542. Kirkden, R.D*., Edwards, J.S.S.+and Broom, D.M.* 2003. A theoretical comparison of the consumer surplus and the elasticities of demand as measures of motivational strength. *Animal Behaviour,* 65, 157-178. ***Department of Clinical Veterinary Medicine, University of Cambridge +Faculty of Economics and Politics, University of Cambridge**

Abstract

The price and income elasticities of demand have been used by ethologists to estimate motivationalstrength. The consumer surplus is an alternative measure of motivation, deriving from microeconomic theory. We made a theoretical assessment of the validity and versatility of these indices. Two factors are expected to compromise the internal validity (veracity) of the elasticity of demand indices: failure to take into account the amount that an animal is required to pay to maintain some level of consumption; and a tendency to confuse its readiness to defend a preferred consumption level with a propensity to becomesatiated. A third factor, expected to compromise the external validity (usefulness) of these indices, is theunrealistic assumption that a single value can be assigned to each resource. None of these problems applies to the consumer surplus index. One further factor, expected to compromise the internal validity of both the consumer surplus and price elasticity indices, is their failure to account for the effects that income has upon consumption. Overall, we conclude that the consumer surplus should be more valid,both internally and externally, than the price elasticity index. The consumer surplus should also be more externally valid than the income elasticity index, but it is unclear, on balance, which of these indices is the more internally valid. Finally, we show that both elasticity indices are considerably less versatile than the consumer surplus, owing to the assumption that a single value can be assigned to each resource.

Introduction

In applied ethology, the measurement of motivational strength is important for assessing the welfare of captive animals. If an animal is strongly motivated to obtain a resource or to perform an activity, but the captive environment does not permit her to do so, then she will suffer in such an environment (Dawkins 1983a, 1988, 1990). (We use the term 'resource' hereafter to refer either to a commodity, or to the opportunity to perform an activity.) Motivation cannot be measured directly. Rather, it must be inferred from behaviour or physiology. Obtaining estimates of motivational strength is therefore a two-step process. First, some kind of response must be measured. For behavioural responses, this is often done by means of a preference test, which is the method this paper is concerned with. Then the observed response must be interpreted, by means of a rule or formula, termed here an 'index of motivational strength'.

The simplest preference tests are designed to establish whether a preference or motivation exists, but give little indication of its strength (Duncan 1978, 1992). They compare resources that satisfy the same motivation to different extents. The resource that is consumed in a larger quantity (e.g. Hughes 1975; Dawkins 1977, 1978, 1981), or on which the subject spends more time, effort or some other currency (e.g. Hughes & Black 1973; Dawkins 1977, 1981; Baldwin & Start 1981, 1985; Dawkins & Beardsley 1986; Nicol 1986; Lagadic & Faure 1987; Blom et al. 1995; Manser et al. 1995; van de Weerd et al. 1998; Harri et al. 2001), is said to be preferred. To estimate motivational strength, the most important modification to the experimental procedure consists of comparing resources that satisfy different motivational systems. An animal can be said to be strongly motivated to obtain a resource only if her motivation is comparable with that for a commodity that satisfies a different motivational system and whose importance is not in question (Dawkins 1980). Many types of preference tests have been used to estimate motivational strength (Broom & Johnson 1993;

Fraser & Matthews 1997), but there is a common principle. All such tests require an animal to make a sacrifice of some kind to gain access to some quantity of a resource or to spend a given amount of time consuming it. Like the simple preference tests, experiments designed to measure motivational strength can be classified into operant tests and choice tests. In operant tests (e.g. Faure 1986; Duncan & Kite 1987; Matthews & Ladewig 1994; Bubier 1996a; Manser et al. 1996; Sherwin 1996; Sherwin & Nicol 1996; Cooper & Mason 1997a, 1999, 2000, 2001; Gunnarsson et al. 2000; Mason et al. 2001; Warburton & Nicol 2001), a cost is imposed upon access to the resource by requiring the subject to perform a task. In choice tests (e.g. Dawkins 1977, 1978, 1981, 1983a, b; van Rooijen 1980; Bubier 1996b; van de Weerd et al. 1998; Harri et al. 2001; Warburton & Nicol 2001), a cost is imposed instead upon consumption, generally requiring the animal to divide a fixed amount of time or a fixed number of choices between alternative resources.

A number of different indices of motivational strength have been used in conjunction with these preference tests. They all work on the assumption that an animal's motivational strength is synonymous with her readiness to make a sacrifice, or, in other words, her willingness to pay. They can be divided into two categories, according to how willingness to pay is assessed: quantity indices and rate of change indices. Of the quantity indices, the simpler ones are identical to the measures used in tests designed to establish whether preferences exist (e.g. Dawkins 1977, 1978, 1981, 1983a; van Rooijen 1980; Faure 1986; van de Weerd et al. 1998; Harri et al. 2001; Warburton & Nicol 2001). It is reasoned that the more an animal consumes, or the more she pays, the stronger her motivation must be. More sophisticated quantity indices infer motivational strength from an estimate of the greatest amount of time, effort or other currency that the animal would be prepared to commit to obtaining a resource, or a specific quantity of a resource. Estimates are derived from observations of the quantities consumed when the cost of consumption is manipulated. These indices include the reservation price index (e.g. Duncan & Kite 1987; Manser et al. 1996; Cooper & Mason 1999, 2000, 2001; Mason et al. 2001) and the consumer surplus (Kirkden 2000; Mason et al. 2001). Rate of change indices, in contrast, infer an animal's readiness to make a sacrifice from the rate at which some measure of consumption or expenditure changes as the cost of consumption is increased (e.g. Dawkins 1983a; Faure & Lagadic 1994; Matthews & Ladewig 1994; Bubier 1996a, b; Sherwin 1996; Sherwin & Nicol 1996; Cooper & Mason 1997a, 1999, 2000, 2001; Gunnarsson et al. 2000; Mason et al. 2001; Warburton & Nicol 2001). The rationale is that if a resource is sufficiently important to an animal, then she should be prepared to maintain consumption at some preferred level, even when it becomes costly to do so (Dawkins 1983a, 1988, 1990). The price elasticity of demand and the income elasticity of demand are rate of change indices.

The simpler quantity indices are frequently unsuitable for comparing resources that satisfy different motivations. This is because different motivations vary not only in their strength, but also in other properties, such as the rate at which they can be satiated and the quantities of their respective resources that are required to satiate them. Measures of interaction time and consumption cannot distinguish these properties from motivational strength and will tend to indicate that resources that are consumed at the slowest rates, or in the largest quantities, are the most highly valued. This problem has generally been regarded as a feature of choice tests, where it is associated with nonexclusive choices (Duncan 1978, 1992; Nicol 1997). However, it also applies to operant tests. The rate of change indices were intended to overcome this problem, by focusing upon the resilience of consumption, not its magnitude (Dawkins 1983a) and have come to be widely used. However, the validity of the price elasticity index has been questioned by Ng (1990) and Houston (1997), both of whom recommended using the consumer surplus instead. Houston (1997) constructed a mathematical argument demonstrating that the price elasticity of demand was not always valid. This argument remains a controversial issue in applied ethology (e.g. Dawkins 1997) and few experimental studies have yet made use of the consumer surplus index (Kirkden 2000; Mason et al. 2001).

Many of the terms used in preference testing have been borrowed from a branch of microeconomics known as consumer demand theory. These include price, expenditure, income, demand, price elasticity of demand, income elasticity of demand, reservation price and consumer surplus. There is a tendency to describe any procedure that draws upon these concepts for the purposes of estimating motivational strength as a 'consumer demand approach'. However, in most cases this is a misnomer. Most indices that have been used by ethologists as measures of motivational strength would never be used by economists for a similar purpose. In fact, the only indices that are underpinned by consumer demand theory are the consumer surplus and the reservation price. The rationale for using elasticities of demand as indices of motivational strength did not originate in economics, but was developed by Dawkins (1983a, 1988, 1990) with reference to existing ideas in theoretical ethology (e.g. Houston & McFarland 1980) and experimental psychology (e.g. Hogan & Roper 1978; Lea 1978; Hursh 1984). Furthermore, ethologists have only recently considered the economic basis of the reservation price index (e.g. Cooper & Mason 1999, 2000, 2001; Mason et al. 2001). In this respect, the recent use of the consumer surplus and reservation price indices represents the first attempt in ethology to apply economic theory to the measurement of motivation.

Economists do not use a concept of motivation, but they do have methods for estimating the value that an individual attaches to a given quantity of a resource, which involve using indices such as the consumer surplus to measure her willingness to pay for it. Because motivation is essentially a construct used to describe the strength or willingness with which an animal engages in an appetitive or consummatory activity (Toates 1986, page 7), there is a close correspondence between what ethologists and economists are attempting to measure.

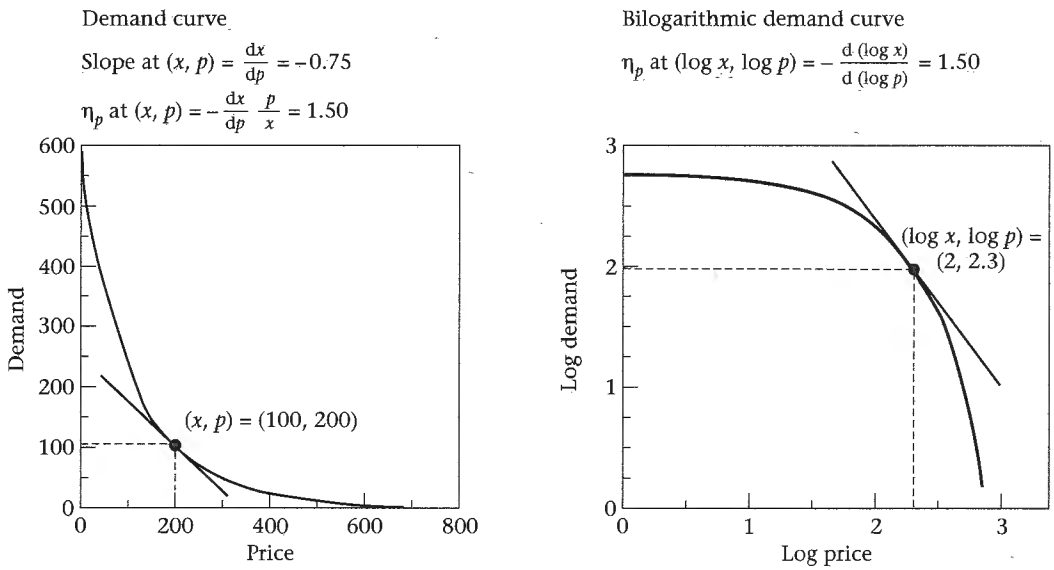

Figure 1. A typical demand curve for food in rats and a bilogarithmic plot of this demand curve (from Hursh et al. 1989), with the price elasticity of demand, η_p, evaluated at a given point on each curve.

