A review article: The case against the use of the sum of compensating variations in cost-benefit analysis

CHARLES BLACKORBY and DAVID DONALDSON
University of British Columbia

Abstract. This paper presents a case against the use of the sum of compensating variations as a cost-benefit test. We argue that: (1) the ethical judgments implied by the test are not defensible; (2) positive sums of compensating variations occur without potential Pareto Improvements, resulting in social preference reversals without simultaneous Scitovsky reversals; (3) when lump-sum transfers are feasible, a positive sum of compensating variations is necessary but not sufficient for a Potential Pareto Improvement; (4) in order to eliminate preference reversals and intransitivities, all households must have almost identical quasi-homothetic preferences – a condition that is not satisfied in real economies.

I. INTRODUCTION

In cost-benefit analysis and other exercises in applied welfare economics, aggregate willingness-to-pay – the simple sum of Hicksian compensating variations, is often used as a test. A positive sum is taken as evidence of a social improvement.

We thank Bob Allen, Soren Lemche, and John Weymark for extensive comments on an earlier version. Thanks also to two anonymous referees. Erwin Diewert originally encouraged us to write this paper, and we are indebted to him for providing us with the impetus.
or an increase in economic efficiency.\(^1\) When Marshallian surpluses are used, their employment is usually justified by their being viewed as approximations of compensating variations (Willig 1976). Occasionally, equivalent variations are used, but the more favoured statistic is willingness-to-pay, measured by the sum of compensating variations.

For a given project or social change the compensating variation for a household is the maximum amount that it would be willing to pay to secure the change. This number is positive if and only if the project moves the (rational) household to a higher indifference surface and negative if and only if it moves the household to a lower one.

At first blush, it seems reasonable to think that a positive sum of compensating variations must enable the gainers to compensate the losers and have something left over for themselves. Such a change is called a Potential Pareto Improvement. When such changes are given social approval and allowance is made for social indifference (the gainers are indifferent after compensating the losers), we call the resulting precept the Potential Pareto Principle.\(^2\) The use of the sum of compensating variations as a cost-benefit test is often justified by appealing to this principle, and a positive sum of compensating variations is thought to characterize a Potential Pareto Improvement.

In this paper we review the known theoretical results that illumine the relationship between sums of Hicksian consumers' surpluses and the Potential Pareto Principle both analytically and geometrically. In addition, we discuss the social ethics implicit in the aggregate willingness-to-pay test. In the process, we believe that we make a strong case against the compensating-variation test.\(^3\)

Our case against it can be summarized as follows:

1. The ethical judgments implied by the compensating-variation test are not defensible. It treats increases in income as equally socially valuable no matter who receives them. Social judgments – revealed by government policy – and the overwhelming majority of individual judgments are not consistent with this indifference toward inequality.

2. A positive sum of compensating variations is not the same thing as an improvement according to the Potential Pareto Principle. For example, a move from one Walrasian equilibrium\(^4\) to another Walrasian equilibrium typically yields a positive sum of compensating variations (the Boadway Paradox) even though no ‘efficiency gain’ has occurred (there is no Potential Pareto Improvement).

---

1 When prices do not change, the compensating variation is equal to income change. Costs, in cost-benefit analysis, represent the value of goods and services forgone when the project is undertaken, and hence a loss of income. Therefore benefits minus costs are equal to the sum of compensating variations in this case.

2 Sometimes it is called the compensation criterion, the Kaldor criterion, or the Hicks-Kaldor criterion.

3 The case against the Potential Pareto Principle itself as a criterion is extensive and is summarized in section VI.

4 All agents are price takers.
3. Although a positive sum of compensating variations is necessary for an improvement according to the Potential Pareto Principle (when lump-sum transfers are feasible), it is not sufficient.

4. Neither the sum of compensating variations nor the Potential Pareto Principle ranks social alternatives in a reasonable way. In order to eliminate intransitivities over consumption-efficient allocations, all households must be assumed to have quasi-homothetic preferences that are identical at the margin. Even then, distorted equilibria are not ranked sensibly.

These criticisms apply with minor changes to the sum of equivalent variations (the equivalent variation is the minimum compensation which the household would accept to forgo the change).5

In the rest of the paper we discuss these reasons, using simple geometric illustrations based on Scitovsky sets for the claims in (2) and (3). Although (1) is more important than (2)–(4), we discuss the latter first, since they inform our discussion of (1).

In section II we define the Hicksian consumers’ surpluses and the Scitovsky sets. In section III we prove the Boadway Paradox for exchange economies (Boadway 1974). In section IV the paradox is extended to production economies (Schweizer 1983). Necessary and sufficient criteria for Potential Pareto Improvements are discussed in section V (Foster 1976; Bruce and Harris 1982). In section VI we compare the rationality properties of the two tests, drawing on Gorman’s (1953, 1955) work on the compensation test and our own (1985) work on the compensating-variation and equivalent-variation tests. Section VII provides a discussion of the social ethics implicit in the compensating-variation test, and section VIII concludes.

II. HICKSIAN CONSUMERS’ SURPLUSES

We consider a private-goods economy with \( H \geq 2 \) consumers or households and \( m \geq 2 \) goods. The \( h \)th household \((h = 1, \ldots, H)\) has preferences represented by a direct utility function \( U^h \), whose image is

\[
u^h = U^h(x^h)
\]

for all \( x^h \in \mathbb{R}^m_+ \), \( x^h \) is household \( h \)'s consumption vector. Given appropriate regularity conditions,6 \( U^h \) can be equivalently represented by the indirect utility function \( V^h \) or the expenditure function \( C^h \) where

\[
V^h(p, y^h) = \max_{x^h} \{ U^h(x^h) \mid p \cdot x^h \leq y^h \},
\]

\[
C^h(u^h, p) = \min_{x^h} \{ p \cdot x^h \mid U^h(x^h) \geq u^h \},
\]

5 The equivalent variation has an advantage over the compensating variation for a single household. If the equivalent variation for one change is greater than the equivalent variation for another, then the household prefers the outcome of the first to the outcome of the second. The compensating variation does not share this property; see Hause (1975) or Pauwels (1978).

