[6] The change in expected consumer's surplus from replacing F by G is

$$\pi_{cs}(G, F, m) = \int cs(p, m)d(G(p) - F(p)). \tag{2}$$

Although d and V are defined for all positive incomes and prices, we confine income to an interval, $M = (\underline{m}, \overline{m})$, where $0 < \underline{m} < \overline{m}$. We assume that $d(p, m) > 0$ on $P \times M$, which implies that $V(\cdot, m)$ is strictly decreasing on P for every $m \in M$.

Our goal is to determine how well π_{cs} approximates π. Since π_{cs} is the difference between two expectations, $\pi_{cs}(\cdot, F, m)$ is linear on D in the sense that $\pi_{cs}(\lambda G + (1-\lambda)H, F, m) = \lambda\pi_{cs}(G, F, m) + (1-\lambda)\pi_{cs}(H, F, m)$ for all $G, H \in D$ and $\lambda \in [0, 1]$. In general, however, $\pi(\cdot, F, m)$ is not linear in the probabilities.[7] So our problem is to determine how well the linear functional π_{cs} approximates (the possibly non-linear functional) π. Machina (1982) considers a related question: when do non-expected utility preferences preserve properties that hold for expected utility? He shows that, if a preference representation is smooth in the sense of being Fréchet differentiable in the probabilities (with respect to the L_1 norm), then it will behave locally as an expected utility representation. Although we shall exploit tools developed in the literature on smooth

[5] Since nonsatiation and strict convexity of preferences implies local nonsatiation, V is strictly increasing in income. Since V is also continuous, the number $\pi(G, F, m)$ exists and is unique.

[6] If the integral $\int_p^{\infty} d(\omega, m)d\omega$ does not exist, then consumer's surplus is not defined. The *change* in consumer's surplus between any two prices, however, is.

[7] Let $\Lambda(G, m) = \int V(p, m)dG(p)$. Then $\pi(G, F, m) = m - \Lambda^{-1}(G; \int V(p, m)dF)$, which is clearly not always linear.

non-expected utility preferences, our question is different: both π and π_{cs} each represent expected utility preferences; we ask instead how much the two utility representations π and π_{cs} differ in magnitude (and as a fraction of π_{cs}).

Since π_{cs} is linear on D (and $cs(\cdot, m)$ is absolutely continuous in p), it is L_1-Fréchet differentiable on D. If π_{cs} approximates π well for small changes in the distribution then evidently $\pi(\cdot, F, m)$ is L_1-Fréchet differentiable at F and it has the same derivative as $\pi_{cs}(\cdot, F, m)$. This observation is the starting point for our approximation results. The following regularity condition will ensure the differentiability of π.

Definition 1 *V is* ***regular*** *if it is continuously differentiable on $P \times M$ with $V_2 > 0$.*[8]

3 Expected Consumer's Surplus as a Welfare Measure

Two related preliminary questions are, first, when does expected consumer's surplus represent the consumer's preferences over D; and, second, when is π_{cs} precisely equal to π?

3.1 Expected consumer's surplus as a representation of preferences

Rogerson (1980, Theorem 1) shows that expected consumer's surplus represents the consumer's preferences over price distributions if and only if the marginal utility of income is independent of price. This condition will play an important role in our analysis.

Lemma 1 *Let V be regular. The following two assertions are equivalent.*[9]

(a) For any $G, F \in D$ and $m \in M$, $\pi_c(G, F, m) \geq 0$ if and only if $\int V(p, m)dG(p) \geq \int V(p, m)dF(p)$.

(b) V is additively separable in (p, m) on $P \times M$.

If V is additively separable (and twice continuously differentiable with $V_2 > 0$), then the income elasticity of demand for good 1, $\eta_m = (\partial d/\partial m)m/d$, equals relative risk aversion over income gambles, $r = -mV_{22}/V_2$ (Turnovsky, et. al. 1980, p. 141).

[8]Numerical subscripts will denote partial derivatives. Recall that preferences over commodities are strictly convex and nonsatiated. If, in addition, we impose two conditions on the agent's von Neumann-Morgenstern utility function over commodities, then V will be regular: it is concave; and it is differentiable in the first commodity with a positive derivative everywhere. (This fact follows from Corollary 5 in Milgrom and Segal [2000]).

[9]Rogerson imposes stronger differentiability assumptions then we do. Lemma 1, however, simply requires Roy's identity to hold, namely, that $d(p, m) = -V_1(p, m)/V_2(p, m)$. If V is differentiable, and preferences over commodity bundles are strictly convex and nonsatiated, then the identity holds (Mas-Colell, et. al., 1995, p. 73).

From Lemma 1, it is easy to see that expected consumer's surplus always equals the willingness to pay only if the indirect utility function is quasilinear in income: Equality between the two measures certainly implies that expected consumer's surplus represents the consumer's preferences on D, so, by Lemma 1, V takes the additively separable form $V(p,m) = a(p) + b(m)$; but since $\pi_{cs}(\cdot, F, m)$ is linear in the probabilities, so is $\pi(\cdot, F, m)$, implying in turn that $b(\cdot)$ is affine on M. If V is quasilinear, of course, then the demand for good 1 is independent of income and the consumer is risk neutral over income gambles.

3.2 Expected consumer's surplus as an approximate welfare measure

If V is not quasilinear in m, then $\pi(\cdot, F, m)$ is not linear in the probabilities. If, however, V is regular, then at least $\pi(\cdot, F, m)$ can be approximated well by some linear functional. Let $||\cdot||$ denote the L_1 norm: for any integrable function f on P, $||f|| = \int |f| dp$.

Lemma 2 *Let V be regular. For any $m \in M$ and $F \in D$,*

$$\pi(G, F, m) = \frac{\int V(p,m) d(G-F)}{\int V_2(\omega, m) dF(\omega)} + o(||G-F||),$$

where $o(\cdot)$ is a real-valued function satisfying $o(0) = 0$ and $\lim_{h \to 0} o(h)/h = 0$.[10]

Proof: Appendix. ■

3.2.1 Small distribution changes

Theorem 1 *Let V be regular. The following two assertions are equivalent:*

(a) For every $(F,m) \in D \times M$, $\pi(G,F,m) - \pi_{cs}(G,F,m) = o(||G-F||)$.

(b) V is additively separable on $P \times M$.

Proof: Suppose that (b) holds, so that V takes the form $V(p,m) = f(p) + g(m)$. Applying Lemma 2, we get

$$\pi(G, F, m) = \int \frac{f(p)}{g'(m)} d(G-F) + o(||G-F||).$$

By Roy's Identity (see footnote 9), $d(p,m) = -V_1/V_2 = -f'(p)/g'(m)$ on $P \times M$. Thus

$$cs(p,m) = \int_p^{\overline{p}} d(\omega, m) d\omega = \frac{-1}{g'(m)} \int_p^{\overline{p}} f'(\omega) d\omega = \frac{f(p) - f(\overline{p})}{g'(m)},$$

[10] Formally, if V is regular, then $\pi(\cdot, F, m)$ is L_1-Fréchet differentiable at F; its derivative at F is given by the linear functional $L(G-F; F) = \int V d(G-F) / \int V_2 dF$, where $L(\cdot; F)$ is defined on the linear space spanned by D endowed with the L_1 norm.

so that

$$\pi_{cs}(G,F,m) = \int cs(p,m)d(G-F) = \int \frac{f(p)}{g'(m)}d(G-F). \tag{3}$$

Consequently,

$$\pi(G,F,m) - \pi_{cs}(G,F,m) = o(||G-F||).$$

Suppose now that (a) holds: for every $(F,m) \in D \times M$, $\pi(G,F,m) = \int cs(p,m)d(G-F) + o(||G-F||)$. In particular,

$$\pi(\alpha G + (1-\alpha)F, F, m) = \alpha \int cs(p,m)d(G-F) + o(\alpha||G-F||)$$

so that

$$\lim_{\alpha\to 0^+} \frac{\pi(\alpha G + (1-\alpha)F, F, m)}{\alpha} = \int cs(p,m)d(G-F).$$

But by Lemma 2 this limit also equals $\int V(p,m)d(G-F)/\int V_2 dF$. Thus

$$\int cs(p,m)d(G-F) = \frac{\int V(p,m)d(G-F)}{\int V_2(\omega,m)dF(\omega)}$$

and $\int csd(G-F) > 0$ if and only if $\int Vd(G-F) > 0$, so that ECS represents the consumer's preferences on $D \times \{m\}$ for every $m \in M$. By Lemma 1, V is additively separable. ■

We now turn to the percentage error from using ECS. Additive separability also implies that this percentage error tends to zero as the change in distribution tends to zero.