Our principal objective in this paper is to investigate the validity of the consumer surplus and elasticities of demand as indices of motivational strength. The validity of the reservation price index, which is closely related to the consumer surplus, is also considered. Theoretical arguments are presented that have a wider application than Houston's (1997). The majority of these arguments are developed from first principles, without drawing upon consumer demand theory. The only argument that derives from economic theory pertains to a shortcoming of the consumer surplus index. This paper is not an economic critique of methods for the measurement of motivation, but an attempt to evaluate alternative approaches without a theoretical bias. Our second objective is to compare the versatility of the various indices. A third objective is to indicate how the consumer surplus index might be estimated in practice.

There are two types of validity: internal and external. We consider both types in this paper, although most arguments pertain to internal validity. In general terms, an index is said to be internally valid if it measures what it purports to measure (Howell 1997; Kline 1998). Internal validity can be defined more precisely if the required level of measurement (Siegel & Castellan 1988, page 23) is specified. The minimum level of measurement required to describe preference relations is an ordinal scale. Thus, a valid index of motivational strength should rank resources, or specified quantities of resources, in an order that reflects the subject's preferences. When an index is invalid, this is sometimes because it fails to account for a factor that is relevant to what it purports to measure. Insofar as some indices fail to account for more than one factor, one may speak of degrees of validity, not merely of its presence or absence.

An index is said to be externally valid if what it purports to measure is useful (Mason et al. 1997). For ethologists, this means that results obtained in the laboratory should be relevant and meaningful in the environments in which they are applied.

The accuracy of an index refers to the magnitude of error associated with its estimation (Kline 1998). If the magnitude of the error depends on the identity of the resource, then the values obtained for a series of resources may not have an ordinal relation to their true values, with the consequence that the error also compromises the internal validity of the index.

The versatility of an index refers, in this paper, to its scope for addressing practical questions of interest to ethologists, both in terms of the range of questions that can be addressed and the ease with which this can be done.

DEFINITIONS OF THE ELASTICITIES OF DEMAND

Price Elasticity of Demand

The price elasticity of demand for a good, η_p, is defined as the proportional rate at which a subject's consumption of the good, also known as demand, changes with the price of the good (Varian 1996, page 266),

$$\eta_p = -\frac{(dx/x)}{(dp/p)} = -\frac{dx}{dp}\frac{p}{x} \qquad (1)$$

where x denotes the initial consumption level, p the initial price level, dx and dp infinitesimal changes in consumption and price, respectively, and dx/dp the instantaneous rate at which consumption changes with price. In other words, the price elasticity of demand is approximately equal to the percentage change in consumption that occurs when the price of a resource is increased or reduced by 1%, starting at a given price or consumption level.

dx/dp corresponds to the slope at a point on a demand curve (Fig. 1), which depicts the relation between consumption and price. Because consumption usually declines as price increases, dx/dp is normally negative. The purpose of the minus sign in the formula is to render the price elasticity of demand positive, thereby making it easier to interpret. The division of the dx and dp terms by x and p, respectively, converts an absolute rate of change into a relative one, thereby rendering the price elasticity of demand independent of the units in which price and consumption are measured. Thus, the price elasticity of demand for a foodstuff, for example, would not depend on whether consumption was measured in pounds or kilograms.

It can be demonstrated mathematically (Chiang 1984, page 304) that the price elasticity of demand is also equal to the negative of the slope at a point on a bilogarithmic plot of the demand curve (Fig. 1):

$$\eta_p = -\frac{d(\log x)}{d(\log p)}. \quad (2)$$

The demand for a resource is said to be price elastic at a given price or consumption level if the price elasticity of demand exceeds 1, that is if the percentage decline in demand exceeds the percentage increase in price. Demand is said to be price inelastic if the price elasticity is less than 1, and it is said to be of unitary price elasticity if the price elasticity of demand is equal to 1.

A property of the price elasticity of demand is that its magnitude indicates whether a subject's expenditure on a resource increases or decreases when price is increased. When demand is price elastic ($\eta_p>1$), a price increase results in reduced expenditure because demand goes down proportionally more quickly than price goes up. Conversely, when demand is price inelastic ($\eta_p<1$), expenditure increases with price. Finally, when demand is of unitary price elasticity, expenditure is unaffected by price changes. However, the rate of change in expenditure cannot be predicted from the magnitude of the price elasticity of demand, because it also depends upon the quantity of the resource that is consumed before the price change (Varian 1996, page 270):

$$\frac{de}{dp} = x[1-|\eta_p|], \quad (3)$$

where de/dp denotes the instantaneous rate at which expenditure changes with price, x the quantity consumed before the price change, and $|\eta_p|$ the absolute value of the price elasticity of demand.

This relation between the price elasticity of demand and the rate of change in expenditure can be exploited to derive analogues of the price elasticity index for use in choice tests (e.g. Bubier 1996a, b), which usually yield measures of expenditure. The rate of change in expenditure must be divided by initial consumption to cancel out the x term on the right and render the rate of change in expenditure proportional to the price elasticity of demand.

Income Elasticity of Demand

The income elasticity of demand, η_m, is defined similarly as the proportional rate at which a subject's consumption changes with her income (Varian 1996, page 276),

$$\eta_m = \frac{(dx/x)}{(dm/m)} = \frac{dx}{dm}\frac{m}{x} \quad (4)$$

where m denotes the initial income level, dm an infinitesimal change in income, and dx/dm the instantaneous rate at which consumption changes with income. In other words, the income elasticity of demand is approximately equal to the percentage change in consumption that occurs when the subject's income is increased or reduced by 1%, starting at a given income or consumption level. dx/dm corresponds to the slope at a point on an Engel curve and the income elasticity of demand is equal to the slope at a point on a bilogarithmic Engel curve (Fig. 2):

$$\eta_m = \frac{d(\log x)}{d(\log m)}. \quad (5)$$

The demand for a resource is said to be income elastic, inelastic or of unitary income elasticity at a given income or consumption level when the income elasticity of demand for it is greater than 1, less than 1 or equal to 1, respectively. The magnitude of the income elasticity of demand also indicates whether a subject's expenditure on a resource, expressed as a percentage of her income, increases or declines when income is reduced. It declines when demand is income elastic, increases when demand is income inelastic and remains constant when demand is of unitary income elasticity. However, the rate of change in expenditure as a proportion of income cannot be predicted from the magnitude of the income elasticity of demand, since it also depends upon the quantity spent on the resource, as a proportion of income, before the income change:

$$\frac{dS}{dm} = \frac{S}{m}[\eta_m - 1], \quad (6)$$

where $S=px/m$ denotes expenditure as a proportion of income, or the expenditure share, before the income change, dS/dm the instantaneous rate at which the expenditure share changes with income, and η_m the income elasticity of demand. This equation is derived from first principles in the Appendix.

To obtain an analogue of the income elasticity index, for use in choice tests, the rate of change in the budget share must be divided by S/m, thereby rendering it proportional to the income elasticity of demand.

DEFINITION OF THE CONSUMER SURPLUS

The consumer surplus is a measure of the difference between the largest amount of money, or some other currency, that a subject with a fixed income would be

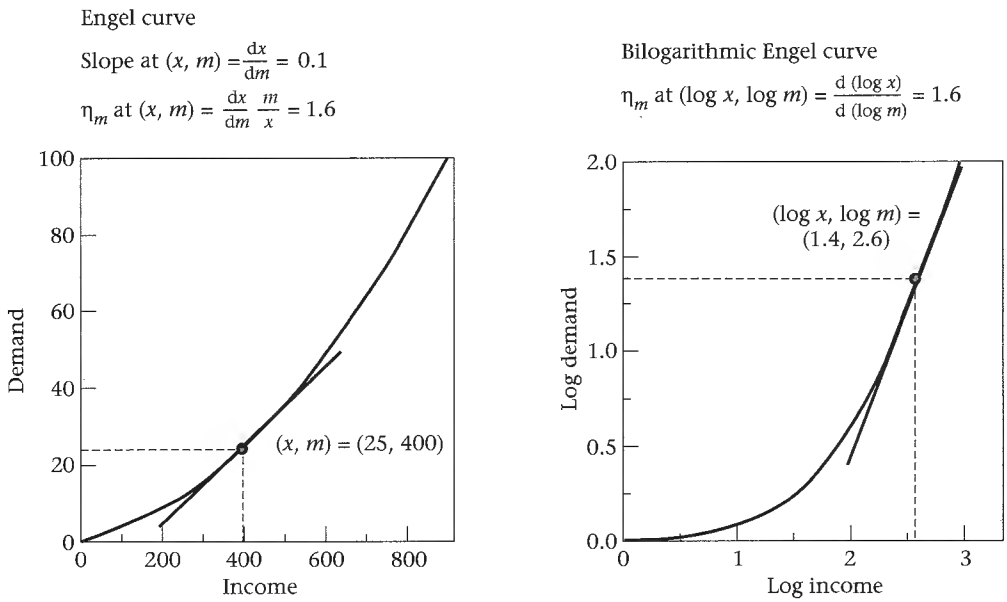

Figure 2. An Engel curve for a hypothetical resource and a bilogarithmic plot of this Engel curve, with the income elasticity of demand, η_m, evaluated at a given point on each curve.