6 See Blackorby, Primont, and Russell (1978, chap. 2 and appendix) for a full treatment. For our purposes it is sufficient to assume that \( U^h \) is quasi-concave, non-decreasing, locally non-satiated, and continuous.
and \(p \in E^m_{++}\) and \(y_h \in E_+\) are prices and household income or consumption expenditures. \(V_h^h\) and \(C_h\) are related by

\[
V_h^h(p, y_h) = u_h \leftrightarrow C_h(u_h, p) = y_h. \tag{4}
\]

A project moves the economy from one general equilibrium to another. Prices change from \(p^b \in E^m_{++}\) to \(p^a \in E^m_{++}\) (the superscripts stand for ‘before’ and ‘after’) and incomes change from the vector \(y^b \in E^H_+\) to \(y^a \in E^H_+\).

The *compensating variation* or *willingness to pay* for household \(h\) (\(s^c_h\)) is the maximum amount (positive or negative) that the household would pay to secure the project. Thus, if household \(h\) were actually to give up its compensating variation in state \(a\), it would be indifferent between the two states. \(s^c_h\) is defined implicitly by

\[
V_h^h(p^a, y^a_h - s^c_h) = V_h^h(p^b, y^b_h) = u^b_h, \tag{5}
\]

where \(u^b_h\) is the household’s utility before the project. Using (4),

\[
C_h(u^a_h, p^a) = y^a_h - s^c_h, \tag{6}
\]
or, since \(y^a_h = C_h(u^a_h, p^a)\),

\[
s^c_h = C_h(u^a_h, p^a) - C_h(u^b_h, p^a). \tag{7}
\]

Alternatively, by subtracting and adding the household’s base-period income in (7), we can rewrite the compensating variation as

\[
s^c_h = [y^a_h - y^b_h] + [C_h(u^b_h, p^b) - C_h(u^b_h, p^a)]. \tag{8}
\]

This latter formulation is of interest for two reasons. First, it shows that, in the absence of price changes, the compensating variation is the difference in incomes between the two states. Second, if there is no income change (a common assumption in the consumer’s-surplus literature), the compensating variation is given by the second bracketed term in (8).

Since \(C_h\) is increasing in \(u_h\), (7) demonstrates that the compensating variation \(s^c_h\) is an exact index of welfare change for households \(h\); that is,

\[
u^b_h \geq u^a_h \iff s^c_h \geq 0. \tag{9}
\]

The *equivalent variation*, or *willingness to accept*, is the minimum payment that would induce household \(h\) to forgo the project. It is implicitly defined by

\[
u^b_h = V_h^h(p^a, y^a_h) = V_h^h(p^b, y^b_h + s^e_h), \tag{10}
\]

7 We assume positive prices for convenience, but the arguments can be generalized to allow some free goods. One of the prices may be chosen as a numeraire and set equal to one, or prices may be constrained to add to one. Compensating and equivalent variations can be defined using (5), (7), (8), (10), (11), and (12) when some or all prices are household-specific.

8 A strict inequality on either side of the equivalence symbol ‘\(\leftrightarrow\)’ implies a strict inequality on the other.
where $u^a_h$ is the household's utility after the project. Using (4),

$$s^e_h = C^h(u^a_h, p^b) - C^h(u^b_h, p^b).$$  \hspace{1cm} (11)

Alternatively, by adding and subtracting $y^a_h = C(u^a_h, p^a)$, (11) can be rewritten as

$$s^e_h = [y^a_h - y^b_h] + [C^h(u^a_h, p^b) - C^h(u^a_h, p^a)].$$  \hspace{1cm} (12)

The comments made about (8) are equally relevant here. Again, $s^e_h$ is an exact index of welfare change for household $h$, and

$$u^a_h \geq u^b_h \Leftrightarrow s^e_h \geq 0.$$  \hspace{1cm} (13)

Comparing (11) and (7), it is easy to see that the equivalent variation associated with the move from $a$ to $b$ is minus the compensating variation for the move from $b$ to $a$.

The Scitovsky Set $S(u)$ for utility vector $u = (u_1, \ldots, u_H)$ is the set of aggregate consumption vectors $x = \Sigma x^h$ which can, by appropriate distributions to the $H$ households, provide each household with a utility level at least as great as its utility level in the vector $u$. Thus, if the aggregate vector $x$ is in $S(u)$ there exists a distribution $X = (x^1, \ldots, x^H)$ such that $U^h(x^h)$ is at least $u^h$, $h = 1, \ldots, H$. More formally, the Scitovsky Set $S(u)$ is given by

$$S(u) = \left\{ x \in E^m_+ \mid x = \sum_h x^h, U^h(x^h) \geq u^h, h = 1, \ldots, H \right\}.$$  \hspace{1cm} (14)

The definition of $S(u)$ assumes, implicitly, that any aggregate consumption vector $x$ can be distributed to the $H$ households with costless lump-sum transfers of goods.\(^9\)

The Scitovsky Set has another useful characterization. Denote the ‘no-worse-than’ set for household $h$ at utility level $u^h$ by

$$N^h(u^h) = \{ x^h \in E^m_+ \mid U^h(x^h) \geq u^h \}.$$  \hspace{1cm} (15)

Then, using the rules for set summation,\(^10\) the Scitovsky Set (14) can be written as

$$S(u) = \sum_h N^h(u^h).$$  \hspace{1cm} (16)

If an aggregate consumption vector $x$ is on the boundary of $S(u)$, then, any distribution $(x^1, \ldots, x^H)$ of it with $U^h(x^h) \geq u^h, h = 1, \ldots, H$ necessarily results

\(^9\) Transfers are costless if none of $x$ used up as administration or other costs in the transfer process, and lump-sum if the amount received by any household is not influenced by the household’s behaviour. Taxes are negative transfers.

\(^10\) The sum of any two sets $S_1$ and $S_2$ in $E^n$ is the set $(S_1 + S_2)$ in $E^n$, defined by $S_1 + S_2 := \{ x \in E^n \mid x = x^1 + x^2, x^1 \in S_1, \text{and} x^2 \in S_2 \}$.
in $U^h(x^h) = u_h$ for each $h$. Thus, all such distributions are consumption-efficient. This in turn implies that if every household consumes positive quantities and utility functions are differentiable at the point in question, then the marginal rates of substitution are equal for all households.