Proposition 1 *If V is regular and additively separable on $P \times M$, then for any sequence $\{G_n\}$ in D converging to F with $\pi_{cs}(G_n,F,m) \neq 0$ for all n,*

$$\lim_{n\to\infty} \frac{\pi(G_n,F,m) - \pi_{cs}(G_n,F,m)}{\pi_{cs}(G_n,F,m)} = 0.$$

Proof: Let $V(p,m) = f(p) + g(m)$ and consider any sequence $\{G_n\}$ satisfying the conditions of the Proposition. Since g is differentiable on M , the Mean Value Theorem implies that, for n sufficiently large, there is number t_n between m and $m-$ $\pi(G_n,F,m)$ such that $g(m - \pi(G_n,F,m)) - g(m) = -\pi(G_n,F,m)g'(t_n)$. Thus $\pi(G_n,F,m) = \int f(p)d(G-F)/g'(t_n)$ for large enough n. Since g' is continuous and positive on M, we have (using (3)) that

$$\lim_{n\to\infty} \frac{\pi(G_n,F,m) - \pi_{cs}(G_n,F,m)}{\pi_{cs}(G_n,F,m)} = \lim_{n\to\infty} \frac{g'(m) - g'(t_n)}{g'(t_n)} = 0. \blacksquare$$

If V is not additively separable, however, then the percentage error from using expected consumer's surplus is unbounded. Moreover, if the percentage error does tend to zero for a particular path of convergence to a non-degenerate c.d.f. F, then a small perturbation of that path will result in a nonzero limit: thus a zero limit for the percentage error is a knife-edge phenomenon. To demonstrate the second point, we use the notion of smooth paths on D. A *path on D to F*, which we denote by $\{H(\cdot,\alpha)|\alpha \in [0,1]\}$, is a function from $[0,1]$ to D with $H(\cdot,0) = F$. A path $\{H(\cdot,\alpha)\}$ to F is *smooth* (at F) if $H_2(p,\alpha)$ exists and is bounded on $P \times [0,1]$ and if the family of functions $\{(H(p,\alpha) - F(p))/\alpha|\alpha \in [0,1]\}$ of p is of bounded variation uniformly in α (Wang, 1993, p. 145).[11] In part (b) of the following result, we assume that the marginal utility of income, V_2, is not F-*a.e.* constant as a function of p. This implies that V is not additively separable and requires F to be a nondegenerate c.d.f.

Theorem 2 *Let V be regular.*

(a) If V is not additively separable on $P \times M$, then there is an $(F,m) \in D \times M$, and a sequence $\{G_n\}$ in D converging to F such that $\pi_{cs}(G_n,F,m) \neq 0$ for all n and

$$\lim_{n\to\infty} \left| \frac{\pi(G_n,F,m) - \pi_{cs}(G_n,F,m)}{\pi_{cs}(G_n,F,m)} \right| = \infty.$$

(b) Let F be nondegenerate and let $\{\widehat{H}(\cdot,\alpha)\}$ be any smooth path to F with $\pi_{cs}(\widehat{H}(\cdot,\alpha),F,m) \neq 0$ for all $\alpha \in (0,1]$. If $V_2(\cdot,m)$ is not F–a.e. constant in p, then for any $\varepsilon > 0$, there is a smooth path $\{H(\cdot,\alpha)\}$ to F with $||H(\cdot,\alpha) - \widehat{H}(\cdot,\alpha)|| < \varepsilon$ for all $\alpha \in [0,1]$ and $||H_2(\cdot,0) - \widehat{H}_2(\cdot,0)|| < \varepsilon$ such that

$$\lim_{\alpha\to 0} \frac{\pi(H(\cdot,\alpha),F,m) - \pi_{cs}(H(\cdot,\alpha),F,m)}{\pi_{cs}(H(\cdot,\alpha),F,m)} \neq 0.$$

Proof: Suppose throughout that V is regular. For part (a), since V is not additively separable, Lemma 1 implies that ECS does not represent preferences over D. Since, in addition, $V(\cdot,m)$ and $cs(\cdot,m)$ are both continuous and strictly decreasing, it is easy to show that there are c.d.f.'s $G^*, G, F \in D$ and an income $m \in M$ such that $\int V(p,m)d(G-F) \geq 0$ and $\int V(p,m)d(G^*-F) > 0$ but $\int cs(p,m)d(G-F) < 0$ and $\int cs(p,m)d(G^*-F) = 0$. (See Figure 1.) Define $G_n = \frac{1}{n}\left(\frac{1}{n}G + (1-\frac{1}{n})G^*\right) + (1-\frac{1}{n})F$. Clearly $||G_n - F||$ converges to zero and $\pi_{cs}(G_n,F,m) \neq 0$ for all n. Moreover, letting $\Omega = \int V_2(\omega,m)dF(\omega)$, we have

$$\left| \frac{\pi(G_n,F,m) - \pi_{cs}(G_n,F,m)}{\pi_{cs}(G_n,F,m)} \right| =$$

$$\left| \frac{\frac{1}{n}\int \left(\frac{V(p,m)}{\Omega} - cs(p,m)\right) d(G-G^*) + \int \left(\frac{V(p,m)}{\Omega} - cs(p,m)\right) d(G^*-F) + no(||G_n - F||)}{\frac{1}{n}\int cs(p,m)d(G-G^*)} \right|.$$

[11] These conditions ensure that we can pass a derivative with respect to α under the integral and that $H_2(\cdot,0)$ is of bounded variation, so that it defines a finite signed measure. In what follows $H_2(\cdot,0)$ should be understood to be a one-sided derivative.

The denominator on the right side of the equality converges to zero, as does the first term in the numerator. Since $G_n - F = \frac{1}{n}\left(\frac{1}{n}(G - G^*) + G^* - F\right)$, the third term in the numerator converges to zero. Since the second term in the numerator is nonzero, the entire expression diverges to $+\infty$ as $n \to \infty$.