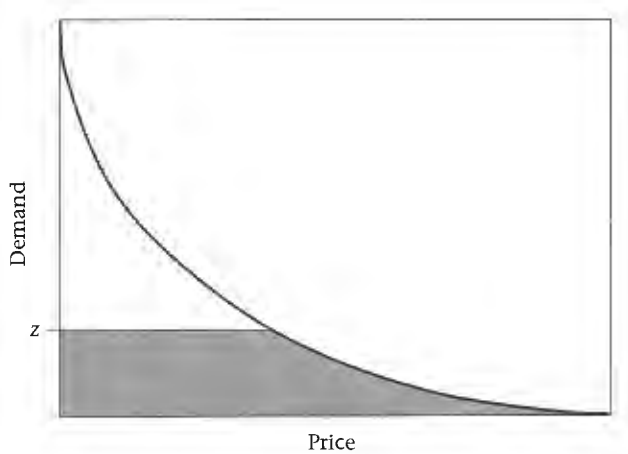

Figure 3. The area under a demand curve that corresponds to the consumer surplus of the quantity *z*.

prepared to spend on a given quantity of a resource and the amount that she actually has to pay (Katz & Rosen 1998, page 110; Mansfield & Yohe 2000, page 106). It is a surplus in the sense that it measures 'left over' willingness to pay. Ethologists want to know about an animal's willingness to pay for a resource when it is made available free, that is, at a price of zero. In this case, the consumer surplus is equal to the largest amount that the subject would be prepared to spend on a given quantity of a resource.

Unlike the elasticity indices, which measure a rate of change of demand, the consumer surplus measures an amount of expenditure. It corresponds to an area beneath the demand curve, not usually the entire area. The area in Fig. 3 has a boundary at a particular level of consumption, *z*. It is a consumer surplus measure of how much the subject is prepared to spend on this quantity of the resource.

It is not hard to understand why an area under the demand curve measures an amount of expenditure. Expenditure is just the product of price and consumption. Any area on a plot of consumption against price will clearly be a measure of expenditure of some kind. However, it may be less clear why consumer surplus areas should measure the largest amount that an animal is prepared to spend on a given quantity of a resource. This can be explained in fairly simple terms. Similar explanations can be found in several textbooks (e.g. Katz & Rosen 1998, page 108; Mansfield & Yohe 2000, page 104).

The key to understanding the consumer surplus lies in the meaning of the inverse demand curve, which is a plot of price against demand, instead of demand against price. This is the way in which economists normally plot the relation between demand and price. The inverse demand curve describes what the price of a resource would have to be for the subject to choose to purchase a given quantity of it. In Fig. 4a, the subject would choose to purchase a quantity, *z*, of the resource when its price was equal to *p*, but would purchase less than *z* if the price was any higher. If the resource was divisible into discrete units (Fig. 4b), then we could infer that *p* is the highest price that the subject is willing to pay for the *z*th unit of the resource. She would pay more for preceding units and less for subsequent ones.

The height of each column in Fig. 4b represents the highest price that will be paid for a particular unit, known as its reservation price (Varian 1996, page 109). The area of each column therefore corresponds to the largest amount that the subject is prepared to spend on a particular unit. It follows that the largest amount she is prepared to spend on all units of the resource up to *z* must be equal to the sum of the column areas up to *z*. This area, shaded in Fig. 4b, corresponds to the consumer surplus of the first *z* units. When the demand curve is continuous, as in Fig. 4a, the consumer surplus area can be calculated

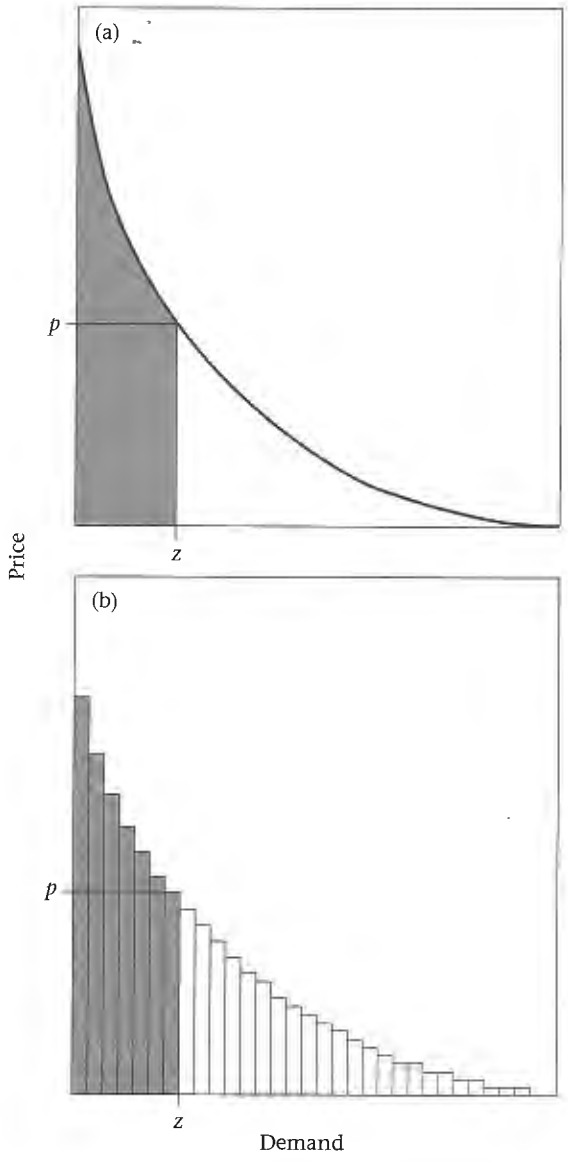

Figure 4. The area under an inverse demand curve that corresponds to the consumer surplus of the quantity z. (a) A continuous resource; (b) a resource that can be divided into discrete units.

by integration of the inverse demand function between the relevant boundaries, the inverse demand function being the equation describing the inverse demand curve. The consumer surplus (CS) of the first z units is equal to the integral of the inverse demand function, $p(x)$, with respect to demand, between the bounds of zero and z:

$$CS=\int_0^z p(x)\mathrm{d}x \qquad (7)$$

DEFINITION OF THE RESERVATION PRICE INDEX

Another important quantity index, used by an increasing number of authors (e.g. Duncan & Kite 1987; Manser et al. 1996; Cooper & Mason 1999, 2000, 2001; Mason et al. 2001; Olsson & Keeling 2002; Olsson et al. 2002), has been called the maximum price paid or reservation

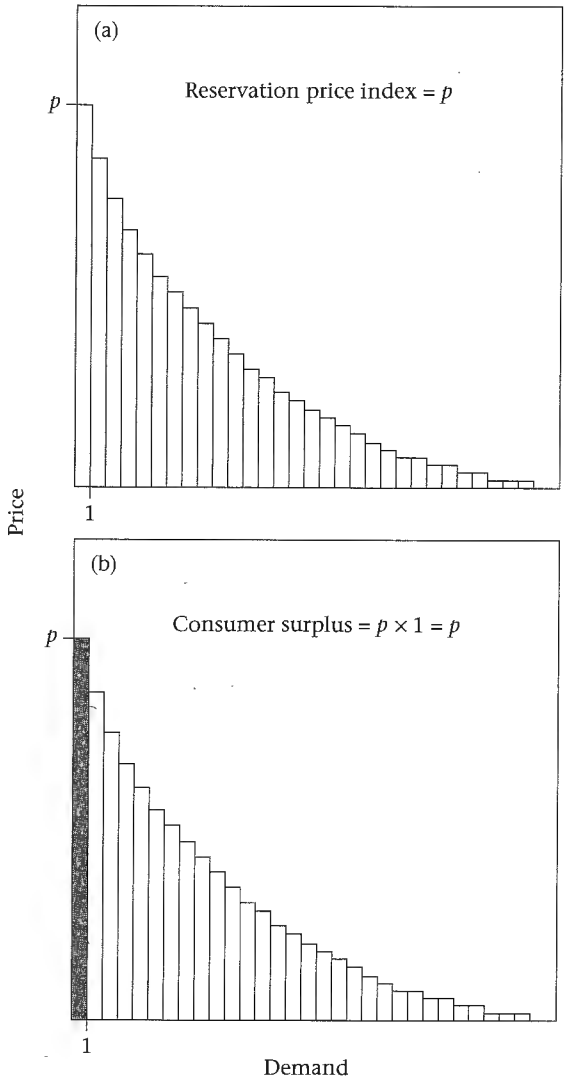

Figure 5. Comparison of (a) the reservation price index (Cooper & Mason 1999; Mason et al. 2001) with (b) the consumer surplus of the first unit of a resource.

price (Cooper & Mason 1999, 2000, 2001; Mason et al. 2001), reservation price being an economic term (Varian 1996, page 109). In this approach, a steadily increasing cost is imposed upon access to a resource and the price at which demand just falls from one to zero is recorded (Fig. 5). This is interpreted as the highest price the subject is prepared to pay for access to the resource.

In economics, this price would be known more precisely as the reservation price of the first unit of the resource, to distinguish it from the reservation prices of subsequent units (Varian 1996, page 109). When the quantity that the subject can consume on each visit is unrestricted, the magnitude of this unit corresponds to the quantity that the subject chooses to consume, on average, when she gains access to the resource only once during the period over which demand is measured. Alternatively, this index might loosely be called the reservation price of the first visit to a resource. Figure 5 shows that when the resource is a discrete good, the

reservation price of the first unit is equal to the consumer surplus of the first unit (Varian 1996, page 245). It is therefore a special case of the consumer surplus index. We discuss its validity alongside that of the consumer surplus.

