Consumer surplus: the first hundred years

Robert B. Ekelund, Jr., and Robert F. Hébert

O happiness! our being's end and aim!
Good, pleasure, ease, content! Whate'er thy name:
That something still which prompts the eternal sigh,
For which we bear to live, or dare to die.

—Alexander Pope, An Essay on Man

Introduction

Some economic ideas may be likened to volcanos—they are certain to erupt periodically. The cause of an ideational ‘eruption’ may be ‘environmental’—a reaction to recurring economic problems—or it may be a more fundamental assault on scientific definition. One such idea is the concept of consumer surplus, the Krakatoa of economic theory. Its long and spotty history has been marked by three major eruptions: the first at its inception; the second in consequence of the peak performance in cardinal-utility/demand theory; and the third in conjunction with the ordinal reconstruction of modern demand analysis.

On the one hand, the ‘doctrine of maximum satisfaction’ has not been and can never be made entirely ‘scientific’ or objective despite periodic counterclaims by some economists. On the other hand, economics makes little sense without it. Because economics deals with maximizing behavior under scarcity constraints, the measurement of satisfaction will always intrigue and frustrate economists. Such has been the case with the definition and measurement of consumer surplus. Diverse, often ambivalent, arguments appear in modern economic literature, as demonstrated by this sampling of recent titles: ‘The ambiguity of the consumer’s surplus measure of welfare change’ (Foster and Neuburger 1974); ‘Consumer’s surplus without apology’ (Willig 1976); ‘The plain truth about consumer surplus’ (Mishan 1977); ‘The ugly truth about consumer surplus’ (Foster and Neuburger 1978); and ‘The three consumer’s surpluses’ (Dixit and Weller 1979). The historical record will show, however, that debate and controversy are not new to the doctrine of consumer surplus. Past and present intellectual turmoil on the subject merely points up a continuing fascination with the
idea. Like it or not, consumer surplus theory, as cost-benefit analysis, is the bread and butter of the practicing economist.

The purpose of this article is to chronicle, analyze, and evaluate the first one hundred years of debate on the matter of consumer surplus, a debate that originated upon the very invention of the idea. The present history of the subject is, by and large, a fractured one. Previous studies have tended to slight various historical aspects that greatly illuminate the received doctrine. Of those papers dealing with early debates on the subject, neither Houghton (1958) nor Button (1979) have plumbed the true measure of Bordas's early contribution (1847). By contrast, Ahmed (1966) and Dooley (1983) virtually ignore the early development of the subject in order to concentrate on Marshall and his critics. Having a broader scope and purpose, Mishan's 1960 classic survey on welfare economics devotes relatively little space to consumer surplus. Its chief value to the matter at hand is that it provides a useful bibliography on the subject prior to 1960. Currie, Murphy, and Schmitz (1971) have surveyed the field of surplus concepts in general, concentrating almost exclusively on recent applications to international trade, taxation, and other areas of economic analysis. Despite some overlap in their treatment and ours of Marshall and Hicks, the Currie et al. study omits a number of important historical contributions that weigh heavily on the origin of the doctrine and its subsequent evolution. Moreover, our focus is exclusively on consumer surplus, to the neglect of other forms of economic rent.

This survey concentrates on the development of the concept from its initial formulation by Jules Dupuit (1844) through its 'rehabilitation' by J. R. Hicks (1941; 1942; 1943; 1946) roughly one hundred years later. This is followed by a modest and necessarily brief review of the present state of the literature. The conclusion of this lengthy investigation is that, instead of being the albatross of economic theory, the principle of consumer surplus is a highly useful mechanism in a world where purely scientific methods fail to accurately measure what we 'know' exists.

**Dupuit and His Critics**

*Origins of consumer surplus*

The theory of consumer surplus emerged simultaneously with the discovery of marginal utility and its application to demand theory. Although Cournot (1838 [1897, 78–81]) developed an adequate measure of producer surplus his method failed to produce the same result for consumer surplus. Cournot always measured the cost to consumers by the extra expenditure of those who continue to consume at a higher price rather than that amount plus the loss of those who stop consuming. Furthermore, he refused to identify utility with demand, thereby denying any operational
Ekelund and Hébert  . Consumer surplus  421

measure of psychic gain. The modern idea of consumer surplus, and the explicit conjunction of utility and demand that supports the idea, originated in the writings of Jules Dupuit. In a series of famous papers, Dupuit (1844; 1849a; 1849b; 1853) attacked the classical value-in-use/value-in-exchange dichotomy, substituting an improved theory of value in which price became the independent and simultaneous product of the forces of scarcity and marginal utility.

Dupuit (1853,7) unraveled the water-diamond paradox in a telling example of a city receiving ample water from a stream flowing through it. Owing to its abundance, water would have no value in exchange. In the face of scarcity, however (either natural or contrived), water takes on a value that is reflected in progressively higher prices as the quantity available for all uses declines relative to the demand for it. A city under siege, for example, may have its water supply so reduced by the enemy that none of the inhabitants would be willing to give up a liter of water, even though a diamond be offered in exchange for it. From such logic, and from observation of the markets for public works with which he was involved, Dupuit developed a workable theory of demand in which the marginal utility curve for any product or service is the demand curve for that good. The important corollaries that follow from this fusion are that (i) the area under the demand curve must equal the total utility of the good up to that point, and (ii) when price is zero, total utility is maximized.

Figure 1 depicts the fusion of demand and marginal utility in the form of Dupuit's courbe de consommation. Dupuit argued that the total utility (l'utilité absolue) of Or'' articles is equal to the area Or''n''P under the demand curve. From this he derived relative utility, or what is now called consumer surplus, by subtracting total costs of production, Or''n''p''. With reference to Figure 1, consumer surplus is equal to the area of the curvi-

\[ \text{Figure 1. The demand curve as a welfare measure} \]
linear) triangle, $p''n''P$. The remaining area $r''n''N$, Dupuit called "lost utility" (utilité perdue) in the sense that it could not be claimed by either consumers or producers for a market the size of $Or''$.

According to Dupuit, a change in relative utility (consumer surplus) could be calculated in the following manner. Suppose the price falls from $p''$ to $p'$ owing to a decrease in production costs, so that the quantity taken increases from $r''$ to $r'$. This raises absolute utility to $Or'n'P$. Subtracting costs of production $Op'n'nr'$ from this amount yields a total consumer surplus of $p'Pn'$. The net gain in consumer surplus is consequently measured by $p'n'n''p''$. In this way Dupuit developed a money measure of the benefit of public works and of goods in general, thereby forging the most important single tool of welfare economics. It was a significant breakthrough, developed in the peculiar milieu of the civil engineer forced to confront practical economic problems. But like all pioneer efforts, it was far from perfect.

Insofar as Dupuit's demand curve is a horizontal summation of individual demand curves, it presents an immediate problem. Interpersonal utility comparisons inevitably intrude on a market demand curve that is used to depict the utility surplus enjoyed by consumers of the product. A price may not represent the same utility to different individuals, since the price one would pay for a given quantity of a good depends not only on the utility afforded him by the good but on the income he possesses as well. In other words, the maximum price an individual is willing to pay for any unit varies with the amount of income he holds as well as with the utility the good provides. Thus we have the 'problem of the apostrophe.' If the concept under consideration is (aggregate) consumers' surplus, interpersonal utility comparisons are unavoidable; but the problem does not occur in the notion of a single individual's consumer surplus. Dupuit's discussions involved both concepts, but he put the greatest emphasis on consumers' surplus. Strictly speaking, then, differences in income distribution prohibit a legitimate utility summation; but as we shall soon see, Dupuit assumed away this problem.

A second problem in Dupuit's approach is the tacit assumption that utility is a measurable quantity. He regarded the true measure of the utility of an object as the "maximum sacrifice expressed in money that one is willing to make in order to procure it" (Dupuit 1849b, 177). Indeed, relative utility is defined as the difference between the maximum amount (price) the consumer would be willing to pay for each unit in his entire stock and what he must in fact pay for the entire stock. As stated above, it is the area under the demand curve above the total expenditures rectangle, and it is a money measure. But this measure cannot be a valid one if the marginal utility of money expenditures is allowed to change as price changes. The problem is one of distinguishing marginal utility curves on the one hand...
from demand curves on the other. Dupuit failed to make the distinction, with the result that his money measure, in all cases save one, tends to misstate true utility.¹

Another objection to Dupuit’s money measure of consumers’ surplus arises if the demand curve does not intersect the price axis. In such a case the offer price for the first unit(s) of the commodity is infinite, and consumers’ surplus is therefore unmeasurable. Dupuit (1853,26) attempted to skirt this problem by recognizing the limits to human knowledge. He observed:

when one cannot know something it is already quite a lot to know the limits of one’s knowledge. . . . We may not know that the utility of a canal will be only 5 million, but we might know that it will not be six, and that consequently we should forgo its construction; we may not know that the utility of a bridge will be 120,000F, but we could determine perhaps that it will be more than 80,000F and that may be sufficient to show that it will be very beneficial.

