Journal of the History of Economic Thought, Volume 22, Number 1, 2000

MATHEMATICAL FITNESS IN THE EVOLUTION OF THE UTILITY CONCEPT FROM BENTHAM TO JEVONS TO MARSHALL

BY

TOM WARKE

I. INTRODUCTION

This paper seeks an answer to the following question: By what process did the utility concept in economics evolve from its Benthamite to its modern form? Jeremy Bentham applied his principle of utility to ethical, legal, and constitutional issues, and only after more than a century of adaptation did it become the dominant explanation for consumer choice. The paper identifies fitness for mathematical analysis as the underlying principle of selection for this evolutionary success, exemplified in particular by the mutations of utility between Bentham and W. S. Jevons, and between Jevons and Alfred Marshall.

Exactly fifty years ago, George J. Stigler rounded off his history of the utility concept from 1776 to 1915 with a "theory of economic theories," a recapitulation of his findings "with a view to isolating the characteristics of successful (and hence of unsuccessful) theories, where success is measured in terms of acceptance by leading economists" (Stigler 1950, p. 392). Stigler identified three criteria for acceptance which, he was careful to point out, did not necessarily imply genuine conceptual progress, judged retrospectively. The first criterion was *generality*, occasionally attained by weakening the assumptions to reach a given conclusion but more commonly shown by a successful theory encompassing a wider range of phenomena than its competitors. *Manageability* was seen by Stigler as a necessary condition for acceptance, but one that frequently also acted as a brake on progress: "Manageability should mean the ability to bring the theory to bear on specific economic problems, not ease of manipulation. The economist has no right to expect of the universe he explores that its laws are discoverable by the indolent and the unlearned" (ibid., p. 394). And finally, *congruence with reality* had amounted in practice to "casual observation and general knowledge" but "should have been sharpened—sharpened into the

Department of Economics, University of Sheffield, United Kingdom. I wish to acknowledge advice and encouragement from William Baumol and Mark Blaug, as well as very helpful comments from Philip Schofield of the Bentham Project.

ISSN 1042-7716 Print; ISSN 1469-9656 Online/00/010005-23

insistence that theories be examined for their implications for observable behavior, and these specific implications compared with observable behavior" (ibid., p. 395). Stigler cited several cases where economists had shown "at least a lack of enterprise … [and] also a lack of imagination" (ibid., p. 394) in avoiding the joint challenges of falsification tests and mathematical complexity. These examples included the lengthy adherence to an additive utility function despite its manifestly false implication that there are no inferior consumer goods; the reluctance to abandon cardinality despite the demonstration that marginal utility equations are not uniquely integrable; and the refusal "to include in the individual's utility function the consumption of other individuals, although this extension was clearly unimportant only in the social life of Oxford" (ibid., p. 393).[1]

Whatever the stumbles and hesitations, developments in utility theory from Smith to Slutsky clearly correlate with the spread of a presumption within economics that the laws of human behavior, like the Laws of Nature, are written in the language of mathematics. It is from this perspective that we may jointly describe Stigler's criteria of generality and manageability as "fitness for mathematical analysis." The level of generality achieved by Slutsky permitted utility to be shorn of its psychological trappings and to be treated simply as an ordinal index of desirability. As Stigler put it: "With Slutsky's development, introspection no longer plays a significant role in utility theory. There is postulated a function which the consumer seeks to maximize, and the function is given the characteristics necessary to permit a maximum" (ibid., pp. 382–83). So far as manageability is concerned, the mathematical sophistication of economists has evolved in tandem with that of their theories, so that modern economists—at least "leading" ones—can no longer be accused of avoiding mathematical complexity *per se*. There is, however, a distinction between the complexity of mathematically sophisticated representations and the complexity of what, for want of a better term, we might call mathematically messy representations. It is the latter which lack fitness for mathematical analysis, thereby severely prejudicing their survival in economic theory irrespective of their congruence with reality. One type of messiness occurs when a representation imposes large information requirements, beyond the standard market constraints, in order to determine individual choice. It is this, I suggest, that explains why the lone survivor among Stigler's illustrations of pre-Slutsky laxity is the reluctance to include in the individual's utility function the consumption of other individuals, even though utility function independence implies, among other things, that business decision-makers are irrationally wasting a substantial share of their advertising expenditures.

Although interdependence among individuals' utilities played an important role in Bentham's ethical system, it is a second type of mathematical messiness,

[1] Stigler's reference is to statistics gathered by Edgeworth from "a certain Oxford College" in response to Pigou's specification of an individual's utility function that was dependent upon consumption by others (Pigou 1903). As reported by Stigler, Edgeworth concluded that the "size of the party" had no significant effect upon "the depth of the potations" (per capita consumption of wine) (Stigler 1950, p. 324).

also present in Bentham and also since excised axiomatically from standard economic theory, which is the main subject of this paper. Unlike preference set independence, the congruence with reality in evading this second type of messiness was not questioned by Stigler, despite its more fundamental role in modern decision theory. Whereas an axiom of preference set independence serves to simplify the constraint specification required to determine individual choice, axioms of completeness and transitivity are necessary in order for a rank ordering to exist in the first place.[2] The corresponding requirement on utility (whatever we believe it to measure) is that it have but a single dimension, for the "index number problem" tells us that no unique and continuous ordering exists for a multi-dimensional entity except in the trivial cases where all dimensions save one remain constant or where all dimensions vary in the same proportion. In Section II of this paper, I show that Bentham's original utility construct was irreducibly multi-dimensional, with a distinct dimension for each type of pleasure and for each type of pain. It had, therefore, to be stripped of its multi-dimensionality before it could serve as a suitable maximand for determinate choice. Section III shows that the method by which Jevons reduced utility to a single dimension depended not only upon a dichotomy between psychological hedonism and ethical hedonism that Bentham had explicitly denied, but also upon a misrepresentation of Bentham's method of measuring utility. In Section IV, I argue that it was a dubious conflation of "indecision" with "indifference," under the spellbinding rhetoric of Alfred Marshall, which persuaded economists that mathematical analysis could be applied to choice theory without regard either to Jevons's struggles with utility units or to his distinction between duty to self and duty to others. My conclusions are presented in Section V. The primary purpose of the paper is not to debate the degree to which determinate choice is congruent with reality, but rather to describe the steps by which determinate choice was incorporated so seamlessly into economic theory that the process could remain unremarked by so astute an historian as George Stigler.

II. JEREMY BENTHAM'S UTILITARIANISM

The one thing on which Bentham scholars agree is that he has been badly misunderstood.[3] In particular, the depth of thought and subtle sophistication with which Bentham handled the utility concept are vastly underrated, and not only by economists.[4] Many social historians and philosophers, drawing upon G. E.

[2] In a current graduate-level textbook on competitive equilibrium, Brian Ellickson characterizes "a world in which preferences need be neither transitive nor complete" as "terrain alien to most of economics" since "without transitivity or completeness, utility functions do not exist and indifference contours lose their meaning" (Ellickson 1993, pp. 314–15).

[3] The main reason for misunderstanding Bentham is that he wrote much but published little during his lifetime, and was rather ill-served by his literary executor, John Bowring. The ongoing effort of the Bentham Project at University College London, at work since 1959 on Bentham's legacy of tens of thousands of uncollated pages of barely legible manuscript scattered haphazardly in nearly two hundred boxes, is slowly but surely making Bentham's writings available to the general public.

[4] The typical economist's view of Bentham is reinforced by political theorist John Plamenatz in his Preface to a re-issue of Elie Halévy's *The Growth of Philosophic Radicalism*: "Social 'scientists' today

https://doi.org/10.1080/104277100112527 Published online by Cambridge University Press

Moore's critique of John Stuart Mill's version of utilitarianism (Moore 1903), are dismissive of Bentham for having overlooked conflict between "egoistic psychological hedonism" and "universal ethical hedonism" when in fact Bentham's very purpose in adopting a morally neutral utility concept was to *resolve* the conflict. When the twenty-year-old Bentham discovered utility in the writings of Hume, Helvétius, and Beccaria, he embraced it as the alternative he had been seeking to the "fictions" of natural rights or divine revelation on the one hand and social contract theory on the other, for he believed that all such moral criteria were merely artifices for propagating their proponents' own "sympathies" and "antipathies":

> The various systems that have been formed concerning the standard of right and wrong, may all be reduced to the principle of sympathy and antipathy. One account may serve for all of them. They consist all of them in so many contrivances for avoiding the obligation of appealing to any external standard, and for prevailing upon the reader to accept of the author's sentiment or opinion as a reason and that a sufficient one for itself (Bentham 1789/1996, pp. 25–26).