It is important to distinguish the reservation price index from the break point index (e.g. Hutson & van Mourik 1981; Hutson 1988; Lawrence & Illius 1989; Hutson & Haskell 1990). The procedure for estimating the reservation price imposes a single price in each session, at least when demand is approaching zero, so the subject's motivation to obtain the first unit is evaluated. In contrast, the procedure for estimating the break point (Hodos 1961) imposes a series of prices in each session, with the consequence that the unit for which the maximum price is paid is arbitrary and may differ between subjects and resources. The break point index has no clear economic meaning.

VALIDITY OF THE ELASTICITY OF DEMAND INDICES

Dawkins (1983a, 1988, 1990) regarded the price and income elasticities of demand as interchangeable, assuming, like Houston & McFarland (1980), that demand would respond in a similar way to price rises and income reductions. Her rationale for using the elasticities of demand as indices of motivational strength proceeded as follows. She began by assuming that the mechanism of motivation is partly homeostatic and that animals strive to maintain their consumption of some resources close to preferred levels. Dawkins went on to propose that the more strongly motivated an animal is to consume a resource, the more determined she will be to maintain consumption in the face of increasing access or consumption costs and the less rapidly consumption will decline as price is increased or income is reduced. In other words, the more value the subject attaches to a resource, the less elastic demand for the resource should be with respect to price or income.

When Dawkins (1983a, 1988, 1990) proposed using the elasticities of demand as indices of motivational strength, she did not differentiate elasticities of demand from other rate of change measures, including the slope of the demand curve or Engel curve. As a result, researchers have used a variety of rate of change indices, including: the rate at which demand changes with price (e.g. Cooper & Appleby 1995, 1996, 1997; Matthews et al. 1995; Bubier 1996a; Sherwin 1996, 1998; Sherwin & Nicol 1996; Warburton & Nicol 2001); the rate at which expenditure changes with price (e.g. Bubier 1996b); the rate at which expenditure or the expenditure share change with income (e.g. Dawkins 1983a; Warburton & Nicol 2001); the slope of a bilogarithmic demand curve, which corresponds to the price elasticity of demand (e.g. Matthews & Ladewig 1994; Cooper & Mason 1999, 2000; Sherwin & Nicol 1997; Gunnarsson et al. 2000; Mason et al. 2001); and the slope of a bilogarithmic plot of expenditure against income (e.g. Cooper & Mason 1997a). These indices have generally been regarded as interchangeable, since their underlying rationale is the same.

To obtain a demand or Engel curve, which is the starting point for all of these indices, the cost of access or consumption must be manipulated. This has been done in several ways. In operant tests, the amount the subject is required to pay for access is manipulated by varying the duration, difficulty or unpleasantness of the task. Where completion of the task gives access to a fixed quantity of the resource, these variables constitute measures of the price of the resource. Demand is defined as the number of accesses earned. A demand curve may be obtained from the observed relation between price and demand. In choice tests, the cost of consumption has been manipulated in two ways. The first consists of altering the returns per unit time (e.g. Bubier 1996b), the identity (e.g. Bubier 1996b) or the importance (e.g. Dawkins 1983a, experiment 1) of one or more of the available resources, or the intensity of an unpleasant stimulus to which the animal is subjected while consuming the resource of interest (e.g. Faure & Lagadic 1994). It constitutes a change in the relative price of the resource of interest. The second proceeds instead by varying the time available to the subject (e.g. Dawkins 1983a, experiment 2; Cooper & Mason 1997a; Warburton & Nicol 2001). The less time that is available, the less will remain for consumption of the alternative resource after allocating a given amount of time to the resource of interest, so the more costly this amount of consumption time will be. In economic terms, the time available to the animal is a measure of her income.

Many choice tests yield estimates of expenditure on the resource of interest, defined as the time spent interacting with it. Analogues of the price and income elasticities of demand can be derived from the observed relation between price and expenditure, or between income and expenditure. However, in some circumstances, where it is possible to distinguish demand from expenditure, the price elasticity of demand may be measured directly; for example, when the return per unit time is manipulated, or when the subject is exposed to an unpleasant stimulus while consuming the resource of interest.

We now describe four shortcomings of the elasticity of demand indices. They apply equally to allied rate of change measures, such as the slope of the demand curve or Engel curve. The first arises from an auxiliary assumption that ethologists make when they estimate demand elasticities, and it compromises the external validity of these indices. The second and third shortcomings are intrinsic to the indices themselves and compromise their internal validity. The fourth problem impairs the external validity of the indices. It also compromises the internal validity of the price elasticity index, but not that of the income elasticity index.

Assigning a Single Value to a Resource

When ethologists use the price or income elasticities of demand as indices of motivational strength, they invariably assume that a single elasticity value can be assigned to each resource. This permits them to argue that one resource is more important to an animal than another. Some ethologists have calculated what amounts to an

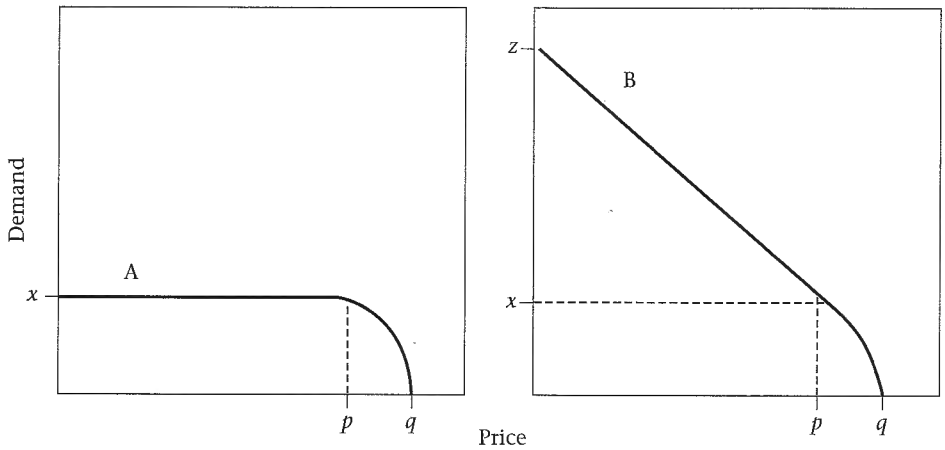

Figure 6. Demand curves for two hypothetical resources, A and B. See text for explanation.

average price elasticity value by fitting a straight line to the bilogarithmic plot of demand against price (e.g. Matthews & Ladewig 1994; Sherwin & Nicol 1997; Gunnarsson et al. 2000). A similar approach has been taken by some ethologists using the slope of the demand curve as a rate of change index (e.g. Matthews et al. 1995; Sherwin 1996). Other researchers have simply ascertained, using a statistical test, whether demand or expenditure declines significantly over the range of price or income levels used (e.g. Dawkins 1983a; Cooper & Appleby 1995, 1996, 1997; Sherwin & Nicol 1996; Sherwin 1998; Warburton & Nicol 2001), or have restricted their observations to the effect of a single price or income change (e.g. Bubier 1996a, b).

The assumption is reasonable only if elasticities of demand are more or less constant across all consumption levels. Preference tests conducted by ethologists have rarely used a sufficient number of price or income changes over a wide enough range to ascertain whether demand elasticities are in fact constant. However, where this has been done, the price elasticity of demand appears to increase with price (Gunnarsson et al. 2000). Studies conducted by experimental psychologists have had similar findings (Hogan & Roper 1978). Furthermore, an experimental study by Findley (1962), in which analogues of Engel curves were generated (Lea 1978), revealed that the income elasticities of demand for several resources increased with income. In such cases, when single elasticity values are assigned to resources whose elasticities of demand are not constant, these values will at best be inaccurate as indices of motivational strength.

Furthermore, it is not meaningful to assign a single value to a given commodity. This is because, in general, successive units of a resource are not worth the same as one another. As a rule, the more of a resource that a subject has already consumed, the less an additional unit is worth to her. From this point of view, it makes little sense to ask how much a resource is worth. Instead, we should ask how much a particular quantity of a resource is worth. What the price and income elasticity indices purport to measure has little meaning in the real world and may be of limited practical value. The assumption that a single elasticity value can be assigned to each resource therefore compromises the external validity of these indices.

Defence of Initial Consumption Level

When a subject has a low price or income elasticity of demand for a resource, this has been said to indicate a readiness to 'defend' some initial consumption level against increasing access or consumption costs (Hursh 1984), or to 'compensate' for the increased costs by working more (Hogan & Roper 1978). This interpretation of demand elasticities is central to Dawkins' (1983a, 1988, 1990) rationale, but it is not the only interpretation. A low price or income elasticity of demand might just as easily reflect a tendency to become rapidly satiated with the resource.

Figure 6 depicts demand curves for two hypothetical resources, A and B. The quantities of these resources that are consumed when their prices at zero are x and z, respectively. As price is increased to p, demand for resource A remains constant, but demand for resource B declines. In this region of the demand curve, the price elasticity of demand for resource A is equal to zero. The price elasticity of demand for resource B is equal to zero when its price is zero, but increases steadily as its price is increased. Beyond a price of p, the demand curves for the two resources are approximately the same.

Where the price elasticity of demand is seen as reflecting the subject's readiness to defend an initial consumption level, the demand curves would be interpreted as follows. A price elasticity of demand for resource A of zero would indicate that the subject was highly motivated to maintain her consumption of this resource at x. The higher average price elasticity of demand for resource B would show that the subject was less prepared to defend her initial consumption level of this resource against price rises. Below a price of p, consumption of resource B would be said to be less important to the subject than consumption of resource A. The same conclusion would be drawn if price elasticity was averaged across all price levels.