A dual representation of the Scitovsky Set is often useful. We call the support function of the Scitovsky Set the Scitovsky Expenditure Function. It is the minimum total income needed to bring each household to utility level $u_h$, $h = 1, \ldots, H$, when prices are $p$, and is defined by

$$C(u, p) = \min_x \{ p \cdot x \mid x \in S(u) \}. \quad (17)$$

using (16),

$$C(u, p) = \min_x \{ p \cdot \sum_h x^h \mid x^h \in N^h(u_h), h = 1, \ldots, H \}, \quad (18)$$

and, using (3) and (15),

$$C(u, p) = \sum_h C^h(u_h, p). \quad (19)$$

The Scitovsky expenditure function is the simple sum of individual expenditure functions evaluated at the appropriate utility levels.

III. THE SUM OF COMPENSATING OR EQUIVALENT VARIATIONS IN AN EXCHANGE ECONOMY: THE BOADWAY PARADOX

Boadway discovered (1974) that the sum of compensating variations is non-negative, and usually positive, for any move in an exchange economy from one efficient allocation to any other efficient allocation. In this section we provide a general proof of his proposition together with a simple geometric argument using Scitovsky sets.

There is an initial endowment given by $Q = (\omega^1, \ldots, \omega^H)$ where $\omega^h$ is the endowment of household $h$. A Walrasian (price-taking) equilibrium is characterized by an equilibrium price vector $p \in E^n$ and an allocation $X^e = (x^{1e}, \ldots, x^{He})$ such that

$$p \cdot \omega = \sum_h p \cdot \omega^h = \sum_h p \cdot x^{he}. \quad (20)$$

For each $h$, $x^{he}$ maximizes $u_h = U^h(x^h)$ over the household’s budget set.

The first theorem of welfare economics tells us that any Walrasian equilibrium with $\sum \omega^h = \omega$ is a Pareto optimum in the exchange economy with aggregate endowment $\omega$. Further, for convex preferences the second theorem guarantees that every Pareto-optimum in the exchange economy can be decentralized as a Walrasian
equilibrium for some $\Omega = (\omega^1, \ldots, \omega^H)$ with $\sum \omega^h = \omega$. Consequently, moves between Pareto-optima are equivalent to moves between Walrasian equilibria with different distributions of $\omega$.

Hence, we consider two different distributions $(\omega^{lb}, \ldots, \omega^{Hb})$ and $(\omega^{la}, \ldots, \omega^{Ha})$ of the aggregate endowment $\omega$ with

$$\sum_h \omega^{hb} = \sum_h \omega^{ha} = \omega. \quad (21)$$

Let the associated equilibrium price and allocation vectors be represented by $(p^b, p^a)$ and $X^b = (x^{1b}, \ldots, x^{Hb})$ and $X^a = (x^{1a}, \ldots, x^{Ha})$ respectively.\(^{11}\) Since this is a move from one Pareto-optimum to another, there is no efficiency gain or loss and it is impossible for the gainers to compensate the losers and still gain. Nevertheless, the sum of compensating variations is always non-negative and usually positive. To demonstrate this, note that using (7) and (19) – the sum of compensating variations is

$$\sum_h s^c_h = \sum_h \left[ C^h(u^a_h, p^a) - C^h(u^b_h, p^b) \right]$$

$$= \sum_h C^h(u^a_h, p^a) - \sum_h C^h(u^b_h, p^b)$$

$$= C(u^a, p^a) - C(u^b, p^b). \quad (22)$$

Because $x^a = x^b = \omega$, the aggregate consumption vector $x^a = \Sigma x^{ha}$ is in $S(u^a)$, and also in $S(u^b)$. Further, from the definition of the minimization in (17) we know that

$$C(u^a, p^a) = p^a \cdot x^a = p^a \cdot \omega. \quad (23)$$

Similarly, since $x^b$ is in $S(u^a)$, we find, again from the minimization in (17), that

$$C(u^b, p^a) \leq p^a \cdot x^b = p^a \cdot \omega. \quad (24)$$

Hence, from (22)–(24),

$$\sum s^c_h \geq p^a \cdot \omega - p^a \cdot \omega = 0. \quad (25)$$

This completes the proof that, in an exchange economy, the sum of compensating variations is always non-negative. The inequality in (25) will be strict if relative prices in $p^a$ are different from relative prices in $p^b$ and preferences allow some substitution possibilities as relative prices change.

\(^{11}\) The same change could be brought about by lump-sum transfers of purchasing power. The two fundamental theorems of welfare economics guarantee that the set of Walrasian equilibria associated with different distributions of $\Omega$ is the set of Pareto optima.
There are only two general cases where a strictly positive sum of compensating variations will not result from a move between Pareto-optima. The first of these is the case where Gorman’s (1953) condition for the existence of an aggregate consumer is satisfied (see section vi, below). In that case, because all consumers have identical marginal expenditure patterns, changes in endowments do not change excess demands. This in turn implies that the change in the distribution of initial endowments does not induce a price change and hence that the sum of compensating (equivalent) variations will be zero. Alternatively, if relative prices do change and household’s preferences do not allow substitution (to take advantage of the new market signals), the sum of compensating variations will be zero. These two cases are unlikely and are not supported by empirical investigations; therefore, in general, we may expect positive sums of compensating variations.

A similar result may be obtained for the sum of equivalent variations between Walrasian equilibria. Since the equivalent variation from $b$ to $a$ is minus the compensating variation from $a$ to $b$, it follows that the sum of the equivalent variations must be non-positive; that is,

$$\sum_h s^e_h \leq 0. \quad (26)$$

A strictly negative sum results when relative prices change and preferences allow substitution.

This argument is illustrated in figure 1. The project moves the economy from $X^b$ to $X^a$ in the Edgeworth box with aggregate endowment $\omega$. The two Scitovsky sets $S(u^a)$ and $S(u^b)$ contain $x^a = x^b = \omega$, and that point is on the boundary of each. A price line through $\omega$ must support the appropriate Scitovsky Set; that is, a line through $\omega$ whose slope is $(-p_1/p_2)$ must be tangent to the Scitovsky Set $S(u^a)$ at $\omega$; similarly, a line with slope $(-p_1/p_2)$ through $\omega$ must be tangent to $S(u^b)$ at $\omega$.

If prices change when $\omega$ is redistributed, then the new Scitovsky Set will (usually) intersect the old one at $\omega$, and $C(u^b, p^a)$ will be strictly less than $p^a \omega$. For example, in figure 1,

$$C(u^b, p^a) = p^a \cdot \hat{x} < p^a \cdot \omega = C(u^a, p^a), \quad (27)$$

and the sum of compensating variations is strictly positive.