Throughout his long life, Bentham remained adamant that utility, or happiness, for he used the terms interchangeably, is simultaneously the single motive for individual actions ("what is") *and* the single valid criterion for judging the aggregate welfare consequences of those actions ("what ought to be").

Bentham usually referred to his aggregate welfare criterion as "the greatest happiness principle," for by judging an action to be "right" if it generates positive net utility and judging one of two alternative actions to be "better" if it generates the greater aggregate utility, it follows that aggregate utility would be at a maximum in the best of all possible worlds. The central focus of Bentham's work was initially judicial, and ultimately constitutional, reform. Insofar as the consequences of individual actions are confined to their perpetrators, he advocated freedom of choice as the optimal route to a welfare maximum. When the greatest happiness principle identifies types of actions in which pleasures for perpetrators are outweighed by pains imposed on outsiders, however, the ideal utilitarian state would impose political sanctions which alter the pleasure/pain balance for individuals in order to modify the offending behavior.[5] Bentham's utilitarianism in practice therefore requires that the aggregate utility associated

---

no longer speak, as Bentham and his disciples did, of 'maximising happiness,' but they still speak of 'maximising' … Their methods and arguments are similar in type to those of Bentham, though what they seek to measure is not happiness as he conceived of it. They have more sophisticated ideas than he had about measurement, and are much readier than he was to admit that there are limits to what can be measured" (Plamenatz 1972, p. xv). It is the contention of this paper that the "utility" which today's economists, more than ever, "speak of maximising" differs from Bentham's "happiness" primarily because axioms of completeness and transitivity have been imposed on preference sets, and that in the sense of *determinate* measurement, the converse of the final quoted sentence from Plamenatz is nearer the truth.

[5] It is yet another misrepresentation of Bentham to associate him too closely with classical *laissez-faire*. Although he urged the State not to interfere with mutually agreed economic transactions (most famously in *Defence of Usury*, though this was later modified in *Defence of a Maximum*), his utilitarianism led him to advocate a considerable degree of income redistribution, a substantive

with particular acts be measured, as well as alterations in aggregate utility brought about by any proposed sanctions. Nevertheless, and contrary to common opinion, Bentham did not believe that either aggregate or individual utility could be measured with any precision, and in his own applications he invariably confined himself to *qualitative* judgements.[6] The reason for this reticence is found in Bentham's conception of utility.

The demonstration that Bentham perceived of utility as irreducibly multi-dimensional in pleasures and pains is straightforward. Bentham says so explicitly throughout his writings. A typical example is found in a preparatory manuscript for his last major work, the *Codification Proposal*: "The elements of happiness are *pleasures* and *exemptions from pains*: individual pleasures, and exemptions from individual pains" (Bentham 1822/1954, p. 440). Furthermore, he devoted a great deal of effort to classifying types of pleasures and pains into mutually exclusive categories. A good illustration of this taxonomy occurs in Chapter V of his best-known work, *An Introduction to the Principles of Morals and Legislation* (hereafter *IPML*). His purpose, he tells us, is to reduce the "interesting perceptions" of "exciting causes" down to the "simple" pleasures and pains of which they are composed, and "which cannot any one of them be resolved into more" (Bentham 1789/1996, p. 42). For example, "The pleasures of amity, or self-recommendation, are the pleasures that accompany the persuasion of a man's being in the acquisition or the possession of the good-will of such or such assignable person or persons in particular" (ibid., p. 43).

Bentham thus started from the premise that aggregate utility for a community of n agents experiencing m different types of pleasure and pain is an nm-dimensional object. Using modern terminology, an index number *measure* of this entity, for any particular act having specified pleasure/pain consequences, requires n distinct sets of *intra*personal utility weights (one for each agent's personal allocation over the m pleasures and pains) plus one set of *inter*personal utility weights over the n agents. With regard to interpersonal weights, Bentham set the stage for modern utilitarians by imposing egalitarianism as an ethical rule:[7] "Everybody to count for one, nobody for more than

subsistence safety net and, in general, State action whenever it could be justified by externalities and spillovers from private actions. Frank Petrella (1977) has argued that Bentham was by far the strongest single influence in a transition of European social theory from eighteenth century *Ordnungspolitik* (concern with optimal institutional structures) to nineteenth century *Prozesspolitik* (concern with optimal intervention).

[6] As Lord Robbins put it: "There is much talk in the Benthamite literature of a felicific calculus; and the term naturally suggests a most pretentious apparatus of measurement and computation. But, in fact, this is all shop window … . [Benthamite] use of the felicific calculus lay in quite another direction—in rough judgements of the expediency of particular items of the penal law, in general estimates of the suitability of existing institutions or the desirability of other institutions to take their place. It was not necessary for all this that they should have used such a pretentious label" (Robbins 1978, p. 181).

[7] R. M. Hare, for example, uses egalitarian weights to *define* moral reasoning, thereby, in his view, attaining a more secure logical basis for impartiality than does, for example, the non-utilitarian John Rawls: "I try to base myself, unlike Rawls, entirely on the formal properties of the moral concepts as revealed by the logical study of moral language; and in particular on the features of prescriptivity and universalisability which I think moral judgements … all have … . In this position, I am

https://doi.org/10.1080/104277100112527 Published online by Cambridge University Press

one."[8] This impartiality then implies that the appropriate weights on pleasures and pains, to be used in making utilitarian ethical judgements, judicial decisions, and legislative enactments, are their respective means across all individuals whose interests are at stake. The problem for utilitarian practice is, of course, that these mean values remain unknown to utilitarian ethicists, judges, and legislators. Rigorously impartial utilitarians can legitimately disagree as to what the representative utility weights on pleasures and pains might be, thus coming to a different rank ordering among alternative acts, and hence to differing judgements about right and wrong in any given situation.

This ambiguity, as manifested in a utilitarian legislature, can be illustrated by considering another of Bentham's sub-categories of pleasure: "The pleasures of malevolence are the pleasures resulting from the view of any pain supposed to be suffered by the beings who may have become the objects of malevolence: to wit, 1. Human beings. 2. Other animals" (ibid., p. 44). Strict avoidance of what Bentham called "*ipse dixitism*" ("he himself said it") precludes legislators acting on personal judgements that malevolent pleasures are *inherently* less worthy than benevolent ones. Instead, State intervention to suppress malevolence must be justified by the greatest happiness principle: if and only if the pleasure value gained by the perpetrator of a malevolent act is less than the pain value suffered by victims will total happiness be increased by a political sanction which deters the act.

The initial legislative task is to decide which acts are to be made illegal. Legislators may be confident that pleasures for political terrorists are outweighed by pains of the target population, but how about pleasures for players of practical jokes compared to pains of their hapless foils? And note that Bentham's definition of malevolence also requires inter*species* weights (presumably *not* "everybody to count for one"?), for utilitarian legislators must compare the pleasures of persons, displacing anger at their bosses or spouses, with the pains of the cats that they kick. Any given set of weights will produce a complete and transitive ordering of malevolent acts, along with a designated boundary between legality and illegality, yet anything short of legislative unanimity demonstrates ambiguity about the weights that were used.

Once illegal acts are defined, the next step is to determine the penalties to be incurred. If there were no variation in intrapersonal utility weights among persons, legislators could totally eliminate illegal malevolent acts by introspection (or experimentally), simply raising penalties until their pains, as anticipated

---

prescribing universally for all situations just like the one I am considering, and thus for all such situations, whatever role, among those in the situations, I might myself occupy. I shall therefore give equal weight to the equal interest of the occupants of all the roles in the situation; and, since any of these occupants might be myself, this weight will be positive. Thus the impartiality which is the purpose of Rawls's 'veil of ignorance' is achieved by purely formal means; and so is the purpose of his insistence that his contractors be rational, i.e. prudent" (Hare 1982, p. 25).

