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 August 4, 2003

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. If, however, the distribution change is smooth on the space of random variables, and either the initial price is nonrandom or state-contingent payments are possible, then the change in ECS might approximate the willingness to pay well. Unfortunately, this smoothness condition necessarily fails in some important applications of ECS.

JEL Classifications: D6, D8
Keywords: Expected Consumer's Surplus, Welfare under Uncertainty, Local Cost-Benefit Analysis

---

[1] I thank Hector Chade, Glenn Ellison, Peter Hammond, Manuel Santos and seminar participants at Arizona State and Yale 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. Its popularity is easy to explain: it allows economists to evaluate consumer welfare under uncertainty using only demands, without directly specifying the preferences and endowments of consumers.

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 distributions 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). In many applications, however, expected consumer's surplus is not simply used to rank random alternatives for consumers: it is added to the change in expected profit to evaluate overall welfare. For this *sum* to be a valid welfare index in the sense that it rises if and only if the change is a potential Pareto improvement (money can be redistributed to make everyone better off after the change) expected consumer's surplus must not only represent each consumer's preferences, but must equal each consumer's *ex ante* willingness to pay for the change, a much more demanding requirement (see our Remark 1). Indeed, expected consumer's surplus equals the willingness to pay if and only if the indirect utility is quasilinear in income (e. g. Stennek, 1999), implying both that the demand for good is independent of income and that the consumer is risk neutral over income gambles.

Although income effects might often be small, risk aversion over income gambles is not. To illustrate how risk aversion can dramatically affect welfare results, consider the regulation of a monopolist who has private information about its costs. If we use expected total surplus to evaluate outcomes, then the highest possible price under the optimal policy can exceed the highest price for an unregulated monopoly (Baron and Myerson, 1982: 922). Suppose that this is the case and imagine that we take a concave transformation of each consumer's indirect utility function. This transformation does not affect demand or ECS but makes each consumer income risk averse. If all consumers are sufficiently risk averse, then all of them will prefer no regulation at all to the 'optimal' regulatory policy: the worst possible outcome for consumers is under the policy that evaluates outcomes by expected total surplus.

A natural question is how well the change in expected consumer's surplus approximates the willingness to pay when consumers are risk averse and income affects demand. For nonrandom prices, Willig (1976; 1979) 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.

Weitzman (1988) derives a complementary approximation result for multiple price changes. His summary (p. 552) of the issue for price changes under certainty is a useful starting point for us. Paraphrasing slightly: *As a situation changes smoothly from one price-income configuration to another, consumer's surplus is likely to approximate a consumer's willingness to pay well.* We pose two questions: first, in what sense does this assertion remain true for price changes under uncertainty; and, second, can this sense provide a foundation for using ECS to evaluate welfare under uncertainty in applied microeconomic models?[2]

To begin, we show that some natural extensions of this approximation argument fail if prices are random. First, the difference between the change in ECS and willingness to pay is of higher order than the L_1 distance between the cumulative distribution functions (c.d.f.'s) of price 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 2). If, however, additive separability fails, then the percentage error from using ECS is unbounded for small distribution changes; and 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 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 demand studies (e.g. Blundell et. al., 1993).

We also show that ECS might be a good approximation in two cases: if the initial price is not random (Theorem 3); or if state-contingent payments are possible (Proposition 3). Each result assumes that the price change is smooth on the space of random variables. Unfortunately, this smoothness condition fails in some important applications of ECS (Lemma 3). Moreover, we argue that the appeal to state-contingent payments as a defense of ECS is problematic. We illustrate these points with two examples from the literature: information acquisition and sharing in oligopoly; and minimum resale price maintenance (Section 4).

The reader may wonder why we emphasize a *zero* limiting percentage error for ECS. If the limiting percentage error is small, then perhaps the added benefit from modeling consumers more carefully does not justify the cost. But as Hausman (1981) points out, the percentage error in calculating the overall welfare change can be large even if the percentage error in the change in consumer

[2]Vives (1987, Proposition 1) shows that, under strong smoothness and curvature assumptions, the percentage error from using consumer's surplus is at most of order $1/\sqrt{\ell}$, where ℓ is the number of goods. Since increasing the number of goods, by itself, implies nothing about risk attitudes, it is clear that extending this result to price changes under uncertainty requires adding an assumption to ensure that risk aversion vanishes as ℓ increases.

welfare is small. Essentially the only way to ensure a small limiting percentage error in the the overall welfare change is a zero limiting error in measuring the change to consumers.[3]

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 consumption 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 consumer is endowed with income m. The price of good 1 is random, taking on values in a 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

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

where V is the consumer's von Neumann-Morgenstern indirect utility function. Recall that V is continuous, increasing in $(-p, m)$, and quasiconvex. 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 consumption bundles; and risk-preferences, which depend on the particular representation of preferences over non-random bundles.

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 :

$$\int V(p, m)dF = \int V(p, m - \pi(G, F, m))dG.^4 \tag{1}$$

The number $\pi(G, F, m)$ is sometimes called the *ex ante* compensating variation for the change. Clearly, $\pi(G, F, m) \geq 0$ if and only if the consumer weakly prefers G to F at income m, and, for any $G \in D$, $-\pi(G, \cdot, m)$ represents the

[3] If the derivative of both consumer and total surplus with respect to the policy parameter is *non*zero and the limiting percentage error from using consumer surplus to measure consumer welfare is zero, then the limiting percentage error in measuring the total welfare change is zero. In Hausman's (1981: 672-73) example of the welfare loss of a tax, the derivative of total surplus with respect to the tax level is zero (at a tax of zero), and the percentage error in calculating the overall welfare change can remain large as the tax tends to zero. Our Theorem 3(b) is an uncertainty analogue of this phenomenon.

[4] 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.

consumer's preferences on $D \times \{m\}$. (Note, however, that $\pi(\cdot, F, m)$ need not represent preferences on $D \times \{m\}$.)

Let $(p, m) \mapsto d(p, m)$ denote the consumer's demand function for good 1. Since we are only interested in changes in consumer's surplus and prices lie in $P \equiv [\underline{p}, \overline{p}]$, we lose nothing by setting consumer's surplus at p equal to $\int_p^{\overline{p}} d(\omega, m) d\omega \equiv cs(p, m)$ for any $p \in P$.[5] 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. 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 exploit tools used in the literature on smooth 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 condition ensures that π is differentiable.

Definition 1 *V is* ***regular*** *if it is continuously differentiable on $P \times M$ with $V_2 > 0$.*[6]

We assume that the only uncertainty that a consumer faces is over the price of the good. In particular, consumers know their own preferences over goods

[5]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.

[6]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 [2002].)

when the price distribution changes. Whether this assumption is reasonable depends on why prices are random. Most applications of ECS consider one of two possibilities: the firms' costs are uncertain; or demand is uncertain. Pure cost uncertainty poses no problem for our assumption of known preferences; demand uncertainty obviously might, since it is often interpreted as preference uncertainty. If, however, aggregate demand is uncertain to each consumer and firm simply because consumer preferences are *private information*–the consumer and *only* the consumer knows his own preferences–then our results apply. In this case, we must interpret the c.d.f. of price to be the consumer's posterior belief about prices, conditional on knowing his own preferences. What we exclude is that consumer preferences depend on some event that no one knows when the price distribution changes. (See Section 5.3.)

3 Expected Consumer's Surplus as a Welfare Measure

Three related preliminary questions are, first, when does expected consumer's surplus represent the consumer's preferences over D; second, when is π_{cs} precisely equal to π; and, third, what are the grounds for using aggregate ECS as a welfare index for consumers as a whole?

3.1 Expected consumer surplus as a representation of individual and social preferences

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

Lemma 1 *Let V be regular. The following two assertions are equivalent.*[7]

(a) *For any $G, F \in D$ and $m \in M$, $\pi_{cs}(G, F, m) \geq 0$ 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, $r = -mV_{22}/V_2$.

It is immediate from Lemma 1 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

[7] 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: 73).

consumer's surplus represents the consumer's preferences on D, so 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 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.

Rogerson (1980) also considers aggregate expected consumers' surplus as a representation of social preferences in the sense that it is *Pareto consistent*: if *all* consumers prefer G to F, then aggregate expected consumers' surplus is higher under G than F. He shows (Theorem 2) that if each consumer's indirect utility function is additively separability in price and income, then aggregate expected consumers' surplus is Pareto consistent.

Pareto consistency is a weak normative requirement: it is satisfied if just one consumer is better off whenever aggregate ECS rises. Most applications of ECS usually demand that consumers "as a whole" are better off in some sense when aggregate expected consumers' surplus rises. A usual formalization of this requirement is *Kaldor consistency*: the change in aggregate ECS is positive if and only if the change is a potential Pareto improvement (income can be redistributed among consumers so that they all prefer the change). It turns out that Kaldor consistency of aggregate ECS is essentially equivalent to quasilinear utility. To show this, let there be a continuum of consumers, indexed by $i \in [0,1]$, and a finite number of von Neumann-Morgenstern indirect utility functions indexed by $\tau \in T$. The distribution of preferences and incomes (attributes) is given by a (Borel) measurable function $a : [0,1] \to T \times (\underline{m}, \overline{m})$. We let $\mathcal{A}$ denote the set of such measurable functions. Let $\pi_{cs}(G, F, m, \tau)$ be the change in expected consumer's surplus and $\pi(G, F, m, \tau)$ the willingness to pay for a change in the price distribution from F to G for a consumer with indirect utility function $V(\cdot, \cdot; \tau)$ and income m. We say that aggregate expected consumers' surplus is Kaldor consistent if for all G, F in D and all distribution of attributes $a \in \mathcal{A}$, we have $\int_{[0,1]} \pi_{cs}(G, F, a(i))di \geq 0$ if and only if $\int_{[0,1]} \pi(G, F, a(i))di \geq 0$.

Proposition 1 *Suppose that for each $\tau \in T$, $V(\cdot, \cdot, \tau)$ is regular and twice differentiable in m on $(\underline{m}, \overline{m})$. Aggregate expected consumers' surplus is Kaldor consistent if and only if for each $\tau \in T$, $V(\cdot, \cdot; \tau)$ is additively separable in p and m on $P \times M$ and one of the following conditions holds:*

- *(a) Each consumer $i \in [0,1]$ has the same preferences on D.*
- *(b) For every $\tau \in T$, $V(\cdot, \cdot; \tau)$ is affine in income: $V(p, m, \tau) = f(p, \tau) + m$ for some real-valued function f on $P \times T$.*

Proof: appendix.