For part (b) let $\{\widehat{H}(\cdot, \alpha)\}$ be a smooth path to F satisfying the hypotheses. For any $\alpha > 0$ we have

$$\frac{\pi(\widehat{H}(\cdot,\alpha), F, m) - \pi_{cs}(\widehat{H}(\cdot,\alpha), F, m)}{\pi_{cs}(\widehat{H}(\cdot,\alpha), F, m)} = \frac{\left(\pi(\widehat{H}(\cdot,\alpha), F, m) - \pi_{cs}(\widehat{H}(\cdot,\alpha), F, m)\right)/\alpha}{\pi_{cs}(\widehat{H}(\cdot,\alpha), F, m)/\alpha}.$$

Since $\{\widehat{H}(\cdot, \alpha)\}$ is smooth, the limit of the denominator is $\int cs(p,m) d\widehat{H}_2(p,0)$ (Wang, 1993, Lemma 4(b)). Similarly, using Lemma 2, the limit of the numerator is

$$\frac{\int V(p,m) d\widehat{H}_2(p,0)}{\int V_2(\omega,m) dF(\omega)} - \int cs(p,m) d\widehat{H}_2(p,0).$$

Integrating each integral involving $\widehat{H}_2$ by parts and rearranging, this expression equals

$$\int_P \left(-\frac{V_1(p,m)}{V_2(p,m)} \frac{V_2(p,m)}{\int V_2(\omega,m) dF(\omega)} - d(p,m)\right) \widehat{H}_2(p,0) dp,$$

or, using Roy's Identity,

$$\int_P \left[d(p,m)\left(\frac{V_2(p,m)}{\int V_2(\omega,m) dF(\omega)} - 1\right) \widehat{H}_2(p,0)\right] dp. \tag{4}$$

If this integral is nonzero, then there is nothing to prove. So suppose that it is zero. Let $A_+ = \{p \in P | V_2 > \int V_2 dF\}$ and $A_- = \{p \in P | V_2 < \int V_2 dF\}$. Since $V_2(\cdot, m)$ is not F-*a.e.* constant, $\int_{A_+} dF > 0$ and $\int_{A_-} dF > 0$. And since V_2 is continuous, both A_+ and A_- are open relative to P (that is, each is the intersection of P and an open set) and hence each is the union of a countable collection of disjoint intervals that are open relative to P. Consequently, there are intervals $I_+ \in A_+$ and $I_- \in A_-$ from these countable collections with $\int_{I_+} dF > 0$ and $\int_{I_-} dF > 0$ and at most one of these intervals is closed on the left (which can happen only if the left endpoint is $\underline{p}$). If I_+ is open on the left, then let $G \in D$ satisfy $G = F$ for all $p \notin I_+$ and $G \geq F$ on I_+ with the inequality strict on a set of positive Lebesgue measure. (Such a G exists since I_+ is open on the left and F is not constant on I_+.) If I_+ is closed on the left, then let G satisfy $G = F$ for all $p \notin I_-$ and $G \geq F$ on I_- with the inequality strict on a set of positive Lebesgue measure.

In either case, for every $\lambda \in (0,1)$, define a smooth path to F by $H^\lambda = (1-\lambda)\widehat{H} + \lambda(\alpha G + (1-\alpha)F)$. Since $d > 0$, we have (replace $\widehat{H}$ with H^λ in (4))

$$\lim_{\alpha \to 0} \frac{\pi(H^\lambda(\cdot,\alpha), F, m) - \pi_{cs}(H^\lambda(\cdot,\alpha), F, m)}{\alpha} \neq 0$$

for any $\lambda \in (0,1)$. For λ small enough, $||H^\lambda(\cdot,\alpha) - \widehat{H}(\cdot,\alpha)|| < \varepsilon$ for all $\alpha \in [0,1]$ and $||H_2^\lambda(\cdot,0) - \widehat{H}_2(\cdot,0)|| < \varepsilon$. Set $H = H^\lambda$ for such a λ. Noting that $\pi_{cs}(H(\cdot,\alpha),F,m) \neq 0$ except for at most one value of $\alpha > 0$, we have

$$\lim_{\alpha\to 0} \frac{\pi(H(\cdot,\alpha),F,m) - \pi_{cs}(H(\cdot,\alpha),F,m)}{\pi_{cs}(H(\cdot,\alpha),F,m)} \neq 0. \quad \blacksquare$$

Figure 1 illustrates the argument for part (a). If V is not additively separable, then ECS does not represent preferences over price distributions: there are distributions G and F such that ECS ranks them differently than expected utility. If the path $\{H(\cdot,\alpha)\}$ from G to F is such that ECS ranks $H(\cdot,\alpha)$ and F differently than expected utility for all $\alpha \in (0,1]$ and is tangent to the indifference set of ECS at F, then π_{cs} tends to zero faster than $\pi - \pi_{cs}$ and hence the percentage error from using ECS tends to $+\infty$.

Theorem 2 contrasts markedly with the case of a single price change under certainty: as the change in price under certainty tends to zero, the percentage error from using consumer's surplus to measure willingness to pay always tends to zero.[12] Theorem 2(b) implies that a zero limiting percentage error is a knife-edge phenomenon when additive separability fails. As noted already, additive separability of the indirect utility is a strong assumption even under certainty, implying that own-price elasticities are independent of income. Hence we cannot reasonably justify the change in expected consumer's surplus as a measure of the willingness to pay even for local cost-benefit analysis.

Remark 1 Theorem 2(b) requires that the initial distribution be nondegenerate.. We can dispense even with that condition if $V_2(\cdot,m)$ is strictly monotone on P. Let $F = \delta_{p^*}$ for some $p^* \in P$. If $p^* > \underline{p}$ then modify G in the paragraph following (4) to satisfy $G(p) > 0$ for $p < p^*$ with $G(p) = 1$ for $p \geq p^*$. If $p^* = \underline{p}$, let $0 < G < 1$ on $[\underline{p},\overline{p})$. In either case, define H^λ as in the proof and the rest of the proof goes through as before. The example in Section 4.2 illustrates this point.

In Theorem 2, we used paths that are smooth on D to show that ECS is generally a poor approximation to a consumer's willingness to pay. If we consider other paths of approach, then we can avoid at least some of the grim consequences of Theorem 2. Besides paths that are linear on the space of c.d.f.'s, the most commonly used paths in the economics of uncertainty are linear on the space of random variables. (See Section 4.1 for an economic example.) Let $\widetilde{p}$ denote a random variable with c.d.f. F and let $\widetilde{\varepsilon}$ be real-valued random variable with range in some bounded interval containing zero.[13] $H(\cdot|p)$ here stands for the c.d.f. of $\widetilde{\varepsilon}$ conditional on p, and $\mu(p)$ denotes the mean and $\sigma(p)^2$ the variance of $\widetilde{\varepsilon}$ conditional on p. Let G_x denote the c.d.f. of the random variable

[12] That is, as long as preferences are strictly convex and nonsatiated, and demand is positive. (Use Willig, (1976), eq. (20) or apply the Integral Mean Value Theorem directly to the formula for the percentage error).

[13] We want $p + \alpha\varepsilon$ to lie in P for α small.

$\widetilde{x}$. Given the initial price distribution F, we will consider final distributions of the form $\{G_{p+\alpha\varepsilon}|\alpha \geq 0\}$. It follows readily that $\lim_{\alpha\to 0} ||G_{p+\alpha\varepsilon} - F|| = 0$. Define $\Pi(\alpha) = \pi(G_{p+\alpha\varepsilon}, F, m)$ and $\Pi_{cs}(\alpha) = \pi_{cs}(G_{p+\alpha\varepsilon}, F, m)$. The next result summarizes some consequences of this constraint on paths.

Theorem 3 *Let V be regular.*

(a) If F is degenerate at p and $\mu(p) \neq 0$, then $\lim\limits_{\alpha\to 0} \frac{\Pi(\alpha)-\Pi_{cs}(\alpha)}{\Pi_{cs}(\alpha)} = 0$.

(b) If V is twice continuously differentiable, F is degenerate at p, $\mu(p) = 0$ and $d_1(p,m)\sigma(p)^2 \neq 0$, then

$$\lim_{\alpha\to 0} \frac{\Pi(\alpha) - \Pi_{cs}(\alpha)}{\Pi_{cs}(\alpha)} = -s\left[\eta_m - r\right]/\eta_p,$$

where η_m is the income elasticity of demand, η_p the price elasticity of demand, r the measure of relative risk aversion $(-mV_{22}/V_2)$,and s the budget share for good 1(all evaluated at (p,m)).

In each case, the limit is zero for all $(F,m) \in D \times M$ and all conditional distributions $H(\cdot|\cdot)$ if and only if V is additively separable.