[8] This particular expression for egalitarian interpersonal weights was attributed to Bentham by John Stuart Mill. After many years of searching, the original source has been located by Philip Schofield in the 1827 edition of Bentham's *Rationale of Judicial Evidence*, which was edited by Mill when he was a teenager.

by potential offenders, exceed anticipated malevolent pleasures. But so long as utility weights vary within the general population, "normal" deterrents fail for persons especially susceptible to malevolent pleasure or especially insusceptible to penalty pain. Therefore, in order for utilitarian legislators to determine the optimal harshness, further weighting judgements are required to balance the trade-off, as harshness rises, between less victim pain from fewer undeterred malevolent acts and more punishment pain for residual offenders, while taking account of the malevolent pleasures of the residual victims in seeing their tormentors punished. Once again, any given set of weights will determine the penalties in a utilitarian criminal code, but any legislative dissent or indecision is evidence of ambiguity about the weights that were used.

There can be no doubt that Bentham himself was fully aware that the greatest happiness principle is a very blunt instrument. In an early manuscript, he analyses a two-dimensional criminal act in which "x" denotes the pain of privation (losing, say, a gold watch) and "y" denotes the pain of apprehension (being threatened, say with a gun):

> Call the species of misery produced by any one action in a single person, x, and that produced by another, y. Now whether x or y be the greater, is a matter of conjecture and opinion, but that x + y is greater than either x or y alone, is a matter of demonstration … . In this manner it is a matter of demonstration that Robbery is worse than Theft … . Figure or the local position of parts with respect to each other is out of the question. It is only the very first principles of mathematics that have anything here to do (Bentham *ca.* 1773/1962, p. 117).

Thus, according to Bentham, it is possible to rank-order robbery and theft for any *given* loss of property, but the rank order of, say, a $100 petty theft and a $20 armed robbery would be problematic. In the language of economics, the penultimate sentence of the quotation rules out an indifference map in this two-pain space.

For modern public choice economists, an acknowledged intransigency in identifying and aggregating preferences has nothing to do with ambiguity on the part of the individuals who possess them. Yet if utility consists of irreducibly multi-dimensional pleasures and pains, as Bentham assumed, the index number problem applies to private as well as social optimal choice. An individual faced with alternative mixes of pleasures and pains has no natural common denominator—no *generic* pleasure—by which to compare and rank order them unambiguously. Moreover, this was not the only measurement problem so far as Bentham was concerned, for he found imprecision at even the most microscopic level—the value of a single pleasure or pain to a single person. "To a person, considered *by himself*," he tells us in *IPML*, "the value of a pleasure or pain, considered *by itself*, will be greater or less, according to the four following circumstances: (1) Its *intensity*. (2) Its *duration*. (3) Its *certainty* or *uncertainty*. (4) Its *propinquity* or *remoteness*" (Bentham 1789/1996, p. 38). In an earlier more detailed analysis, Bentham had discussed measurement scales for each of these circumstances. He treated the dose of a particular pleasure as its duration discounted for its degree of uncertainty and remoteness, and represented as a unit degree of intensity "the faintest of any that can be distinguished" (Bentham 1782/1901, p. 398). In the

https://doi.org/10.1080/104277100112527 Published online by Cambridge University Press

end, however, he acknowledged that any such intensity unit must remain imprecise:

> The elements … of value (it has been seen) are four: *intensity*, *duration*, *propinquity*, *certainty*; add, if in a political community, *extent*. Of these five, the first, it is true, is not susceptible of precise expression: it not being susceptible of measurement. But the four others *are* (Bentham 1822/1954, p. 443).

W. S. Jevons, an ardent admirer of Bentham, accepted as definitive these same elements of value, yet managed to measure in a determinate single dimension the utility flow from an act of consumption. My next task is to examine how he got there from here.

III. W. S. JEVONS: THE HOMOGENIZATION OF UTILITY

The half century separating the first edition of Jevons's *Theory of Political Economy* and the final edition of Marshall's *Principles of Economics* saw the instigation, triumph, and consolidation of the marginal revolution. Marginal analysis in economics means constrained optimization, and unambiguous optimization requires a one-dimensional maximand. The circumstances of the threefold discovery of marginalist exchange equilibrium during the 1860s and 1870s have been well explored. Perhaps the point of greatest relevance for this paper is that Jevons, like Léon Walras, became convinced very early in his study of economics that its progress as a science required a complete re-working, involving in particular the application of mathematical analysis to the determination of value, and that he would be the one to provide it.[9]

The distinction that makes Jevons part of my story, and excludes Walras, is that Jevons attributed his basic concepts directly to Bentham.[10] He tells us in the

[9] Given an early private viewing of Jevons's diary for 1860, since published (Black & Könekamp 1973), J. A. La Nauze was able to pinpoint the weekend (3–5 February) that Jevons began his mathematical reconstruction of economics, aged twenty-four and in the first year of his resumed B.A. studies at University College, London, after five years working in Australia. The relevant entry includes the statement that "value is to be established *on the basis of labour* and the problems of rent wages interest etc., to be solved as mathematical functions." (La Nauze 1953, p. 357, emphasis added) Jevons's *Eureka!* day followed on February 19th, for that is when he abandoned the classical labor-cost concept: "At home all day and working chiefly at Economy, arriving as I suppose at a true comprehension of *Value* regarding which I have lately very much blundered" (ibid). Jevons's progress by the 1st of June 1860, is revealed in a letter to his brother Herbert (then a homesteader in Minnesota): "One of the most important axioms is that as the quantity of any commodity, for instance plain food, which a man has to consume increases, so the utility or benefit derived from the last portion used decreases in degree … . And I assume that on an average the *ratio of utility* is some continuous mathematical function of the quantity of commodity. This law of utility has in fact always been assumed by Pol. Econ. under the more complex form and name of the Law of Supply & Demand. But once fairly stated in its simple form it opens up the whole of the subject" (Black 1973, II, p. 410).

[10] Ross Robertson has expressed the view that, with respect to Bentham, "Jevons is simply going through the motions of citing an unquestioned authority before proceeding to an altogether different kind of analysis" (Robertson 1951, p. 233). Although this paper argues that Jevons departed significantly from Bentham in three directions, two taken very deliberately and the third less consciously, there can be little doubt that Jevons saw *himself* as faithfully applying a Benthamite analysis to an activity, namely consumer choice, that Bentham had failed to explore. R.D.C. Black

Preface to the first edition of his book, "In this work I have attempted to treat Economy as a Calculus of Pleasure and Pain, and have sketched out, almost irrespective of previous opinions, the form which the science, as it seems to me, must ultimately take" (Jevons 1871, p. vii). This form, as we know, was to "treat pleasure and pain as positive and negative quantities are treated in algebra … . Our object will always be to maximise the resulting sum in the direction of pleasure, which we may fairly call the positive direction" (ibid., p. 38).

As Stigler aptly described it, "Jevons' attack on the problem of measurability was characteristically frank and confused" (Stigler 1950, p. 317). The confusion arose from the inconsistency between Jevons's *definition* of utility as composed of pleasures and pains, and his desire, for the sake of mathematical fitness, to *measure* (marginal) utility as a real number uniquely determined for a given agent at a given moment by the rate of consumption of a commodity (that is, the continuous mathematical function he had described to his brother eleven years earlier). His exposition was frank as well as confused because, although his own awareness of his inconsistency was far from complete, he did not gloss over (as Marshall surely would have done) the doubts that continued to nag him.