Remark 1 *Proposition 1 just considers consumer welfare. As mentioned already, the change in aggregate ECS is often added to the change in expected profit (minus any other cost) from a policy change to evaluate overall welfare. If this sum is to be Kaldor consistent for a rich enough class of environments, then condition (b) of Proposition 1 must hold. To see this, suppose that the good* *is produced by a competitive industry under constant returns, but that the (common) unit cost is unknown to the firms when they choose how much to produce. In this case expected profit is always zero and the price distribution is the same as the unit-cost distribution. Suppose that by spending C dollars, the government can change the distribution of unit cost, and that consumers' aggregate willingness to pay for this change in the cost distribution is positive. If the aggregate willingness to pay for the change does not equal the change in aggregate expected consumers' surplus, then just set C to be in between those two numbers to generate a Kaldor inconsistency from using ECS. Hence the two measures must be equal. The requirement that $\int_{[0,1]} \pi_{cs}(G, F, a(i))di = \int_{[0,1]} \pi(G, F, a(i))di$ for all $a \in \mathcal{A}$ then implies that condition (b) holds.*

Proposition 1 and Remark 1 imply that additive separability cannot in general rationalize aggregate ECS as a measure of consumer welfare for arbitrary policy changes. We emphasize small policy changes in what follows.

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, $||f|| = \int_P |f|dp$. (All integrals with respect to price will be over the interval P, a dependence we will sometimes omit.)

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

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

where $o(\cdot)$ is a real-valued function satisfying $o(0) = 0$ and $\lim_{h \to 0} o(h)/h = 0$.[8]

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 $V(p, m) = f(p) + g(m)$. From Lemma 2,

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

[8]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.

By Roy's Identity (see footnote 7), $d(p,m) = -V_1/V_2 = -f'(p)/g'(m)$ on $P \times M$. Thus

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

so that

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

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, for any $\alpha \in [0, 1]$,

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

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

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

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

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

Proposition 2 *If V is regular and additively separable on $P \times M$, then for any sequence $\langle G_n \rangle$ 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 let $\langle G_n \rangle$ be any sequence satisfying the conditions of the Proposition. By the Mean Value Theorem there is (for n large enough) a 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 $\pi(G_n, F, m) = \int f(p) d(G - F)/g'(t_n)$. Since g' is continuous and positive on M, we have (using (3)) that

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

If V is not additively separable, then the percentage error from using expected consumer's surplus is unbounded. Moreover, if the percentage error does tend to zero for a particular sequence of c.d.f.'s with a nondegenerate limit F, then a small perturbation of the 'direction' that the sequence approaches its limit will result in a nonzero limit.

To formalize this second point, we consider smooth paths on D. A *path* on D to F is a continuous mapping from $[0, 1]$ into D which equals F at 0. We denote a path by $\langle H(\cdot, \alpha) : \alpha \in [0, 1]\rangle$ (or by $\langle H(\cdot, \alpha)\rangle$ for short). We say that a path $\langle H(\cdot, \alpha) : \alpha \in [0, 1]\rangle$ is *smooth at* $\alpha = 0$ if for every $p \in P$ the derivative $H_2(p, 0)$ exists and the family of functions $\langle (H(\cdot, \alpha) - H(\cdot, 0)) / \alpha : \alpha \in (0, 1]\rangle$ is bounded uniformly on $P \times (0, 1]$. (Since $0 \leq \alpha \leq 1$, it should be understood that all limits as α tends to 0 are from the right.) In part (b) of the next result, we assume that the marginal utility of income, $V_2(\cdot, m)$, is not F-*a.e.* constant. This implies that V is not additively separable and that F is 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 path $\langle H(\cdot, \alpha)\rangle$ in D to F such that $\pi_{cs}(H(\cdot, \alpha), F, m) \neq 0$ for all $\alpha \in [0, 1]$ and

$$\lim_{\alpha \to 0} \left| \frac{\pi(H(\cdot, \alpha), F, m) - \pi_{cs}(H(\cdot, \alpha), F, m)}{\pi_{cs}(H(\cdot, \alpha), F, m)} \right| = \infty. \tag{4}$$

(b) Let F be nondegenerate and let $\langle \widehat{H}(\cdot, \alpha)\rangle$ be a 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 $\langle H(\cdot, \alpha)\rangle$ to F with $||\widehat{H}_2(\cdot, 0) - H_2(\cdot, 0)|| < \varepsilon$ such that

$$\frac{\pi(H(\cdot, \alpha), F, m) - \pi_{cs}(H(\cdot, \alpha), F, m)}{\pi_{cs}(H(\cdot, \alpha), F, m)} \tag{5}$$

does not tend to zero as $\alpha \to 0$.

Proof: appendix.

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 $\langle H(\cdot, \alpha)\rangle$ 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 $H_2(p, 0)$ lies in the tangent space 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.[9] Theorem 2(b) implies that a zero limiting percentage error is a knife-edge phenomenon when additive separability fails. Indeed we can say something a little stronger than this.

Remark 2 ***(Genericity)*** *Fix an $(F, m) \in D \times M$ and suppose that V is regular, but that $V_2(\cdot, m)$ is not F-a.e. constant in p. Consider the following pseudo-metric, ρ, on the space of smooth paths to F: $\rho(\widehat{H}(\cdot, \alpha), H(\cdot, \alpha)) = \|\widehat{H}_2(\cdot, 0) - H_2(\cdot, 0))\|$. If we treat as equivalent any two paths whose derivative with respect to α at 0 is Lebesgue a.e. equal, then the space of smooth paths to F with metric ρ becomes a metric space. It is then straightforward to modify the proof of Theorem 2 to show that the set of smooth paths for which the percentage error from using ECS does not tend to zero contains an open dense subset of the set of smooth paths to F.*

3.2.2 Smooth paths on the space of random variables

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 Theorem 2 suggests that 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. Smooth paths on D, however, exclude price changes under certainty. If we consider other classes of paths, then we can avoid at least some of the grim consequences of Theorem 2.

Besides paths on the space of c.d.f.'s, the most commonly used paths in the economics of uncertainty are those on the space of random variables. Let $(\Omega, \mathcal{B}, \nu)$ be a probability space and $\langle q(\cdot, \alpha) : \alpha \in [0, 1]\rangle$ a family of random variables defined on it with range in P (which we will often denote simply by $\langle q(\cdot, \alpha)\rangle$). If $q(\omega, \cdot)$ is continuous at 0 for each $\omega \in \Omega$, then the family is a *path* to $q(\cdot, 0)$. The path $\langle q(\cdot, \alpha)\rangle$ is *smooth at* $\alpha = 0$ if the derivative $q_2(\omega, 0)$ exists for each $\omega \in \Omega$ and if the family of functions $\langle (q(\cdot, \alpha) - q(\cdot, 0)) / \alpha : \alpha \in (0, 1]\rangle$ is bounded uniformly on $\Omega \times (0, 1]$. Let G_x denote the c.d.f. of the random variable x. Given the initial price distribution F (defined by $F(p) = \int_{\{\omega | q(\omega, 0) \leq p\}} d\nu(\omega)$ for all p) we consider final distributions of the form $G_{q(\cdot, \alpha)}$. Define $\Pi(\alpha) = \pi(G_{q(\cdot, \alpha)}, F, m)$ and $\Pi_{cs}(\alpha) = \pi_{cs}(G_{q(\cdot, \alpha)}, F, m)$. And let $\mu(\alpha) = \int q(\omega, \alpha) d\omega$, the mean price at α. If a path is smooth at 0, then $\mu(\cdot)$ is differentiable at 0 and $\mu'(0) = \int q_2(\omega, 0) d\omega$. Let $\sigma^2 = \int (q_2(\omega, 0))^2 d\omega - (\mu'(0))^2$, the variance of the first-order effect of α on the price. Finally for a path that is regular at zero define (in an abuse of notation)

$$\mu''(0) = \lim_{\alpha \to 0} \frac{\int (q(\omega, \alpha) - q(\omega, 0)) / \alpha d\nu(\omega) - \mu'(0)}{\alpha}$$

whenever the limit exists. The next result gives some consequences of this notion of smoothness under the assumption that the initial price is not random.

[9] 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.)

Theorem 3 *Let V be regular, let $p \in P$, and suppose that the path $\langle q(\cdot, \alpha)\rangle$ is smooth at $\alpha = 0$ with $q(\omega, 0) = p$ ν-a.e. and $\Pi_{cs}(\alpha) \neq 0$ for all $\alpha \in (0, 1]$.*

(a) If $\mu'(0) \neq 0$, then

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

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

$$\lim_{\alpha \to 0} \frac{\Pi(\alpha) - \Pi_{cs}(\alpha)}{\Pi_{cs}(\alpha)} = \frac{s\,[\eta_m - r]}{-2p\frac{\mu''(0)}{\sigma^2} - \eta_p},$$

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)). If $\mu''(0) = 0$, then the formula for the percentage error simplifies to $-s\,[\eta_m - r]\,/\eta_p$.

Proof: Fix $(p, m) \in P \times M$. By the Mean Value Theorem, for each $(\alpha, \omega) \in (0, 1] \times \Omega$, there exists a point $\xi(\alpha, \omega) \in P \times M$ on the line segment between $(q(\omega, \alpha), m - \Pi(\alpha))$ and (p, m) such that

$$V(q(\alpha, \omega), m - \Pi(\alpha)) = V(p, m) + V_1(\xi(\alpha, \omega))(q(\omega, \alpha) - p) - V_2(\xi(\alpha, \omega))\Pi(\alpha).$$

After integrating, rearranging and dividing by α we get

$$\frac{\Pi(\alpha)}{\alpha} = \frac{\int V_1(\xi(\alpha, \omega))(\frac{q(\omega,\alpha) - p}{\alpha})d\nu(\omega)}{\int V_2(\xi(\alpha, \omega))d\nu(\omega)}.$$

Since each integrand is bounded and each has a limit as α tends to zero, the Lebesgue Dominated Convergence Theorem[10] (LDCT) implies that

$$\Pi'(0) = -d(p, m)\mu'(0). \tag{6}$$

Moreover, a similar application of the Mean Value Theorem and the LDCT implies that Π_{cs} is differentiable at $\alpha = 0$ with

$$\Pi'_{cs}(0) = -\int d(q(\omega, 0), m)q_2(\omega, 0)d\nu(\omega) = -d(p, m)\mu'(0). \tag{7}$$

Since $\mu'(0) \neq 0$, we have $\Pi'_{cs}(0) \neq 0$, so by L'Hôpital's Rule,

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

For part (b), $\mu'(0) = 0$ implies that $\Pi'(0) = \Pi'_{cs}(0) = 0$. Since V is twice continuously differentiable on $P \times M$, by Taylor's Theorem for each $(\alpha, \omega) \in$

[10] E.g. Billingsley (1986: Theorem 16.8(i)).