However, if the price elasticity of demand is seen instead as reflecting the subject's tendency to become

satiated with the resources, then the demand curves would be interpreted differently. The clarity of this interpretation is assisted by considering how a subject would respond to price reductions, rather than to price rises. Starting at the highest price, q, when price is reduced, demand for both resources increases at the same rate until it reaches x. At this point, the subject is satiated for resource A and is not interested in purchasing more than this quantity at any price. The price elasticity of demand for resource A beyond this point is therefore equal to zero. However, the subject is not satiated by x units of resource B and is prepared to purchase more of it as its price declines further. The price elasticity of demand for resource B beyond this point is therefore greater than zero. The principal difference between the two resources, according to this interpretation, is the rate at which the subject becomes satiated with them. There is nothing to suggest that resource A is more important to the subject than resource B.

A pair of resources that might show this sort of relationship in a human economy would be bread and cake. In terms of hedonic value and nutritional benefit, a small amount of cake is no less valuable for most people than the same amount of bread. However, because cake is sweet, satiation generally occurs more rapidly for cake than it does for bread.

The satiation interpretation is readily extended to the general case where two demand curves differ in average elasticity. There is no need for complete satiation to occur for one of the goods at a price above zero, as is the case for resource A. Successive units of most goods are of diminishing value and resources differ in the rate at which the value of successive units declines, i.e. in the rate at which the subject approaches satiation for them. The more rapidly this occurs, the more that price must decline before an additional unit is purchased, and the lower the price elasticity of demand will be. This means that the satiation interpretation is a competing explanation in all cases where resources differ in their average price elasticities of demand, not merely in cases where one resource has a price elasticity of zero.

It is likely that a tendency to defend a preferred consumption level and a tendency to become satiated are jointly responsible for the price elasticity of many demand curves. Attempts to infer a subject's readiness to defend a consumption level against price rises from the price elasticity of demand will therefore be confounded in many cases by the effects of satiation. The index will tend to overestimate the value of resources for which satiation occurs rapidly.

The income elasticity of demand can also be interpreted in terms of the rate at which the subject becomes satiated with a resource. The principle is the same: as income is increased, more units become affordable, and the rate at which demand increases depends upon the rate at which the value of successive units declines. Hence, attempts to infer motivational strength from the income elasticity of demand will be confounded in many cases by the effects of satiation. Like the price elasticity of demand, it will tend to overestimate the value of resources for which satiation occurs rapidly.

Amount Paid

A common principle underpins Dawkins' rationale for the elasticity of demand indices and the economic rationale for the consumer surplus index. They both maintain that the value that a subject attaches to a resource, or a specific quantity of a resource, is revealed by her readiness to pay for it. The difference between the two rationales is how the subject's readiness to pay is assessed. In Dawkins' rationale, readiness is inferred from the rate at which demand changes with price or income, but in the economic rationale the actual quantity that the subject is willing to pay is estimated using the consumer surplus.

We argue here that readiness to pay cannot generally be inferred from the price or income elasticities of demand, because the subject's decision to maintain consumption at a given level is influenced not only by her motivation, but also by how much it would cost her to do so. Hence, when two resources have equal price or income elasticities of demand, this does not generally mean that the subject is equally willing to pay for them. We show below that when the elasticities of demand are used as indices of motivational strength, using the rule that 'equal elasticities means equal motivational strengths', they fail to take into account the amount that the subject is required to pay when she maintains consumption at some percentage of initial demand. There is no way that readiness to pay can be inferred from the rate at which demand changes with price or income, without taking this into account.

The amount that the subject is required to pay to maintain consumption at a given percentage of initial demand, or the 'costliness' of doing so, refers ultimately to the quantity of other goods that this renders unaffordable, although in practice, it is more easily measured as a reduction in the amount of money, time or energy that remains to be spent on other goods. An animal in an experimental economy would probably be unaware of what was affordable, but her consumption decisions would be influenced by the costliness of her previous choices. Thus, once demand had stabilized, the magnitude of the price or income elasticity of demand would be influenced by the costliness of a set of consumption possibilities, including that of the chosen consumption level.

The costliness of 'maintaining' consumption at a given level, x^1, which may be less than or equal to the initial level of consumption, x^0, is defined as follows. When price changes from p^0 to p^1, but income remains constant, costliness is simply defined as the difference between expenditure on the resource of interest before the price change and what expenditure would be after the change: $p^1x^1 - p^0x^0$. When income changes, it is necessary to take into account the change in what the subject can afford, as well as the change in expenditure. This can be done by expressing expenditure as a percentage of income, to give a measure of the proportion of income that is allocated to the resource of interest, known in economics as the budget share. Hence, costliness is defined as the difference between the budget share before the income change and what the budget share would be

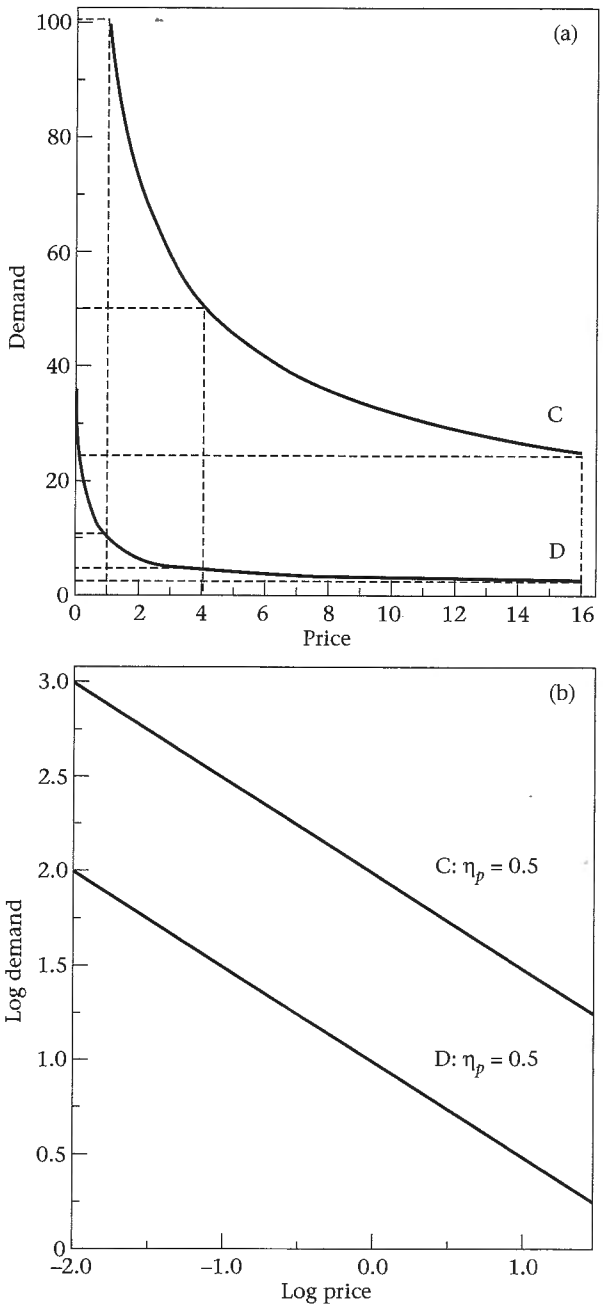

Figure 7. (a) Demand curves and (b) bilogarithmic demand curves for two hypothetical resources, C and D, with equal price elasticities, η_p.

after the change: $p^1x^1/m^1 - p^0x^0/m^0$. The simpler of these two definitions of costliness, in which costliness is expressed as a difference between expenditures, is in fact a special case of the budget share definition, in which $m^1=m^0=1$.

Price elasticity of demand

The argument against the price elasticity of demand is as follows. Figure 7 depicts demand curves for two hypothetical resources, C and D. The bilogarithmic plots of these curves show that both have the same price elasticity of 0.5. The price elasticity index therefore indicates that the resources are of equal value. The principal difference between the two demand curves is that, at a given price, a larger quantity of resource C is demanded than of resource D. What a price elasticity of 0.5 tells us is that if price is increased fourfold, then demand will be reduced by half. Thus, when price is increased from 1 to 4 currency units, demand for resource C declines from 100 to 50, and demand for resource D declines from 10 to 5.

Let us say that our consumer is human and that the initial price of both resources is £1. At this price, demand for resource C is 100 and demand for resource D is 10. It follows that our subject's expenditure on resource C at this price will be £100, but expenditure on resource D will be £10.

Now we increase the price. Since the price elasticity is 0.5, increasing the price 16-fold reduces demand for both resources to one-fourth of its initial level. At a price of £16, demand for resource C is 25, at a cost of £400, and demand for resource D is 2.5, at a cost of £40.

This means that, when we increase the price from £1 to £16, it costs the consumer an additional £300 to maintain her consumption of resource C at 25% of its initial level, but only an additional £30 to maintain consumption of resource D at 25% of its initial level. If her income was, say, £500, then a substantial sacrifice would be required, in terms of lost consumption of other goods, to maintain her consumption of resource C at this level, but only a small sacrifice for resource D.

The price elasticity index fails to account for this difference in costliness. This is because the rule that equal demand elasticities means equal motivational strengths is simplistic. It takes no account of how large the subject's demand is at the outset and therefore no account of how much she would have to spend to maintain her consumption in the face of price rises. The only circumstance under which the price elasticity would be a valid comparator of motivational strengths is where the initial consumption levels of two resources were equal. Under all other circumstances, the price elasticity index will tend to underestimate the relative value of resources whose initial consumption levels are high, because the costliness of maintaining consumption of these resources will also be high.

A pair of resources that might present this sort of problem in a human economy would be bread and salt. Because salt is consumed in much smaller quantities than bread, it would be much less costly to maintain consumption of salt, in the face of a given percentage price rise, than it would be to maintain consumption of bread. Hence, demand for salt would probably be considerably more inelastic than demand for bread, not because salt is more important to people than bread, but because it is consumed in smaller quantities when cheap.