Proof: Consider the function ψ defined by $\psi(\alpha, m) = \int\int V(p+\alpha\varepsilon, m)dH(\varepsilon|p)dF(p)$. Since V is regular, and $P \times M$ is bounded, there is an open interval A containing 0 such that ψ is continuously differentiable on $A \times M$; moreover $\psi_1(\alpha, m) = \int\int V_1(p+\alpha\varepsilon, m)\varepsilon dH(\varepsilon|p)dF(p)$ and $\psi_2(\alpha, m) = \int\int V_2(p+\alpha\varepsilon, m)dH(\varepsilon|p)dF(p)$. (See e.g. Bartle, 1995, Corollaries 5.8 and 5.9.) Hence, by the Implicit Function Theorem, Π is continuously differentiable on an open interval containing 0 with

$$\Pi'(\alpha) = \frac{\int\int V_1(p+\alpha\varepsilon, m-\Pi(\alpha))\varepsilon dH(\varepsilon|p)dF(p)}{\int\int V_2(p+\alpha\varepsilon, m-\Pi(\alpha))dH(\varepsilon|p)dF(p)}, \tag{5}$$

and

$$\Pi'(0) = \frac{\int V_1(p,m)\mu(p)dF(p)}{\int V_2(p,m)dF(p)}. \tag{6}$$

Moreover, Π_c is continuously differentiable with

$$\Pi'_{cs}(\alpha) = -\int\int d(p+\alpha\varepsilon, m)\varepsilon dH(\varepsilon|p)dF(p), \tag{7}$$

and

$$\Pi'_{cs}(0) = -\int d(p,m)\mu(p)dF(p). \tag{8}$$

If F is degenerate at some p, then Roy's Identity implies that $\Pi'_{cs}(0) = \Pi'(0) = -d(p,m)\mu(p)$, so (a) follows from L'Hôpital's Rule.

Now let F be nondegenerate and suppose that $\mu(p)$ is nonzero and constant: $\mu(p) = \widehat{\mu} \neq 0$ for all p. From (6) and (8) we have that

$$\Pi'(0) - \Pi'_{cs}(0) = \widehat{\mu}\int\left(\frac{V_1}{\int V_2 dF} - \frac{V_1}{V_2}\right)dF = \widehat{\mu}\int\frac{V_1}{V_2}\left(\frac{V_2}{\int V_2 dF} - 1\right)dF.$$

Clearly, the last expression is zero for all F if and only if V_2 is independent of p for all m : if $V_2(p', m) \neq V_2(p'', m)$ for some $m \in M$ and p', p'' in P, choose F to have two-point support on $\{p', p''\}$; since V_1/V_2 is strictly increasing, the covariance inequality then implies that $\Pi'(0) - \Pi'_{cs}(0) \neq 0$. This proves the last sentence of Theorem 3 for case (a).

For (b), since $\mu(p) = 0$, $\Pi'(0) = \Pi'_{cs}(0) = 0$. And since V is twice continuously differentiable on $P \times M$, it follows that Π and Π_{cs} are twice continuously differentiable on an open interval containing zero, so that (provided $\Pi''_{cs}(0) \neq 0$),

$$\lim_{\alpha \to 0} \frac{\Pi(\alpha) - \Pi_{cs}(\alpha)}{\Pi_{cs}(\alpha)} = \frac{\Pi''(0) - \Pi''_{cs}(0)}{\Pi''_{cs}(0)}. \tag{9}$$

Straightforward calculations reveal that

$$\Pi''(0) = \frac{V_{11}(p,m)}{V_2(p,m)} \sigma(p)^2 \tag{10}$$

and

$$\Pi''_{cs}(0) = -d_1(p,m)\sigma(p)^2 = \frac{V_{11}(p,m)}{V_2(p,m)}\sigma(p)^2 - \frac{V_{21}(p,m)V_1(p,m)}{\left(V_2(p,m)\right)^2}\sigma(p)^2. \tag{11}$$

Manipulation of Roy's Identity yields that $V_{12} = (\eta_m - r)V_1/m$. Substituting this term into (11), the resulting expression and (10) into (9) and rearranging yields (b). The equivalence between additive separability and a zero limit is immediate. ■

Theorem 3(a) requires nothing other than regularity of V. Hence, for this class of changes, the change in ECS approximates the willingness to pay well if the distribution change is small. Indeed, if $\widetilde{\varepsilon}$ is degenerate, then the change collapses to the familiar case of a price change under certainty. Thus part (a) can be viewed as a stochastic generalization of that case. Theorem 3 also shows, however, that if the initial price distribution is nondegenerate, then this generalization fails unless V is additively separable.

Remark 2 We now give a geometric interpretation of Theorem 3. Figure 2 depicts possible indifference sets for expected utility and ECS over prices viewed as random variables for the two-state case. The slope of any indifference curve along the certainty line equals the relative state probabilities. (Compare with Figure 1.) If the change does not preserve the mean ($\mu(p) \neq 0$), then $\Pi'_{cs}(0) \neq 0$. Hence (a) holds if and only if $\Pi'(0) = \Pi'_{cs}(0) \neq 0$, which implies that $\Pi'(0)$ and $\Pi'_{cs}(0)$ agree in sign: for α close to zero, ECS represents preferences over values of α. But since ECS and V are differentiable and $\mu(p) \neq 0$, we have that $\Pi(\alpha)$ and $\Pi_{cs}(\alpha)$ do agree in sign for α small enough (e.g. the path ab). If, however, we start from a nonrandom price, then the slopes of the indifference curves sometimes differ, as at point d (unless V is additively separable). Hence we can find a linear path such that $\Pi'(0)$ and $\Pi'_c(0)$ have different signs and (a) fails. If $\mu = 0$, then $\Pi'_{cs}(0) = 0 = \Pi'(0)$ and the property of equal slopes along the certainty line no longer suffices for conclusion (a): we must consider the *curvature* of the indifference curves at the initial distribution (path ac). And, unless V is additively separable, the curvature of the two families of indifference curves differ on the certainty line.

Remark 3 Since smooth paths are locally linear, one can easily extend the conclusion of Theorem 3(a) to a class of smooth paths on the space of random variables. Remark 1, however, demonstrates that the conclusion fails for smooth paths on the space of c.d.f.'s to a degenerate c.d.f.

3.2.2 Small income effects and risk aversion

Willig (1976, p. 594, eq. (20)) constructs percentage error bounds from using consumer's surplus under certainty. The upper and lower bounds depend on the size of the change in consumer's surplus and the maximum and minimum values of the income elasticity of demand. One implication of his inequality is that, if we fix income, and consider a sequence of consumers whose income elasticity converges to zero uniformly, then the percentage error from using consumer's surplus converges to zero uniformly in the price change (if prices lie in a positive, compact interval)[14]: a small departure from the condition for the exact validity of consumer's surplus leads to a small percentage error. We can use Theorem 3(b) to show that a natural analog of this result fails for ECS under uncertainty: we can construct a sequence of consumers whose income elasticity and coefficient of relative risk aversion both tend to zero uniformly, yet the percentage error from using ECS does not converge to zero uniformly in the distribution change.

Two examples

Example 1 *(Stone-Geary) Let preferences over sure consumption bundles be represented by the Stone-Geary utility function*

$$u(x;\theta) = \theta \ln(x_1 - \gamma) + (1-\theta)(\ell - 1)^{-1} \sum\nolimits_2^{\ell} \ln x_j,$$

where $\gamma > 0$ and $0 < \theta < 1$. Let p^* denote the sure initial price in Theorem 3(b). The demand for good 1 is $d(p, m, \theta) = \theta m/p + \gamma(1-\theta)$. Thus $\eta_m = -\eta_p = \theta m/(\theta m + p\gamma(1-\theta))$, and, as θ tends to zero, both these functions converge uniformly to zero. The budget share, however, tends to $p\gamma/m > 0$. These properties are unaffected by taking an increasing transformation $T(\cdot;\theta)$ of u. One can choose T such that relative risk aversion equals θ at $p = p^*$ and tends to zero uniformly (in p and m) as θ tends to zero.[15] *In this case $-s\left[\eta_m - r\right]/\eta_p$*

[14]If prices are confined to a positive, compact interval and the income level is fixed, then change in consumer's surplus under certainty is bounded by a number that does not depend on the consumer. The convergence claim in the text follows from this fact and the continuity of Willig's upper and lower bounds in his eq. (20).