As an example, let the preferences of two agents be given by

$$u_1 = U^1(x^1) = (x^1_1)^{1/2}(x^1_2)^{1/2}, \quad (28)$$

and

$$u_2 = U^2(x^2) = x^2_1. \quad (29)$$

Their expenditure functions are given by

$$C^1(u_1, p) = 2p_1^{1/2}p_2^{1/2}u_1 \quad (30)$$
and

\[ C^2(u_2, p) = p_1 u_2, \] (31)

respectively. Letting the initial endowments be

\[ \omega^b_1 = (8, 0), \ \omega^b_2 = (8, 4), \ \omega^a_1 = (16, 4), \ \text{and} \ \omega^a_2 = (0, 0), \] (32)

equilibrium prices are

\[ p^b = (1, 1), \ \text{and} \ p^a = \left( \frac{1}{4}, 1 \right). \] (33)

The compensating and equivalent variations can be shown to be

\[ s^c_1 = 4, \ s^c_2 = -3, \ s^e_1 = 8, \ \text{and} \ s^e_2 = -12. \] (34)
The sums of the compensating and equivalent variations are, respectively, \( s'_1 + s'_2 = +1 \) and \( s'_3 + s'_4 = -4 \).

This result is called the Boadway Paradox (from Boadway 1974). It means that neither the sum of compensating variations nor the sum of equivalent variations is a 'pure efficiency' index. Further, if there are multiple equilibria for a given \( \Omega \), the sum of compensating variations may be positive without a change of endowments. If social preferences are based on positive sums of compensating variations, then social preference reversals or asymmetries (\( X^a \) preferred to \( X^b \) and \( X^b \) preferred to \( X^a \)) may occur without Scitovsky preference reversals (reversals according to the Potential Pareto Principle or compensation test).

IV. THE SUM OF COMPENSATING OR EQUIVALENT VARIATIONS IN A PRODUCTION ECONOMY

In order to extend Boadway’s result to a production economy, we postulate \( F \) competitive firms, \( f = 1, \ldots , F \). \( T_f \) is firm \( f \)'s production set: it is the set of feasible input-output vectors for the firm, with inputs measured as negative numbers. The economy’s production set is

\[
T := \sum_{f} T_f, \quad (35)
\]

and the set of feasible aggregate consumption vectors is given by

\[
\tilde{T} := (T + \{\omega\}) \cap E^m_+, \quad (36)
\]

where \( \omega \) is the aggregate endowment; \( \tilde{T} \) is the production possibility set. Price-taking, profit-maximizing behaviour will ensure that, for any equilibrium prices \( p \in E^m_+ \), \( x = \Sigma_x^h \) will be chosen to maximize \( p \cdot x \) over the set \( \tilde{T} \).

We again consider a move from one Walrasian equilibrium to another, noting that in this case \( x^b \) and \( x^a \) may be different.\(^{12}\)

The sum of compensating variations is given by

\[
\sum_{h} s^c_h = \sum_{h} C(h, a, p^a) - \sum_{h} C(h, b, p^a) = C(u^a, p^a) - C(u^b, p^a). \quad (37)
\]

We know, of course, that

\[
C(u^a, p^a) = p^a \cdot x^a, \quad (38)
\]

and, since \( U^h(x^h) = u^h \) for \( h = 1, \ldots , H \),

\[
C(u^b, p^a) \leq p^a \cdot x^b. \quad (39)
\]

\(^{12}\) If there is a linear technology, then no price change is possible and the sum of compensating variations will always be zero.
Hence,

$$\sum_h s_h^c \geq p^a \cdot x^a - p^a \cdot x^b.$$  \hspace{1cm} (40)

$x^a$ and $x^b$ both belong to $\tilde{T}$, and, since $x^a$ is chosen at prices $p^a$, profit maximization ensures that

$$p^a \cdot x^a \geq p^a \cdot x^b.$$  \hspace{1cm} (41)

It follows that

$$\sum_h s_h^c \geq 0.$$  \hspace{1cm} (42)

This result is demonstrated in figure 2. Letting $\bar{x}$ minimize $p^a \cdot x$ subject to $x \in S(u^b)$, Figure 2 shows clearly that $p^a \cdot x^a \geq p^a \cdot \bar{x}$. 

FIGURE 2
A similar argument guarantees that
\[ \sum_{h} s_{h}^{e} \leq 0 \]
(43)
for the same change.

Expression (42) does not depend on state b’s being a Walrasian equilibrium; it is true for a change from any ‘distorted’ equilibrium (with consumer prices \( p_{b}^{c} \)) to any Walrasian equilibrium.\(^{13}\) If preferences and technologies are convex, (42) is true for any move to an efficient allocation. Similarly, (43) is true for any move away from a Walrasian equilibrium or (with convexity) away from an efficient allocation.

It follows that the sum of compensating variations exhibits an upward bias relative to an ideal efficiency measure, and the sum of equivalent variations exhibits a downward bias. In fact, it is clear that the sum of compensating variations is ‘more likely’ to be positive\(^{14}\) for efficient moves in an economy with production than in an exchange economy.

V. WILLINGNESS-TO-PAY AND THE POTENTIAL PARETO PRINCIPLE

The \textit{Potential Pareto Principle} (hereafter, the \textit{PPP}) evaluates state a as ‘at least as good’ (represented by \( R^{P} \)) as state b if and only if it is possible to rearrange the allocation in a so that the resulting allocation weakly Pareto dominates the allocation in b. Writing \( X \) for the allocation \((x^{1}, \ldots, x^{H})\),

\[ X^{a} R^{P} X^{b} \iff \sum_{h} x^{ha} = x^{a} \in S(u^{b}). \]
(44)

Letting \( P^{P} \) stand for the strict preference relation\(^{15}\) yields

\[ X^{a} P^{P} X^{b} \iff x^{a} \in \text{int} S(u^{b}). \]
(45)

\( X^{a} P^{P} X^{b} \) means that it is possible, through redistribution of \( X^{a} \), to find a strict Pareto improvement over \( X^{b} \). Thus \( X^{a} P^{P} X^{b} \) if and only if \( X^{a} \) is a \textit{Potential Pareto Improvement} on \( X^{b} \).

The \textit{PPP} implicitly assumes (in (44)) the possibility of costless lump-sum transfers of goods to individual households.