The problems of the constancy of the marginal utility of expenditures, and all that this implies, together with the interpersonal utility comparisons associated with a market demand curve (although later sidestepped by Dupuit) subsequently proved troublesome in the history of consumer surplus. But this did not void Dupuit’s use of the demand curve as an approximation of this surplus, nor did it render the definition of consumer surplus invalid. In fact, the idea persisted, and Dupuit was certainly not the last economist to proceed in such a fashion.

The Bordas offensive

Within a short time, Dupuit’s attack on established notions of utility elicited a major rebuttal. Like Dupuit, Louis Bordas was an engineer of considerable economic sophistication. However, his response to Dupuit consisted mostly of a mélange of confusions on the meaning of the word utility. Bordas (1847) defended Say’s theory of value, which confused utility with costs of production. Caught in a quagmire of terminology, he made some rather ill-advised statements on utility. At one point, for example, Bordas (1847,252) stated that “current price . . . depends on the intrinsic value of the monetary measure and on that of the object given in exchange.” At another (1847,258) he maintained that “the utility of . . .

¹ A deeper problem lurks in the conventional exposition of consumer surplus: the consumer is faced with a fixed budget; therefore he will not pay a higher price for ‘earlier’ units if he consumes up to the point where his demand curve meets the market price. This problem exercises modern welfare specialists, like Michael Burns (1973; 1975; 1977), who therefore tend to use consumer surplus in the context of a market demand curve, with each consumer taking only one unit.
tea is inherent to this substance and . . . does not at all depend on the price at which it is sold." These statements show that Bordas had no appreciation for Dupuit's marginal utility theory or for the solution it provided to the value-in-use/value-in-exchange dichotomy.

Nevertheless, Bordas brought out important and relevant points in his assessments of Dupuit's consumers' surplus concept. These criticisms, it turns out, echoed repeatedly against the doctrine. For example, J. S. Nicholson (1893; 1894; 1903) raised them in his attack on Marshall's formulation. Bordas admitted some connection between the utility of a certain quantity of a good and the maximum sacrifice which an individual would be willing to make for it, but the point he emphasized is that the sacrifice depends on a person's income and on the price of other goods as well. As Bordas (1847,278–79) stated:

Let us suppose that it is a matter of evaluating a kilogram of meat and that a person is asked to state the sacrifice that he is ready to make to procure it. Can this person answer categorically? Evidently not. Indeed, doesn't this sacrifice depend on the means of this person as well as the current price of other alimentary products which are capable of being substituted for the meat? . . . Therefore, what theory can one establish on so variable a basis and which depends on the taste as well as the means of each consumer?

Bordas (1847,282) pressed the argument further in reference to Dupuit's method of determining the utility and consumers' surplus of the quarry rock used in road-building:

insofar as the rock is taxed at a progressive rate, is it necessary to sell [its substitute] brick at its original price or at a new price? The result will be quite different according to what is done.

Bordas's argument asserted that if the price of brick is not held constant, Dupuit's measure of consumer surplus is rendered inoperable, because the demand curve for rock would shift erratically under such circumstances. Moreover, Bordas implied that since the necessary assumption of 'other things equal' generally does not hold in any concrete case, Dupuit's measure of consumers' surplus is practically useless. He was within the bounds of legitimate criticism on the former point, since Dupuit failed to invoke the explicit assumption of constancy of the prices of related goods.

Bordas also cast a jaundiced eye on Dupuit's tacit interpersonal utility comparisons. In ascertaining the desirability of bridges and other public projects, Dupuit sought to compare the project's utility with its costs. The utility of the project was measured, in the case of a bridge, for example, by placing incrementally increasing tolls in a fashion that revealed the
resulting use and its accompanying consumer surplus. Bordas (1847,283) objected to this calculation on the grounds that

... it is necessary, before applying it, to logically establish the relationship which connects the taxpayer's revenue loss to the sum of the relative utilities yielded by this approach. ... .This connection seems, in effect, very difficult, for the quantities to be matched or compared, although expressed in money, are altogether of a different kind.

The basis of Bordas's argument is that the marginal utility of a dollar collected from the taxpayer does not necessarily equal the marginal utility received from a dollar spent on any particular public project. In fact, Dupuit did ignore this problem, thereby leaving himself open to criticism of this sort.

The finance of public projects usually involves taxation and consequently a redistribution of income. Judgments about such redistribution require an illegitimate interpersonal welfare pronouncement. If the marginal utility of money was the same (and constant) for every individual in the economy, or alternatively, if the distribution of income were of no concern to the economist, it might be concluded that welfare is increased by a transfer, provided the increase in consumer surplus (in money terms) exceeds the money amount of the subsidy. Under such conditions a net increase in the money measure of utility is all that is needed. But if some such assumption is not invoked, it does not necessarily follow that welfare is increased by redistributing income from personal consumption to public projects, even if the money measure of the increase in consumer surplus is greater than the money amount of taxation required. Conceivably, such a transfer may involve a diminution in aggregate utility in spite of a net money-measure increase. This would occur if the utility decrease surrounding the tax receipts exceeded the utility increase to the consumers of the public good (i.e., the money measure of the increase in consumer surplus). Bordas (1847,284) correctly pointed out that “The whole question consists in knowing on what side the difference lies.”

It is not clear whether Dupuit fully appreciated the problems posed by the distribution-of-income question, but he may have had an inkling of them, because he tried to sidestep the issue from the outset. In his first article, Dupuit (1844,98–99) maintained that income distribution did not matter with respect to utility calculations, “because the losses and gains [from taxation and public works construction] offset each other.” Further, by declaring the matter of income distribution to be the province of the state rather than political economy, Dupuit apparently thought that he had cleared a major obstacle. Bordas was not so easily satisfied, and although he failed to acknowledge that Dupuit even recognized the problem, Bordas
was on the verge of unlinking demand curves from utility curves, and close
to the discovery of a Slutsky-type income-compensation principle.

To understand Bordas’s argument in detail, consider his example in which
a new manufacturing process reduces the price of stockings from 6 francs
to 3 francs. If the consumer has a fixed ‘stocking budget,’ he will be able
to buy eight pairs of stockings at the new price instead of the four pairs
previously purchased. But, according to Bordas (1847,260):

In order to consume as much as before, the individual must set aside
48 francs for the acquisition of this product, and reduce his other
consumptions by 24 francs. Compared to his starting position, it is as
though he had an annual gain of 24 francs, or that his income had
been increased by this sum. If, instead of consuming 8 pairs of stock-
ings, he only consumed 7 and used the 3 francs left over to buy other
things whose prices have not changed, his relative gain [on stockings]
would be no more than 21 francs.

In the first part of the above passage, the money expenditures on stock-
ings do not change; consequently, the marginal utility of money expendi-
tures is invariant. Letting \( x \) represent the quantity of stockings, and \( M \)
marginal utility, the mathematical expression for the price \( P \) of stockings
is \( P_x = M_x/M_e \). If expenditures on \( x \) remain constant when \( P_x \) falls, the
marginal utility of total expenditures remains constant, and the increased
purchases of \( x \) lead to a decline in \( M_x \). Ironically, Bordas’s first example
allowed the identification of utility and demand. This may be called Du-
puit’s case, since the individual’s demand curve can be identified with the
marginal utility curve for \( x \), and declines in price can be associated with
proportional reductions in marginal utility.