$(0,1] \times \Omega$, there is a point $\psi(\alpha,\omega) \in P \times M$ on the line segment between $(q(\omega,\alpha), m - \Pi(\alpha))$ and (p,m) such that

$$\begin{aligned} V(q(\alpha,\omega), m-\Pi(\alpha)) &= V(p,m) + V_1(p,m)(q(\omega,\alpha)-p) - V_2(p,m)\Pi(\alpha) \\ &\quad + \frac{1}{2}V_{11}(\xi(\alpha,\omega))(q(\omega,\alpha)-p)^2 - V_{12}(\xi(\alpha,\omega))\Pi(\alpha)(q(\omega,\alpha)-p) \\ &\quad + \frac{1}{2}V_{22}(\xi(\alpha,\omega))(\Pi(\alpha))^2. \end{aligned}$$

After integrating both sides of this equation, rearranging and dividing by α^2 we get

$$\begin{aligned} \frac{\Pi(\alpha)}{\alpha^2} &= -d(p,m)\frac{\mu(\alpha)-p}{\alpha^2} + \frac{1}{2}\int \frac{V_{11}(\xi)}{V_2(p,m)}(\frac{q(\omega,\alpha)-p}{\alpha})^2 d\nu(\omega) \\ &\quad - \frac{\Pi(\alpha)}{\alpha}\int \frac{V_{12}(\xi)}{V_2(p,m)}(\frac{q(\omega,\alpha)-p}{\alpha})d\nu(\omega) + \frac{1}{2}\left(\frac{\Pi(\alpha)}{\alpha}\right)^2 \int \frac{V_{22}(\xi)}{V_2(p,m)}d\nu(\omega). \end{aligned}$$

Similarly, for every $(\omega,\alpha) \in \Omega \times [0,1]$, there is a number $\theta(\omega,\alpha)$ between $q(\omega,\alpha)$ and p such that

$$\begin{aligned} \frac{\Pi_{cs}(\alpha)}{\alpha^2} &= -d(p,m)\frac{\mu(\alpha)-p}{\alpha^2} - \frac{1}{2}\int d_1(\theta(\omega,\alpha),m)(\frac{q(\omega,\alpha)-p}{\alpha})^2 d\nu(\omega) \qquad (8) \\ &= -d(p,m)\frac{\mu(\alpha)-p}{\alpha^2} \\ &\quad + \frac{1}{2}\int \left(\frac{V_{11}(\theta(\omega,\alpha),m)}{V_2(\theta(\omega,\alpha),m)} - \frac{V_{21}(\theta(\omega,\alpha),m)V_1(\theta(\omega,\alpha),m)}{V_2(\theta(\omega,\alpha),m)^2}\right)(\frac{q(\omega,\alpha)-p}{\alpha})^2 d\nu(\omega). \end{aligned}$$

It follows from differentiating Roy's Identity with respect to income that $V_{12} = (\eta_m - r)V_1/m$. Using this fact and the LDCT, we have

$$\begin{aligned} \lim_{\alpha\to 0}\left(\frac{\Pi(\alpha)}{\alpha^2} - \frac{\Pi_{cs}(\alpha)}{\alpha^2}\right) &= \frac{V_{21}(p,m)V_1(p,m)}{2V_2(p,m)^2}\sigma^2 \\ &= \frac{(\eta_m - r)d^2}{2m}\sigma^2. \qquad (9) \end{aligned}$$

By equation (8) and the assumption that $\mu''(0)$ exists,

$$\lim_{\alpha\to 0}\frac{\Pi_{cs}(\alpha)}{\alpha^2} = -d(p,m)\mu''(0) - d_1(p,m)\sigma^2/2 \neq 0. \qquad (10)$$

Since

$$\lim_{\alpha\to 0}\frac{\Pi(\alpha)-\Pi_{cs}(\alpha)}{\Pi_{cs}(\alpha)} = \lim_{\alpha\to 0}\frac{(\Pi(\alpha)-\Pi_{cs}(\alpha))/\alpha^2}{\Pi_{cs}(\alpha)/\alpha^2}$$

part (b) follows from combining (10) and (9). ■

Theorem 3(a) requires the initial price to be nonrandom and that $\mu'(0) \neq 0$— a small change in α from 0 has a first-order effect on the mean price. But part

(a) imposes no conditions on V other than regularity. Hence, the change in ECS approximates the willingness to pay well for this class of changes if the distribution change is small. Indeed, if $q(\cdot, \alpha)$ is a degenerate random variable for each $\alpha \in [0, 1]$, then the change collapses to the familiar case of a price change under certainty. Thus part (a) is a stochastic generalization of that case. Conceivably the initial price could be known either because it is regulated or because the initial environment is known; if so, then this result could sometimes justify ECS as a welfare measure for local cost-benefit analysis. As our analysis of information acquisition in Section 4.1 shows, however, the hypotheses of Theorem 3(a) sometimes fail even if the initial price is not random.

Even if the path is smooth, the conclusion of Theorem 3(a) generally fails unless V is additively separable: by Theorem 3(b) it fails when there is no first-order effect on the mean price (unless $\eta_m = r$);[11] and it generally fails if the initial price distribution is nondegenerate.

Remark 3 *Fix an initial price $q(\cdot, 0)$ defined on a probability space $(\Omega, \mathcal{B}, \nu)$ and suppose that V is regular, but $V_2(q(\cdot, 0), m)$ is not ν-a.e. constant on Ω. Consider all paths to $q(\cdot, 0)$ defined on $(\Omega, \mathcal{B}, \nu)$ that are smooth (at zero) and define a pseudo-metric ρ on this space by $\rho(\widehat{q}(\cdot, 0), q(\cdot, 0)) = \int |\widehat{q}_2(\omega, 0) - q_2(\omega, 0)| \, d\nu(\omega)$. If we identify any two smooth paths to $q(\cdot, 0)$ whose derivatives at $\alpha = 0$ are ν-a.e. equal, then the set of such paths endowed with ρ is a metric space. One can show that the set of smooth paths on $(\Omega, \mathcal{B}, \nu)$ to $q(\cdot, 0)$ for which the percentage error does not tend to zero contains an open dense subset of all such paths.*

Remark 4 *The condition in part (a) that the path be smooth at 0 with $\mu'(0) \neq 0$ can be replaced with various weaker conditions. For example, if (i) $q(\cdot, \alpha)$ converges to $q(\cdot, 0)$* uniformly*; and (ii) $\liminf_{\alpha \to 0} |\Pi_{cs}(\alpha)/\Lambda(\alpha)| > 0$, where $\Lambda(\alpha) = \sup_{\omega \in \Omega} |q(\omega, \alpha) - q(\omega, 0)|$, then the conclusion of Theorem 3(a) holds if $q(\cdot, 0)$ is constant. (Consider the path given by $q(\omega, \alpha) = f(\omega) + g(\alpha)\varepsilon(\omega)$ for all $(\omega, \alpha) \in \Omega \times [0, 1]$ where $f(\cdot)$, $g(\cdot)$ and $\varepsilon(\cdot)$ are bounded measurable functions, with $\int \varepsilon(\omega) d\nu(\omega) \neq 0$ and g is continuous at 0 with $g(0) = 0$. This path satifies (i) and (ii), but need not be smooth at $\alpha = 0$.) What is **not** sufficient, however, is simply the condition that $q(\cdot, \alpha)$ converge to p pointwise (Theorem 2), or even that $q(\omega, \cdot)$ be differentiable at $\alpha = 0$ for each $\omega \in \Omega$, since neither forces the supports to converge.*[12]

We now consider a simple example which illustrates why smoothness of paths on the space of random variables can fail in applications of ECS. For every

[11] If the derivative of expected *total* surplus is zero at $\alpha = 0$ under the hypotheses of Theorem 3, and the second derivative is nonzero, then the limiting percentage error from using ECS to measure the total welfare change is zero if and only if $\eta_m = r$ at the initial price. This result is one extension of Hausman's (1981) analysis of small taxes under certainty (note 3).

[12] Consider the path $\langle q(\cdot, \alpha) \rangle$ on the unit interval endowed with Lebesgue measure given by $q(\omega, \alpha) = 1$ if $0 < \omega < \alpha$, $q(\omega, \alpha) = 2$ otherwise. For each $\omega \in \Omega$, $q(\omega, \cdot)$ is differentiable at $\alpha = 0$ (the derivative is 0 for all ω), $\Pi'_{cs}(0) \neq 0$, yet the limiting percentage error is not zero unless V is additively separable: since this path is linear on D, a simple application of Lemma 2 gives the percentage error to be $((\Delta V / V_2(2, m)) - \Delta cs) / \Delta cs$, where $\Delta V = V(1, m) - V(2, m)$ and $\Delta cs = cs(1, m) - cs(2, m)$; if $V_2(\cdot, m)$ is strictly monotonic on P, this limiting percentage error is nonzero.

$\alpha \in [0,1]$, let the price take on just two values, $Q_2(\alpha)$ with probability $\lambda(\alpha)$ and $Q_1(\alpha)$ with probability $1-\lambda(\alpha)$. Suppose that $\lambda(\cdot)$ and $Q_j(\cdot)$, for $j=1,2$ are all continuous functions that are differentiable at $\alpha=0$. The question here is whether there is a probability space $(\Omega, \mathcal{B}, \nu)$ and a path defined on it with the same probability law as the family of distributions just described, and which is smooth at $\alpha=0$. In general the answer is 'no'; indeed the path will not even converge uniformly to the distribution at $\alpha=0$.

Lemma 3 *Suppose that $\lambda'(0) \neq 0$.*

(a) If $Q_2(0) \neq Q_1(0)$, then there is no probability space $(\Omega, \mathcal{B}, \nu)$ with path $\langle q(\cdot,\alpha)\rangle$ defined on it with the following properties: (i) $\nu\{\omega | q(\omega,\alpha) = Q_2(\alpha)\} = \lambda(\alpha)$ for all $\alpha \in [0,1]$; and (ii) $q(\cdot,\alpha)$ converges to $q(\cdot,0)$ uniformly as α tends to zero.