[15]The indirect utility function is of the form $V(p, m, \theta) = \ln(m - \gamma p) + k(\theta, p)$ where $k(\theta, \cdot)$ is continuous. Define $T(\xi, \theta) = \left(e^{\xi - k(\theta, p^*)} + \gamma p\right)^{1-\theta}$. Then relative risk aversion of the function $v(p, m, \theta) = T(V(p, m, \theta), \theta)$ equals θ at $p = p^*$and it converges to zero uniformly as θ tends to 0.

from Theorem 3(b) tends to $(\gamma - 1)p^*\gamma/m$, *which is nonzero if* $(\gamma - 1)\gamma \neq 0$. *By Theorem 3(b), the percentage error from using expected consumer's surplus does not tend to zero uniformly (in the distribution change) even though income elasticity and relative risk aversion converge to zero uniformly.*

In Example 1, the budget share for good 1 is bounded away from zero. Since consumer's surplus under certainty is often defended by invoking small budget shares, we now present a linear demand example in which the budget share, income elasticity and risk aversion measure all tend to zero at the initial price, yet the percentage error from using ECS does not vanish. (See also the example in Section 4.1.)

Example 2 *(Linear Demands) Let there be two goods,* x *and* y, *and let preferences over sure consumption bundles be represented by* $u(x, y, \theta) = x\theta^{-2} - \frac{1}{2}x^2\theta^{-3} + y$ *for* $\theta < 1$. *If the price of good* y *is unity and* $p < 1$, *we have* $d(p, m, \theta) = \theta - \theta^3 p$, $V(p, m) = \frac{1}{2\theta} - \theta p + \frac{1}{2}\theta^3 p^2 + m$, $\eta_m = 0$, *and* $\eta_p = -\theta^3 p(\theta - \theta^3 p)^{-1}$. *Clearly* $\lim_{\theta \to 0} pd(p, m, \theta)/m = 0$ *uniformly. But* $s/\eta_p = -\theta^{-1} + 2\theta p - \theta^3 p^2$ *which diverges to* $-\infty$. *As in Example 1, we can find an increasing transformation* T *of* V *so that relative risk aversion of* $T(V(p, m, \theta); \theta)$ *equals* θ *at the initial sure price* p^* *and converges to 0 uniformly as* θ *tends to 0. In this case,* $\lim_{\theta \to 0} s\left[\eta_m - r\right]/\eta_p = 1$ *at* $p = p^*$.[16]

The price elasticity of demand plays a critical role in Examples 1 and 2 and Theorem 3(b) but none at all in Theorem 3(a) (or for consumer's surplus approximations under certainty). A moment's reflection shows why. Suppose for illustration that the 'limiting' consumer has a zero price elasticity of demand. For Theorem 3(a), if α is small, then the change in expected consumer's surplus is nonzero for this limiting consumer. If, however, the price distributions have the same mean ($\mu = 0$) , then the change in ECS is zero for the limiting consumer. Consumers with low values of η_m and r with $\eta_m \neq r$ are not generally indifferent about small, mean preserving changes in the price distribution, so the percentage error from using ECS can be large for consumers whose behavior is nevertheless 'close' to a consumer with quasilinear utility.[17]

[16] This example can also be used to show that the percentage error does not converge to zero for a *given* change in the price distribution (rather than simply failure of the error to converge to zero uniformly in the change): let the initial price be nonrandom, choose the transformation T so that the agent has constant absolute risk aversion equal to θ, set $\alpha = 1$, and let ε have two point support with mean zero. In this case both Π_c and Π can be calculated explicitly to show that the percentage error does not always tend to zero as θ tends to zero.

[17] One might argue that constraining the price elasticity of demand to be bounded away from zero would restore the approximation result. First, that doesn't affect the qualitative point that small violations of the conditions for the exact equality between the change in ECS and the willingness to pay can lead to large percentage errors in the approximation: exact equality between the two measures is consistent with any magnitude for the price elasticity of demand. Second, the formula in (b) holds only for small changes from an initially degenerate distribution along a linear path on the space of random variables. If the initial distribution and path of approach are arbitrary, then the percentage error need not converge to zero uniformly as η_m and r tend to zero uniformly, even of η_p is bounded away from zero. (Figure 2 helps visualize how to construct the argument.)

Further remarks

1. If the indirect utility is C^2 and additively separable (so that $\eta_m \equiv r$), then the percentage error does tend to zero uniformly in the distribution change as relative risk aversion tends to zero uniformly. To prove this (recalling that relative risk aversion is just the elasticity of marginal utility of income), apply Willig's (1976) eq. (18) to bound relative marginal utilities of income at any two income levels; then apply these bounds to the formula for the percentage error in the last line of the proof of Proposition 1 (without taking the indicated limits), whence the result follows.

2. As mentioned already, Vives (1987, Proposition 1) shows that, under smoothness and curvature restrictions on utility, if the number of commodities ℓ is large, then budget shares and income effects are small and the percentage error from using consumer's surplus under certainty is likewise small. Since increasing ℓ by itself has no implications for risk aversion, it is perhaps not surprising that the percentage error from using ECS need not vanish as ℓ grows. Let preferences be symmetric Cobb-Douglas and, following Vives, normalize income to equal the number of goods. This example satisfies the hypotheses of his Proposition 1 and the limit in Theorem 3(b) is $(r(\ell) - 1)/\ell$. If the consumer has constant absolute risk aversion for income gambles equal to λ, then $\lim_{\ell \to \infty}(r(\ell) - 1)/\ell = \lambda$.

3. Hausman (1981) argues that a more important question than the percentage error in measuring the willingness to pay is the error in calculating the *deadweight loss*. Of course, Theorem 1, which deals with differences, is unaffected by this consideration. For the percentage error, we need to add the change in expected profit to the denominator of our calculations. We haven't explicitly modelled the source of the uncertainty or the supply side of the market. Since, however, the change in expected profit could be zero, our qualitative conclusions stand: the percentage error from using ECS to calculate the deadweight loss can be large even if the distribution change is small or if income effects and risk aversion are small.

4 Two Applications to Investment under Uncertainty

We now illustrate our findings with two applications to investment under uncertainty: investment in cost-reduction; and investment in information about cost. Although income effects on demand might sometimes be negligible, risk aversion surely is not. In these applications, we assume away income effects and focus instead on how risk aversion affects the accuracy of expected consumer's surplus for small policy changes.