Figure 2 illustrates why the demand curve may represent a utility mea-
sure in this case. Assume that a consumer of stockings is in initial equilib-
rium at A. Bordas implied that the money proxy for the welfare gain is
given by an amount \( \Delta p \Delta q \), which is 24 francs in his example. Note that in
Figure 2 this quantity of income could be removed from the consumer
after the price decline, so that he or she would move to a new equilibrium
at C. The same quantity of stockings (8 pairs), in other words, would be
purchased (at points B and C) when the substitution effect is isolated from
the income effect. In this case, and in this case only (i.e., when demand
elasticity is \(-1\) and income elasticity is 0), the marginal utility curve may
be identified with the demand curve. Thus, 24 francs correctly measures
the change in welfare, since the whole increase in real income is used to
purchase additional stockings, and no part of the real-income increase is
devoted to expenditures on other goods. Money expenditures on the good
remain constant after the price decline, indicating, of course, that the de-
mand curve is of unit elasticity. Therefore, in this special case, a money measure under the demand curve may represent consumer surplus.

However, it is in discovering the other alternative open to the consumer that Bordas exposed the principal flaw in Dupuit’s (and Marshall’s) consumer-surplus theory. In the latter part of the above example, the consumer buys only 7 pairs of stockings at the lower price of 3 francs. The analytics of this ‘Bordas variation’ are presented in Figure 3.

Initially, the consumer is in equilibrium at A’. When the price of stockings is reduced to 3 francs per pair, the budget line of the consumer shifts outward. The new point of tangency with indifference curve I₁ is at point B’, after equilibrium is re-established and all effects have been accounted for. The new quantity taken, q₁, can be explained by both income and substitution effects in the following manner. Remove an amount of money income from the consumer equivalent to the increase in real income. The consumer would then choose combination C’ of money income and stockings. Thus, owing to the decrease in price alone, the consumer purchases additional stockings in the amount q₀q₂. The simultaneous price decrease/real-income increase, however, caused him to increase his purchases to q₁, and in equilibrium at B’, total expenditures on stockings have declined, as shown by the reduction from Y₀r to Y₀m. Alternatively, expenditures on

![Figure 2. The Dupuit-Marshall case](image-url)
all other goods have increased from Or to Om. Thus a part of the increase in real income is not realized by stocking gains, but by gains in other goods. Consequently, the demand curve for stockings cannot depict consumer surplus, for several reasons. In the first place, part of the increase in real income resulting from the price decline is spent on other commodities. This is a part of consumer surplus that the demand curve for stockings does not reveal. Moreover, since expenditures on other goods have increased, the marginal utility of money expenditures has decreased vis-à-vis the price decline. Given the formulation, \( P_x = M_x/M_e \), the change in the marginal utility of \( x \) can no longer be assumed proportionate to the change in the price of \( x \). The 'traditional' demand curve, where both income and substitution effects vary with price and quantity selections, cannot accurately measure the change in consumer surplus.

Although Bordas did not draw any of these implications from his discussion of income effects, it is to his credit that he suggested their existence. He did see that the entire real-income increase caused by a price decrease may not be spent entirely on additional units of the same commodity, and that the additional expenditures would disturb the demands for other goods. Had Bordas carried the argument a step further and shown that such 'income effects' may disturb the marginal utility of income or money expenditures, he would have presented the most convincing theoretical argument to date against the use of demand curves to measure con-

![Figure 3. The Bordas variation](image-url)
sumer surplus. Although he did not adhere to a marginal utility theory of value, his discussion, as is, could at least be said to presage the theoretical concerns of Slutsky (1915) and Hicks (1934; 1943; 1946). In any case, Bordas should be considered in the vanguard of the critics of consumers' surplus theory.

Dupuit's defense

The latent promise in Bordas's critique was unfortunately aborted by Dupuit's rejoinder (1849b), which ignored the problem of interpersonal utility comparisons and the pregnant suggestion that price changes may have 'income effects.' Dupuit took aim at easier targets: he castigated Bordas for repeating the errors of his predecessors and for adding new ones of his own invention; he denounced Bordas's multiple and ambiguous use of the term utility; and he rejected his critic's claim that utility is unmeasurable, citing Bordas's lack of proof for the assertion. With a measure of subtle irony, Dupuit enlisted Say as his ally against Bordas, reminding his fellow engineer that Say had also thought utility measurable despite its subjective and variable nature. In Dupuit's judgment, exposing the inaccuracies of Say's measure of utility was obviously one thing, whereas denying the prospect of measuring utility was quite another.

In retrospect, the one issue on which Dupuit capitulated seems less significant than those he ignored, but it nevertheless influenced later treatments of demand, especially Alfred Marshall's. In 1844 Dupuit failed to specify those 'determinants' of demand that serve to fix each individual's 'maximum sacrifice.' Bordas correctly chided Dupuit for this omission, citing the relevance of income, tastes, and the prices of related goods. Dupuit subsequently acknowledged the importance of these determinants, but cavalierly dismissed Bordas's complaint by declaring the ceteris paribus assumption implicit in his approach. The evidence for this is contained in Dupuit's (1849b, 184) answer to Bordas's query concerning the maximum sacrifice a consumer of meat would be willing to make (p. 424 above):

Would this price be the same for all persons? Evidently not. Because not only does this price depend on the wealth of that person, as Mr. Bordas observes, but on his taste for meat, on his hunger, on the

2. In his review of consumer surplus theory, Houghton (1958) reviewed two of Bordas's criticisms, but ignored the 'income effects' passage. Furthermore, in referring to a like criticism made later by Walras, he makes a rather poor assessment of this point. According to Houghton (1958, 52): "Dupuit's implied confusion [identification?] of demand and utility curves was of course a much less serious blunder [abstraction?] than Walras believed." This conclusion is untenable. The presence of a real-income effect and of a varying marginal utility of money expenditures puts an end to demand and utility curve identification and, therefore, to the use of demand curves to measure a 'utility' surplus. Since Dupuit did not hedge his theory with protective assumptions, his use of demand curves for such measurement is theoretically illegitimate, except in some rather restrictive circumstances.
prices of other food products and on a *thousand other circumstances* impossible to enumerate in complete fashion; but all these circumstances do not mean that this price does not exist for each object, each person and at each instant [emphasis supplied].

Whatever victory this provided Dupuit, it was a minor one, because he seems to have missed the flavor of Bordas’s criticism, which was that Dupuit’s measure of consumer surplus becomes suspect in any concrete case where the determinants of demand do in fact change.

As a whole, Dupuit’s rejoinder was disappointing. Had the issue of the marginal utility of money been clarified at this early date, later theorists would have been spared considerable confusion. But Dupuit turned a deaf ear to several of Bordas’s criticisms. In the end, Dupuit (1849b,205) steadfastly affirmed his original position, declaring: “I persist in the ideas on utility that I developed in 1844; I do not wish to change the formula that I gave for the measure of utility.”

**Anglo-Austrian extensions and continental criticism**

Although Dupuit’s rejoinder was clearly not the last word on the subject, the issue of consumer surplus made little further impact on economic literature until Léon Walras called attention to Dupuit’s measure in 1874. A few years before, unbeknownst to almost everyone, Fleeming Jenkin independently rediscovered Dupuit’s basic measure of consumer surplus and used it to determine the incidence of various taxes. The fundamental distinction between Jenkin and Dupuit is that the former eschewed utility considerations in developing a graphical measure of consumer surplus. It cannot be said, therefore, that Jenkin improved on Dupuit’s earlier performance by unlinking demand and utility curves. He never linked them in the first place, nor did he think such a linkage held much promise. Jenkin (1871,229) noted that (Jevons’ definition of) utility “admits of no practical measurement”; thus he opted for

a numerical estimate in money of the value of any given trade, which might be approximately determined by observing the effect of a change of prices on the trade; the [demand and supply] curves could certainly not, in most cases, be determined by experiment, but statistics gathered through a few years would show approximately the steepness of each curve near the market price, . . .[which] is the most important information.

Of course a purely statistical measure such as Jenkin proposed does not avoid all of the problems inherent in Dupuit’s original concept. Chief among the problems it does not confront is the existence of income and substi-
tion effects. Jenkin might have profited from the Dupuit-Bordas exchange, but he was apparently unaware of prior attempts to develop a welfare measure similar to his own. Ironically, Jenkin, like Dupuit, was a practicing engineer.