Quite early in his book, Jevons spelled out his crucial transformation of Bentham's utility concept. Although "pleasure and pain are undoubtedly the ultimate objects of the Calculus of Economy … it is convenient to transfer our attention as soon as possible to the physical objects or actions which are the source to us of pleasures or pains … [and to] employ the word *utility* to denote the abstract quality whereby an object serves our purposes, and becomes entitled to rank as a commodity" (Jevons 1871, pp. 44–45). Inconsistent with this transformed usage, however, there followed a standard Benthamite definition: "Utility must be considered as measured by, or even as actually identical with, the addition made to a person's happiness. It is a convenient name for the aggregate of the favourable balance of feeling produced—the sum of the pleasure created and the pain prevented" (ibid., pp. 53–54). It is then this "sum," attributed to the consumption of food, that Jevons mapped onto the real number line: "The law of the variation of the degree of utility of food may thus be represented by a continuous curve … and the perpendicular height of each point of the curve … represents the degree of utility of the commodity when a certain amount has been consumed" (ibid., p. 57). Since, according to Jevons, "We shall seldom need to consider the degree of utility except as regards the last increment which is consumed, [for which] I shall … use the expression *final degree of utility*," the remainder of his book serves to support his sweeping assertion that "the final degree of utility is that function upon which the whole Theory of Economy will be found to turn" (ibid., pp. 61–62).

Although marginal utility equilibria in their various modern manifestations are indeed the pivots upon which the "whole Theory of Economy" still turns, to the

---

comments as follows on Robertson: "To my mind this illustrates an understandable, but regrettable, vice of the intellectual historian—that of trying to explain the thinking of his subject, not in relation to the thought of the subject's own time, but in relation to the ideas of the present time" (Black 1972, p. 123).

https://doi.org/10.1080/104277100112527 Published online by Cambridge University Press

modern reader Jevons's pathbreaking book is irredeemably dated, not just by its assertion of utility's cardinal measurability but also by its expression of cardinality in terms of pleasure and pain. As we have seen, any single-valued measure of Bentham's utility must be an index number. An index number can, of course, be represented legitimately by a continuous surface—a functional relationship jointly determined by the values of its component parts *and* by the values of the weights allocated to them. But when Jevons came to "consider how pleasure and pain can be estimated as magnitudes" in consumer choice, he evaded any consideration of weights on utility's components by associating only *one* type of pleasure with each type of commodity consumed, and then went on, largely implicitly, to treat all consumption pleasures as homogeneous. Before we examine in detail Jevons's procedure for measuring utility, however, it is important to note that he very explicitly rejected what subsequently became Alfred Marshall's approach to the measurement problem, in which utility (Marshall's "value in use") is directly identified with "willingness to pay" at the margin of demand, hence producing a functional relationship between an individual's rate of consumption for any given commodity and a uniquely determined amount of money.[11]

In Marshall's view, Jevons promulgated a "systematic confusion" between "hedonics and economics" in "applying to utility propositions that are only true of price" (Guillebaud 1961, p. 260). Jevons's failure to proceed along Marshall's direct route to individual quantity-price demand curves was not because he missed the signposts. References to market prices as external measures for desire are scattered throughout *The Theory of Political Economy*, beginning on page 14 with: "The will is our pendulum, and its oscillations are minutely registered in all the price lists of the markets." Insofar as relationships between quantities and prices appear in *The Theory of Political Economy*, however, they are treated as aggregate outcomes rather than explanations of individual choice. Black suggests that Jevons's reliance on hedonic motivation was "proof of his allegiance to Bentham" (Black 1972, p. 127), and this may be true in more than one sense. Jevons also failed to follow Walras's *indirect* route to demand curves in deriving them from marginal utility equilibria. An immediate explanation is that Jevons never adopted Walras's general equilibrium perspective, but the underlying reason for that fact may have been Jevons's residual adherence to a Benthamite indeterminism in individual choice. Though his *theory* makes maximization of one-dimensional utility the sole motive for consumer choice, Jevons nonetheless regarded individual choice in practice to be subject to "numerous" motives (and conditions). Thus, for Jevons, observed market equilibria were consequences of the law of large numbers in which individuals' departures from equilibrium

[11] The history of Marshall's theorizing about demand is thoroughly explored in Aldrich (1996), where the relationship between price willingly paid and value in use is traced back at least as far as the unpublished *Essay on Value*, written around 1870. Whereas Jevons was unaware of Marshall's work when writing *The Theory of Political Economy*, Aldrich shows that the influence in the other direction was very strong … more so than Marshall acknowledged: "Of the utility theorists mentioned in *Principles*, Jevons is the only one who matters; his results are cited and his views criticized. The argument also drew on his work in uncredited ways and not just for its 'form' … As he went Jevonian, Marshall distanced himself from the man" (Aldrich 1996, p. 185).

cancel out, not a consequence of a Walrasian aggregation of individuals, each of whom is in personal equilibrium:

> Practically, however, it is quite impossible to detect the operation of general laws of this kind in the actions of one or a few individuals. The motives and conditions are so numerous and complicated, that the resulting actions have the appearance of caprice, and are beyond the analysis and prediction of science … The use of an average, or what is the same, an aggregate result depends upon the high probability that accidental and disturbing causes will operate, in the long run, as often in one direction as the other, so as to neutralise each other … Accordingly, questions which appear, and perhaps are quite, indeterminate as regards individuals, may be capable of exact investigation and solution in regard to great masses and wide averages (Jevons 1871, pp. 22–23).

For Jevons, it was the hedonic motivation hidden behind the complications and caprice of observable behavior that was significant, acting as erratically yet as remorselessly as the gravitational field of force behind the descent of a feather. Since the purpose of *The Theory of Political Economy* was to uncover these hedonic foundations, it was upon them rather than their market manifestations that Jevons chose to concentrate.

A final caveat by Jevons about the use of price-quantity relationships for measuring final degree of utility should also be noted, not the least for its prescience. Unlike Jevons's concern with individual caprice, it is now a familiar component of consumer behavior theory, for it applies under the modern presumption that each consumer's equilibrium is continually maintained. What Jevons saw immediately (and Marshall only very slowly and reluctantly acknowledged), is that the income effects of price changes make willingness to pay for a commodity an unreliable gauge for the utility increment that the commodity purveys:

> The price of a commodity is the only test we have of the utility of the commodity to the purchaser; and if we could tell exactly how much people reduce their consumption of each important article when the price rises, we could determine, at least approximately, the variation of the final degree of utility—the all-important element in Economy … The method of determining the function of utility explained above will hardly apply, however, to the main elements of expenditure. The price of bread, for instance, cannot be properly brought under the equation used, because, when the price of bread rises much, the resources of poor persons are strained, money becomes scarcer with them, and … the utility of money rises … Great difficulty is thrown in the way of all such inquiries by the vast differences in the conditions of persons (ibid., pp. 140–42).

Having examined Jevons's reasons for adhering to a hedonic explanation for optimal choice, I next turn to the three mutations by which he transformed Bentham's utility concept into an entirely different animal. First, Jevons departed from Bentham regarding the validity of deriving aggregate welfare conclusions from the greatest happiness principle. Second, Jevons departed from Bentham by endowing humankind with an innate moral sense. Third, Jevons departed from

Bentham by treating the array of pleasures and pains from acts of consumption as a single homogeneous sensation (pain being "negative pleasure"), varying only in intensity and duration. Let us take each of these three points in turn.

The most explicit of Jevons's departures from Bentham did not involve the meaning or measurement of utility, but rather its applicability for aggregate welfare calculations. Not only did Jevons avoid *intra*personal utility weights by treating pleasure as a homogeneous sensation, he also repudiated *inter*personal utility weights on the grounds that they are unknowable:

> The reader will find, again, that there is never, in a single instance, an attempt made to compare the amount of feeling in one mind with that in another. I see no means by which such comparison can ever be accomplished. The susceptibility of one mind may, for what we know, be a thousand times greater than that of another … Every mind is thus unscrutable to every other mind, and no common denominator of feeling is possible (ibid., p. 21).

In fact, as we have seen, Bentham imposed equal interpersonal weights on ethical rather than empirical grounds. But Bentham also realized, as Jevons apparently did not, that it is not possible to use an aggregate utility concept without *some* form of interpersonal comparison:[12]

> 'Tis in vain to talk of adding quantities which after the addition will continue distinct as they were before, one man's happiness will never be another man's happiness … This addibility of the happiness of different subjects, however, when considered rigorously, it may appear fictitious, is a postulatum without the allowance of which all political reasoning is at a stand (Bentham 1781/1972, p. 495).