(b) If $Q_2(0) = Q_1(0)$, then there is a probability space and a path defined on it satisfying (i) which is smooth at $\alpha=0$ (so (ii) holds) with $\mu'(0) = Q_2'(0)(1-\lambda(0)) + Q_1'(0)\lambda(0)$.

Proof: appendix.

Since we allow $\lambda(0)$ to be 0 or 1, the price can be nonrandom at $\alpha=0$ even when $Q_2(0) \neq Q_1(0)$. The failure of uniform convergence has no counterpart in the certainty case and is one important reason why ECS may poorly approximate a consumer's willingness to pay. (The example in note 12 fits Lemma 3(a) for $\lambda(0)=1$, yet the percentage error does not tend to zero.) What drives failure of smoothness when $Q_2(0) \neq Q_1(0)$ is the that the *likelihood* of the prices, $\lambda(\alpha)$, varies with α. This leads to the conjecture that Theorem 3(a) (and its extension to random initial prices, Proposition 3) might not apply to policies that change the likelihood of prices, such as acquiring or spreading information or policies which make the realization of some supply states less likely (for example flood prevention in agriculture).

4 Two Applications

Our results have been in terms of abstract paths of c.d.f.'s or random variables. We now show how the issues stressed in the formal results arise naturally in applications that use ECS.

4.1 Information Acquisition and Sharing

Consider an industry in which firms are uncertain about either cost or demand. Firms can either acquire additional information, or share the information that they have. The information sharing literature is mainly concerned with two questions: how much information will firms acquire or share; and what is the welfare effect of the firms' decisions?[13]

[13]Novshek and Sonnenschein (1982) is an early contribution; Vives (1999: chapter 8) surveys this large literature; and Raith (1996) provides some unifying results.

Assume, as is common in this literature, that the industry's (single) good is produced under constant returns and that demand is linear. Suppose that the demand is known, but that the unit cost is unknown, having support in $[0, 1]$; denote it by $\widetilde{k}$.[14] Let the demand be $d(p, m) = a - p$, where $a > 1$, which comes from a single consumer with von Neumann-Morgenstern utility

$$u(x, y) = T(ax - (1/2)x^2 + y), \tag{11}$$

where x is the quantity of the good, y the expenditure on all other goods and $T(\cdot)$ is a differentiable, strictly concave function with $T' > 0$. The indirect utility function is $V(p, m) = T((a - p)^2/2 + m)$. Note that demand is independent of income (if income is high enough).

4.1.1 Information Acquisition

To begin, suppose that the good is produced by a monopolist. Before choosing its output (or equivalently price), the firm observes the realization of one out of a family of random variables, $\langle \widetilde{s}_\alpha : \alpha \in \mathbb{R}_+ \rangle$. Most of the information sharing literature assumes that information is *affine*. Specifically, for every $\alpha \geq 0$, $E[\widetilde{s}_\alpha | k] = k$ for all realizations of k, and the the posterior expectation of k, conditional on s, is affine in s: $E[\widetilde{k}|s] = A_\alpha + B_\alpha s$ for all realizations s of $\widetilde{s}_\alpha$ for some numbers A_α and B_α. ($E[\cdot]$ denotes the expectations operator.) Let v denote the prior precision of $\widetilde{k}$ and the let the index α equal the precision of the signal: $\alpha = \Big(E[Var(\widetilde{s}_\alpha|\widetilde{k})]\Big)^{-1}$.[15] Suppose that the firm can choose the precision of the signal at a cost, with more precise signals costing more. Under affine information,

$$E[\widetilde{k}|s] = \frac{\alpha}{\alpha + v}s + \frac{v}{\alpha + v}E[\widetilde{k}].$$[16]

The firm's output conditional on the signal realization s is $(a - E[\widetilde{k}|s])/2$ and so the price is $(a + E[\widetilde{k}|s])/2$. From an *ex ante* viewpoint (before the signal is realized) the output has mean $(a - E[\widetilde{k}])/2$ and variance $\alpha/4v(\alpha + v)$; the mean price is $(a + E[\widetilde{k}])/2$, which is independent of α. When $\alpha = 0$, both output and price are nonrandom, so at least one assumption of Theorem 3 holds.

Ex post consumer's surplus equals $x^2/2$, where x is the firm's output. It follows that

$$\Pi_{cs}(\alpha) = \frac{\alpha}{8v(\alpha + v)},$$

which is increasing in the information index α with

$$\Pi'_{cs}(0) = \frac{1}{8v^2} > 0.$$

[14]Since we do not use it, we do not specify the underlying probability space.

[15] As an example, suppose that the cost takes on two equally likely values, 0 and 1. Conditional on the cost being 0, the signal equals 0 for sure. If the cost equals 1, the signal equals $(2 + \alpha)/\alpha$ with probability $\alpha/(2 + \alpha)$ and 0 with the remaining probability. One can verify that the average precision is α, and that this information structure is affine.

[16]Ericson (1969); Li (1985: 523).

Since a small amount of information has a positive first-order effect on ECS, even though information does not affect the mean price, the path of prices cannot be smooth at $\alpha = 0$ (by equation (7)). In particular, the price distribution does not generally even converge uniformly to the price with no information (see Remark 4 and Lemma 3).[17]

Now consider the consumer's willingness to pay for better information. Let T in equation (11) be given by $T(z) = \sqrt{z - \xi}$, with $\xi = m$, the consumer's income. Then Π is given implicitly by

$$E\left[\sqrt{\frac{((a - E[\widetilde{k}|\widetilde{s}]))^2}{8} - \Pi(\alpha)}\right] = \sqrt{\frac{((a - E[\widetilde{k}]))^2}{8}},$$

so that $\Pi(\alpha) = 0$ for all α and $\Pi'(0) = 0$: the error from using ECS to measure the willingness to pay is 100% for a small increase in information starting from null information. Ignoring information costs, the firm's profit is

$$\frac{((a - E[\widetilde{k}]))^2}{4} + \frac{\alpha}{4(\alpha + v)}$$

which is increasing in α. If the cost of information with precision α always lies in between $\alpha/4v(\alpha + v)$ and $3\alpha/8v(\alpha + v)$, then the firm would acquire no information–which is the socially optimal decision. If we had used ECS to measure consumer welfare, however, then we would have concluded that the firm should acquire some information.

4.1.2 Information Sharing

Sometimes information sharing in an industry works much like information acquisition of a monopolist (Vives, 1999: 248). Suppose there are two firms with a common unit cost of k that is unknown. Each firm observes a private signal of this common cost. The signal distributions are identical and independent, conditional on the cost k, and the signals are affine, as before. Suppose that each firm can simply choose whether to share its information or not. In keeping with (almost all) the information sharing literature, we assume that the sharing decision is made before the signals are realized and that a this choice cannot be revoked afterwards. Let the firms have a common prior about the cost with precision v, and let α denote the precision of the pooled signal. If the firms unilaterally decide whether or not to share their information, the unique Nash equilibrium is that neither firm shares (Li, 1985, Proposition 3); moreover industry profit is lower when both firms share than when they do not (Li, 1985, Proposition 2). The industry output when the firms share information is

$$\frac{2}{3}\left(a - E[\widetilde{k}] + \frac{\alpha}{\alpha + v}(s - E[\widetilde{k}])\right)$$

[17] As α tends to zero, the average variance of the signals grows without bound. In the example in note 15, for each $\alpha > 0$, there is a positive probability that the firm's posterior expected cost equals 1, but the prior expected cost is 1/2.

for each s and when they do not share it is

$$\frac{2}{3}\left(a - E[\widetilde{k}] + \frac{3\alpha}{3\alpha + 4v}(s - E[\widetilde{k}])\right)$$

for each s (Li [1985: 529-30]). Li notes (1985: 534) that expected consumers' surplus and expected total surplus are both higher when information is shared: from the perspective of expected total surplus, firms do not share enough information.

A risk-averse consumer, however, may not prefer that firms share information: If we again set $T(z) = \sqrt{z - \xi}$, with $\xi = m$, then the consumer is indifferent about the firms' sharing decision: all the consumer cares about is the mean output, which is the same whether the firms share or not. Here the equilibrium outcome of no information sharing is optimal. Indeed, expected total surplus is Pareto inconsistent: it is higher when firms share information, but the firms strictly prefer not to share and the consumer is indifferent. (If we replace T in this example with a function that is more concave than it, then the consumer will prefers that the firms not share information.)

4.1.3 Extension to Other Information Structures

We noted in Section 4.1.1 that the path to the (nonrandom) price at null information was not smooth, or even uniformly convergent, under affine information. The failure of smoothness is not peculiar to affine information or the initial point being null information (a point relevant to our treatment of state-contingent payments in Section 5.1).

Suppose as before that a monopoly firm produces the good under constant returns with an unknown cost. Let the unit cost take on two (initially) equally likely distinct values, k_1 or k_2. As before we index the available information structures by $\alpha \in [0, 1]$. Suppose for simplicity that there are just two possible signal realizations, z_1 and z_2. Let $r_{ij}(\alpha)$ denote the probability that signal z_j is drawn if the cost is k_i, and suppose that each $r_{ij}(\cdot)$ is differentiable at $\alpha = 0$. (We do not require that $\alpha = 0$ corresponds to null information.) The prior probability of observing signal z_j is $\lambda_{z_j}(\alpha) = r_{1j}(\alpha)/2 + r_{2j}(\alpha)/2$. Let $Q_{z_j}(\alpha)$ denote the equilibrium price conditional on observing signal z_j from information structure indexed by α. Suppose that the firm's revenue (as a function of output) is C^2 with a negative second derivative and that it is active at both cost realizations. Then $Q_{z_j}(\cdot)$ is differentiable at $\alpha = 0$ for $j = 1, 2$. This situation fits Lemma 3.