A 'mature' Walras (1874, 1926 [1954,445]) considered Dupuit's doctrine fallacious, but his reasoning—except in one important respect—was unoriginal. For example, his complaints—that Dupuit neither considered the effect of the utility and price of other goods on the 'maximum sacrifice,' nor understood that the 'consumer's means' also contributed to the determination of this sacrifice—were clearly misplaced in the light of Bordas's comment and Dupuit's subsequent rejoinder.3

A second point made by Walras was more significant, however. In the general equilibrium framework which he pioneered, income or 'wealth' is measured in terms of a numéraire commodity, one of constant purchasing power. This numéraire is also the commodity in terms of which all other prices are expressed. Walras (1874, 1926 [1954,445]) held that "Dupuit failed to see that the maximum pecuniary sacrifice in question depends in part . . . on the quantity of the wealth (measured in terms of a numéraire which the consumer possesses.)" In other words, the maximum sacrifice is determined not only by the utilities of all other goods in the consumer's array, but also by the quantity of wealth he holds in terms of the numéraire commodity. In the Walrasian system, however, each participant's marginal utility function for each commodity is a function of the quantity of this commodity alone. Since the demand curve is determined by the quantity of a consumer's wealth together with other variables (e.g., prices of related goods), Walras (1874, 1926 [1954,446]) indicted Dupuit for his "complete failure to distinguish between utility or want curves on the one hand, and demand curves on the other." At a later date, Walras (1874, 1926 [1954,486]) raised the same objection against the work of Auspitz and Lieben.4

Walras's intolerance masked the substantive contribution of the two Austrians. Although their graphical apparatus appears cumbersome by modern comparison, Auspitz and Lieben (1889) nevertheless clearly distinguished between the individual concept of consumer profit (Dupuit's money measure of relative utility) and the aggregate notion of consumer surplus. They

---

3. Walras referred to both of Dupuit's major articles in his Eléments, but there is no indication that he was acquainted with Bordas's comment or the salient parts of Dupuit's rejoinder.

4. Jaffé (1972,395–96) notes that in his Geneva lectures of 1871, Walras taught Dupuit's doctrine of consumers' surplus, but without any mention of Dupuit. In view of his frequent and fervid denials of the practice of identifying demand curves with utility curves, this discovery means that an earlier 'unregenerate' Walras was guilty of the same sin—even worse, since he identified the utility curve not with an individual's demand curve, but with a market demand curve.
also specified the explicit assumptions necessary to validate the analysis, namely: (i) constancy of the marginal utility of money; (ii) invariance of other prices; and (iii) unchanged tastes. Houghton (1958,57), for one, judged the Austrian contribution superior to Marshall’s, concluding: “Many of the difficulties and confusions which gave the concept a bad name during much of the twentieth century might perhaps have been avoided if the Austrian treatment had been given a share of the close attention that was lavished on Marshall’s work.”

Maffeo Pantaleoni (1889) was more generous than Walras in his praise of Dupuit, even though in the end he apparently accepted Walras’s fundamental criticism. The English translation of Pantaleoni’s *Manuale di economia pura* contains no fewer than six references to Dupuit’s preeminence in utility theory, including a detailed reference to the Dupuit-Bordas controversy. Nevertheless, Pantaleoni (1889 [1898,15511]) credited Walras with the vital distinction between utility curves and demand curves and took note of a further criticism by Pareto (1896) concerning the legitimacy of the constancy-of-the-marginal-utility-of-money assumption employed by Marshall. Moreover, from Walras’s correspondence (Jaffé 1965, 2:343–47) it appears that Pantaleoni accepted Walras’s critique of Dupuit. Shortly after the publication of his *Manuale*, Pantaleoni wrote to Walras seeking clarification of the difference between Walras and Dupuit on the measure of consumer surplus. Walras responded at length, basing his objection, as in the *Eléments*, on the illegitimacy of identifying utility with demand in a general equilibrium world. There is no record of further correspondence on this issue, and in 1889 Pantaleoni called attention to Walras’s ‘contribution’ without further comment.

Walras’s criticisms are important for the consumer-surplus doctrine, although some of them would not have been necessary had he given Dupuit’s works a more careful reading. Moreover, Walras’s view of the economic system was unparalleled in his time, whereas Dupuit’s frame of reference was more modest. Dupuit’s theoretic objective was simply to find a standard by which public projects could be evaluated, a problem not inherently suggestive of the interdependencies of the general equilibrium system. The tools that Dupuit developed were partial-equilibrium concepts, and should be evaluated as such. Walras, after all, had very few reservations concerning the measurability of utility, yet he made no progress whatsoever in his *Eléments* toward developing a ‘correct’ measure of consumer surplus.

These points are not offered as apologetics for Dupuit, because Walras’s criticisms were, in point of fact, fertile. But equally important for the history of economic theory is the fact that these criticisms were largely ignored by English economists before Alfred Marshall renewed the controversy late in the nineteenth century. Jevons failed to develop the doctrine even though he became aware of the Dupuit-Bordas controversy in 1879.
Marshall may, in fact, have become acquainted with Dupuit through Jevons, although we find the evidence on this point unconvincing.\(^5\)

**Marshall and His Critics**

*The early writings, 1867–1879*

Marshall gave 1867 as the year in which he began to study economics, but it is unclear when he first employed diagrams in his analysis or the extent of his debt to Dupuit, if any. By claiming Cournot, von Thünen, and Bentham as his mentors, Marshall (1890 [1961, 2:263]) implicitly denied Dupuit’s influence, although Marshall’s remarks on predecessors—particularly in the *Principles*—must be regarded with caution. His only reference to Dupuit’s priority was curiously dropped from the fourth (1898) and subsequent editions of the *Principles*. Whitaker (1975, 2:240) and Dooley (1983,27) have corroborated Cournot’s influence without comment on the deleted footnote. Marshall first read Cournot’s work in 1868. He exercised a graphical measure of consumer surplus several years later. The year is uncertain because of problems in dating, but probably in the early 1870s (Whitaker, 2:281–83) Marshall set down an example of consumer surplus that bears an uncanny resemblance to the earlier preoccupation and method of Dupuit. Marshall’s notebook entry, abbreviated here, is recorded in Whitaker.

[In Figure 4] when the toll equals \(PM\) let \(OM\) tolls be paid on a certain bridge. The amount levied will be greatest when \(OM\cdot MP\) is greatest (i.e., if a rectangular hyperbola with \(O_x\) and \(O_y\) as axes touch[es] the curve in \(P\), the amount levied will be greatest when the toll equals \(PM\)). Let the equation to the locus of \(P\) be \(y = f(x)\).

When \(OM\) carriages pass over the bridge let the damage done by each of them equal \(PP'\). Let the equation to the locus of \(P'\) be \(y = f'(x) - \phi(x)\), i.e. \(PP' = \phi(x)\). A toll, should now be levied such as to make \(OM\cdot mQ\) a maximum, i.e. \(Q\) should be chosen so that at \(Q\) the curve touches one of [the] above series of hyperbolas.

The number of people who would pay a toll \(BD\), but not a toll \(AC\)

---

5. Marshall wrote on the subject of consumer surplus as early as 1879, in his privately printed *Pure theory of domestic values* (Whitaker 1975, 2:212–36), but Whitaker (1975, 2:279–83) found evidence that Marshall had mastered the concept sometime earlier, probably between 1867 and 1872. Pantaleoni (1889 [1898,78n]) asserts that Marshall taught the theory of ‘residual utility’ [consumer surplus] at Cambridge as far back as 1869. Under these circumstances, it is hard to believe that Marshall first learned of the doctrine through Jevons. Moreover, the second of Marshall’s examples [on tolls] from his mathematical notebook (Whitaker 1975, 2:281–83) is so like Dupuit’s in both form and content that it is equally difficult to accept Marshall’s express denial of Dupuit’s influence (see below). Still, the general view is that Dupuit’s work was completely unknown in England until Jevons discovered it in the late 1870s.
Figure 4. Marshall's early measure of consumer surplus

is equal to $CD$ when $CD$ is very small; and the loss to those people in consequence of the tolls being greater than they will pay is $ACDB$; thus the whole loss which people who do not pay the toll $PM$ undergo is equal to $PMS$ [Dupuit's *utilité perdue*]. We may suppose that this loss causes to the state a loss equal in amount to $n$ times it, where $n$ is less than unity but dependent for its value on $OM$. Make $P'N'N''P'' = nPMS$. Then the net gain to the state resulting from a toll $PM$ is $OMP''N''$. The toll should be levied so as to make this a maximum.

The total advantage which people gain from the bridge after deducting the tolls which they pay is $TPN$ [Dupuit's relative utility], when the toll is $PM$. As before let the state gain from this an advantage $n$ times its amount. Then if $(abOE) + n(Tru)$ is greater than the interest on the bridge's cost (allowing for its being perishable) the bridge ought to be built.