Income elasticity of demand

The argument against the income elasticity of demand is similar. Figure 8 depicts Engel curves for two hypothetical resources, E and F, and bilogarithmic plots of these curves. Both resources have the same income elasticity of

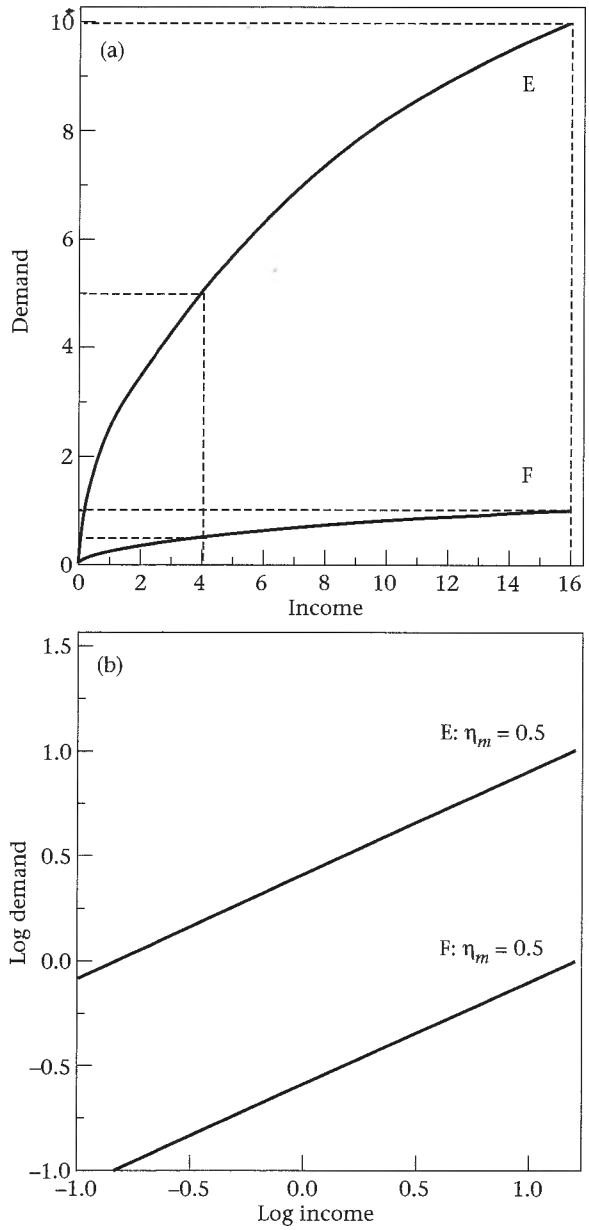

Figure 8. (a) Engel curves and (b) bilogarithmic Engel curves for two hypothetical resources, E and F, with equal income elasticities, η_m.

0.5, so the income elasticity index indicates that they are of equal value. The main difference between the two Engel curves is that at a given income level, a larger quantity of resource E is demanded than of resource F. An income elasticity of 0.5 means that a fourfold reduction in income causes demand to be reduced by half. Thus, when income is reduced from 16 to 4 currency units, demand for resource E declines from 10 to 5 and demand for resource F declines from 1 to 0.5.

Because income is not constant, we cannot simply use the changes in expenditure on resources E and F as measures of costliness. The change in what the subject can afford must also be taken into account. This is done by expressing expenditure as a percentage of income.

Let us say that the prices of resources E and F are constant at a level of £0.50. When income is £16, demand for resource E is equal to 10, so expenditure on this resource will be £5. This constitutes 31.25% of the subject's income. Meanwhile, demand for resource F at this income level is equal to 1, meaning that expenditure on this resource will be £0.50, comprising 3.125% of income. Now, if income is reduced to £4, the new level of expenditure on resource E will be £2.50, constituting 62.5% of the subject's reduced income, and the new level of expenditure on resource F will be £0.25, comprising 6.25% of income. Hence, when income is reduced from £16 to £4, the consumer must give up an additional 31.25% of her income to maintain her consumption of resource F at 50% of its initial level, but only an additional 3.125% of her income to maintain consumption of resource E at 50% of its initial level.

The income elasticity index fails to account for this difference in costliness. The rule that equal demand elasticities means equal motivational strengths ignores the fact that the subject would be required to pay more to maintain consumption of the resource with the higher initial consumption level at a given percentage of initial demand. This index will be a valid comparator of motivational strengths only in the special case where the initial consumption levels of two resources are equal. Under all other circumstances, it will tend to underestimate the relative value of resources whose initial consumption levels are high.

Effect of Income on Demand

The price elasticity of demand is at best an approximate index of the value that animals attach to commodities, because it does not account for the effect of the subject's income on her purchasing decisions. The problem arises from the fact that it is obtained from a demand curve, which describes the relation between price and consumption when income is held constant. Because the quantity that is demanded at a given price depends upon the subject's income, her income level will influence the height of the demand curve at that price. Furthermore, unless income exerts the same effect upon demand at all price levels, which is unlikely, her income will influence the slope of the demand curve and hence the price elasticity of demand.

The effect of income on demand has two implications for the price elasticity of demand and other indices derived from demand curves. The first may compromise their external validity, unless suitable precautions are taken when designing preference tests (see Discussion). The second will compromise their internal validity.

The first implication is obvious: the price elasticity of demand is specific to the income level at which the demand curve was generated. For example, it might estimate how much a subject with a time budget of 8 h per day was prepared to pay, but not how much a subject with 4 h per day would pay. How well this index can be generalized to other income levels depends upon the sensitivity of the subject's consumption of the resource in question to income, i.e. upon the income elasticity of her demand for it.

The second implication is less obvious but potentially more serious. It follows from the fact that any change in the price of a resource, while holding income and the prices of other resources constant, changes the purchasing power of the subject's income, defined as the highest level of utility or welfare that it allows her to attain. For example, a price reduction will permit a subject to purchase the same bundle of commodities as she did before the reduction, while having some money left over. This means that she can afford to buy more with the same income and hence reach a higher level of welfare. Insofar as the subject's income has no intrinsic value, but is worth only as much as what can be purchased with it, an increase in purchasing power will have exactly the same effect upon consumption as an increase in income. The subject's real income is said to have been increased by the price reduction. This means that every time the price of a resource is changed during the generation of a demand curve, the subject's real income will also change. The effect that this has upon consumption is known as the income effect of the price change (Varian 1996, page 140). It is as if every point on the demand curve was estimated at a different income level.

The influence of the income effect of a price change upon the price elasticity of demand is relatively straightforward. In the case of normal goods, for which demand increases with income (Varian 1996, page 96), the reduction in real income that results from a price rise will cause demand to decline, increasing the slope of the demand curve and thereby increasing the price elasticity of demand. The price elasticity index will therefore tend to underestimate the value of such resources. The converse will be true in the case of inferior goods, for which an increase in income causes a reduction in demand.

In practice, the resources used in operant tests and most choice tests will be normal goods. Inferior goods are low-quality commodities, which the consumer consumes less of as her income increases because higher-quality substitutes, which satisfy the same motivation, become affordable. Because most preference tests designed to measure motivational strength set out to compare the strength of different motivations, substitutes are not usually made available.

The magnitude of the income effect of a price change, Δx^m, is determined by two factors: the money value, or other currency value, of the change in real income, Δm; and the rate at which demand changes with income, $\Delta x/\Delta m$. The currency value of the change in real income depends in turn upon the size of the price change, Δp, and the magnitude of consumption before the price change, x. The greater the change in the price of a resource and the more of it that was being purchased before its price changed, the more the price change will affect the quantity of all goods, and hence the level of welfare, that is affordable. The relation between these factors is defined in the following equation (Varian 1996, page 145).

$$\Delta x^m = \frac{\Delta x}{\Delta m}\Delta m \approx \frac{\Delta x}{\Delta m} x \Delta p. \qquad (8)$$

The influence of the income effect upon the price elasticity of demand will also depend on how large the currency value of the change in real income, Δm, is in relation to the subject's income, m. If the change in real income is comparatively small, it will have little effect upon how the subject allocates her income amongst resources, and hence upon demand.

Because the price elasticity index attempts to describe the entire demand curve, the relevant price change, Δp, is from zero to the price at which demand falls to zero. This price change will often be large and will vary a great deal from one commodity to another. This means that the price elasticity index may substantially underestimate the value of some normal goods, especially those for which a subject is prepared to pay a high price. Furthermore, the influence of the subject's initial consumption level, x, upon the income effect means that the price elasticity index will tend to underestimate the value of normal goods that are consumed in large quantities when they are cheap. Finally, the influence of the rate at which demand changes with income, $\Delta x/\Delta m$, means that the more sensitive demand for a normal good is to income, the more that the price elasticity index will underestimate its value.

Because the magnitude of the income effect of a price change varies from one resource to another, the error associated with the price elasticity index will also vary, with the consequence that the values obtained for a series of resources will not necessarily have an ordinal relation to their true values. In other words, the failure of the price elasticity index to account for the income effects of price changes will compromise its internal validity.

There is a superficial similarity between the income effect of a price change and the effect of the costliness of maintaining consumption upon the price elasticity of demand. Both phenomena cause this index to underestimate the value of resources (specifically normal goods, in the case of the income effect) whose initial consumption levels are high. However, the two phenomena are independent, because negating the influence of one does not negate the influence of the other.