Although Jevons's belief in the existence of an innate moral sense, his second major departure from Bentham, was not specifically expounded in *The Theory of Political Economy*, a perusal of his journal and personal letters (Black 1973–77) shows how thoroughly it permeated his general thinking. The underlying belief surfaces most clearly in the book when Jevons spells out a hierarchy of *duties*:

> Starting with the lowest stage—it is a man's duty, as it is his natural inclination, to earn sufficient food and whatever else may best satisfy his proper and moderate desires. If the claims of a family or of friends fall upon him, it may become desirable that he should deny his own desires and even his physical needs their full customary gratification. But the claims of a family are only a step to a higher grade of duties. The safety of a nation, the welfare of great populations, may happen to depend upon his exertions, if he be a soldier or a statesman; claims of a very strong kind may now be over-balanced by claims of a still stronger kind. Nor should I venture to say that, at any point, we have reached the highest rank—the supreme motives which should guide the mind (Jevons 1871, p. 30).

Jevons's concern with duty undoubtedly reflected the time and place in which he

[12] A referee has pointed out that, despite his disclaimer, Jevons did treat the aggregate utility of "trading bodies" in *The Theory of Political Economy*.

https://doi.org/10.1080/104277100112527 Published online by Cambridge University Press

lived, but it was very much at odds with Bentham's utilitarian ethics. As we have already seen, except insofar as it is a person's "natural inclination" (that is, insofar as it gives that person pleasure), appeal to "duty" was for Bentham just another "fiction," used "for prevailing upon the reader to accept of the author's sentiment or opinion as a reason and that a sufficient one for itself." Although the greatest happiness principle will identify some actions in which a net pain value for the perpetrator (that is, pain which exceeds the pleasures of benevolence, amity, and reputation) is outweighed by net pleasure values for other people, Bentham regarded it as *impossible* for an individual to choose such an action voluntarily. Obviously aggregate happiness is enhanced the more that benevolence, amity, and desire for reputation exist, and Bentham was optimistic that environment, education, and example could reshape human nature significantly in those directions, but his ethical system imposes no "duties" that require the *denial* of natural inclinations, in whatever current state they happen to be.[13]

What is more significant for our present story, however, is that Jevons used his distinction between "supreme motives" and individual hedonism to limit drastically the *scope* of economic analysis:

> My present purpose is accomplished in pointing out this hierarchy of feeling, and assigning a proper place to the pleasures and pains with which the economist deals. It is the lowest rank of feelings which we here treat. The calculus of utility aims at supplying the ordinary wants of man at the least cost of labour. Each labourer, in the absence of other motives, is supposed to devote his energy to the accumulation of wealth. A higher calculus of moral right and wrong would be needed to show how he may best employ that wealth for the good of others as well as himself. But when that higher calculus gives no prohibition, we need the lower calculus to gain us the utmost good in matters of moral indifference (ibid., p. 32).

Jevons's third departure from Bentham was his implicit presumption that all self-regarding pleasures—those relevant for the lower calculus—are the *same* sensation, differing only in intensity and duration. Since homogeneous utility is the counterpart of the well-ordered preferences that economists accept without qualms today, this particular mutation of Bentham's utility concept has considerable importance in the history of economic thought. It also accords well with the notion of mathematical fitness as an external guiding principle, operating beyond the awareness of the propagators of change, that Jevons himself did not recognize any discontinuity between that which he inherited and that which he passed on: "Proceeding to consider how pleasure and pain can be estimated as magnitudes, we must undoubtedly accept what Bentham has laid down upon this

[13] For a much more extensive discussion of Bentham's ethical system, and in particular the extent to which it was accepted by John Stuart Mill, see T. Warke, "Multi-Dimensional Utility and the Index Number Problem: Jeremy Bentham, J. S. Mill and Qualitative Hedonism," *Utilitas* (forthcoming).

subject" (Jevons 1879/1970, p. 94).[14] Let us then see just how Jevons in fact distorted significantly that which Bentham had laid down.

In the chapter entitled "Theory of Pleasure and Pain," Jevons quotes from Bentham's *IPML* and proceeds to implement its instructions to take account of intensity, duration, certainty or uncertainty, and propinquity or remoteness. Given a duration base (say one minute), Jevons followed Bentham in using a probability weight to measure degree of certainty, and improved upon Bentham's exposition with a time preference weight to discount remoteness.[15] Had he continued to follow Bentham, Jevons's "final degree of utility" would then be the intensity of a minute's worth of a particular *pleasure or pain*, immediately and with certainty realized. Instead, Jevons applied his term to the consumption of a particular commodity. For Bentham, the act of consumption would be an "exciting cause" of a *complex* pleasure, which is by definition "resolvable into divers simple ones" (Bentham 1789/1996, p. 42). The intensity of a complex pleasure clearly varies with the intensities of its component parts and the weights allocated to them. Jevons simply ignored this index number problem, for his measurement method expresses the multiple pleasures of, say, a shared bottle of wine over a romantic dinner as the intensity value of a *single* sensation (in this case, in units such as "degree of utility per millilitre of wine per minute" which, when multiplied by a rate of wine consumption, yields a corresponding degree of one-dimensional utility)[16] (Jevons 1879/1970, p. 121).

Jevons's distortion of Bentham is exacerbated when he discusses three additional characteristics which Bentham had assigned to acts to take account of interrelationships between their various mixes of pleasures and pains. Quoting

[14] I have hitherto quoted from the first (1871) edition of the *Theory of Political Economy*, but it seems better on this particular topic to present Jevons's position as it had developed by 1879, since by his own testimony the question of measuring pleasure and pain continued to perplex him. As one example of the changes, Jevons's 1871 statement of the law of the variation of the degree of utility is now supplemented by a five-page section on "the theory of dimensions of economic quantities" (Jevons 1879/1970, pp. 117–21).

[15] Stigler's interpretation of the same passage from Bentham seems far less perceptive than that of Jevons: "The first two dimensions [intensity and duration] are clearly relevant to the measurement of a pleasure, but the latter two are better treated as two of the factors which influence an individual's response to a particular pleasure or pain" (Stigler 1950, p. 309). The "factors" in question being "How certain is it?" and "How soon is it?" Jevons's incorporation of these influences within the measure accords exactly with modern practice. However modern his method, though, Jevons's treatment is dated by his proviso that time preference *ought* to be zero, reflecting his belief that the most reliable index of a progressive society is the forward-lookingness of its members: "To secure a maximum of benefit in life, all future events, all future pleasures and pains, should act upon us with the same force as if they were present, allowance being made for their uncertainty. The factor expressing the effect of remoteness should, in short, always be unity, so that time should have no influence." Unfortunately, since "no human mind is constituted in this perfect way," the fact that "a future feeling is always less influential than a present one" had got to be taken into account (Jevons 1879/1970, p. 124).

[16] This may be the place to emphasize that the index number problem of multi-dimensional utility is not resolved by the characteristics approach to consumer theory introduced by Lancaster (1966). Lancaster takes it for granted that each consumer's preferences are well-ordered over bundles of characteristics, implying that the utility associated with different *characteristics* is homogeneous (the fixed weights which link characteristics to commodities are, of course, determined by technology in Lancaster's approach, not by preferences).