4.1 Investment in cost reduction

Suppose that a good is produced by a competitive industry with constant returns to scale. From an *ex ante* viewpoint, the unit cost is random, having support on an open interval $(0, a)$; denote it by $\widetilde{k}$. At the time prices are set, however, all cost uncertainty is resolved. (For example, some input prices may initially uncertain, but that uncertainty could be resolved by the time firms choose output.) By spending $C(\alpha)$ dollars the unit cost of every firm can be reduced to $(1-\alpha)\widetilde{k}$ for $\alpha \in [0,1]$. Under competition, the price will be $(1-\alpha)k$. Let the demand for the good be $d(p,m) = a - p$, which arises from a single consumer with von Neumann-Morgenstern utility $u(q,y) = T(aq - \frac{1}{2}q^2 + y)$, where q is the quantity of the good, y is expenditure on all other goods and $T(\cdot)$ is a differentiable, strictly concave function with $T' > 0$. The indirect utility function is $V(p,m) = T((a-p)^2/2+m)$. Let $\Pi(\alpha)$ denote the *ex ante* willingness to pay for reducing the cost by fraction α and $\Pi_{cs}(\alpha)$ the resulting change in expected consumer's surplus. Straightforward calculations reveal that

$$\Pi'(0) = \frac{E\left[T'(\frac{1}{2}(a-\widetilde{k})^2) + m)(a-\widetilde{k})\widetilde{k}\right]}{E\left[T'(\frac{1}{2}(a-\widetilde{k})^2) + m)\right]}$$

and

$$\Pi'_{cs}(0) = E\left[(a-\widetilde{k})\widetilde{k}\right].$$

Since $(a-k)k$ is increasing in k on $(0, a/2)$ and decreasing on $(a/2, a)$, the covariance inequality implies that $\Pi'(0) > \Pi'_{cs}(0)$ if the support of $\widetilde{k}$ lies in $(0, a/2)$, with the inequality reversed if its support lies in $(a/2, a)$. Moreover, the more risk averse the consumer, the larger the percentage error in each case. To illustrate let the support of $\widetilde{k}$ be $(0, a/2]$ and its mean $a/4$. As the consumer becomes infinitely risk averse (i.e. the Arrow-Pratt risk aversion measure increases uniformly without bound), the percentage error for a small increase in α from 0 must exceed 133%. Note that the budget share of the good will be small if income is large or if the support of $\widetilde{k}$ is concentrated around a or 0.

4.2 Investment in cost information

Let the demand side of the economy remain the same as in Section 4.1, but let the good be produced by a risk-neutral, price-setting monopolist who faces an uncertain unit cost. Before choosing a price, the firm can buy information: by spending $C(\alpha)$ dollars, it will learn the cost for sure with probability $\alpha \in [0,1]$; with probability $1-\alpha$ it learns nothing. Let C be differentiable and strictly convex with $C'(0) = 0$. (This parameterization of information is largely for convenience.) Suppose the unit cost takes on two equally likely values $\underline{k}$ and $\overline{k}$, with $0 \leq \underline{k} < \overline{k} < a$. If the firm learns the cost is k, it will set a price of $(a+k)/2$. If it learns nothing, the price will be $(a+k_0)/2$, where $k_0 = (\underline{k}+\overline{k})/2$ is the prior expected cost. Note that the price is not random at $\alpha = 0$. Let

$\Pi(\alpha)$ denote the *ex ante* willingness to pay for chance α of learning the cost and $\Pi_{cs}(\alpha)$ the resulting change in ECS. Here

$$\Pi'(0) = \frac{\frac{1}{2}T(\frac{1}{8}(a-\overline{k})^2)+m) + \frac{1}{2}T(\frac{1}{8}(a-\underline{k})^2)+m) - T(\frac{1}{8}(a-k_0)^2)+m)}{T'(\frac{1}{8}(a-k_0)^2)+m)},$$

where the numerator is the expected value of perfect information for the consumer. Moreover,

$$\Pi'_{cs}(0) = \frac{1}{8}\left[\frac{1}{2}(a-\overline{k})^2 + \frac{1}{2}(a-\underline{k})^2 - (a-k_0)^2\right] > 0.$$

Indeed, ECS always rises with better information with linear demand and constant returns: better information results in a mean-preserving increase in price risk; since consumer's surplus is convex in price, ECS always rises. Not so for expected utility. Since T is strictly concave, $\Pi'(0) < \Pi'_{cs}(0)$; if T is concave enough, $\Pi'(0) < 0$; indeed we can make $\Pi'(0)$ as negative as we like by increasing the concavity of T. (As the consumer's risk aversion tends uniformly to $+\infty$, Π converges to $(2a - \overline{k} - k_0)(k_0 - \overline{k})/2$ for $\alpha > 0$ and to 0 if $\alpha = 0$.) Since $C'(0) = 0$ and the marginal benefit of information to the firm is positive at $\alpha = 0$, the firm's net expected profit is increasing in α in a neighborhood of 0. A sufficiently risk averse consumer, however, would be willing to pay the firm enough not to acquire any information: even a small change that seems to be a Pareto improvement using ECS to measure consumer welfare can *lower* the sum of expected profit and willingness-to-pay of consumers. Although our calculation assumes a price-setting firm, the same conclusion holds for a quantity-setter.

This reversal is relevant for the literature on information-sharing in oligopoly. Several papers have found that total expected surplus rises when firms share information (see Vives, 1999, section 8.3, for a summary). Part of the intuition for this finding is that sharing of information among firms is akin to an increase in information from the viewpoint of consumers. Our example suggests that consumer risk aversion could overturn some of these benign welfare conclusions.

5 Extensions

Thus far we have assumed that agents satisfy the expected utility hypothesis and that only the price of good 1 is random. We briefly consider extensions of our results when these restrictions are relaxed.

5.1 Nonexpected utility preferences

Suppose that a consumer's preferences over $D \times M$ are represented by a continuous real-valued functional v on $D \times M$ that is strictly increasing in $(-F, m)$ (where the c.d.f.'s are partially ordered by the first order stochastic dominance relation).[18] Here π is defined implicitly by $v(F, m) = v(G, \pi(G, F, m))$. Even if v

[18]For c.d.f.'s defined on a real interval I, G first order stochastically dominates F if $G(x) \leq F(x)$ for all x in I with a strict inequality for some x in I.

violates the independence axiom of expected utility theory,[19] Machina's (1982) analysis implies that, if $v(\cdot, m)$ is L_1-Fréchet differentiable on D, then it will behave locally as an expected utility functional: in our application, there is a absolutely continuous function $U(\cdot; m, F)$ on P such that

$$v(G, m) - v(F, m) = \int_P U(p; m, F)d(G - F) + o(||G - F||).$$

The function U is called the *local utility function* of $v(\cdot, m)$ at F. Since we consider small changes in the price distribution, one might conjecture that a version of Theorem 1 holds for non-expected utility preferences by imposing a condition such as additive separability on each local utility function of v. This conjecture, however, is false. The role of additive separability of the indirect utility function is to ensure that ECS ranks distributions the same way as the consumer (Lemma 1). If the agent violates independence, then no restriction (short of independence) can restore that agreement. Indeed, an adaptation of the proof that (a) implies (b) in Theorem 1 yields the following result.

Corollary 1 *Let v be a real-valued functional v on $D \times M$ that is strictly increasing in $(-F, m)$. If $v(\cdot, m)$ violates independence for some $m \in M$, then there is an $F \in D$ such that $\pi(G, F, m) - \pi_{cs}(G, F, m) \neq o(||G - F||)$.*

Similarly the conclusion of Theorem 2 applies to all nonexpected utility preferences as well. Interestingly, however, the conclusion of Theorem 3(a) does carry over to smooth nonexpected utility preferences: for paths linear on the space of random variables and a nonrandom initial price, the percentage error from using the change in ECS to measure the willingness to pay tends to zero as the change in distribution tends to zero.

Corollary 2 *In addition to the hypotheses of Corollary 1, suppose that v is continuously differentiable on $D \times M$, with $v_2(F, m) > 0$, and that the local utility function of $v(\cdot, m)$ is continuously differentiable in the price. Then the conclusion of Theorem 3(a) obtains.*[20]

5.2 Multivariate risk

Rogerson (1980) and Turnovsky et. al. (1980) each allow other arguments of the consumer's indirect utility function to be random variables.[21] In this more

[19] A functional f on a convex set C of c.d.f.'s satisfies the Independence Axiom if for any F, G, H in C, and any real number $\lambda \in [0, 1)$, $f(F) \geq f(G)$ if and only if $f(\lambda F + (1 - \lambda)H) \geq f(\lambda G + (1 - \lambda)H)$.