Given the biases of sums of compensating and equivalent variations, one might not be surprised by the claim that a positive sum of compensating variations is necessary (but not sufficient) for such an improvement, while a positive sum of

\(^{13}\) This result is proven by Schweizer (1983) and by Diewert (1985). Diewert considers the ‘Hicks-Boiteux measure of waste or deadweight loss’ which is minus the sum of equivalent variations \textit{from} a selected Pareto optimum to a distorted equilibrium. This is equal to the sum of compensating variations for the reverse move. Diewert’s proof applies to a small open economy.

\(^{14}\) And more likely to be negative for equivalent variations.

\(^{15}\) \( P^{P} \) is \textit{not} the asymmetric factor of \( R^{P} \). That is, if \( P^{*} \) is defined by \( X^{a} R^{P} X^{b} \) and not \( X^{b} R^{P} X^{a} \), then \( P^{*} \) is not equal to \( P^{P} \). Scitovsky reversals show that \( P^{P} \) can fail to be asymmetric.
equivalent variations is sufficient (but not necessary). But such pleasant symmetries are not to be found – a positive sum of compensating variations is necessary for an improvement according to the PPP, but it is not sufficient; worse still, a positive sum of equivalent variations is neither necessary nor sufficient for a Potential Pareto Improvement.¹⁶

In this section we allow (but do not require) the technology sets \{T_f\} and endowments \{\omega^h\} (but not preferences) to change. Thus, prices \(p^a\) and \(p^b\) may take on any values in \(E_++^m\). These changes may be due to technical change, government provision of public inputs, etc. The arguments are completely general as long as everyone faces the same prices. No assumptions concerning competition or lack of distortions is needed for the following argument.

First, we show that \(\Sigma \delta^c_h > 0\) is a necessary condition for a Potential Pareto Improvement. Suppose that \(X^a\) is a Potential Pareto Improvement on \(X^b\), that is, \(x^a \in \text{int} S(u^b)\). Referring to figure 3(a), let \(\bar{x}\) minimize \(p^a \cdot x\) over \(S(u^b)\) so that

\[ p^a \cdot \bar{x} < p^a \cdot x^a. \]  
(46)

Clearly,

\[ C(u^b, p^a) = p^a \cdot \bar{x}, \]  
(47)

and by definition

\[ C(u^a, p^a) = p^a \cdot x^a, \]  
(48)

so that

\[ \sum \delta^c_h = C(u^a, p^a) - C(u^b, p^a) > 0, \]  
(49)

which completes the argument.

This demonstrates that a Scitovsky preference reversal (according to the PPP, \(X^a\) is an improvement on \(X^b\) and \(X^b\) is an improvement on \(X^a\)) implies that the sum of compensating variations is positive in both directions. However, the Bovdway Paradox shows that ‘compensating-variation reversals’ are more common than Scitovsky reversals.

Next we show, as might be surmised from (43), that a positive sum of equivalent variations is not necessary for an improvement according to the Potential Pareto Principle. If prices change, the sum of equivalent variations is normally negative when moving from one competitive equilibrium to another. A small additional move into the interior of \(S(u^b)\) can be made without reversing the sign. Figure 3(b)

¹⁶ Since the sum of equivalent variations is minus the sum of compensating variations for the move from \(a\) to \(b\), a positive sum of equivalent variations for the move from \(a\) to \(b\) is sufficient for the reverse move not to be a Potential Pareto Improvement.
shows a Potential Pareto Improvement, \(x^a \in \text{int } S(u^b)\); let \(\hat{x}\) minimize \(p^b \cdot x\) over \(S(u^a)\) so that
\[
p^b \cdot \hat{x} - p^b \cdot x^b < 0. \tag{50}
\]
Since,
\[
C(u^a, p^b) = p^b \cdot \hat{x}, \tag{51}
\]
and, by definition
\[
C(u^b, p^b) = p^b \cdot x^b, \tag{52}
\]
it follows that
\[
\sum_h s^e_h = C(u^a, p^b) - C(u^b, p^b) < 0, \tag{53}
\]
which demonstrates the claim.

Finally we show that \(\sum s^e_h > 0\) is not a sufficient condition for a Potential Pareto Improvement. This result is illustrated in figure 3(c). There, \(x^b \not\in S(u^a)\) and \(\bar{x}\) minimizes \(p^b \cdot x\) over \(S(u^a)\); clearly,
\[
C(u^a, p^b) = p^b \cdot \bar{x} > p^b \cdot x^b = C(u^b, p^b) \tag{54}
\]
so that
\[
\sum_h s^e_h = C(u^a, p^b) - C(u^b, p^b) > 0. \tag{55}
\]

This demonstrates that although a positive sum of compensating variations is necessary for a Potential Pareto Improvement, it is not sufficient; furthermore, a positive sum of equivalent variations is neither necessary nor sufficient.

If these measures will not do, what will? We can define a consumer's surplus in terms of a numeraire good as follows. \(s^{nc}_h\) is \(h\)'s numeraire-based compensating variation for the change if and only if
\[
U^h(x^{ha} - s^{nc}_h e_1) = u^b_h, \tag{56}
\]
where \(e_1 = (1, \ 0, \ \ldots, \ 0)\). Good 1 is the numeraire chosen, and each \(U^h\) must be strictly increasing in its first argument. \(s^{nc}_h\) is the maximum amount of the numeraire that household \(h\) would give up in state \(a\), given that consumption of other goods is unchanged, to secure the change. \(\Sigma s^{nc}_h > 0\) is a sufficient but not a necessary condition for a Potential Pareto Improvement (Schweizer 1983). To see this, imagine taking away exactly \(s^{nc}_h\) units of the numeraire from household \(h\) (if \(s^{nc}_h < 0\), household \(h\) gets a transfer). Then household \(h\) will have arrived at the
utility level \( u_h^b (h = 1, \ldots, H). \) Since \( \Sigma s_h^{sc} > 0, \) an equal share can be given to each household, making them all better off. Hence \( x^a \) is a Potential Pareto Improvement over \( x^b. \)

To see that the above test is not necessary, define the statistic\(^{17} \sigma \) by

\[
\sigma = \max \left\{ s \mid (x^a - se_1) \in S(u^b) \right\}.
\]

Clearly \( \sigma > 0 \) if and only if \( x^a \in \text{int} \ S(u^b); \) so \( \sigma > 0 \) is necessary and sufficient for a Potential Pareto Improvement. It is clear from the definitions that