Bentham accurately, Jevons tells us that the *fecundity* of an act refers to the chances of its pleasures being followed by other pleasures, and that the *purity* of an act refers to the chances of its pleasures *not* being followed by pains.[17] The consumption of wine is fecund, for example, if its pleasures increase the probability of (or enhance) sexual pleasure to follow, and impure if its pleasures are likely to be followed by a hangover. The third characteristic of an act, its *extent*, refers to externalities—the number of persons other than the perpetrator that the act affects. By accepting a very narrow scope for the "lower calculus" Jevons might argue that its applicability must be confined to acts with zero extent (that is, that "moral indifference" is breached if any externality whatsoever exists). But, since consumption with zero extent is still likely to be fecund and impure to the consumers themselves, it can only have been Jevons's unconscious dissembling, in the face of Bentham's awkward multiplicity of pleasures and pains, that led him to deny the relevance to economics of *all three* act-characteristics: "These three last circumstances are of high importance as regards the theory of morals; but they will not enter into the more simple and restricted problem which we attempt to solve in economics" (ibid., pp. 94–95).[18]

Having thus misrepresented Bentham by reducing his many-hued utility to a single (though unspecified) wavelength, Jevons proceeded in his theory of exchange to treat this pure sensation as differing only in its intensity—its force per time unit—among all acts of consumption and (with a negative sign) all types of work effort by any single agent. So regarded, as we now know, it was not necessary that the measure of this intensity be cardinal. So long as it diminishes continuously with the rate of each activity, it was straightforward for Jevons to arrive at the now-familiar conditions for an agent attaining equilibrium at the margin:

> I never attempt to estimate the whole pleasure gained by purchasing a commodity; the theory merely expresses that, when a man has purchased enough, he would derive equal pleasure from the possession of a small quantity more as he would from the money price of it. Similarly, the whole amount of pleasure that a man gains by a day's labour hardly enters into the question; it is when a man is doubtful whether to increase his hours of labour or not, that we discover an equality between the pain of that extension and the pleasure of the increase of possessions derived from it (ibid., p. 85).

To summarize the story thus far, we can contrast two distinct ways to reach a

[17] Stigler's misreading of Bentham is particularly blatant at this point, and also illustrates his complete though unconscious assimilation of Jevons's utility mutation, as opposed to the Benthamite original upon which he supposed himself to be commenting: "In addition, two further 'dimensions' were added for the appraisal of the total satisfaction of an 'act': the consumption of a loaf of bread might be the pleasure to which the first four dimensions refer; the theft of the loaf might be the act" (Stigler 1950, p. 309n).

[18] Because fecundity and impurity encompass all interdependencies between pleasures and pains of different acts, they include any complementarity and substitutability in consumption. If the exclusion of fecundity and impurity from consumer choice theory were valid, it would therefore justify Jevons's use of an *additive* utility function (however utility is to be measured). This was surely not Jevons's intention, however, and we must take him at his word that his purpose was to exclude moral issues, not to specify zero values for the off-diagonal terms in the Jacobean matrix of consumer choice.

mathematically tractable theory of choice. The more modern route begins with axioms of well-ordered preferences, so that it is *synonymous with rationality* that agents facing well-defined constraints arrive at a uniquely determined choice, acting "as if" they were maximizing a utility function with a one-dimensional domain. The route charted by Jevons begins from the other side of rational choice by establishing utility as a one-dimensional object, thus providing agents with a monotonic criterion by which to *carry out* the ordering of the alternative outcomes they face. But there is a pitfall in Jevons's approach to rational choice determinacy that the modern method rules out, and which Jevons himself was careful to acknowledge. Fully rational agents, according to Jevons, often lack the *capacity* to grasp fully the dictates of this theoretically unambiguous standard: "It is true that the mind often hesitates and is perplexed in making a choice of great importance; this indicates either varying estimates of the motives, or a feeling of incapacity to grasp the quantities concerned" (ibid, p. 84).[19]

Jevons's caveat brings us to a third method of attaining a mathematical formalization of choice theory, the method promulgated by Marshall and the one still most frequently used when undergraduates are introduced to microeconomics: if the mind hesitates over a choice but is *neither* perplexed *nor* views it of great importance, we are entitled to presume that the cause of hesitation is "indifference," which further implies that the agent *has* grasped the marginal utility quantities for each alternative with sufficient precision to regard them as "equal" *whether or not they are qualitatively identical*. From equality follows "more" and "less" and, if indifference can be traced across the entire choice space, we attain our well-ordered preferences (or, at least, an indifference mapping). Before proceeding to Marshall, though, we should note that Jevons had also moved along this route between 1871 and 1879. Even in 1871 he had made a connection between indifference and equilibrium, for example when discussing the market for consols: "When the price of the funds is very steady and the market quiescent, it means that the stocks are distributed among holders in such a way that the exchange of more or less at the prevailing price is a matter of indifference" (Jevons 1871, p. 109). But by 1879, indifference had become a "law."

In the second edition, Jevons *introduces* his law of indifference as a conclusion, not about equilibrium, but rather about the uniformity of price at any one time in a fully informed market for a "perfectly uniform or homogeneous commodity" (Jevons 1879/1970, p. 136). If this conclusion seems far too innocuous to justify Jevons's claim that "the principle above expressed is a general law of the utmost importance in economics," on the very next page his usage becomes considerably more potent. Indifference is no longer a condition that defines arbitrage. It is now applied to *pairs* of commodities that *serve a similar purpose*:

[19] The first source of mental perplexity in the quotation—varying estimates of the motives—is read most consistently as a reference to Jevons's concern about conflict between duty to self and duty to others, since within the "lower calculus" there should be only the *single* motive of maximizing perceived personal pleasure. The second source of perplexity can then be read as a difficulty in *perceiving* the precise utility ranking among alternative sources of personal pleasure.

> I propose to call it the *law of indifference*, meaning that, when two objects or commodities are subject to no important difference as regards the purpose in view, they will either of them be taken instead of the other with perfect indifference by a purchaser. Every such act of indifferent choice gives rise to an equation of degrees of utility, so that in this principle of indifference we have one of the central pivots of the theory (ibid., p. 137).

Jevons's failure to distinguish between a law of uniform price in a single market on the one hand, and his marginal utility equilibrium condition on the other, where it is the quantities of two different commodities already in possession of a particular agent that makes an additional unit of each a matter of indifference, is with modern hindsight a very considerable degree of confusion. Yet there is an explanation for Jevons's thinking, and it lies in his ongoing doubts about the homogeneity of utility/pleasure *unless* the choice alternatives be subject to "no important difference as regards the purpose in view." Modern theory, of course, pays no heed to any such restriction and neither, as we shall see, did Marshall. But Jevons arrived at his expanded "law of indifference" from his work on scientific method, to which he devoted a great deal of time and effort during the 1870s and which came to fruition in *The Principles of Science* (the second edition of which was published posthumously in 1887). According to Jevons, the unifying theme of all science is what he called "The Substitution of Similars":

> The one supreme rule of inference consists, as I have said, in the direction to affirm of anything whatever is known of its like, equal or equivalent. The Substitution of Similars is a phrase which seems aptly to express the capacity of mutual replacement existing in any two objects which are like or equivalent to a sufficient degree. It is a matter for further investigation to ascertain when and for what purposes a degree of similarity less than complete identity is sufficient to warrant substitution (Jevons 1887, p. 17).

I am not the first to point out that Jevons's expanded "law of indifference" is simply his "substitution of similars" in the context of choice, with objects being subject "to no important difference as regards the purpose in view" corresponding to objects being "like or equivalent to a sufficient degree."[20] What I want to highlight here, however, is that Jevons would then make it "a matter for further investigation" to ascertain when differences among a purchaser's purposes are sufficiently important not to warrant indifference at the margin. In these choice situations, the Jevonian agent would not be indifferent but *indecisive*, either because the objects of choice involve conflicting motives of duty, or because the nature of the differences among alternatives invokes "a feeling of incapacity to grasp the [utility] quantities concerned."

One way to sum up the mathematically tractable choice theory that Jevons bequeathed to his successors is by reviewing the three hedges with which he surrounded it. First, the theory describes only an underlying tendency toward individual equilibrium, whereas actual decision-making processes are subject to "motives and conditions … so numerous and complicated, that the resulting

[20] Robertson attributes the initial observation to Wicksteed, though without a specific reference (Robertson 1951, p. 243).

https://doi.org/10.1080/104277100112527 Published online by Cambridge University Press

actions have the appearance of caprice, and are beyond the analysis and prediction of science." Second, the theory applies only to "the lowest rank of feelings … in the absence of other motives," which led Jevons to preclude actions involving externalities (and, though inadvertently, all other utility interdependencies as well). Third, even where the theory is applicable, if objects of choice are insufficiently alike, the mind will remain perplexed, unable to grasp their utilities with sufficient precision to locate equality (indifference) among them. It required someone with fewer qualms than Jevons about plowing over these hedges, with the rhetorical skills to cover their traces and with pervasive influence upon other economists, before one-dimensional utility would be established as the solitary explanation for rational choice. Such a one was Alfred Marshall.