Perhaps Marshall should be taken at his word—that Dupuit had no influence on his formulation of consumer surplus—but the similarity between the (circa) 1872 Marshall and the 1844 Dupuit is worthy of notice. Our opinion is that the question of filiation is not historiographically settled. What is undisputed, however, is that ultimately it was Dupuit's theory of consumer surplus that found its way into Marshall's *Principles*, albeit through the back door.

By 1879 Marshall felt confident enough of his measure of consumer surplus...
surplus to go into print, although the work in question, *The pure theory of domestic values*, circulated only privately. Here Marshall (Whitaker 1975, 2:213) defined consumer surplus precisely the way Dupuit had before him: as the ‘economic measure’ of “that which a person would be just willing to pay for any satisfaction rather than go without it.” The unveiling of the concept was tentative, yet simultaneously hopeful, Marshall observing: “It is somewhat difficult to discern clearly the nature of this surplus satisfaction and of its economic measure: but when this difficulty has been overcome, the apparatus of diagrams that is here applied will be found to be easily handled, and to be capable of achieving important new results.” No mention was made at this time of either Dupuit or Jenkin.

The example Marshall developed in *The pure theory of domestic values* was incorporated fully into the *Principles* eleven years later. Furthermore, Marshall had put the concept through its paces in the earlier work, investigating the effects of taxes and subsidies. In sum, the concept was highly advanced in Marshall’s mind more than a decade before the *Principles* appeared. Moreover, he began to surround the analysis with protective assumptions early on, showing alertness to the pitfall of interpersonal utility comparisons. Thus in 1879 Marshall (Whitaker 1975, 2:215) cautioned that the measure of human satisfaction captured by consumer surplus

... is indeed a rough measure. For in this as in many other portions of economic reasoning it is necessary, as a first approximation, to treat a pleasure that is worth a shilling to one man as equivalent to a pleasure that is worth a shilling to any other man. Assumptions of this nature have indeed to be made in almost every branch of statistical science. For all social and therefore all economic statistics deal with aggregates of human feelings and affections. It is not possible to add together arithmetically any two pleasures without some more or less arbitrary mode of measuring them. Now the economic measure of the satisfaction which a man derives from any source is ... the amount of money which he will just give in order to obtain it. The economic measures ... may be used in establishing economic laws. But such laws will contain only a portion of the whole truth of the matter to which they relate. And before deductions from these laws can be used for practical purposes, allowance must be made for the fact that a satisfaction which a rich man values at a shilling is slight in comparison with one for which a poor man will be willing to pay a shilling.

*The ‘Principles’ and its aftermath*

The enormous popularity of the *Principles*, attested to by its eight editions over a thirty-year span, gave the notion of consumer surplus much
more exposure than either Dupuit, Jenkin, or Auspitz and Lieben could provide. Its first statement and illustration in the *Principles*, as noted above, was transferred virtually intact from the privately circulated *Pure theory of domestic values*. Initial criticism, primarily from J. S. Nicholson (1893; 1894) and S. N. Patten (1893a; 1893b) induced Marshall to make minor emendations in the third edition (1895), duly noted by Edgeworth (1895,67) in his review in the *Economic Journal*.

While Patten’s critique has been more or less forgotten, Nicholson’s has periodically echoed through the corridors of time. Particular reverberations can be detected in the subsequent criticisms of Hobson (1900), Davenport (1935), Tharakhan (1939), and Knight (1944). Nicholson (1894,344) objected to the measurement of utility by money, observing: “Price is objective, utility is subjective. The price paid depends on one set of causes and the pleasure derived depends upon a different set.” He also questioned the legitimacy of assuming the marginal utility of money to be constant. Nicholson (1894,336) wrote: “A theory of expenditure which neglects the two primary facts that incomes are limited, and that the utility of the money retained increases as it becomes smaller is in my view an unreal theory. It is only applicable to a few careless millionaires.” Richard Lieben (1894) quickly refuted Nicholson’s charge of unrealism, in the process reaffirming the value of Marshall’s ceteris paribus assumption. Edgeworth (1894) and Barone (1894) provided additional defenses that Marshall endorsed, but Marshall also took care to make his assumptions more explicit in the third edition of the *Principles*, which recognized Nicholson’s criticisms.

A brief aside is in order here on Walras and Pareto. As we saw earlier, Walras (1874 [1926,486]) dismissed the idea of consumer surplus because “the definite integral of the demand function does not represent total utility” and therefore cannot measure consumer surplus. His criticism was blunted considerably by his refusal to acknowledge the legitimacy of a partial-equilibrium framework. Furthermore, Walras’ dismissal of the doctrine was undermined by Barone’s (1894) proof that a consumer surplus for one individual in isolation could be determined within a Walrasian system and that it could be reconciled with Marshall’s treatment (see Dooley 1983,33). Pareto, on the other hand, had demonstrated as early as 1892 that the marginal utility of money balances will only remain theoretically

6. Even before publication of the first edition of the *Principles* Marshall (1890 [1961, 2:260]) made it clear in a letter to J. N. Keynes that he regarded consumer surplus as a *sum of money*, not utility. He was very anxious that his doctrine not be confused with Jevons’ notion of total utility. Nevertheless, Marshall did follow Dupuit’s practice of identifying the demand curve with marginal utility, thereby inviting criticisms like Nicholson’s.

At the same point, Marshall showed that he was aware of the income distribution problem, declaring to Keynes: “I can see no connection between the loss of Consumer’s Rent and the loss of Total Utility resulting from a tax, unless it is known whether the commodity taxed is one consumed by the rich, by the poor, or by all classes alike.”
constant provided the composite elasticities of demand for all the other commodities concerned are equal to 1. This constituted a much more serious threat to Marshall's constancy-of-the-marginal-utility-of-money balances assumption, but Marshall took no note of it, despite his awareness of Pareto's work at the time he was preparing the third edition of the *Principles*. 

Simon Nelson Patten's (1893a; 1893b) critique must have been more provocative because it elicited a direct response from Marshall (1893) and was later cited by Pigou (1903,58) as inspiration for his own thoughts on consumer surplus. Patten's objection to Marshall's measure has a decidedly Austrian flavor. He readily accepted the subjective nature of utility and value, claiming, in fact, that Marshall did not go far enough in this regard: "He seeks to measure objectively and indirectly," said Patten, "what I seek to measure subjectively and directly." Specifically, Patten argued that Marshall's measure of consumer surplus overstates consumer welfare because it neglects the interdependence of utilities among commodity "groups," or classes of like goods. The problem Patten identified is analogous (in production space) to the 'imputation' problem that earlier occupied Menger and von Wieser, viz., if we 'remove' successive units of a single item from a commodity bundle and 'observe' the consequent loss of utility, our observations will be untrustworthy because each good in a commodity class depends for part of its utility on the other goods in the class. Adding the separate marginal utilities, therefore, produces an exaggerated sum of welfare. To quote Patten (1893a,422–23), in reference to Marshall:

Nowhere does he try to add together the consumer’s surplus of all the articles consumed by an individual to get the whole consumer’s surplus. . . . If he did he would see an error, for the parts will not add . . . [because] he estimates the surplus not from a given situation of the consumer, but from a series of situations representing different stages of supply.

For his part, Marshall (1893,619) complained that Patten misunderstood, or failed to appreciate, the significance of the *ceteris paribus* assumption, a contention in which he was later supported by Pigou (1903). Nevertheless, Patten scored some points in the skirmish, and Marshall (1890 [1961, 1:131–32]) acknowledged Patten’s criticism in the third edition of the *Principles*, admitting that "when the total utilities of two commodities which contribute to the same purpose are calculated on this plan,

---

7. A partial summary of Pareto's work was provided by Sanger (1895), whose review was cited by Marshall (1890 [1961, 1:132n]) in the third edition of the *Principles*. More recently, Abouchar (1982) has argued flatly that Marshall did not hold, and did not need, the assumption of constant marginal utility of money.
we cannot say that the total utility of the two together is equal to the sum
of the total utilities of each separately."