\(^{17} \) This measure was introduced by Dierker and Lenninghaus (1983). See Schweizer (1983) for a good discussion.
\[ \sum s_{hc}^{nc} \leq \sigma \]  
(58)

(see Schweizer 1983 for a careful proof). Consequently, \( \Sigma s_{hc}^{nc} > 0 \) is not necessary.\(^{18}\)

Why not use \( \sigma \)? Its computation requires knowledge of the Scitovsky set, but it can be found from individual (household) preferences. Preferences, in turn, can be found if individual demand functions are known. However, knowledge of preferences is, potentially at least, more useful than knowledge of Scitovsky sets, compensating variations, or \( \sigma \) – all of these can be computed from preferences. On the other hand, investigators may believe that sums of compensating variations can be approximated from aggregate demand data. This is possible only when there is a representative consumer (see section vi, below) and, in that case, there are no Scitovsky reversals or compensating-variation reversals and \( \sigma \) is superfluous – a positive sum of compensating variations is necessary and sufficient for a Potential Pareto Improvement. At the same time, demand studies show that individual preferences do not satisfy the conditions necessary for there to be a representative consumer.

VI. RATIONALITY CONDITIONS AND THE COMPENSATING-VARIATION TEST

Although the Potential Pareto Principle and the aggregate willingness-to-pay criterion (sum of compensating variations) are different, both are flawed; both fail to order alternative allocations\(^{19}\) sensibly.

Consider first the Potential Pareto Principle; it may exhibit Scitovsky (1942) reversals, even when the allocations in question are consumption efficient (total consumption is distributed efficiently). That is, there may exist \( X^a \) and \( X^b \) with \( u^a := (U^1(x^{1a}), \ldots, U^H(x^{Ha}), u^b := (U^1(x^{1b}), \ldots, U^H(x^{Hb}), x^a \in \text{boundary } S(u^a) \) and \( x^b \in \text{boundary } S(u^b) \) such that\(^{20}\)

\[ X^a \not\equiv P^X X^b \]  
and \( X^b \not\equiv P^X X^a. \)  
(59)

Further, Gorman (1955) showed that, if attention were limited to consumption-efficient bundles that were not subject to Scitovsky reversals, \( P^P \) could be intransitive. In fact, combining the results in Gorman (1953, 1955, 1961), the Potential

---

18 Using analogous equations to the numeraire-based equivalent variations \( \{s_{hc}^{ne}\} \) can be defined. If their sum is non-positive, then \( X^a \) is not a Potential Pareto Improvement on \( X^b \). Thus, \( \Sigma s_{hc}^{ne} > 0 \) and \( \Sigma s_{hc}^{ne} \leq 0 \) are together sufficient (but not necessary) for the prevention of a Scitovsky reversal (Schweizer 1983).

19 An ordering \( R \) is a binary ‘no-worse-than’ relation that is reflexive, transitive and complete. The corresponding strict preference relation \( P \) is defined by \( xPy \iff [xRx \text{ and not } yRx] \). \( P \) must be asymmetric; that is, \( xPy \not\leftrightarrow \text{not } yPx \). Scitovsky reversals occur because the \( PPP \)'s strict preference relation is not its asymmetric factor.

20 Remember that allocations are efficiently distributed if and only if they are in the boundary of the Scitovsky set.
Pareto Principle *orders* consumption-efficient allocations if and only if preferences are quasi-homothetic and identical at the margin. In that case, the indirect utility functions can be written as

\[ V^h(p, y_h) \equiv \alpha(p)y_h + \beta^h(p), \]  

where \( \alpha \) is homogeneous of degree minus one and common across households, and \( \beta^h \) is homogeneous of degree zero; the symbol \( \equiv \) means 'is ordinally equivalent to.' At every price vector, agents' Engel curves for each good are parallel straight lines. In this case there is an 'aggregate consumer' whose preferences generate aggregate demands when income is equal to \( \Sigma y_h \). The no-worse-than sets of this aggregate consumer are the Scitovsky sets. They are the same for each consumption-efficient allocation corresponding to a given aggregate consumption \( x \); that is, through every point in the space there is but one Scitovsky indifference curve – the boundary of a Scitovsky set.

However, even this extremely strong restriction on preferences is not enough to rank consumption-inefficient allocations in a reasonable way. Suppose that, for a given \( x \in E_+^n \), there is a consumption-inefficient allocation. Then, there is a second consumption-inefficient allocation, given our assumptions, and according to the Potential Pareto Principle *each* is an improvement on the other! This occurs because the definition of consumption-inefficiency implies that there exist distributions of \( x \) that are Pareto superior to each. Therefore, preference reversal occurs. In fact, these two inefficient allocations may be Pareto ranked against each other, and the reversal still occurs.\(^{21}\) If this is to be ruled out, *all* distributions of a given \( x \) must be consumption efficient (the contract curve must fill the Edgeworth box). This can occur only when indifference surfaces are hyperplanes; this too can be found in Gorman (1955).

These results make it clear that the Potential Pareto Principle is not a good tool for cost-benefit analysis. A modification of the PPI has been analysed by Chipman and Moore (1971, 1973) which they call the Kaldor-Hicks-Samuelson Criterion. According to this criterion \( X^a \) is at least as good as \( X^b \) if and only if \( x^a \) belongs to *all* the Scitovsky sets for \( X^b \). It has the advantage of being a quasi-ordering.\(^{22}\) Unfortunately, it is, in most cases, a less comprehensive criterion than the PPI, failing to rank some Potential Pareto Improvements.\(^{23}\)

On the other hand, cost-benefit analysts often use compensating variations or their approximation by Marshallian surpluses. The ranking generated by the sum of

\(^{21}\) Imagine moving from a consumption-inefficient allocation in a two-household Edgeworth box to another allocation where both households are worse off. There are no gainers, but the two losers can compensate themselves for their losses by redistributing their total consumption efficiently. Thus the Pareto worsening is also a Potential Pareto Improvement.

\(^{22}\) That is, reflexive and transitive, but not necessarily complete.

\(^{23}\) In a recent paper Ruiz-Castillo (1987) shows (theorem 1) that if a project is at least as good as the status quo by the KHS Criterion, then the sum of the compensating variations is non-negative. This follows our necessity argument above as KHS implies PPI. R-C also shows that if preferences are identical and homothetic, then the sum of the cvs is equivalent to KHS. This follows from the Gorman result referred to above.
compensating variations is clearly different from the ranking generated by the PPP; so one might hope that it would behave differently. Unfortunately, its differences serve to make its performance even worse than that of the PPP.