Corollary 1 *Suppose that $Q_{z_2}(0) \neq Q_{z_1}(0)$, and that $r'_{22}(0) \neq -r'_{12}(0)$. Then there is no probability space $(\Omega, \mathcal{B}, \nu)$ with path $\langle q(\cdot, \alpha)\rangle$ defined on it with the following properties: $\nu\{\omega | q(\omega, \alpha) = Q_{z_j}(\alpha)\} = \lambda_{z_j}(\alpha)$ for all $\alpha \in [0, 1]$ and $j = 1, 2$; $q(\cdot, \alpha)$ converges to $q(\cdot, 0)$ uniformly.*

There are two ways that $Q_{z_2}(0) \neq Q_{z_1}(0)$. First, the structure could be informative at $\alpha = 0$; second, it could uninformative at $\alpha = 0$, but one of the two signals has zero probability at $\alpha = 0$, and both have positive probability otherwise. In the second case, a small increase in α starting from 0 provides lots of information; and since the optimal price is differentiable in α the price support 'implodes' at $\alpha = 0$. (See Examples 2 and 6 in Chade and Schlee [2002: 429, 440].) If, however, $\alpha = 0$ corresponds to null information, and $r_{ij}(0) >$ for $i, j = 1, 2$, then $Q_{z_2}(0) = Q_{z_1}(0)$ (Chade and Schlee [2002: 439, Corollary 5]), leading to the hope that Lemma 3(b) and Theorem 3(a) might deliver a zero limiting percentage error. Unfortunately, since the firm's optimal price is just a function of the *expected* cost, $\mu'(0) = Q'_{z_2}(0)\lambda_{z_2}(0) + Q'_{z_1}(0)\lambda_{z_1}(0) = 0$, and we cannot apply Theorem 3(a). (Indeed, here the family of information structures meets the conditions for the Radner-Stiglitz (1984) nonconcavity in the value of information: the marginal value of a little information is zero.) In this case, if the random price meets the conditions of Theorem 3(b), then the percentage error from using ECS tends to zero with α if and only if the income elasticity of demand equals relative risk aversion at the initial price.

4.2 Minimum Resale Price Maintenance

Deneckere, Marvel, and Peck (1997, henceforth DMP) construct a theory of ruinous price competition under demand uncertainty.[18] A good is produced by a monopolist manufacturer and sold by a continuum of identical retailers with zero cost. Retailers must order inventories before the demand uncertainty is resolved; any unsold inventories are worthless. They consider two scenarios. The first is *flexible pricing*: the manufacturer sets a wholesale price, p_w, that each retailer must pay at the time that inventories are ordered. After the demand uncertainty is resolved, the retail price is set so that demand equals aggregate inventory. The second is *minimum resale price maintenance* (RPM): the manufacturer sets both a wholesale price p_w and a minimum price, $p_{\min}$, below which the retail price cannot fall. If demand equals supply at a price above $p_{\min}$, then the market-clearing price prevails. If, however, supply exceeds demand at a price below $p_{\min}$, then the retail price is set at $p_{\min}$ and consumers are rationed among retailers to equalize the *ex ante* probability of a sale.

One striking result is that expected total surplus–and even expected consumers' surplus–can be higher under RPM than under flexible pricing. In this sense price competition can be ruinous. Intuitively, RPM can raise expected consumers' surplus (or not lower it more than expected profit rises) since minimum RPM can lead retailers to hold more inventory: although the price floor raises the retail price when demand is low, the higher output lowers price when demand is high.

DMP (1997: 638, Theorem A1) give sufficient conditions for expected total surplus to rise with the imposition of RPM. One condition is that the demand uncertainty is multiplicative: total demand is given by $\theta D(p)$, where θ is an unknown parameter. The other conditions are that production cost is zero;

[18] See also Deneckere, Marvel, and Peck (1996). Rey and Tirole (1986) similarly use ECS to evaluate vertical restraints, such as resale price maintenance, under demand uncertainty.

the support of θ is positive, compact, nondegenerate but not "too large"; demand is C^2; and a curvature condition on the total revenue function to ensure uniqueness. Under these conditions the equilibrium retail price under RPM is nonrandom (the revenue-maximizing price is unique, and since production cost is zero the manufacturer always wants to set p_{min} low enough so that retailers order enough to meet the highest demand at that price).

How robust is their welfare result when consumers do not have quasilinear utility? The answer depends on how we interpret the demand uncertainty. Since the support of θ is positive and compact, we may without loss of generality let it lie in $(0, 1]$. A natural interpretation of the multiplicative case is uncertainty over market size: a fraction $1 - \theta$ of consumers do not like the good at all (i.e. will not buy any at any positive price) and a fraction θ do like the good; the function D is the total demand if all consumers like the good (so that the distribution of types who like the good is independent of how many consumers like the good). Consumers know their own preferences, but each is uncertain about price because other consumers' preferences are private information. This interpretation fits our assumption that the only uncertainty each consumer faces is over price.

For illustration consider the indirect utility function

$$V(p, m) = T(cs(p, m_0) + g(m)) \tag{12}$$

where T is twice differentiable, concave and strictly increasing, and g is differentiable and strictly increasing with $g'(m_0) = 1$ and $g(m_0) = m_0$. For any such T, g pair, V generates the same demand at $m = m_0$ that consumer's surplus does: $-V_1(p, m_0)/V_2(p, m_0) = d(p, m_0) = -cs_1(p, m_0)$ for all $p \in P$. Thus a consumer with indirect utility function (12) and income m_0 would behave exactly in DMP's model as a consumer with a quasilinear indirect utility function $cs(p, m_0)+m$, but would have different preferences over price distributions. The good is noninferior if and only if g is concave. If T is strictly concave, then V is not additively separable and ECS does not represent preferences over $D\times\{m_0\}$. Nonetheless under plausible conditions on T and g, the aggregate willingness to pay of consumers and firms for RPM is still positive under the conditions of DMP's Theorem A1. Obviously, if g and T are both affine then the willingness to pay for RPM equals the change in ECS. Any increase in the concavity of g (which preserves $g'(m_0) = 1$ and $g(m_0) = m_0$) lowers a consumer's willingness to pay for RPM.[19] But any increase in the concavity of T increases a consumer's willingness to pay for RPM (the certainty equivalent price under flexible pricing rises with any increase in the concavity of T). So if each consumer is sufficiently risk averse over income gambles (where the required degree of risk aversion is higher, the more sensitive the consumer's demand is to income), then the aggregate willingness to pay for RPM over flexible prices will be positive.

[19] Let $\hat{\pi}$ denote the willingness to pay to replace price distribution F with a sure price, p^*: $T(cs(p^*, m_0) + g(m_0 - \hat{\pi})) = \int T(cs(p, m_0) + g(m_0))dF(p)$ If we replace g with a function $\tilde{g}$ which is more concave (but preserves $\widetilde{g}(m_0) = m_0$ and $\widetilde{g}'(m_0) = 1$ then $\widetilde{g}(m_0-\hat{\pi}) \leq g(m_0-\hat{\pi})$, so to restore the equality which defines the willingness to pay, $\hat{\pi}$ must rise.

We can make this point more precise by considering small changes in the distribution of θ starting from no uncertainty. Drop the requirement that each consumer has an indirect utility of the form (12) (but retain the interpretation of the uncertainty as private information). Let $(\Omega, \mathcal{B}, \nu)$ be a probability space and for each $\alpha \in [0,1]$ let $\theta(\cdot, \alpha)$ be a random variable on this space which takes on at most a finite number n of values for each α, is differentiable in α for each $\omega \in \Omega$ with $\theta_2(\cdot, \cdot)$ uniformly bounded on $\Omega \times (0, 1]$, $\theta(\cdot, 0)$ almost everywhere equal to some $t \in (0, 1]$, $\int \theta(\cdot, \alpha) d\nu = t$ for all $\alpha \in [0, 1]$ and $\int \theta_2(\omega, 0)^2 d\nu \neq 0$. (This specification is certainly consistent with the 'small support' assumption of DMP's Theorem A1.) Now let there be a finite number of types of consumers (who can differ in income, demand or risk preferences), and suppose that the other conditions of DMP's Theorem A1 are satisfied. As already mentioned, the equilibrium retail price under RPM is nonrandom; it is also independent of θ. Let $q(\omega, \alpha)$ denote the equilibrium retail price under flexible pricing when the state is ω. If $X(\alpha)$ denotes the (unique) output under flexible pricing, the market-clearing price is given by $X(\alpha) = \theta(\omega, \alpha) D(q(\omega, \alpha))$, so the price path is smooth at $\alpha = 0$. ($X(\cdot)$ is differentiable under DMP's assumptions.) Since consumers have private information about demand, they will update their beliefs about the state of the world; applying Bayes's rule, a consumer who likes the good will have belief given by $\nu_\alpha(A) = \int_A \theta(\omega, \alpha) d\nu(\omega)/t$ for $A \in \mathcal{B}$. If $n = 2$, then by Lemma 3(b) there is a common probability space and path of prices to $q(\cdot, 0) \equiv D^{-1}(X(0)/t)$ which is smooth. The extension of Lemma 3 to $n > 2$ is straightforward (or one can use ν_α in place of ν in the proof of Theorem 3). It is easy to verify that there is no first order effect on the mean price at $\alpha = 0$ (for any consumer) under flexible pricing, and that $\sigma^2 \neq 0$.[20] We therefore cannot apply Theorem 3(a). But we can use equation (9) in the proof of Theorem 3(b) to determine the direction of the bias: if a consumer's relative risk aversion exceeds the income elasticity of demand (at $p = D^{-1}(X(0)/t)$), then the willingness to pay for RPM exceeds the change in ECS. (Recall that Π is the willingness to pay to move *away* from a sure initial price.) The indirect utility function (12) automatically satisfies this condition if T is strictly concave. So DMP's ruinous price competition theorem survives departures from quasilinear utility well.

4.3 A Simple Robustness Test

We used the two applications to illustrate a simple test of how robust welfare conclusions are to relaxing quasilinear utility: introduce risk aversion by taking a concave (strictly increasing) transformation of each consumer's indirect utility function. This test has three things to recommend it: first, it doesn't change demand, so it doesn't change the purely positive parts of any analysis; second,

[20] Apply the implicit function theorem to the first order conditions for the manufacturer's problem under flexible pricing to conclude that $X'(0) = 0$. Integrate both sides of $X(\alpha) = \theta(\alpha, \omega) D(q(\omega, \alpha))$ with respect to ω and differentiate with respect to α (using the assumption that the mean value of θ is independent of α) to conclude that $\mu'(0) = 0$–whether from the perspective of the firm, or of a consumer who likes the good. That $\sigma^2 \neq 0$ follows from $\int \theta_2(\omega, 0)^2 d\nu \neq 0$.

it is often easy to do; third, risk aversion is apt to be quantitatively the most important departure from quasilinear utility.