Like Patten, but on far less substantial ground, Henry Cunynghame
(1892;1905) thought that Marshall's measure overstated true consumer
surplus. Cunynghame asserted that every individual derives smaller incre-
ments of utility from each item consumed as the quantity purchased of that
article by others increases. What this seems to suggest is that Marshall's
'normal' demand is not the appropriate concept to use in measuring con-
sumer surplus. Marshall took no note of this in the Principles, but in a
letter to Edgeworth of 1892, Marshall (1890 [1961, 2:809]) wrote "It is a
free country. I deliberately decided that [Cunynghame's] temporary de-
mand curves (as contrasted with normal demand curves whose shape could
be shifted if need be) would not be of any practical use, and that this would
encumber the reader and divert his attention from more important things." Edgeworth took note of Cunynghame's argument in a later review, but
failed to endorse the notion of successive demand curves as they relate to
the consumer-surplus argument.

Pigou (1904;1910) endorsed and toyed with Marshall's notion of con-
sumer surplus in the ensuing years, supplementing in some respects the
analysis found in the Principles. From the start, however, Pigou (1903,66)
held the view that the measure is inadequate for a summation of total
happiness, but is suitable for more modest applications, e.g., demon-
strating how a monopolist could appropriate consumer surplus as profits through
price discrimination. Indeed, Marshall never asserted any more than this.
In his Economics of welfare, Pigou (1920) eschewed even the partial-
equilibrium notion of consumer surplus, his attention having shifted to
aggregate notions of economic welfare. Most other theoreticians of the
interwar period did not get past the problem of the marginal utility of
money, and Marshall left the concept essentially unchanged from the third
through eighth editions of his Principles.

The interwar years

Marshall was able to successfully defend his notion of consumer surplus
against most critics because he hedged his theory all around with protective
assumptions. His ceteris paribus mechanism included money income, the
tastes and preferences of purchasers, and the prices of all other goods.
Despite his awareness of the inherent difficulties of the concept, however,
Marshall shared Dupuit's beliefs in the measurability of utility and the
tendency of differences in income distribution to cancel out in the aggre-

8. In the fourth edition of the Principles Marshall (1890 [1961, 1:463n]) publicly re-
ferred to Cunynghame's argument as "ingenious," whereas he (1890 [1961, 2:812; 810])
privately wrote to Edgeworth that Cunynghame's work was of "undergraduate rather than
graduate" calibre, and that Cunynghame was "quick but impetuous; . . . all through his life
[he] has constantly supposed himself to know what he means when he does not."
gate. Initially Marshall (1890 [1961, 1:131]) had great hope for consumer surplus as a tool of practical import. But in his later years he confessed to his nephew Claude Guillebaud (1971, 6) that the concept was a major disappointment in his life because it was incapable of being quantified in a meaningful way. He reluctantly concluded that it was a theoretical rather than a practical tool in the economist's workbox.9

Despite its retention through eight editions of the *Principles*, theoretical interest in consumer surplus waned after the turn of the century. Perhaps this was because Marshall's two most able students, Pigou and Keynes, failed to take much interest in the idea. Minor skirmishes and/or attempts to improve the doctrine appeared during the interwar years, but without any real effect on Marshall's theory. A wartime attempt by P. G. Wright (1917) to analyze the principle of consumer surplus under different income distributions drew little or no attention. Shortly after the war, Edwin Cannan (1924) issued a broadside against the doctrine which drew prompt but uninspired rebuttal by D. H. Macgregor (1924) and by A. L. Bowley (1924). Cannan's view was probably indicative of a general feeling among economists that the doctrine had slipped beyond repair. Winch (1965, 401) has aptly pinpointed the reasons for the collective disenchantment:

Use of the Marshallian triangle when the $MUm$ [marginal utility of money] is not constant involves measurement in money, the marginal utility of which changes in the course of measurement. While there are pitfalls in using units of measurement, money, which do not have a constant relationship to the thing being measured, utility, there must also be objections to using any money of constant utility to measure changes in a case where the utility of money is not in fact constant.

In other words, the most serious problem with the money measure of consumer surplus goes back to the early recognition by Bordas that the presence of an 'income effect' (i.e., changing marginal utility of money), tends to misstate the losses and/or gains associated with price changes. Further progress required either a change in the demand curve to account for variability in the utility of money or alternatively a change in the definition of consumer surplus to fit the Marshallian demand curve. After a further hiatus, neoclassical economic theory made tentative advances in both directions.

**Hicks and His Critics**

*Renascence and rehabilitation*

A major theoretic development of the interwar years was the careful and systematic attempt by J. R. Hicks and R. G. D. Allen (1934) to establish

---

a more 'objective' theory of value, which they accomplished by reintroducing concepts originated many years earlier by Edgeworth and Pareto. Hicks and Allen framed their analysis independently of rigid cardinality, proving all of the familiar properties of demand curves by using indifference curves and marginal rates of substitution, instead. In the process, they translated Marshall's marginal utility of money into "exactly definable terms," to wit:

If the marginal utility of commodity \( Y \) is constant, the marginal rate of substitution between \( X \) and \( Y \) must depend on \( X \) only. If the quantity of \( X \) is given, the marginal rate of substitution (or the slope of the indifference curve) is given, too; the tangents to the indifference curves at all points with the same abscissa must be parallel . . . and the income elasticity of demand for \( X \) must be zero.

Hicks followed this early effort with a series of papers in the 1940s reaffirming the value of consumer surplus and amending Marshall's demand curve measure to accommodate it.\(^\text{10}\) The problem Hicks had to overcome is that Marshall's demand curve does not accurately measure consumer surplus in cases where a price change induces substantial income effects. Hicks (1941,109) faced two alternatives: either abandon the demand curve altogether in favor of indifference curves; or "adjust the ordinary demand curve so as to allow for the effects of the changes in real income." He chose the latter.

Hicks's rehabilitation of consumer surplus rests on the following assumptions: (i) that the good demanded is 'normal' with respect to changes in income; (ii) that the prices of other consumer goods remain constant during the course of measurement; and (iii) that the individual possessing a given amount of money income faces given market prices for \( n - 1 \) commodities to which he must confine his purchases. Given these assumptions,\(^\text{11}\) the individual will allocate his income in a particular manner. If a new commodity is introduced with only one unit available, the individual will decide whether to purchase this \( n \)th commodity depending on its price. Hicks (1943,31) maintained that under these conditions there will be some price which serves to separate the high prices, at which the consumer will not purchase, from the low prices at which he is just on the verge of purchasing. He called this price the marginal valuation of the unit, rec-

\(^{10}\) In his enthusiasm for the concept, Hicks (1941,108) proclaimed that the theory Marshall unveiled in the Principles "was immediately recognized as the most striking novelty in the book." The early reviews of the Principles, however, do not support this assertion.

\(^{11}\) Hicks also extended his analysis to the case of an inferior good, but the normal good case is sufficient to illustrate why the Dupuit-Marshall triangle, except in unusual circumstances, cannot be used as a valid measure of consumer surplus.
Ekelund and Hébert

Recognizing that it is the same thing as Marshall’s “marginal utility in terms of money.”

The unit will be purchased if the actual price is less than the marginal valuation. The marginal valuations of all units can be determined once the market price is given. In Figure 5, for example, AV represents a marginal valuation curve corresponding to market price OH. At price OH, quantity HP will be purchased, since all units of the good less than quantity HP have marginal valuations greater than OH. Point P is found by extending a horizontal from price OH to the marginal valuation curve. A new marginal valuational curve, Av, would correspond to a lower price, Oh. In the case of a normal good (as in Figure 5), the increase in real income occasioned by the price decrease will shift the new marginal valuation curve Av to the right and above the one corresponding to the higher price OH. Other things being equal, an increase in income will raise the marginal valuation of any given quantity of the good. This is the Hicksian “income effect,” which he identified with the movement from one curve to the other. By contrast, the substitution effect of a price fall consists of movements along the marginal valuation curves. Finally, the ordinary Marshallian demand curve can be determined by tracing the equilibrium points, e.g., the dotted line APpD. It is clear that when the income effect is of little significance the Marshallian curve approaches the marginal valuation curves. But when this is not the case, Hicks provides alternative measures for consumer surplus.