Consumption-efficient allocations can (given convex preferences) be represented by a price vector, common to all households, and an income vector \( y = (y_1, \ldots, y_H) \). These price-income combinations can be ranked by the no-worse-than relation \( R^c \), where

\[
(p^a, y^a)R^c(p^b, y^b) \leftrightarrow \sum s_h^c \geq 0,
\]

with a strict inequality for strict preference.

A compensating-variation preference reversal occurs when

\[
(p^a, y^a)R^c(p^b, y^b)
\]

and

\[
(p^b, y^b)R^c(p^a, y^a).
\]

The results of sections III–V show that Scitovsky reversals imply compensating-variation reversals, and that the latter occur without the former (the Boadway Paradox). These reversals are ruled out if \( R^c \) is an ordering.

Roberts (1980) and Blackorby and Donaldson (1985) have shown that \( R^c \) is an ordering of efficient allocations of commodities if and only if (60) holds – that is, if and only if an aggregate consumer exists. The binary relation defined in (61) can be extended to consumption-inefficient allocations as well by allowing for household-specific prices. In our 1985 article we show that, in this case, \( R^c \) can never be an ordering.\(^{24}\)

When the PPP and the aggregate willingness-to-pay criteria order consumption-efficient alternatives consistently, they do it in exactly the same way, under the same restrictive conditions on preferences. Both do very badly on consumption-inefficient allocations.

The same problem arises for \( R^c \), the binary relation defined with the sum of equivalent variations in (61). \( R^c \) is an ordering of consumption-efficient allocations if and only if an aggregate consumer exists (60). When (60) holds, \( R^c \) agrees with the Potential Pareto Principle over consumption-efficient allocations.\(^{25}\) If households are allowed to face different prices, consistency is impossible.

---

\(^{24}\) If only some prices can be household specific, then \( \alpha \) in (36) must not depend on these prices. Thus, these goods may not exhibit income effects.

\(^{25}\) For a single household, the compensating variation may not rank two or more alternative projects correctly against each other when surpluses from the status quo are compared (Pauwels 1978). The equivalent variation does do this correctly. However the fact that a positive sum of equivalent variations across two or more households is neither necessary nor sufficient for a Potential Pareto Improvement should make this result of little interest.
When attention is restricted to consumption-efficient allocations and an aggregate consumer exists, $\mathbb{R}^c$, $\mathbb{R}^e$, $\mathbb{R}^p$ (the PPP) are identical and given by

$$X^a \mathbb{R}^c X^b \leftrightarrow X^a \mathbb{R}^e X^b \leftrightarrow X^a \mathbb{R}^p X^b \leftrightarrow \alpha(p^a) \left[ \sum y_h^a \right]$$

$$+ \sum_h \gamma(p^a) \geq \alpha(p^b) \left[ \sum y_h^b \right] + \sum_h \beta(p^b), \quad (64)$$

where $(p^a, y^a)$ and $(p^b, y^b)$ support allocations $X^a$ and $X^b$. (64) makes it clear that consistent ordering with these principles renders investigators indifferent to the distribution of income (for fixed prices).

The orderings in (64) require households to face the same prices – only consumption-efficient allocations are ranked. In many standard cost-benefit exercises however, the existence of public goods, semi-public goods, and different labour qualities rules out identical prices. The impossibility of consistent ordering of alternatives in such cases should count (we believe) very strongly against the compensating-variation (willingness-to-pay) test.

In the case where costless lump-sum transfers of goods are not feasible, the Potential Pareto Principle and the compensating variation test diverge in an interesting way. $\mathbb{R}^c$ – the binary relation based on compensating variations – is not affected by feasibility considerations, as (61) makes clear. But consistent application of the PPP is affected, because the principle requires the existence (but not the achievement) of a feasible state that can be reached by compensation from $X^a$ and is an actual Pareto-improvement on $X^b$. Thus, in a second-best world, $\mathbb{R}^c$ will rank some states of affairs as preferred that would be ranked as preferred by $\mathbb{R}^p$ given feasible costless lump-sum transfers, but are not ranked as preferred, given second-best considerations.

Other second-best situations complicate things a good deal. For example, we have shown (1988a) that, when governments have incomplete information concerning household preferences and rely instead on self-selection, some moves towards second-best Pareto-efficient allocations would be rejected by the compensating-variation (willingness-to-pay) test.

The compensating-variation test can be applied, as well, to projects that change probabilities of death (see, e.g., Jones-Lee 1976). We have investigated the possibilities for consistent aggregation of these compensating variations (1986) and have shown that, when individuals can experience different probabilities of death, consistent and reasonable aggregation is not possible.

VII. THE ETHICS OF WILLINGNESS-TO-PAY

The arguments of previous sections appear to be technical, having to do with simple consistency considerations. We believe, however, that they shed a good

26 If different labour qualities are dealt with by making different people’s leisure time different goods, then (64) cannot be satisfied with the same $\alpha$ for all agents. A similar remark applies to public goods.
deal of light on the ethics implicit in standard cost-benefit analysis and applied welfare economics.

Consider, to take the simplest case, a one-good economy (yams, say). An allocation in such an economy can be described by two statistics: the total amount of yams and individual shares in the total. It is tempting in such a situation to identify the total with efficiency (the ‘size of the pie’) and the shares with distribution. Indeed, in such an economy (if individuals are selfish), the condition for an aggregate consumer is satisfied, the $\text{PPP}$ and the compensating-variation ordering are identical, and

$$X^a \leq X^b \iff \sum x^{ha} \geq \sum x^{hb},$$

where $x^h$ is household $h$’s consumption of yams. In this situation, the separation of social changes into efficiency considerations (the total number of yams available) and distributional considerations makes sense.