It is well-known that a mean-preserving increase in price risk raises ECS (Waugh, 1944). But if the uncertainty in prices comes solely from the supply side of the market, then policies which increase price risk hurt consumers if they are sufficiently risk averse. So conclusions which exploit the fact that ECS rises with price risk–as information acquisition and sharing often do–are not likely to be robust. By contrast, in Deneckere, Peck and Marvel's (1997) model of demand uncertainty, imposing minimum RPM can result in a mean-preserving *decrease* in price risk. It is thus easy to understand why minimum RPM can be more attractive to risk averse than risk neutral consumers.

Before considering some extensions we give one more illustration. Lewis and Sappington (1988) analyze regulation of a monopoly with private information about its demand. If the firm's marginal cost is decreasing, then setting a price that does not vary at all with the firm's demand report winds up maximizing total expected surplus (subject to the usual incentive compatibility and participation constraints [Proposition 2]). Suppose as in the last subsection that each consumer knows its own preferences, but not the aggregate demand. If we simply take a concave transformation of each consumer's (quasilinear) indirect utility function (and use the consumer's certainty equivalent wealth as a measure of his welfare) then the optimal policy is unchanged: the *same* constant price is still optimal.[21]

5 Extensions: state-contingent payments; non-expected utility preferences; multivariate risk

Thus far we have used the *ex ante* willingness to pay to measure the welfare change to consumers; we assumed that consumers 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 State-contingent payments

If state contingent payments are possible, then our *ex ante* willingness to pay criterion might miss some potential Pareto improvements: the aggregate sure willingness to pay for a change can be negative, yet there might be state contingent payments that make all consumers and firms better off *ex ante* with the change (Graham, 1981).

Suppose we adopt the criterion that a change from A to B is desirable if there are state contingent money transfers which would make everyone better off under B than A. For the moment, assume that such transfers are not actually

[21]for any strictly concave, strictly increasing transformation T, we have $T^{-1}\left(\int T(cs(p)+m)dF(p)\right) \leq \int cs(p)dF(p) + m$ with an equality if and only if the price is not random; so if a sure price maximizes the sum of ECS and expected profit, that same sure price will maximize the sum of certainty equivalents and expected profit.

made. As before let $\langle q(\cdot, \alpha) : \alpha \in [0,1] \rangle$ be a path defined on a probability space (Ω, B, ν). Let $\gamma(\cdot, \alpha)$ be a (measurable) real-valued function on Ω for each $\alpha \in [0,1]$. A consumer would just be willing to make state contingent payments of $\gamma(\cdot, \alpha)$ to replace the random price $q(\cdot, 0)$ with $q(\cdot, \alpha)$ if

$$\int V(q(\omega, \alpha), m - \gamma(\omega, \alpha)) d\nu = \int V(q(\omega, 0), m) d\nu. \tag{13}$$

Suppose that V is regular and that $V(p, \cdot)$ is strictly concave for each p, so that the consumer is risk averse, and suppose that firms are risk neutral. If we consider transfers between firms and consumers, these two facts imply that an optimal contingent payment must equate any consumer's marginal utility of income in any two states:

$$V_2(q(\omega, \alpha), m - \gamma(\omega, \alpha)) = V_2(q(\omega', \alpha), m - \gamma(\omega', \alpha)) \tag{14}$$

for (almost-all) (ω, ω') in Ω. If (14) holds for *all* (ω, ω'), then there is a unique $\gamma(\cdot)$ satisfying (14) and (13); and if the path $\langle q(\cdot, \alpha) \rangle$ is smooth at $\alpha = 0$, then so will the path $\langle \gamma(\cdot, \alpha) \rangle$. Let $\Delta\gamma(\omega, \alpha) = \gamma(\omega, \alpha) - \gamma(\omega, 0)$. A measure of the change in consumer welfare is the expected value of $\Delta\gamma(\cdot, \alpha)$ for a path of payments satisfying (13) and (14). For example, if the indirect utility V is additive separable in p and m, then this measure equals the *ex ante* willingness to pay, $\Pi(\alpha)$; if there are no income effects on demand, then it equals the expected compensating variation for the change. Less justifiable is simply to use the expectation of $\Delta\gamma(\cdot, \alpha)$ for a smooth path $\langle \gamma(\cdot, \alpha) \rangle$ satisfying (13). One possibility is to set $\gamma(\omega, \alpha)$ equal to the compensating variation for each state: $V(q(\omega, \alpha), m - \gamma(\omega, \alpha)) = V(q(\omega, 0), m)$ for each $(\omega, \alpha) \in \Omega \times [0,1]$. We include this possibility for completeness.

Proposition 3 *Let V be regular and strictly concave in income for all $p \in P$; the path $\langle q(\cdot, \alpha) \rangle$ be smooth at $\alpha = 0$; and $\Pi'_{cs}(0) \neq 0$. We have*

$$\lim_{\alpha \to 0} \frac{\int \Delta\gamma(\omega, \alpha) d\nu(\omega) - \Pi_{cs}(\alpha)}{\Pi_{cs}(\alpha)} = 0 \tag{15}$$

if either of the following two conditions holds:

- *(a) The path $\langle \gamma(\cdot, \alpha) \rangle$ satisfies (13) and (14) for all $(\omega, \omega', \alpha) \in \Omega^2 \times [0,1]$ and there are no income effects on demand;*
- *(b) For every $(\omega, \alpha) \in \Omega \times [0,1]$, $\gamma(\omega, \alpha)$ equals the compensating variation for the change in price from $q(\omega, 0)$ to $q(\omega, \alpha)$.*

Proof: appendix.

Proposition 3 provides some justification for using ECS in local cost-benefit analysis. Several qualifications are in order, however.

- In some important applications, the path of prices is not smooth, or even uniformly convergent (Lemma 3, Section 4.1 and Corollary 1). If $q(\cdot, \alpha)$ just converges to $q(\cdot, 0)$ pointwise, then the conclusions of Proposition 3 fail.

- Part (a) asserts that there are *no* income effects on demand: if the initial price is random and V is not additively separable, then (15) generally fails if there are income effects (since (14) requires that income vary across states).[22]

- For all the same reasons that markets are incomplete, it may be impossible to make such state-contingent payments.

- If the payments are not made, then the criterion of aggregate state-contingent payments can be Pareto inconsistent. (See e.g. Graham [1981: 721-22].)[23]

- For part (b), if the payments are made and there are income effects on demand then the *positive* analysis in many applied models is incorrect. Most applications–including all those in our reference list–assume that demand is the same before and after the policy change.

This last point raises a further problem: if state-contingent payments can be made, then presumably *consumers* can trade income across states of the world; even if income effects on demand are slight, if risk preferences vary enough across the population, the resulting trades can overturn both the positive and normative predictions of applied partial equilibrium models.[24]

5.2 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).[25] Here π is given implicitly by $v(F, m) = v(G, m - \pi(G, F, m))$. Even if v violates the independence axiom of expected utility theory,[26] if $v(\cdot, m)$ is L_1-Fréchet differentiable on D, then it will behave locally as an expected

[22] If we had just used the mean of γ rather than $\Delta\gamma$, then we would need to impose quasilinear utility to establish (15) under (a), since in general $\int \gamma(\omega, 0) d\nu \neq 0$.

[23] For example, consider the monopoly model of information acquisition from Section 4.1.1 in which the consumer is risk averse but there are no income effects on demand. In this case, $\gamma(\omega, \alpha)$ is the compensating variation for state ω. Suppose that the choice is simply between null and perfect information, and that the cost of information just exceeds the value of information to the firm, but that the sum of expected profit and the expected compensating variation exceeds the cost of information. If payments are not made, then the firm is clearly worse off with information if it bears the cost; and if risk aversion is large enough, then the consumer is worse off as well.

[24] For example, on our interpretation of the demand uncertainty in Deneckere, Marvel and Peck (1997) as private preference information, consumers and firms have different beliefs about the states of the world; even if attitudes towards income risk are the same, there is an incentive to trade. And Schlee (2001) emphasizes how risk sharing among consumers affects the welfare consequences of improved information.

[25] 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.

[26] 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 $f(\lambda F + (1-\lambda)H) \geq f(\lambda G + (1-\lambda)H)$.

utility functional (Machina (1982)): 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 is to ensure that ECS represents the consumer's preferences over price distributions (Lemma 1), which is impossible if the consumer violates independence. An adaptation of the proof that (a) implies (b) in Theorem 1 yields the following result.

Corollary 2 *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 $\pi(G, F, m) - \pi_{cs}(G, F, m) \neq o(||G - F||)$.*

The conclusion of Theorem 3(a), however, does extend to smooth nonexpected utility preferences: for smooth paths on the space of random variables and a nonrandom initial price, the percentage error from using ECS tends to zero as the distribution change tends to zero.

Corollary 3 *In addition to the hypotheses of Corollary 2, 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 Theorem 3(a) holds.*[27]

5.3 Multivariate risk

If variables in the consumer's indirect utility function other than the good's price are random, then ECS ranks changes in the distribution of price 1 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 (Rogerson [1980] and Turnovsky et. al. [1980]). 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. This includes the possibility that consumers are unsure about their tastes when the price distribution changes. (To model preference uncertainty, we let the consumer's indirectly utility function depend directly on the state of the world ω.) If a consumer's marginal utility of income is not independent of the state then even our modest positive results–Proposition 2, Theorem 3(a) and Proposition 3–fail.

[27]To prove the corollary, use equation (8) on p. 296 in Machina (1982) to show that Π_{cs} is differentiable at $\alpha = 0$ with $\Pi'_{cs}(0) = \psi_1(p; m, \delta_p)/v_2(\delta_p, m)\mu'(0)$, where $\psi(\cdot, m, F)$ is the local utility function for v at (m, F). (Since $\langle q(\cdot, \alpha)\rangle$ is smooth, $||G_{q(\cdot,\alpha)} - \delta_p||$ is differentiable in α at zero, as Machina's (8) requires.) The Corollary follows after noting that we can express Roy's Identity as $d(p, m) = -\psi_1(p, m, \delta_p)/v_2(\delta_p, m)$.

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 no doubt 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 for consumer's surplus under certainty. We find that expected consumer's surplus can be a poor approximation to willingness to pay for a small distribution change. It can be a good approximation if we restrict the distribution change to smooth, or at least uniformly convergent, paths on the space of random variables; but this condition necessarily fails in some applications.

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 more. The simple robustness checks that we carried out in Section 4 are a start.