![Figure 5. The Hicksian reconstruction](image-url)
Compensating and equivalent variations

When the marginal utility of money is allowed to change, or identically, when there is an income effect, the gain to the consumer from a price fall can be viewed in several ways, some of which have already been discussed in connection with the Bordas example (see Figure 3). At the core of Hicks's macro-oriented compensation principle is the development of different 'variations' as measures of consumer surplus. Hicks inquired into the amount of money income which, taken from the consumer at the new price $O_h$, would leave him/her no better off than he/she was at the former price $O_H$. This amount is called a price compensating variation, and it is obtained (with reference to Figure 5) in the following fashion: allow the consumer to purchase $H_P$ units at price $O_H$ and, for the following unit, lower price only as far as necessary for him to purchase it. The curve $H_P C$ can be traced out by continuing in this manner. At $C$ on this curve, the consumer is neither better nor worse off than at point $P$. The segment $P_C$ lies above marginal valuation curve $P V$, since the consumer is better off than if he/she were forced to pay $O_H$ for $h_C$ units. But segment $P_C$ is below Marshallian segment $P_p$ because the consumer is in a worse position than if he/she were allowed to purchase all these units at $O_h$, even though the marginal unit can be purchased at that price. At $C$ the consumer is in the same position as if he/she had been allowed to purchase all the units at price $O_h$, but he/she has been forced to part with an amount of income equal to $H_P C h$, which is, simultaneously, the compensating variation and a measure of consumer surplus.

This Hicksian measure can be conveniently contrasted to Marshall's measure which, geometrically, is equal to the area $H_P p h$. Marshall's money measure assumed that the marginal utility of money was the same at positions $P$ and at $p$, a condition which could not possibly obtain with an income effect. The marginal utility of money does in fact vary along the Marshallian curve. A positive income effect would mean that the first cent added to the consumer's income would have a higher marginal utility than the last cent. In order to get the demand curve to express consumer surplus, Marshall had to assume that each cent in the money measure of consumer surplus added a constant amount to total utility. Hicks's compensating variation assumes, more properly, that with an income effect, each cent in the money measure added a diminishing increment to the total utility of the consumer. Hicks's compensating variation takes account of this diminishing marginal utility of money and is therefore less than the area under Marshall's demand curve.

Hicks's "rehabilitation" of consumer surplus made it clear that what is being measured is amounts of money (not utility), and that the marginal utility of money does not have to be constant for the idea to have theoretic
and practical value. Nevertheless, certain ambiguities in Hicks's measure were quickly identified. H. W. Robinson (1939) and A. Kozlick (1941) both argued that Hicks's measure of consumer surplus produces different results *ex ante* than it does *ex post*, a criticism later answered by Mishan (1947). A more durable criticism was made by Henderson (1941), who argued that Hicks's claim notwithstanding, the compensating variation is not the same as Marshall's consumer surplus. Henderson maintained that by Hicks's analysis, there existed four alternative expressions of consumer surplus, depending on the particular problem confronted. Hicks (1943) conceded Henderson's point and shortly thereafter expounded the notions of price (and quantity) *equivalent* variations as well as price (and quantity) *compensating* variations.

Hicks's "price equivalent variation" can be set forth in much the same terms as used to explain the compensating variation. Consider Figure 5 once again. Hicks asked the question, "What amount of money income would be required, in the absence of the price decrease, to raise the individual to the level of satisfaction attained at p?" His method requires asking the consumer, starting at p, to state the maximum price that would induce him to diminish his holdings of the commodity, *seriatim*. The price equivalent variation, area HEph, is yet another measure of consumer surplus. At point E the consumer is no worse off than at p, but he is consuming at price OH. The Hicksian equivalent variation is a larger money sum than Marshall's money measure under the demand curve because the value of money in terms of goods is different in the two situations P and p. The equivalent variation takes account of the increased level of satisfaction attained at p. In order to maintain this new level of satisfaction at price OH, the sum of money given to the consumer would have to be greater than the money amount under the Marshallian curve, since the marginal utility of money would have declined at p.

*Realism and relevance*

In the wake of the Hicks-Allen refinements in value theory and the Hicks-Henderson extensions of consumer surplus, Frank Knight (1944) issued a methodological broadside against the "realism and relevance" of the new theoretic developments. Of concern here are the particular arguments on consumer surplus, which constitute a small part of Knight's broader challenge. Knight (1944,311) derogated the practical significance of the Marshallian concept, declaring it (merely) "useful in bringing out the relations between the individual demand curve and indifference curves, with which it is much confused . . . [and] also useful for the pure theory of monopoly, in connection with perfectly classified monopoly price." On two occasions in the argument, Knight (1944,313n,318n) sided with J. S. Nicholson against Marshall, and in the end he denied any economic mean-
Knight further attacked the Hicks-Henderson analysis, offering in its place a more ‘correct’ measure of consumer surplus based on the intricate notion of a series of ‘indifference-combinations’ curves for quantities of money and good X.

Knight’s critique was countered by R. L. Bishop (1946), who attacked Knight’s analytics, denounced his version of consumer’s surplus as “incongruous,” and declared an unambiguous measure of consumer surplus to be a mere “will-o’-the-wisp.” Bishop cataloged seven measures of consumer surplus then extant in economic literature (including Knight’s), asserting (as Henderson had earlier for a smaller number of concepts) that the appropriate definition “in any one connection depends upon the purpose at hand.”

Hicks (1946,68n) ignored Knight’s criticism, but recognized Bishop’s refutation, apparently finding vindication therein for most of his earlier elaborations on the different measures of consumer surplus. Subsequently, Knight’s critique was reconsidered, and further discredited, by Mishan (1947) and by Pfouts (1953). Mishan (1947,33) narrowed the list of ‘acceptable’ measures of consumer surplus from four to two (the compensating and equivalent variations)12 and explained why no further reduction could be achieved:

The two different measures arise simply from the fact of the diminishing marginal utility of money. It is a distinction between what the consumer would pay (in order to get the lower price, or in order to avoid a higher price), and what the consumer must be paid (to induce him to forego the lower price, or to accept a higher price). For what he would pay or pays is to be considered a subtraction from his money income; what he must be paid or is actually paid an addition to his money income. . . . the difference between the two situations (the difference in utility) is unequivocal, but the sum of money required to express this difference is larger for an addition to an individual’s money income than for a subtraction from it.

The major American contribution of the period came not from Knight, whose concept of consumer surplus was roundly rejected, but from Harold Hotelling (1938), who drew freely from Dupuit’s theoretic wellspring. Hotelling developed a line integral representation of consumer surplus consistent with Dupuit’s definition and with the example from Marshall’s Principles. His consumer surplus is a collective notion, tied to a market demand curve; but while it is an index of total utility, it does not imply that utility is measurable. Since the prospects of obtaining specific utility indicators

---

12. Patinkin (1963) later maintained that Mishan’s analysis holds only for perfectly competitive equilibrium situations.
empirically are not promising, this constitutes a distinct advantage over the Dupuit-Marshall measure. Most welfare theorists lined up behind Hicks rather than Hotelling, however, and it was the Hicksian synthesis of the welfare economics of Marshall, Pigou, and Pareto that carried the day. By 1950 the theory of consumer surplus had reached its third major plateau—the first peak having been scaled by Dupuit, the second by Marshall, and the third by Hicks.

Epilogue: The Modern Era

Mapping the theoretical terrain

Hicks’s ‘improvement’ of the Marshallian measure was genuine in the sense that it rendered consumer surplus theoretically correct, but it was also the catalyst for a proliferation of consumer surplus measures and a new debate over the appropriateness or ‘exactness’ of one measure versus another. As a consequence, theoretical welfare economics in the modern period is in considerable disarray. On the one hand, the concept of consumer surplus is roundly condemned by a small group of economists, most notably Samuelson (1942; 1947), Little (1950), and Graaff (1957). On the other hand, it is favored by many microeconomists, especially those who accept the validity of partial-equilibrium analysis. Its defenders, however, are nowhere near a consensus on what constitutes the ‘correct’ welfare measure. The issue, as it has evolved historically, is intricate and complex.

One point of contention concerns what Marshall ‘really meant’ by consumer surplus. We have seen that Marshall introduced the subject by defining consumer surplus in the same fashion as Dupuit, that is, as an “all or nothing” proposition. Unfortunately, Marshall failed to distinguish this definition from two other concepts discussed in his Principles: (a) the area under a commodity demand curve minus expenditures on that commodity, and (b) the area under the utility curve for a good less marginal utility times the quantity consumed. After Hicks, the question arose as to whether Marshall had in mind a ‘compensated’ or an ‘uncompensated’ demand curve. Friedman (1949) defends the former interpretation, arguing that Marshall constructed his demand schedule on the assumption that the consumer’s level of satisfaction was being held constant.