Cost-benefit analysts and applied welfare economists have attempted to extend this separation to more complex economies by devising principles, such as the Potential Pareto and Willingness-to-Pay principles. For example Harberger (1971, 785) argues that ‘costs and benefits … should normally be added without regard to the individual(s) to whom they accrue.’ This clearly normative statement has strong positive consequences if it is to be applied consistently. Previous sections show clearly that consistent application is either impossible or requires the existence of an aggregate consumer. It is only in this very restrictive case that a many-good economy, with the same prices for all households, can sustain the notion of a level of total income which can be distributed to individual consumers without affecting prices. The reason is that demands are not sensitive to income distribution.27

Harberger does not claim, however, that he is indifferent to income distribution. His claim is rather that economists are not ‘professionally qualified to pronounce’ (785) on such issues. This sentiment is echoed by Mishan (1972, chap. 23), who defends the $\text{PPP}$ with the assertion that such changes can be divided into an actual Pareto improvement and a redistribution of income. But such views require a notion of efficiency that is independent of the distribution of income – an idea that makes no sense in real-world economies. Costless lump-sum transfers (and taxes) are not feasible, and, in addition, governments cannot be counted on to pursue distributive justice rigorously and effectively. As a result, the social ethics implicit in rules such as the compensating-variation test must be subjected to serious scrutiny and judged on their own merit.28

27 Another way to think about this is to ask whether indifference to the distribution of income can be consistent with a Bergson-Samuelson social-welfare function. Roberts (1980) has shown that this requires the existence of an aggregate consumer.

28 Ng (1984, 1037) argues that ‘the objective of achieving a more equal distribution of income is better achieved through income taxation even if disincentive effects are involved since purely equality-oriented preferential policies have efficiency costs.’ Even if governments actively pursue distributive justice through income taxes and transfers, this claim, even in cases where it is true, does not justify disregard for distribution in applied welfare economics. Instead, criteria that exhibit inequality aversion should be applied to the whole range of government activities.
The compensating-variation criterion ranks all distributions of a given total money income (prices constant) as equally good. This is inconsistent with almost everyone's ethical preferences and with social policy. Indeed, in a society of individuals with identical strictly concave utility functions (that is, with diminishing marginal utility of income), this rule requires that the utilities of the rich be regarded as socially more important than the utilities of the poor. Although there is near unanimity about the undesirability of such social ethics, people are much less unanimous in their judgments of the appropriate level of inequality-aversion. Surprisingly, Harberger uses this disagreement to advocate indifference to income-inequality, at least in the work of professional economists. A more reasonable conclusion might be that criteria that exhibit various degrees of inequality aversion should be employed, in the manner of an ethical sensitivity analysis.

Families of social-evaluation functions in which inequality aversion is described by a single parameter can be used for this task. They require the employment of measures of levels of well-being rather than the differences that consumers' surpluses assess, and, because of this, the social evaluations produced are free from rationality problems (social preference reversals and intransitivities). We have described some of these possibilities in (1987) and (1988c).

Of course, as Harberger rightly notes, these criteria should not be used to reject a particular project on distributional grounds if other policies (tax-transfer policies for example) are actually used to offset the distributional losses. Since practical tax-transfer policies are not neutral, distributionally sensitive cost-benefit tests can and should be used to evaluate such combinations of social changes.

VIII. CONCLUSION

We believe that the arguments presented above present an overwhelming case against the use of consumers' surpluses in cost-benefit and general applied welfare analysis. Although these arguments are not new, many economists continue to use consumers' surpluses. How can such behaviour be justified?

Ng (1979, 98) has argued that the Boadway Paradox and (by implication) the failure of the compensating-variation test to produce an ordering of social alternatives can be ignored because 'the payment of compensation is unlikely to change prices significantly.' It should be clear from our analysis however that relative price changes due to projects is what gives rise to social preference reversals - the phenomenon is not limited to the payment of compensation. Further, Ng suggests that 'objective measures [such as compensating variations] can be ... no more than an approximate measure of welfare' (99) because 'the welfare of an individual is subjective state of mind' (98). Noting that this is complicated by data limitations, Ng then argues that 'the problems of ... inconsistencies etc. shrink into insignificance.' This seems misguided to us. If demand behaviour reveals only an approximation of 'real' well-being, it is important that estimated rankings of social

29 A referee suggested that we consider Ng's arguments.
alternatives can be thought of as approximating a social ordering that is free of inconsistencies. Simply supposing that any and all asymmetries and intransitivities are due to approximation error is not justified.

The other side of the compensating-variation test is the ethical judgment that ‘a dollar is a dollar’ – income-inequality is ignored. Under what circumstances might it be correct to adopt such a rule? If there were a government with the power to implement costless lump-sum taxes and transfers, and if this power were used to maximize a social-welfare function, then the marginal social value of a dollar would be the same, regardless of the household or individual that received it. Further, in this circumstance, social welfare could be written as a quasi-concave function of the aggregate consumption vector, and the aggregate compensating variation generated by this function would be equal to the sum of individual surpluses. In this world there would be no need for a cost-benefit test: aggregate preferences (known by the government) would be sufficient. In any other situation – where taxes and transfers are non-neutral or governments are less than completely serious about distributive justice – the social value of a dollar to John will, in most cases, be different from the social value of a dollar to Erwin, and distribution ought not to be ignored.

A third argument for clinging to the willingness-to-pay methodology might be that investigators do not know what to do instead. Such an excuse cannot be seriously advanced, given recent work on alternative cost-benefit methodologies that permit the inclusion of inequality aversion.

We believe that the choice of an alternate methodology should satisfy two criteria:

1. Indexes of household welfare *levels* (rather than of welfare gains and losses) should be used, so that the aggregation rules can be distributionally sensitive;
2. The aggregates employed should *order* the alternatives in a way that is consistent with normal distributional judgments.

Consumers’-surplus-based tests fail to meet both of these criteria. Other social-welfare indexes such as aggregates of money metrics satisfy (i) but may fail to satisfy (ii).30

We have investigated other methods for performing distributionally sensitive cost-benefit analysis. One is the employment of welfare ratios (ratios of household incomes to the household’s poverty lines) as indexes of well-being (Blackorby and Donaldson 1987). Another is the employment of household equivalence scales in estimated utility functions (Blackorby and Donaldson 1988c). the latter method has been used by Jorgenson and Slesnick (1984a, 1984b). In general, the procedure is straightforward. It requires an econometric procedure for estimating household preferences, a way to move from household well-being to individual well-being (such as equivalence scales), and a family of social-welfare functions with a parameter that allows for different degrees of inequality aversion. The impact of the project on incomes and prices must be forecast (with, perhaps, some aggregation into income

---

30 Money metrics are not always concave representations of preferences (Blackorby and Donaldson 1988b).
classes) and the project evaluated with different values of the inequality-aversion parameter. The results of these procedures are approximate, of course, but there is no underlying difficulty with the social ordering, and the ethics are explicit.

REFERENCES