IV. ALFRED MARSHALL: THE UBIQUITY OF INDIFFERENCE

Several recent papers (for example, Aldrich 1996, Hart 1996, and White 1990) have stressed the pragmatic purposes of Marshall's writing and his concomitant carelessness with theoretical rigor or even consistency. With reference to Marshall's treatment of an individual's demand for a single commodity, Aldrich tells us:

> Marshall's arguments for these principles are hard to grasp because the premises are not fully specified and the conclusions are not clearly expressed. An embarrassment of theorems can be proved with assumptions made somewhere in the *Principles* … (Aldrich 1996, p. 189).

It would not be productive, therefore, to proceed with Marshall as I have done with Bentham and Jevons, seeking to trace and explain a coherent theoretical structure underlying their concepts of utility. Instead, I shall merely illustrate Marshall's great power of persuasion, the tone of sweet reason by which he managed to convince the bulk of economists that their profession had found the key to a determinate and wide-ranging theory of individual choice. Marshall begins with an uncompromising disavowal of commensurability when it comes to "affections of the mind," the term itself replacing Jevons's "pleasure and pain":

> It is essential to note that the economist does not claim to measure any affection of the mind in itself, or directly; but only indirectly through its effect. No one can compare and measure accurately against one another even his own mental states at different times: and no one can measure the mental states of another at all except indirectly and conjecturally by their effects (Marshall 1920, p. 15).[21]

Marshall next adopts Jevons's classification of affections according to their moral qualities and, like Jevons but far more subtly, he uses this distinction to

[21] Marshall's most detailed treatment of utility and individual demand for a commodity was in the third (1895) edition of the *Principles*. My reason for quoting from the 1920 edition, which generally retained the 1895 line on these matters, is that Marshall's explanation of individual choice had by then largely prevailed, at least among English-speaking economists.

https://doi.org/10.1080/104277100112527 Published online by Cambridge University Press

isolate one-dimensional utility ("pleasures and pains of the same kind") as the underlying explanation for an individual's market behavior (the observable outcome of the "lower" affections):

> Of course various affections belong to man's higher nature and others to his lower, and are thus different in kind. But, even if we confine our attention to mere physical pleasures and pains of the same kind, we find that they can only be compared indirectly by their effects. In fact, even this comparison is necessarily to some extent conjectural, unless they occur to the same person at the same time (ibid., p. 15).

What appears on the surface as a series of disclaimers in fact asks us to concede that pleasures and pains *can*, beyond conjecture, "be compared indirectly by their effects" (that is to say, revealed preferences *will* be well-ordered), *if* "they occur to the same person at the same time." Jevons's first hedge, that complicated motives and conditions render individual market behavior beyond the analysis and prediction of science, so that consumption regularities can only be observed in the aggregate, has already disappeared. Marshall's performance is so smooth that it is worth investigating more closely the backstage machinery that supports it.

The quotation begins with the acknowledgement of differences in kind among affections, with "man's higher nature" covering Jevons's multi-layered hierarchy of other-regarding duties. Next, "mere physical pleasures and pains of the same kind" are selected as a subset of affections. But the second sentence is deliberately vague about two crucial points. First, are "affections belonging to man's lower nature" and "mere physical pleasures and pains" an identity, or is the latter a *proper* subset of the former? Second, are *all* physical pleasures and pains to be regarded as "of the same kind," or are we to confine our attention to some (unspecified) subset of such pleasures and pains (hunger gratification, perhaps) which can be so regarded? Marshall cagily avoids committing himself, leaving the casual reader to accept unthinkingly the interpretation most congenial to his purpose, namely that all "lower nature" affections *are* of the same kind, so that the effects of all choices involving them can be compared unambiguously (no utility weights being required) by any given person at any given time.

True to his general style, Marshall never explicitly states that he is assuming one-dimensional utility, nor, in the manner of Jevons, does he attempt to discover the measure of utility from first principles. Instead, he mounts his primary campaign for determinate choice from the third angle that we have identified, assuming (while calling it "ordinary usage") that hesitation at the margin of choice necessarily identifies equal pleasures (and by extension, unambiguous comparability among all other combinations of outcomes):

> For instance the pleasures which two persons derive from smoking cannot be directly compared: nor can even those which the same person derives from it at different times. But if we find a man in doubt whether to spend a few pence on a cigar, or a cup of tea, or on riding home instead of walking home, then we may follow ordinary usage, and say that he expects from them equal pleasures (ibid., p. 15).

Having found the suitable bulldozer, Marshall succeeds in leveling Jevons's last two hedges. Jevons's qualification to his law of indifference, that the objects of choice be "subject to no important difference as regards the purpose in view" (here, alleviation of tobacco craving, alleviation of thirst, and avoidance of physical effort) has disappeared without trace. No longer are there circumstances where the mind lacks the capacity to grasp the quantities concerned. Even on moral issues, though philosophers may concern themselves with varying estimates of the motives when the alternatives represent conflict between a person's higher and lower nature, economists see in hesitation the manifestation of *evenly balanced* incentives to action:

> Thus measuring a mental state, as men do in ordinary life, by its motor-force or the incentive which it affords to action, no new difficulty is introduced by the fact that some of the motives of which we have to take account belong to man's higher nature, and others to his lower. For suppose that the person, whom we saw doubting between several little gratifications for himself, had thought after a while of a poor invalid whom he would pass on his way home; and had spent some time in making up his mind whether he would choose a physical gratification for himself, or would do a kindly act and rejoice in another's joy. As his desires turned now towards the one, now the other, there would be change in the quality of his mental states; and the philosopher is bound to study the nature of the change. But the economist studies mental states rather through their manifestations than in themselves; and if he finds they afford evenly balanced incentives to action, he treats them *primâ facie* as for his purpose equal (ibid., p. 16).

Marshall never explains *how* economists know that "a man found in doubt" is facing an even balance rather than hard-to-grasp weights; like quantum uncertainty in Schrödinger's wave function, the mind's perplexity is mysteriously collapsed by outside observation.

For all his rhetorical acuity, however, Marshall's most effective ploy for extending economists' acceptance of well-ordered preferences was largely unintended, consisting as it did of a false trail down which the keenest-nosed hounds of the profession were soon baying. The issue with which Jevons had struggled, commensurability of pleasures or pains felt by *a single agent*, lay neglected in the underbrush as the hunt focused instead on Marshall's further contention, in aid of consumers' surplus, that two outcomes over which *different* persons hesitate must afford equal pleasure, measurable by their time or money cost, provided only that the agents are "in similar circumstances":

> If the desires to secure either of two pleasures will induce people in similar circumstances each to do just an hour's extra work, or will induce men in the same rank of life and with the same means each to pay a shilling for it; we then may say that those pleasures are equal for our purposes, because the desires for them are equally strong incentives to action for persons under similar conditions (ibid., p. 16).

The theoretical bear was eventually treed and shot down by professional consensus that simultaneous equilibria do not imply homogeneity of pleasure among heterogeneous sensibilities just because their external circumstances

happen to be similar. And despite Marshall's camouflage of "common usage," the empirical bear that he sent along the same track was vulnerable to the same objection (over and above the issue of its *factual* accuracy):

> By far the greater number of the events with which economics deals affect in about equal proportions all the different classes of society; so that if the money measures of the happiness caused by the two events are equal, it is reasonable and in accordance with common usage to regard the amounts of the happiness in the two cases as equivalent (ibid., p. 20).