[20] To prove the corollary, follow the steps in the proof of Theorem 3(a), using equation (8) on p. 296 in Machina (1982) to carry out the differentiation of v with respect to α. Note that, if $\psi(\cdot, m, F)$ is the local utility function for v at (m, F), then Roy's Identity may be expressed as $d(p, m) = -\psi_1(p, m, \delta_p)/v_2(\delta_p, m)$. This Corollary is related to the distinction between orders of risk attitudes (e.g. Segal and Spivak, 1997): if v and its local utility function are smooth, then (like ECS) the attitude towards price risk is *not* of order 1.

[21] Rogerson permits other prices, income and preference parameters to be random and treats them as exogenous: as the distribution of the price of good one changes the marginal distribution of these other random variables is held fixed. Turnovsky et. al. do not allow income or preference parameters to be random, but they allow the marginal distribution of any subset of prices to change.

general setting, they find that expected consumer's surplus ranks changes in the distribution of price one the same way as expected utility if and only if the marginal utility of income is independent of *all* random variables entering the indirect utility function. One can show that our Theorem 2 extends to this case as well: additive separability in prices and income is replaced by additive separability in income and the entire vector of random variables. (For multiple price changes under certainty, consumer surplus depends on the path of integration between the two price vectors. Nonetheless, for any path, the percentage error from using consumer surplus under certainty tends to zero as the price change tends to zero–even though consumer's surplus does not in general represent *any* preference relation over price vectors.)

5.2.1 Income uncertainty

The case of income uncertainty is especially interesting. For concreteness, suppose that only income and the price of good one are random. Rogerson's Theorem 1 implies that expected consumer's surplus ranks distributions of price one correctly (for all marginal distributions of income) if and only if the indirect utility function is quasilinear. Thus the extension of our Theorem 1 to this case implies that expected consumer's surplus *approximates* the willingness to pay if and only if it exactly *equals* the willingness to pay.

Now consider an economy of consumers who are endowed with a nonrandom income. In the face price uncertainty, they might have an incentive to open security markets to reallocate risk bearing, thus introducing *endogenous* income uncertainty. If all consumers have additively separable indirect utilities that are concave in income, then preferences over income gambles are state-independent: the nonrandom income endowment is Pareto efficient and thus such markets will not be active. If however some consumers have indirect utilities that are not additively separable, then the income endowment will not generally be efficient, leading to active securities markets. Even consumers with additively separable indirect utilities might participate in these markets, possibly invalidating the approximate validity of expected consumer's surplus (unless utility is quasilinear). Although this argument is not definitive, it suggests one more reason why expected consumer's surplus is apt to be a poor measure of willingness to pay.

5.2.2 Demand-side uncertainty

Rogerson also suggests (Theorem 4) that if consumer demand depends on random preference parameters, then ECS will not generally rank price distributions correctly. This result is noteworthy since many applications of expected consumer's surplus in Industrial Organization consider demand-side uncertainty–e.g. uncertainty about consumer preferences. Let $u(x, \theta)$ denote the von-Neumann Morgenstern utility, where x is the commodity vector and θ a preference parameter. He supposes that, for some $1 \leq k < \ell$, we have that $\partial^2 u/\partial x_i \partial\theta > 0$ for $i = 1, ..., k$; and $\partial^2 u/\partial x_i \partial\theta = 0$ for $i = k+1, ..., \ell$. His Theorem 4 asserts that it is impossible under these conditions for the marginal utility of income to be independent of prices 1 through ℓ and the preference parameter θ.[22] Hence the conditions of his Theorem 1 are apt to fail (and with them, the extension of our approximation results to multivariate risk).

In many applications, however, it is assumed that consumers know their own preferences: aggregate demand, hence prices, are uncertain to all consumers and firms because a consumer's preferences are private information. In this case, a consumer's marginal utility of income need not be independent of his preference parameter for ECS to rank distributions correctly. Moreover, even if a consumer is uncertain about states of the world that affect his preferences, then ECS can still satisfy the Pareto criterion over the different types of that consumer: if ECS rises, then at least one type of that consumer will be better off. (The argument essentially replicates the proof of Rogerson's Theorem 2). For these important applications, it seems that the argument against using expected consumer's surplus must rest on other grounds.

6 Conclusion

Expected consumer's surplus allows economists to evaluate welfare under uncertainty using only demands, without directly specifying the preferences and endowments of consumers. Most scholars who use it would undoubtedly agree that the conditions for its exact validity–risk neutrality over income gambles and a zero income elasticity–are severe; but they would likely justify it by arguing that it provides a useful approximation for cost-benefit analysis, based loosely on results from the literature on consumer's surplus under certainty. Our results imply that expected consumer's surplus can be a poor approximation to willingness to pay for a distribution change even if the change is small or if risk aversion, the income elasticity and the good's budget share are small. Expected consumer's surplus is obviously still useful for counterexamples: if the sum of expected consumer's surplus and expected profit is higher under policy A than B, then B cannot *always* be preferable to A. But the more ambitious and important argument that policy A is socially preferable to B for a range of plausible economic environments requires much more.

[22]Strictly speaking, an additional hypothesis is needed to make the proof valid: the Hessian matrix of the utility function must have a nonzero determinant, which is not implied by his appeal to the implicit function theorem. Consider the following C^2 class of additively separable utilities: $u(x,\theta) = f(x_1, ..., x_k, \theta) + g(x_{k+1}, ..., x_\ell)$, where g is homogenous of degree 1, concave and strictly quasiconcave; and the Hessian of f with respect to $(x_1, ..., x_k)$ is globally negative definite. The indirect utility is both quasilinear in income, and additively separable in (θ, m), so that ECS certainly represents the consumer's preferences. But if $\partial^2 f/\partial x_i \partial\theta > 0$ for all $1 \leq i \leq k$, then the conditions of his Theorem 4 are met.

7 Appendix

Proof of Lemma 2: Since V_2 exists on M, the Mean Value Theorem implies that (for $||G-F||$ small enough) $V(p, m-\pi(G,F,m)) = V(p,m) - V_2(p, m - t(G))\pi(G,F,m)$, where $0 < t(G) < \pi(G,F,m)$. Thus

$$\int V(p, m-\pi(G,F,m))dG = \int V(p,m)dG - \pi(G,F,m)\int V_2(p,m-t(G))dG.$$

Substituting this expression into (1) yields that

$$\pi(G,F,m) = \frac{\int V(p,m)d(G-F)}{\int V_2(p,m-t(G))dG}.$$

Thus

$$\begin{aligned}\pi(G,F,m) &= \frac{\int_P V(p,m)d(G-F)}{\int V_2(\omega,m)dF(\omega)} \qquad (12)\\ &+ \int V(p,m)d(G-F)\left(\frac{1}{\int V_2(p,m-t(G))dG} - \frac{1}{\int V_2(p,m)dF}\right).\end{aligned}$$

Since $V_1(\cdot, m)$ is continuous on the compact set P, V_1 is bounded by some number K. This fact implies that $\int V(p,m)d(G-F)/||G-F||$ is bounded: integrating by parts yields that

$$\left|\frac{\int V(p,m)d(G-F)}{||G-F||}\right| = \frac{\left|\int (G-F)V_1(p,m)dp\right|}{||G-F||} \leq \frac{\int |(G-F)|\,|V_1(p,m)|\,dp}{||G-F||} \leq K.$$

The result follows if the term in parenthesis on the right side of (12) converges to zero as $||G-F||$ tends to zero. Since the L_1 norm on D metrizes the topology of weak convergence (Machina, 1982, Lemma 1), it suffices to show that

$$\lim_{n\to\infty}\int V_2(p, m-t(G_n))dG_n = \int V_2(p,m)dF \qquad (13)$$

for any sequence G_n in D converging (weakly) to F. Since $\pi(\cdot, F, m)$ is continuous, $t(G_n)$ converges to 0. And since V_2 is continuous on the compact set P, $V_2(p, m-t(G_n))$ converges to $V_2(p,m)$ *uniformly* in p. Hence (13) holds[23] and the conclusion follows. ■

[23]Let f_n be a sequence of real-valued functions on P that converges uniformly to the continuous function f and let G_n be a sequence converging to G. We have

$$\int f_n dG_n = \int (f_n - f)dG_n + \int f dG_n.$$

The second integral converges to $\int f dG$ by the definition of weak convergence. And the first integral converges to zero since f_n converges to f uniformly:

$$\left|\int (f_n - f)dG_n\right| \leq \int |f-f_n|dG_n \leq \sup_{p\in P}|f-f_n| \to 0.$$

References

Baron, D. and R. Myerson, "Regulating a Monopolist with Unknown Costs," *Econometrica*, 50 (1982): 911-930.