7 Appendix

Proof of Proposition 1: If each consumer's indirect utility function is additively separable in price and income and either (a) or (b) holds, then aggregate ECS is clearly Kaldor consistent. If the indirect utility function is not additively separable for some type $\tau^* \in T$, then aggregate ECS is not Kaldor consistent: set $a(i) = (m^*, \tau^*)$ for all $i \in [0,1]$, where m^* is any income level for which ECS does not represent the same preferences on D as a type τ^* consumer with income m^* (Lemma 1).

So suppose that $V(\cdot,\cdot,\tau)$ is additively separable for each $\tau \in T$. In particular, let $V(p,m,\tau) = f(p,\tau) + g(m,\tau)$. Suppose that (a) fails and that there are two types τ and τ' in T with different preferences on D and that $g(\cdot,\tau)$ is not affine. We will show that aggregate ECS cannot be Kaldor consistent. Set m such that $g_{11}(m,\tau) \neq 0$ and choose G, F such that $\pi_{cs}(G,F,m,\tau) > 0$ and $\pi_{cs}(G,F,m,\tau') < 0$. We now argue that we can always choose such an (G,F) so that

$$\frac{g_{11}(m,\tau)}{g_1(m,\tau)}\pi_{cs}(G,F,m,\tau) \neq \frac{g_{11}(m,\tau')}{g_1(m,\tau')}\pi_{cs}(G,F,m,\tau'). \tag{16}$$

To begin, note that the left side of (16) is nonzero. If a particular (G,F) violates (16), then replace G with any G' which first-order stochastically dominates G. (Such a G' exists since $\pi_{cs}(G,F,m,\tau') < 0$ implies that $G \neq \delta_{\underline{p}}$.) The left side of (16) must change and the right must either remain unchanged or move in the opposite direction when G' replaces G, and so the two sides of (16) will not be equal at (G',F,m). From now on assume that (16) holds. Let $\theta \in (0,1)$ satisfy

$$\theta\pi_{cs}(G,F,m,\tau) + (1-\theta)\pi_{cs}(G,F,m,\tau') = 0, \tag{17}$$

and let $a(i) = (m,\tau)$ for $i \in [0,\theta]$ and $a(i) = (m,\tau')$ for $i \in (\theta,1]$. If

$$\theta\pi(G,F,m,\tau) + (1-\theta)\pi(G,F,m,\tau') \neq 0$$

then there is nothing to prove, so suppose that

$$\theta\pi(G,F,m,\tau) + (1-\theta)\pi(G,F,m,\tau') = 0.$$

Now replace G with $\alpha G + (1-\alpha)F$. We have

$$\begin{aligned}&\theta\pi_{cs}(\alpha G + (1-\alpha)F, F, m, \tau) + (1-\theta)\pi_{cs}(\alpha G + (1-\alpha)F, F, m, \tau') = \\ &\alpha\left[\theta\pi_{cs}(G,F,m,\tau) + (1-\theta)\pi_{cs}(G,F,m,\tau')\right] = 0\end{aligned} \tag{18}$$

for all $\alpha \in (0,1)$. But for α small enough, we have

$$\theta\pi(\alpha G + (1-\alpha)F, F, m, \tau) + (1-\theta)\pi(\alpha G + (1-\alpha)F, F, m, \tau') \neq 0. \tag{19}$$

To see why (19) must hold, note that by Theorem 1

$$\pi(\alpha G + (1-\alpha)F, F, m, \hat{\tau}) = \alpha\pi_{cs}(G,F,m,\hat{\tau}) + o(\alpha||G-F||)$$

for $\hat{\tau} \in \{\tau, \tau'\}$. Hence

$$\frac{\partial}{\partial \alpha}\theta\pi(\alpha G + (1-\alpha)F, F, m, \tau) + (1-\theta)\pi(\alpha G + (1-\alpha)F, F, m, \tau')\Big|_{\alpha=0+} =$$
$$\theta\pi_{cs}(G, F, m, \tau) + (1-\theta)\pi_{cs}(G, F, m, \tau') = 0.$$

But

$$\frac{\partial^2}{\partial \alpha^2}\left[\theta\pi(\alpha G + (1-\alpha)F, F, m, \tau) + (1-\theta)\pi(\alpha G + (1-\alpha)F, F, m, \tau')\right]\Big|_{\alpha=0+} =$$
$$\theta\frac{g_{11}(m,\tau)}{g_1(m,\tau)}\left(\pi_{cs}(G, F, m, \tau)\right)^2 + (1-\theta)\frac{g_{11}(m,\tau')}{g_1(m,\tau')}\left(\pi_{cs}(G, F, m, \tau')\right)^2 =$$
$$\theta\pi_{cs}(G, F, m, \tau)\left(\frac{g_{11}(m,\tau)}{g_1(m,\tau)}\pi_{cs}(G, F, m, \tau) - \frac{g_{11}(m,\tau')}{g_1(m,\tau')}\pi_{cs}(G, F, m, \tau')\right) \neq 0.$$

The second equality follows from (17). Hence equation (19) holds. Equations (19) and (18) together imply that aggregate ECS is not Kaldor consistent. ■

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

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

Thus

$$\pi(G, F, m) = \frac{\int_P V(p, m)d(G-F)}{\int V_2(\omega, m)dF(\omega)} + \int V(p, m)d(G-F)\left(\frac{1}{\int V_2(p, m - t(G))dG} - \frac{1}{\int V_2(p, m)dF}\right). \tag{20}$$

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 (20) 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

$$\lim_{n\to\infty}\int V_2(p, m - t(G_n))dG_n = \int V_2(p, m)dF \tag{21}$$

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 $V_2(\cdot,\cdot)$ is continuous and P is compact, $V_2(p, m - t(G_n))$ converges to $V_2(p, m)$ *uniformly* in p. Hence (21) holds[28] and the conclusion follows. ■

Proof of Theorem 2: Suppose throughout that V is regular. Consider part (a). Since V is not additively separable, ECS does not represent preferences over D (Lemma 1). 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 $H(\cdot,\alpha) = \alpha\left(\alpha G + (1-\alpha)G^*\right) + (1-\alpha)F$. Clearly $||H(\cdot,\alpha) - F||$ tends to zero as $\alpha \to 0$ and $\pi_{cs}(H(\cdot,\alpha), F, m) \neq 0$ for all $\alpha \in (0,1]$. Moreover, letting

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

we have by Lemma 2 that

$$\left|\frac{\pi(H(\cdot,\alpha),F,m) - \pi_{cs}(H(\cdot,\alpha),F,m)}{\pi_{cs}(H(\cdot,\alpha),F,m)}\right| =$$
$$\left|\frac{\alpha \int K(p,m,F)d(G-G^*) + \int K(p,m,F)d(G^*-F) + \frac{1}{\alpha}o(||H(\cdot,\alpha) - F||)}{\alpha \int cs(p,m)d(G-G^*)}\right|.$$

The first and third terms in the numerator on the right side of the equality tend to zero as α tends to zero. Since the second term in the numerator is nonzero, the entire expression diverges to $+\infty$ as $\alpha \to 0$.

For part (b), let $\langle \widehat{H}(\cdot,\alpha)\rangle$ be any smooth path to F with $\pi_{cs}(\widehat{H}(\cdot,\alpha), F, m) \neq 0$ for all $\alpha \in (0,1]$. We will evaluate the limit of

$$\frac{(\pi(\widehat{H}(\cdot,\alpha),F,m) - \pi_{cs}(\widehat{H}(\cdot,\alpha),F,m))/\alpha}{\pi_{cs}(\widehat{H}(\cdot,\alpha),F,m)/\alpha} \tag{22}$$

as α tends to 0. After integrating by parts we have

$$\frac{\pi_{cs}(\widehat{H}(\cdot,\alpha),F,m)}{\alpha} = \int d(p,m)\left(\frac{\widehat{H}(p,\alpha) - F}{\alpha}\right)dp. \tag{23}$$

[28] 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 on the right 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:

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

Since the integrand in (23) is bounded, the Lebesgue Dominated Convergence Theorem (LDCT) implies that

$$\left.\frac{d\pi_{cs}(\widehat{H}(\cdot,\alpha),F,m))}{d\alpha}\right|_{\alpha=0} = \int d(p,m)H_2(p,0)dp. \tag{24}$$

Moreover, using Lemma 2 and after integrating by parts, we have

$$\pi(\widehat{H}(\cdot,\alpha),F,m) = \frac{-\int V_1(p,m)(\widehat{H}(p,\alpha)-F)dp}{\int V_2(\omega,m)dF(\omega)} + o(||\widehat{H}(\cdot,\alpha)-F||).$$

Thus

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

$$\int_P \left(-\frac{V_1(p,m)}{V_2(p,m)}\frac{V_2(p,m)}{\int V_2(\omega,m)dF(\omega)} - d(p,m)\right)\left(\frac{\widehat{H}(\cdot,\alpha)-F}{\alpha}\right)dp + \frac{o(||\widehat{H}(\cdot,\alpha)-F||)}{\alpha}$$

or, using Roy's Identity, the right side of the last equality becomes

$$\int_P \left[d(p,m)\left(\frac{V_2(p,m)}{\int V_2(\omega,m)dF(\omega)}-1\right)\left(\frac{\widehat{H}(\cdot,\alpha)-F}{\alpha}\right)\right]dp + \frac{o(||\widehat{H}(\cdot,\alpha)-F||)}{\alpha}. \tag{25}$$

Since the integrand in (25) is bounded, the LDCT implies that this expression has limit

$$\int_P \left[d(p,m)\left(\frac{V_2(p,m)}{\int V_2(\omega,m)dF(\omega)}-1\right)\widehat{H}_2(p,0)\right]dp \tag{26}$$

as α tends to zero.