Whether or not Friedman’s interpretation is correct, it is clear that the compensating variation has been pushed to the forefront of discussions on applied welfare economics. Nevertheless, this does not imply that all economists have jumped on the Hicksian bandwagon. In particular, Winch (1965) has rejected Hicks’s rehabilitation in favor of a return to Marshall’s surplus. The latter’s major appeal, according to Winch, is that it is the only measure that can be taken directly from the demand curve. Moreover, its limitations, argues Winch, are no worse than the limitations of Hicks’s
measures. Furthermore, Marshall's measure has the specific advantage of being additive, whereas Hicks's measures are not.

Be that as it may, Foster and Neuburger (1974) caution that considerable care must be exercised to ensure the additivity of Marshall's measure once the analysis extends beyond the single-good partial-equilibrium case. It has been shown that in the case of simultaneous, multiple price changes, the Marshallian surplus is no longer uniquely defined, and that alternative evaluations of a given welfare change depend on the assumed order of price adjustments between the terminal situations being compared. This path-dependence problem was first recognized by Hotelling (1938) and subsequently considered by Mohring (1971), Harberger (1971), Silberberg (1972), Glaister (1974), and Turvey (1974) and in synthetic fashion by Burns (1977). It has come to be a major obstacle to the further development of the consumer surplus concept.

Michael Burns (1973) would circumvent the path-dependency problem by assuming a priori that a specific simultaneous price adjustment process exists. Like Winch, he favors retention of the Marshallian measure. Mohring (1971) and Silberberg (1972) take the opposite position that Marshallian measures should be abandoned in favor of the path-independent compensating or equivalent variations. More recently, Neil Bruce (1977) has denied the operationality of such alternatives and has suggested the economic theory of index numbers as a way of making the path-dependency problem tractable. Bruce Dahlby (1977) has attempted a reconciliation of sorts by setting out the conditions under which the Marshallian measure is path-independent and those conditions under which it is not. His work falls between the efforts of Willig (1976) and Seade (1978) to identify those circumstances in which the Marshallian measure closely approximates Hicks's compensating and equivalent variations. But Hausman (1981) has pointed out certain shortcomings in Willig's approach. At bottom is the fact that Marshall's measure is based on information about uncompensated (market) demand curves whereas Hicks's measures require information on compensated demand functions. In principle it is possible to estimate the latter, but in reality the only data usually available relate to the observable market demand functions.

While the path-dependency problem has been a major recent obstacle to the further advance of the consumer surplus concept, it is only part of a broader concern for the restrictions on preferences that must be specified for some measure of consumer surplus to serve as an exact welfare indicator. This larger concern involves, in addition to the path-dependency problem, the literature on the constancy of the marginal utility of income, the cross-elasticities of demand among products, and the homotheticity of preferences. As such, it takes us beyond the narrow confines of this study.
Obviously, the battle lines are still being drawn on many aspects of consumer surplus analysis. Besides the choice between Marshall and Hicks on the 'appropriate' surplus measure, there is further indecision among the 'Hicksians' as to which of the variations provides an exact welfare indicator. Recently, Chipman and Moore (1980) have analyzed the conditions under which the compensating variation can be validly used as a generalized welfare measure, but McKenzie and Pearce (1982) insist that the equivalent variation (not the compensating variation) is the only directly observable, exact welfare indicator.

**Conclusion**

Despite its inherent ambiguities and difficulties, the Dupuit-Marshall concept of consumer surplus continues to offer some usefulness as a guide to practical policy issues, especially in instances where the Hicksian compensation principle is inapplicable, or the data problems insurmountable. As such it represents the economist's response to the practical imperative of approximating a measure of 'maximum satisfaction' in circumstances where a truly scientific measure is impossible.

When the concept of consumer surplus is generalized to an aggregate welfare measure, it should be noted that its usefulness and its inherent problems are no greater and no less than those involved in measuring changes in national income. Indeed, Harberger (1971), McCloskey (1982, 225–29) and others have demonstrated that changes in national income are the same as changes in aggregate consumer surplus. Estimating the bounds of a change in either requires the application of price indices. Not surprisingly, therefore, since Hicks (1942), the theory of index numbers and the theory of consumer surplus have logically merged.

Harberger and McCloskey argue that because the measure of consumer surplus is an index-number problem, it therefore does not require the usual assumption concerning the constancy of the marginal utility of income. What they do not seem to recognize is that the usual assumption regarding constancy of the marginal utility of income itself implies that some index has been chosen, at least for the individual consumer. The assumption implies that if income effects are significant, the Marshallian demand curve must be adjusted in order to validly register the individual's consumer surplus. In other words, it is precisely because income effects attend price changes that index numbers are customarily invoked. Yet many different measures may obtain, depending on the index chosen to measure the attendant change. Insofar as the constancy of the marginal utility of income assumption implies that a particular index has been chosen, it would appear that Harberger-McCloskey offer a distinction without a difference. Nevertheless, the inherent dependence of consumer's and consumers' sur-
plus measures on index numbers is an indisputable fact, verified by the historical record.

In the final analysis, the history of the concept of consumer surplus reveals a theoretical proposition beset on all sides by challenges to its usefulness and desirability. It has been praised and reviled, expanded and contracted, tinkered with, rehabilitated, ‘improved,’ and above all used over and over again. At the same time, its indispensability has been constantly questioned or denied. As Arvidsson (1974,286) has aptly noted, referring only to its Marshallian (not its French) roots:

It is as though consumers’ surplus not only has to carry the burden of a Victorian past but also is subjected to lingering Victorian double standards in quite the same way as the classical object of those double standards; in other words, the concept is generally used but held in low esteem by respectable people.

Much of the rejection of the concept itself stems from a desire to make economics something it is not and probably never will be: a ‘pure’ science akin to physics. Specifically, the attempt to ‘objectivize’ economics into a science via the theory of revealed preference has been largely unsuccessful. In its defense, Bergson (1975,43) has argued that the revealed-preference approach may be more accurate in cases where the subject of analysis is the individual household, or a collection of households with identical tastes. But when tastes differ, income redistribution occurs, or aggregate measures of consumer surplus are required, Samuelson’s revealed preference measure is, for all practical purposes, computationally impossible. Thus, Bergson’s (1980) recent attempt to integrate consumer surplus with the social welfare function in a general equilibrium context ignores the revealed preference approach.

The very durability of consumer surplus, not to mention its recent extension into other microeconomic areas such as option demand (Lindsay 1969; Byerlee 1971; Cichetti and Freeman 1971; Schmalensee 1972, 1975; Bohm 1975) and ‘full-price’ demand (Lyon 1978), belies Samuelson’s (1947,195) judgment that “the concept is of [mere] historical and doctrinal interest, with a limited appeal as a purely mathematical puzzle.” On the contrary, the idea invented by Dupuit over a century ago is of continuing importance and concern to economists. A doctrine possessed of such a long and interesting history, not to mention its continuous use, will not easily retire to the historical scrapheap of ‘superfluous’ theories, notwithstanding Samuelson’s (1947,195ff) judgment that it belongs there. Nor will it likely be cast aside as a “totally useless theoretical toy” (Little 1950,175). Despite a difficult birth, a troublesome adolescence, and an uncertain adulthood, the Dupuit-Marshall theory of consumer surplus has
survived and prospered through periodic trials of criticism and doubt. Any idea which brings the premier economic aim of 'maximum satisfaction' into full focus, especially within the context of general demand theory, most assuredly has a future, however turbulent, in the annals of economic theory and practice.

We would like to acknowledge the helpful suggestions of William Stober, Denis O'Brien, Randall Holcombe, and two anonymous referees of this journal. The authors alone are responsible for any errors that remain.

**Bibliography, in Chronological Order**


1953. Pfouts, R. W. "A Critique of Some Recent Contributions to the Theory of


1978. Foster, C. D., and Neuburger, H. L. I. "The Ugly Truth About Consum-
1978. Lyon, K. S. “Consumer’s Surplus When Consumers Are Subject to a Time
1978. Seade, James. “Consumer’s Surplus and Linearity of Engel Curves.” Eco-
1979. Cigno, A. “Search and Consumer’s Surplus: A Generalization.” Bulletin of
1980. Bergson, Abram. “Consumer’s Surplus and Income Redistribution.” Jour-
nal of Public Economics, 14:31–47.
1981. Hausman, J. A. “Exact Consumer’s Surplus and Deadweight Loss.” Amer-
Money.” Eastern Economic Journal, 8:79–82.
Geography, 58:286–287.
1983. Dooley, P. C. “Consumer’s Surplus: Marshall and His Critics.” Canadian