(1) Equalizing the costliness of maintaining consumption of two resources is not sufficient to resolve the problem associated with income effects. For a given price change, the costliness of maintaining consumption of two resources at a given percentage of initial demand, $p^1x^1 - p^0x^0$, can be equalized by ensuring that both goods have the same initial consumption level, x^0. Ensuring that initial consumption level is the same for both goods will simultaneously render the change in purchasing power, $x^0\Delta p$, which accompanies the price change, equal for the two resources. However, this is not sufficient to avoid income effects, or even to render them equal for the two goods, because the income effect also depends on the rate at which demand changes with income, $\Delta x/\Delta m$. Thus, a given change in purchasing power influences the price elasticity of demand differently for different goods, when demand for the goods responds differently to income changes. It follows that the sensitivity of the price elasticity of demand to changes in real income compromises its ability to make comparisons between goods,

even when there is no difference in the costliness of maintaining consumption at a given percentage of initial demand.

(2) Holding purchasing power constant is not sufficient to equalize the costliness of maintaining consumption of two resources. When the price of a resource goes up, the problem associated with income effects can be negated by increasing the subject's income sufficiently to hold purchasing power constant. Both the change in purchasing power and the income effect of the price change will then be equal to zero. For example, consider resources C and D in Fig. 7. When the price goes up from £1 to £16, purchasing power would fall by £1500 for resource C ($x^0\Delta p=100 \times £15$) and by £150 for resource D ($x^0\Delta p=10 \times £15$). If the subject's income before the price change was £500, and she was compensated sufficiently to hold purchasing power constant, then her income after the price change would be £2000 for resource C and £650 for resource D. The costliness of maintaining consumption at 25% of initial demand after the price change is given by the formula, $p^1x^1/m^1 - p^0x^0/m^0$. It would cost the subject an additional $20 - 20\%=0\%$ of her income to maintain consumption of resource C at 25% of initial demand, when compensated for the price change, but only an additional $6.2 - 2\%=4.2\%$ of her income to maintain consumption of resource D at 25% of initial demand. We conclude that holding purchasing power constant does not equalize the costliness of maintaining two resources at a given percentage of initial demand and therefore fails to resolve the problem that differential costliness causes for the price elasticity index.

Effect of Price on Demand

The income elasticity of demand index clearly does not fail to take into account the effect of income upon demand. However, because prices are held constant during the generation of an Engel curve, this index will be unable to account for the effect of price upon demand. Its magnitude will be specific to the price at which the Engel curve was generated.

There is no analogue of the income effect of a price change that might compromise the accuracy and validity of the income elasticity of demand. Income changes are much more straightforward than price changes. They alter what the subject can afford, but have no effect upon the relative prices of commodities. An income change does not cause a change in purchasing power, rather it is a change in purchasing power. Because the income elasticity index is not susceptible to an analogue of the income effect of a price change, it is more internally valid than the price elasticity and consumer surplus indices in this respect.

VALIDITY OF THE CONSUMER SURPLUS INDEX

The economic rationale for using the consumer surplus as a measure of the value a subject attaches to a given quantity of a resource is straightforward. Consumer demand theory makes no assumptions about the mechanism of motivation. It simply postulates that the more value a subject attaches to a given quantity of a resource, the more she will be willing to pay for it. The consumer surplus is a direct estimate of this amount.

Only two studies have attempted to estimate the consumer surplus index (Kirkden 2000; Mason et al. 2001). Because the consumer surplus, like the price elasticity of demand, is obtained from a demand curve, the requirements for experimental design are similar. Any experiment that yields a demand curve can be used to estimate both indices. The only factor limiting the range of experimental approaches that can be used to obtain the consumer surplus is that it cannot be derived from a plot of expenditure versus price.

We now discuss the four shortcomings of the elasticity of demand indices in relation to the consumer surplus index. The first three problems do not affect the consumer surplus. However, the fourth problem is expected to compromise its external and internal validity.

Assigning a Single Value to a Resource

When economists use the consumer surplus index, they do not attempt to assign a single value to a given commodity. Instead, they evaluate specific quantities, corresponding to particular areas beneath the inverse demand curve. In this respect, the consumer surplus index has greater external validity than the elasticity of demand indices.

The fact that the consumer surplus index is designed to evaluate specific quantities of resources also renders it substantially more versatile than the price and income elasticity indices. Applied ethologists have used indices of motivational strength to perform two types of ranking operation: establishing that one resource is worth more than another, and identifying an alternative resource that is worth the same as a resource of interest. In the latter operation, the value of the alternative is typically manipulated by varying the subject's deprivation level until it is the same as the value of the resource of interest (e.g. Dawkins 1983a; Duncan & Kite 1987). The corresponding ranking operations that the consumer surplus index permits are establishing that a given quantity of one resource is worth more than a given quantity of another, and identifying the quantity of an alternative that is worth the same as a given quantity of a particular resource.

Ranking quantities of resources differently

The ranking of resources, or of specific quantities of resources, is a straightforward procedure for any index of motivational strength, but because the consumer surplus evaluates specific quantities of resources, it extracts much more information from demand curves than the elasticity of demand indices.

For example, the consumer surplus can be used to compare the value of the first unit of one resource with the first unit of another (e.g. Kirkden 2000). This gives an idea of the relative importance of having the two resources available at all. Thus, in Fig. 9, the consumer

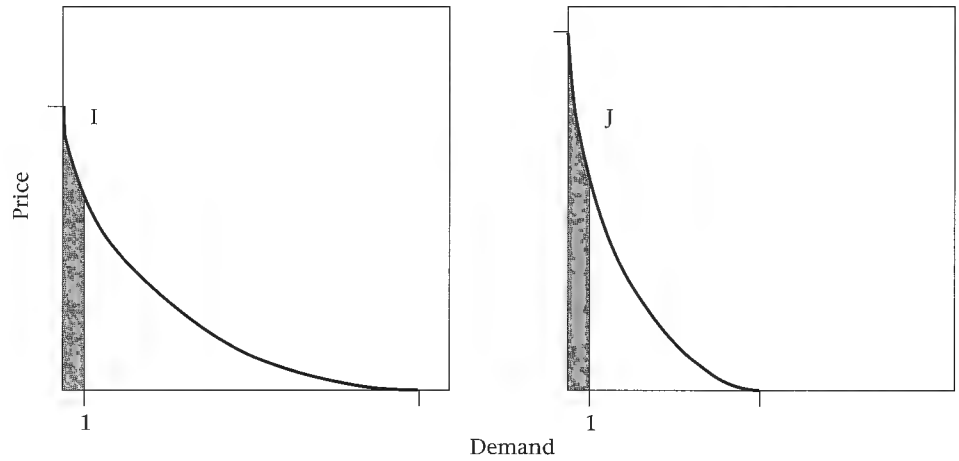

Figure 9. Consumer surplus estimates of the value of the first unit of each of two hypothetical resources, I and J.

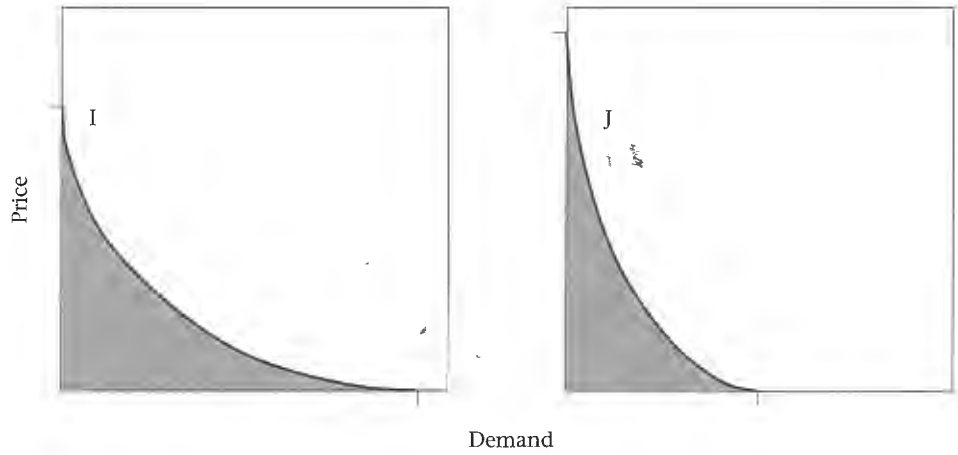

Figure 10. Consumer surplus estimates of the value of the quantity of each of two hypothetical resources, I and J, that would be required for satiation.

surplus areas reveal that the first unit of resource J is worth more to the subject than the first unit of resource I. A similar comparison could be carried out using the reservation price index (e.g. Kirkden 2000; Mason et al. 2001).

The consumer surplus can also be used to compare the quantity of one resource that a subject requires to reach satiation with the quantity of another that she requires to reach satiation (e.g. Kirkden 2000; Mason et al. 2001). This gives an idea of the relative importance of having the two resources available ad libitum. Figure 10 shows the relevant consumer surplus areas for resources I and J, and shows that a satiating quantity of resource I is worth more to the subject than a satiating quantity of resource J.

Where the two resources are substitutes, i.e. where they both satisfy the same motivation, such comparisons could provide the basis for an economic cost–benefit analysis of alternative ways to satisfy the motivation. Say that resources I and J cost a farmer the same amount to procure, and that she could afford to provide her animals with a single unit of either one resource or the other each day. Figure 9 indicates that the animals would receive a greater benefit from one unit of resource J than from one unit of resource I, so resource J is the one that the farmer should purchase. If, on the other hand, resource I was only half the price of resource J, so that the farmer's choice was between two units of I per day and one of J, then resource I would be the better option. If money were no object and the farmer intended to provide one of the resources ad libitum, resource I would again be preferable to resource J (Fig. 10). We add a qualification concerning the use of the word benefit here. Strictly, what is being estimated is value, not benefit. Value refers to a subject's motivation to obtain a resource, but benefit refers to the consequences of obtaining it. The two do not always agree (Duncan 1978, 1992; van Rooijen 1982). Because value and benefit are not the same, motivational strength should never be the sole basis for making decisions about the provision of resources (Dawkins 1988, 1990). Nevertheless, the consumer surplus index could offer unique guidance to the farmer, zookeeper or laboratory technician who has a limited amount of money to spend on providing for an animal's needs and must decide what to purchase.