If the integral in (26) is not zero, then equation (22) does not tend to zero with α. So suppose that (26) equals zero. Let $P_+ = \{p \in P | V_2 > \int V_2 dF\}$ and $P_- = \{p \in P | V_2 < \int V_2 dF\}$. Since $V_2(\cdot,m)$ is not F-*a.e.* constant, $\int_{P_+} dF > 0$ and $\int_{P_-} dF > 0$. And since V_2 is continuous, both P_+ and P_- 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 $I_+ \in P_+$ and $I_- \in P_-$ 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 $H^{\lambda}(\cdot,\alpha) = (1-\lambda)\widehat{H}(\cdot,\alpha) + \lambda(\alpha G + (1-\alpha)F)$. Since $d > 0$, we have (replace $\widehat{H}(\cdot,\alpha)$ with $H^{\lambda}(\cdot,\alpha)$ in (26)

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

for any $\lambda \in (0,1)$. Fix $\varepsilon > 0$. For λ small enough, $||(H_2^{\lambda}(\cdot,0) - \widehat{H}_2^{\lambda}(\cdot,0))|| < \varepsilon$. And since F first-order stochastically dominates G, it follows that, for almost all λ,

$$\left.\frac{d(\pi_{cs}(H^{\lambda}(\cdot,\alpha),F,m))}{d\alpha}\right|_{\alpha=0} \neq 0.$$

For some such λ, set $H(\cdot,\alpha) = H^{\lambda}(\cdot,\alpha)$ for all $\alpha \in [0,1]$. Noting that $\pi_{cs}(H^{\lambda}(\cdot,\alpha),F,m) \neq 0$ except for at most one value of α, (5) has a nonzero limit. ■

Proof of Lemma 3: For part (a), suppose that $Q_1(0) \neq Q_2(0)$ but that there is a path $\langle q(\cdot,\alpha)\rangle$ which satisfies (i) and (ii). We will show that $\lambda'(0) = 0$. Consider any sequence $\langle\alpha_n\rangle$ in $(0,1]$ with limit zero. Let $E_0 = \{\omega \in \Omega | q(\omega,0) \notin \{Q_1(0),Q_2(0)\}\}$, $E_n = \{\omega \in \Omega | q(\omega,\alpha_n) \notin \{Q_1(\alpha_n),Q_2(\alpha_n)\}\}$ for each n and $E = \cup_{i=0}^{\infty}E_i$. By hypothesis, $\nu(E_i) = 0$ for $i = 0,1,2...$, so that $\nu(E) = 0$. Since convergence is uniform, each $Q_j(\cdot)$ is continuous, and $Q_2(0) \neq Q_1(0)$, there is an n^* such that for every $\omega \in \Omega/E$ either $q(\omega,\alpha_n) = Q_1(\alpha)$ for all $n \geq n^*$ or $q(\omega,\alpha_n) = Q_2(\alpha)$ for all $n \geq n^*$. Define a sequence of functions on Ω by $\psi_n(\omega) = (q(\omega,\alpha_n) - q(\omega,0))/\alpha_n$ for all $\omega \in \Omega$ and all $n \geq 1$; and let $\psi(\omega) = Q_2'(0)$ if $q(\omega,0) = Q_2(0)$ and $\psi(\omega) = Q_1'(0)$ for all other $\omega \in \Omega$. Since each Q_j is differentiable at 0, $\psi_n \to \psi$ a.e. and the sequence $\langle(Q_j(\alpha_n) - Q_j(0))/\alpha_n\rangle$ is bounded by some number K_j for $j = 1,2$. Thus the sequence $\langle\psi_n\rangle$ is bounded uniformly on Ω/E by the largest of $\max\{K_1,K_2\}$ and the real number

$$\sup\{\,|q(\omega,\alpha_n) - q(\omega,0)|/\alpha_n |\, \omega \in \Omega/E, n \leq n^*\}\,.$$

Let f be any strictly increasing, C^1 function on $\mathbb{R}_+$. By the Mean Value Theorem, for every n and $\omega \in \Omega$ there is a number $\theta_n(\omega)$ between $q(\omega,0)$ and $q(\omega,\alpha_n)$ such that $f(q(\omega,\alpha_n)) - f(q(\omega,0)) = \alpha_n f'(\theta_n(\omega))\psi_n(\omega)$, so that

$$\int_{\Omega}\left(\frac{f(q(\omega,\alpha_n)) - f(q(\omega,0))}{\alpha_n}\right)d\nu(\omega) = \int_{\Omega} f'(\theta_n(\omega))\psi_n(\omega)d\nu(\omega). \tag{27}$$

Since $f'(\theta_n(\omega))$ is (essentially) bounded, the LDCT (Billingsley, 1986: Theorem 16.4) implies that the right side of (27) converges to

$$f'(Q_1(0))Q_1'(0)(1-\lambda(0)) + f'(Q_2(0))Q_2'(0)\lambda(0) \tag{28}$$

as n tends to infinity. Since the sequence $\langle\alpha_n\rangle$ was arbitrary, this argument shows that the function $\Gamma(\alpha) = \int f(\omega,\alpha)d\nu(\omega)$ of α on $[0,1]$ is differentiable at 0 and $\Gamma'(0)$ equals (28). But $\Gamma'(0)$ must also equal

$$\frac{d}{d\alpha}\, f(Q_1(\alpha))(1-\lambda(\alpha)) + f(Q_2(\alpha))\lambda(\alpha)|_{\alpha=0} =$$
$$f'(Q_1(0))Q_1'(0)(1-\lambda(0)) + f'(Q_2(0))Q_2'(0)\lambda(0) + \lambda'(0)(f(Q_2(0)) - f(Q_1(0))),$$

which implies that $\lambda'(0) = 0$ if $Q_2(0) \neq Q_1(0)$. This proves part (a).

For part (b), let $\Omega = [0,1]$, $\mathcal{B}$ be the Borel subsets of [0,1], and ν Lebesgue measure. Suppose that $\lambda'(0) > 0$. (The case of is $\lambda'(0) < 0$ handled similarly.) Define the following path on $(\Omega, \mathcal{B}, \nu)$: for each $\alpha \in [0,1]$, $q(\omega,\alpha) = Q_2(\alpha)$ if $\omega \leq \lambda(\alpha)$ and $q(\omega,\alpha) = Q_1(\alpha)$ otherwise. Obviously, this path satisfies condition (i). Since the Q_i's are differentiable at $\alpha = 0$, so is $q(\omega,\cdot)$ for each $\omega \in [0,1]$. Let $Q(0)$ denote the sure price at $\alpha = 0$. Since

$$\frac{q(\omega,\alpha) - q(\omega,0)}{\alpha} \in \left\{ \frac{Q_2(\alpha) - Q(0)}{\alpha}, \frac{Q_1(\alpha) - Q(0)}{\alpha} \right\}, \tag{29}$$

for each $(\omega,\alpha) \in \Omega \times (0,1]$ and each $(Q_j(\alpha) - Q(0))/\alpha$ is bounded in α, the path $\langle q(\cdot,\alpha)\rangle$ is smooth at 0, and hence converges to $q(\cdot,0)$ uniformly: if K is the uniform bound for $|q(\omega,\alpha) - q(\omega,0)|/\alpha$, then $\sup_\omega |q(\omega,\alpha) - q(\omega,0)| \leq \alpha K \to 0$. The last sentence of the Lemma follows from the equality between $\Gamma'(0)$ and (28) with f set equal to the identity function. ■

Proof of Proposition 3: Under either (a) or (b), the path $\langle\gamma(\cdot,\alpha)\rangle$ must be smooth at $\alpha = 0$. Applying an argument similar to the one leading to equation (6) in the proof of Theorem 3(a) to equation (13) gives us

$$\begin{aligned} 0 &= \int (V_1(q(\omega,0), m - \gamma(\omega,0))q_2(\omega,0) - V_2(q(\omega,0), m - \gamma(\omega,0))\gamma_2(\omega,0))d\nu(\omega) \\ &= -\int V_2(q, m-\gamma)\left(d(q, m-\gamma)q_2(\omega,0) + \gamma_2(\omega,0)\right)d\nu(\omega) \end{aligned} \tag{30}$$

where $q(\cdot)$ and $\gamma(\cdot)$ on the right side of the second equality are evaluated at $(\omega,0)$. Let (a) hold. Since (14) implies that $V_2(q(\omega,0), m - \gamma(\omega,0))$ is independent of ω, equation (30) implies that

$$\int \gamma_2(\omega,0)d\nu(\omega) = \int d(q(\omega,0), m - \gamma(\omega,0))q_2(\omega,0)d\nu(\omega). \tag{31}$$

A slight adaptation of the argument leading to equation (7) in the proof of Theorem 3(a) implies that

$$\Pi'_{cs}(0) = \int d(q(\omega,0), m)q_2(\omega,0)d\nu(\omega). \tag{32}$$

Since there are no income effects on demand, (15) follows from (31) and (32).

Suppose now that (b) holds. It is easy to show that $\gamma(\omega,0) = 0$ and $\gamma_2(\omega,0) = -d(q(\omega,0), m)q_2(\omega,0)$ for all $\omega \in \Omega$. Hence (31) holds with $\gamma(\omega,0) = 0$ for all $\omega \in \Omega$, from which (15) follows. ■

References

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

Billingsley, P., *Probability and Measure*, 2nd edition. New York: John Wiley and Sons, 1986.

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.

Chade, H. and E. Schlee, "Another Look at the Radner-Stiglitz Nonconcavity in the Value of Information," *Journal of Economic Theory,* 107 (2002): 421-452.

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

Deneckere, R. H. Marvel and J. Peck, "Demand Uncertainty, Inventories and Resale Price Maintenance," *Quarterly Journal of Economics,* 111 (1996): 885-914.

Ericson, W., "A Note on the Posterior Mean of a Population Mean," *Journal of the Royal Statistical Society*, 31 (1969): 332-334.

Graham, D., "Cost-Benefit Analysis under Uncertainty," *American Economic Review,* 87 (1981): 715-725.

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.

Li, L., "Cournot Oligopoly with Information Sharing," *RAND Journal of Economics,* 16 (1985): 521-536

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," *Econometrica* 70 (2002): 583-602.

Novshek, W. and H. Sonnenschein, "Fulfilled Expectations Cournot Duopoly with Information Acquisition and Release," *Bell Journal of Economics* 13 (1982): 214-18.

Radner, R. and J.E. Stiglitz, 'A Nonconcavity in the Value of Information', in M. Boyer and R. Kihlstrom (eds.) *Bayesian Models of Economic Theory.* Elsevier: Amsterdam, 1984.

Raith, M., "A General Model of Information Sharing in Oligopoly," *Journal of Economic Theory*, 71 (1996): 260-288.

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.

Schlee, E., "The Value of Information in Efficient Risk Sharing Arrangements," *American Economic Review* 91 (2001): 509-524.

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.

Waugh, F, "Does the Consumer Benefit from Price Instability," *Quarterly Journal of Economics*, 58 (1944): 589-597.

Weitzman, M. "Consumer's Surplus as an Exact Approximation when Prices are Appropriately Deflated," *Quarterly Journal of Economics*, 103 (1988): 543-553.

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

Willig, R.D. "Consumer's Surplus Without Apology: Reply," *American Economic Review*, 69 (1979): 467-474.