A recounting of the debate over the commensurability of pleasures and pains felt by different persons is not needed for this paper. In the end, according to his nephew, even Marshall was forced to recognize, as "a major disappointment in his life," that "his concept of consumer's surplus was devoid of important practical application, because it was not capable of being quantified in a meaningful way" (Guillebaud 1970, p. 6). My point here is that the length and intensity of the debate served to divert attention from the parallel question about commensurability of pleasures and pains felt "by the same person at the same time." The eventual acknowledgement that the utility of a community consisting of n agents experiencing *m* different types of pleasure and pain is *at least* *n*-dimensional was somehow transmuted into a largely unquestioned assumption that it is *no more than* *n*-dimensional. Mainly by default, it seems, Marshall succeeded in convincing most economists that a single agent at any given time has well-ordered preferences, either because he or she is comparing pleasures and pains of one kind only, or failing that, because his or her hesitation at the margin of choice implies indifference—an absence of perplexity about the personal utility weights that apply to heterogeneous pleasures and pains. The profession was primed to welcome the news, soon to arrive from more rigorous analysts, that simply by taking well-ordered preferences as initial axioms in choice theory, the association of utility with pleasure and pain, or affections of the mind, could be dispensed with altogether. Whether well-ordered preferences are congruent with reality has not been of major concern.

V. CONCLUSION

This paper has traced a part of the process by which the utility concept in economics has mutated from Bentham's multi-dimensional array of pleasures and pains to the modern one-dimensional index of desirability. My exposition of Bentham's perception of utility has emphasized its measurement imprecision and the consequent ambiguity of optimal choice, well recognized by Bentham at both the aggregate and individual levels. My account of subsequent revisions to Bentham's utility concept has been confined to the two authors who, in my view, most effectively illustrate the issues at stake if agents are to be presumed capable of a complete and transitive preference ordering over all conceivable outcomes that they face.

W. S. Jevons retained a cardinal concept of utility throughout his tragically truncated creative life. Compatible with his allegiance to utilitarian hedonism, Jevons's cardinal utility was meant to measure a quantity of pleasure which,

following Bentham as to its "elements," he expressed as the intensity of a given duration of pleasure, immediately and with certainty received. Jevons's theory of choice then makes an agent's decision depend upon the differing intensities of pleasures expected from alternative outcomes together with their terms of trade, and in order for these choices to be *determinate*, the pleasures must be otherwise homogeneous. Jevons's equation of exchange illustrates consumer equilibrium when this condition is met. It is clear from his narrative discussion, however, that Jevons did not presume automatic consumer equilibrium in practice, for he tells us that "the mind often hesitates and is perplexed in making a choice of great importance." I have highlighted two restrictions that Jevons therefore imposed on the applicability of his theory, each of which limits the importance of the choices to an agent and increases the comparability among the pleasures they involve. First, he limited his theory to self-regarding choices, thus eliminating conflicts between agents' personal pleasures and their sense of duty to others. Secondly, though less coherently, he suggested that in order for pleasure quantities from alternative commodity increments to be judged equal, such that his "law of indifference" applies, the commodities should be "subject to no important difference as regards the purpose in view."

Alfred Marshall's utility concept was recognizably modern, in the sense that the marginal utility of any one commodity can be measured by a precise rate at which the agent is willing to exchange it for any other commodity. Marshall did not, however, deduce this conclusion from the modern assumption that agents operate with complete and transitive preferences. Instead, and quite differently from Jevons, Marshall took it for granted that observed agents are *always* in a state of consumer equilibrium, and that this equilibrium *necessarily* represents evenly balanced incentives to action. If only because the implications are the same as assuming complete and transitive preferences from the start, it is worth reminding ourselves what this choice determinism says about human states of mind. Firstly, any agent facing well-specified and desirable alternatives A and B must have one of only three possible responses: A is strictly preferred to B; B is strictly preferred to A; or the agent is exactly indifferent between A and B. Secondly, if for example A is strictly preferred to B, there is some precise cost saving in choosing B (measured in time, money, or anything else) beyond which the agent's scale abruptly tilts toward B; at one penny less, the agent sticks with A.

By way of contrast, an agent living in the more ambiguous world of Bentham and Jevons could have any one of a *continuous* range of responses to desirable alternatives A and B, from strict preference for A, through a declining yet still prevailing A preference (for example, a 60–40 "leaning" toward A) up to indifference, and on through a rising strength of preference for B until strict preference is reached in that direction.[22] Furthermore, there would be no penny-precise cost differential between the two alternatives at which an agent's choice abruptly changes sides. This "incapacity to grasp the quantities concerned" would indeed give actions "the appearance of caprice" and, hence,

[22] Although this paper is not the place to develop it further, readers may recognize an affinity between my description of ambiguous preferences and the fuzzy sets introduced in Zadeh (1965).

https://doi.org/10.1080/104277100112527 Published online by Cambridge University Press

mathematical messiness, but as Stigler proclaimed, that in itself is no reason to avoid comparing these implications with observable behavior.

REFERENCES

Aldrich, J. 1996. "The Course of Marshall's Theorizing about Demand." *History of Political Economy* 28: 171–217.

Bentham, J. *ca.* 1773. In Mary P. Mack, *Jeremy Bentham: An Odyssey of Ideas*. London: Heineman, 1962.

——. 1781. In Elie Halévy, *The Growth of Philosophic Radicalism*, trans. M. Morris. London: Faber and Faber, 1972.

——. 1782. In Elie Halévy, *La Jeunesse de Bentham*. Paris: Felic Alcan, 1901.

——. 1789. *An Introduction to the Principles of Morals and Legislation*, edited by J.H. Burns and H.L.A. Hart. Oxford: Clarendon Press, 1996.

——. 1822. *Codification Proposal*. In W. Stark, ed., *Jeremy Bentham's Economic Writings*, vol. 3. London: George Allen & Unwin, 1954.

Black, R.D.C. 1972. "Jevons, Bentham and De Morgan." *Economica* 39: 119–34.

——, ed. 1973–77. *Papers and Correspondence of William Stanley Jevons*, vols. II-IV. London: Macmillan.

Black, R.D.C. and R. Könekamp, eds. 1973. *Papers and Correspondence of William Stanley Jevons*, vol. I. London: Macmillan

Ellickson, B. 1993. *Competitive Equilibrium*. New York: Cambridge University Press.

Guillebaud, C.W. 1961. *Marshall's Principles of Economics: Notes*. London: Macmillan.

——. 1970. "Some Personal Reminiscences of Alfred Marshall." *History of Political Economy* 1: 1–8.

Hare, R.M. 1982. "Ethical Theory and Utilitarianism." In A. Sen and B. Williams, eds., *Utilitarianism and Beyond*. Cambridge: Cambridge University Press.

Hart, N. 1996 "Equilibrium and Time: Marshall's Dilemma." *Journal of Economic Methodology* 3: 285–306.

Jevons, W.S. 1871. *The Theory of Political Economy*. London: Macmillan and Co.

——. 1879. *The Theory of Political Economy*. London: Penguin Books, 1970.

——. 1887. *The Principles of Science*. London: Macmillan.

La Nauze, J.A. 1953. "The Conception of Jevons's Utility Theory." *Economica* 20: 356–58.

Lancaster, K. 1966. "A New Approach to Consumer Theory." *Journal of Political Economy* 74: 132–57.

Marshall, A. 1920. *Principles of Economics*, 8th Edn. London: Macmillan and Co.

Moore, G.E. 1903. *Principia Ethica*. Cambridge: The University Press, 1960.

Petrella, F. 1977. "Benthamism and the Demise of Classical Economic *Ordnungspolitik*." *History of Political Economy* 9: 215–36.

Pigou, A.C. 1903. "Some Remarks on Utility." *Economic Journal* 13: 58–68.

Plamenatz, J. 1972. In Elie Halévy, *The Growth of Philosophic Radicalism*, trans. M. Morris. London: Faber and Faber.

Robertson, R.M. 1951. "Jevons and his Precursors." *Econometrica* 19: 229–49.

Robbins, L. 1978. *The Theory of Economic Policy in English Classical Political Economy*, 2nd Edn. London: Macmillan.

Stigler, George J. 1950. "The Development of Utility Theory." *Journal of Political Economy* 58: 307–27, 373–96.

Warke, T. W. Forthcoming. "Multi-Dimensional Utility and the Index Number Problem: Jeremy Bentham, J. S. Mill and Qualitative Hedonism." *Utilitas*.

White, M.V. 1990. "Invention in the Face of Necessity: Marshallian Rhetoric and the Giffen Good(s)." *Economic Record* 66: 1–11.

Zadeh, Lofti. 1965. "Fuzzy Sets." *Information and Control* 8: 338–53.