Bartle, R., *The Elements of Integration and Lebesgue Measure*. New York: John Wiley and Sons, 1995.

Blundell, R., P. Pashardes and G. Weber, "What Do we Learn About Consumer Demand Patterns from Micro Data?," *American Economic Review*, 83 (1993): 570-597.

Deneckere, R. H. Marvel and J. Peck, "Demand Uncertainty and Price Maintenance: Markdowns as Destructive Competition," *American Economic Review,* 87 (1997): 619-641.

Hausman, J., "Exact Consumer's Surplus and the Deadweight Loss," *American Economic Review*, 71 (1981): 662-676.

Lewis, T. and D. Sappington, "Regulating a Monopolist with Unknown Demand," *American Economic Review*, 78 (1988): 986-998.

Machina, M., "'Expected Utility' Analysis without the Independence Axiom," *Econometrica* 50 (1982): 277-323.

Mas-colell, A. M. Whinston and J. Green. *Microeconomic Theory* . New York: Oxford University Press, 1995.

Milgrom, P. and I. Segal, "Envelope Theorems for Arbitrary Choice Sets," Stanford University Working Paper (2000).

Rey, P. and J. Tirole, "The Logic of Vertical Restraints," *American Economic Review*, 76 (1986): 921-939.

Rogerson, W.P. "Aggregate Expected Consumer Surplus as a Welfare Index with an Application to Price Stabilization," *Econometrica,* 48 (1980): 423-436.

Segal, U. and A. Spivak, "First-Order Risk Aversion and Non-differentiability," *Economic Theory,* 9 (1997): 179-183.

Stennek, J. "The Expected Consumer's Surplus as a Welfare Measure," *Journal of Public Economics* 73 (1999) 265–288.

Turnovsky, S. H. Shalit, and A. Schmitz, "Consumer's Surplus, Price Instability and Consumer Welfare," *Econometrica,* 48 (1980): 135-152.

Vives, X. "Small Income Effects: A Marshallian Theory of Consumer Surplus and Downward Sloping Demand," *Review of Economic Studies,* 54 (1987): 87-103.

Vives, X. *Oligopoly Pricing: Old Ideas and New Tools.* Cambridge: MIT Press, 1999.

Wang, T. "L_p-Fréchet Differentiable Preferences and 'Local Utility' Analysis," *Journal of Economic Theory*, 61 (1993): 139-159.

Willig, R.D. "Consumer's Surplus Without Apology," *American Economic Review*, 66 (1976): 589-597.

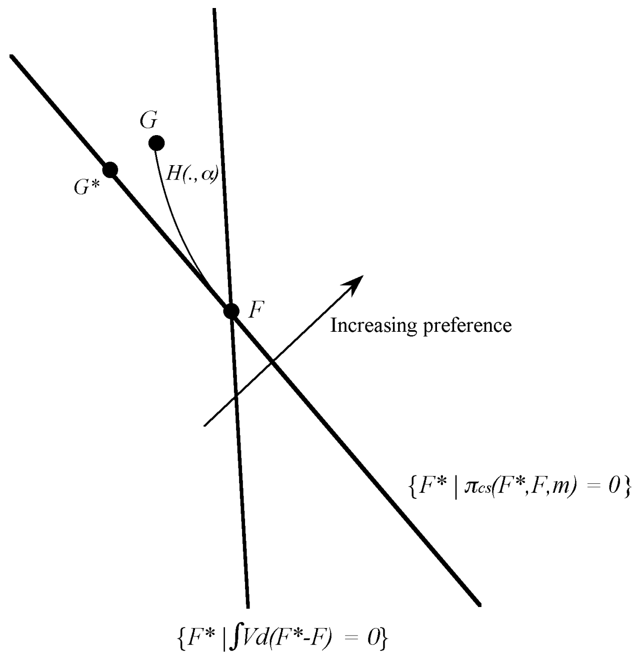

Figure 1. Why the percentage error from using ECS is unbounded: for any sequence G_n along the path H, π_{cs} tends to 0 faster than $\pi - \pi_{cs}$.

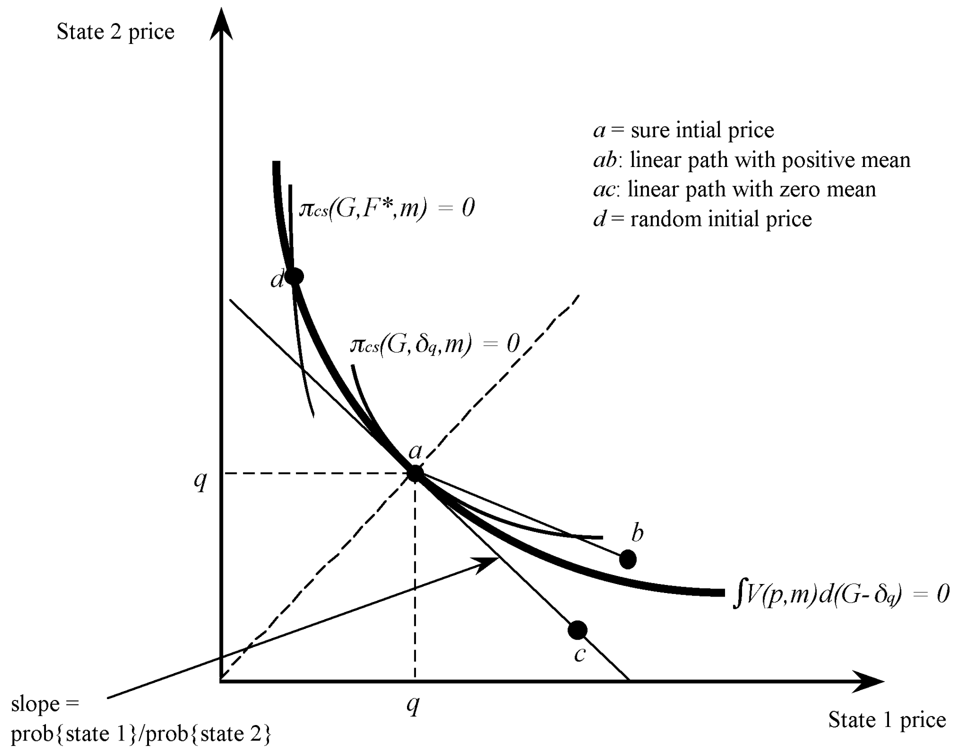

Figure 2: Illustration of Theorem 3

If the initial nonrandom price is a and the path, such as ab, is not mean-preserving, then ECS represents the consumer's preferences along the path close to a. If the indirect utility is not additively separable, then the indifference curves will cross somewhere off the certainty line, such as d. If the initial distribution is d, then we can construct a linear path to d such that ECS ranks path elements the same way as *minus* expected utility; hence the percentage error from ECS does not vanish along such a path close to d. Along the certainty line, the slopes of the indifference curves are equal. Hence for mean preserving paths (such as ca), the relative *curvature* of the indifference curves matters.