The elasticity of demand indices would not support a cost–benefit analysis of this type, because they do not

evaluate specific quantities of resources. A cost–benefit analysis cannot proceed without a measure of benefit that is linked to the quantity consumed.

Ranking quantities of resources equally

The elasticity of demand indices are poorly suited to identifying an alternative resource that is worth the same as a resource of interest. A systematic approach to this ranking operation has only been taken using much simpler indices, such as the reservation price of a single visit to a resource (Duncan & Kite 1987) or the quantity consumed (Dawkins 1983a). In both cases, the procedure involved estimating the relative value of the alternative at a series of different deprivation levels until the index equally ranked the alternative and the resource of interest. Essentially a process of trial and error, this was only practical because the indices were quick and easy to estimate. It would not be feasible to use the elasticity of demand indices in this way, because an entire demand or Engel curve would have to be generated for the alternative resource at every deprivation level, in search of one with the same average elasticity as the resource of interest.

The consumer surplus index is better suited to this type of ranking operation than the elasticities of demand. Although it is still necessary to generate a demand curve for the alternative resource and for the resource of interest before any comparisons can be made, there is no need to repeat this procedure at different deprivation levels. This is because the demand curve yields consumer surplus estimates for numerous quantities of the alternative, giving us a different way to manipulate its value: we can identify the quantity of the alternative that is worth the same as a given quantity of the resource of interest (e.g. Kirkden 2000). This involves finding the consumption level at which the area beneath the demand curve for the alternative is equal to the relevant area beneath the demand curve for the resource of interest. This approach is no less informative and has two advantages over the more usual ethological approach. First, it replaces a vague notion of the value of a resource with a more precise one of the value of a particular quantity of a resource. Second, it does away with trial and error.

An estimate of the value of a given quantity of a resource, in terms of an equivalent quantity of a comparator, such as food, could be very useful in applied ethology. It is a means of quantifying motivational strength, relative to something that is known to be important to animals.

Defence of an Initial Consumption Level

The consumer surplus index, in contrast to the elasticities of demand, is unambiguous. The subject is not assumed to be defending a particular consumption level, only paying for it. In Fig. 6, the consumer surplus index indicates that the subject attaches approximately the same value to x units of resource A as she does to x units of resource B. Furthermore, it indicates that she attaches an additional, diminishing value to subsequent units of resource B, up to the quantity z, but no value to subsequent units of resource A.

Amount Paid

The validity of the consumer surplus index is not compromised by differences in the costliness of consuming different resources. It does not need to take into account the amount a subject is required to pay. This is because it is a direct estimate of willingness to pay. The price and income elasticities of demand are measures of behaviour, not of motivation. Willingness to pay is inferred from these measures, not estimated by them. Because the elasticities of demand measure rates of change in demand, it is necessary that they take into account the costliness of maintaining demand if they are to tell us anything about willingness to pay. The consumer surplus index, in contrast, does not measure anything that is observable. It is a measure of motivation. It estimates the largest amount of time, effort or an alternative resource that the subject is willing to pay for a specific quantity of a resource. The amount that the subject is required to pay is simply not relevant.

Compared to the price elasticity of demand, the consumer surplus represents a very different approach to dealing with the problem of resources having different consumption levels when prices are low. By inverting the demand curve, information about the magnitude of consumption at a given price is transformed into information about the price at which the subject would choose to purchase a given quantity, from which it is possible to estimate how much she is willing to pay for a specific quantity. Thus, a confounding factor is incorporated into the index in a way that not only solves the problem, but also increases the external validity of the index.

Effect of Income on Demand

Economists recognize that the consumer surplus is only an approximate index of the value that people attach to quantities of commodities, because it fails to account for the effect of a consumer's income upon her purchasing decisions. This is a consequence of the fact that, like the price elasticity index, it is obtained from a demand curve. Because the subject's income influences the price at which she would choose to purchase a given quantity of a resource, it also affects the height of the inverse demand curve and hence the magnitude of the consumer surplus.

The first implication of this shortcoming is that the consumer surplus is specific to the income level at which the demand curve was generated. This may compromise its external validity.

The second implication is that the accuracy and internal validity of the consumer surplus index will be compromised by the income effects of price changes. The influence of income effects upon the consumer surplus is less straightforward than is the case with the price elasticity index, because the consumer surplus measures an amount that the consumer is prepared to spend, not a change in demand.

Randall & Stoll (1980) identified three factors that determine the accuracy of the consumer surplus, via an analogue of the income effect applicable to inverse demand curves. The first is the price flexibility of income, ζ, an elasticity measure defined as the proportional rate at which the price a consumer is prepared to pay for a given quantity of a resource changes with income. It is equal to the ratio of the income elasticity of demand and the elasticity of substitution, σ, between the resource of interest and other goods (Hanemann 1991)

$$\zeta = \frac{\eta_m}{\sigma}. \tag{9}$$

The elasticity of substitution is a measure of the consumer's tendency to exchange the resource of interest for other goods when there is a change in their relative price. When a substitute for the resource of interest is not available in the environment, the elasticity of substitution between this good and others will be low and the consumer surplus index will therefore tend to be inaccurate.

The second factor is the value of the quantity of the resource that is being assessed. Where the subject attaches a high value to a given quantity of a resource, a price rise that renders this quantity unaffordable will reduce her welfare substantially and will consequently result in a considerable reduction in the purchasing power of her income. The third factor is the magnitude of the subject's income, m. When income is high, the currency value of a given change in real income will be comparatively small and will have little effect upon how the subject allocates her income among resources. The relation between these factors and the accuracy of the consumer surplus index is defined by the following equation (Randall & Stoll 1980).

$$\varepsilon \approx \frac{\zeta |CS|}{2m}, \tag{10}$$

where ε denotes the percentage error of the consumer surplus estimate, defined as its percentage difference from the equivalent surplus or compensating surplus (see below), and $|CS|$ denotes the absolute value of the consumer surplus, which is approximately equal to the value of the relevant quantity of the resource.

Because several of the factors influencing the accuracy of the consumer surplus vary in magnitude from one resource to another, the values obtained for a series of quantities of resources may not have an ordinal relation to their true values, and the internal validity of the index will be compromised by its inaccuracy. The elasticity of substitution is not one of the factors expected to vary in magnitude between resources, in the context of most preference tests. Because it is normal practice to ensure that there are no substitutes in the experimental environment, the elasticity of substitution will be of a similar, low magnitude for all resources evaluated.

Whether the consumer surplus index tends to underestimate or overestimate the value of a given quantity of a resource depends partly upon whether the resource is a normal or an inferior good. However, it also depends upon the real income or utility level that one chooses as

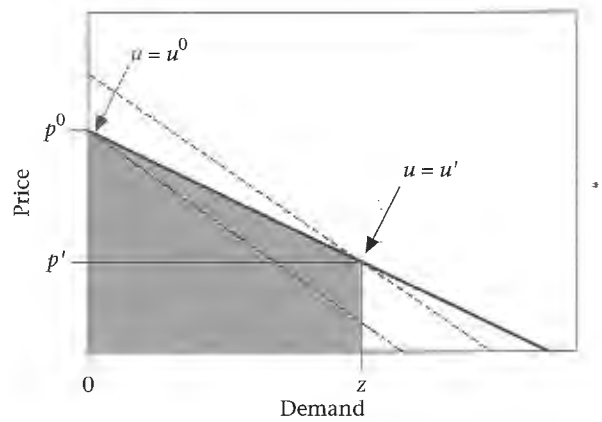

Figure 11. The relation between the inverse demand curve observed for a hypothetical, normal good, where real income increases from $u=u^0$ to $u=u'$ as price is reduced, and the inverse demand curves that would be expected if real income were held constant, either at u^0 (– · · – · · –) or u' (– – –). The consumer surplus for the first z units (shaded), obtained from the observed inverse demand curve, is greater (less) than the consumer surplus that would have been obtained if real income were constant at u^0 (u').

one's point of reference. There are several reference points one might select. Referring to the inverse demand curve (solid line) in Fig. 11, one possibility is the level of real income, u^0, which obtains when price is equal to p^0 and consumption is equal to zero. Another is the level, u', which obtains when price is equal to p' and consumption is equal to z. The broken lines in the figure represent the inverse demand curves that would be expected if real income was constant at u^0 or u'.

The broken-line curves are mathematical constructs, known as inverse compensated demand curves. The areas under these curves, bounded on the demand axis by the quantities zero and z, are economic indices of the value of the first z units, known as the equivalent surplus and the compensating surplus (Randall & Stoll 1980), as long as the quantity evaluated is made available free in the environment to which the results are applied (Lankford 1988). These indices are more accurate and valid than the consumer surplus (see below) and have precise definitions, consistent with the levels of utility they use as their reference points. The equivalent surplus, which corresponds to an area under the $u=u'$ inverse compensated demand curve, measures the amount of currency that must be given to a consumer, in lieu of an increase in consumption from zero to z, to leave her as well off as she would have been if the increase in consumption had occurred (Just et al. 1982, page 139). In other words, it measures how much currency she is willing to accept to forego consuming the quantity z. The compensating surplus, corresponding to an area under the $u=u^0$ curve, measures the amount of currency that must be taken away from a consumer after an increase in consumption from zero to z to restore her to her original welfare level (Just et al. 1982, page 139). In other words, it indicates how much she is willing to pay to secure z units of the resource. The consumer surplus is intermediate in magnitude between the equivalent and compensating surpluses

(Randall & Stoll 1980) and can be thought of as approximating either of these indices. It underestimates the amount the subject is willing to accept to forego consuming z units, while overestimating the amount she is willing to pay to secure this quantity.

Whether the consumer surplus underestimates or overestimates the value of a given quantity of a resource therefore depends upon how valuation is defined. There are no general grounds on which to prefer one definition to the other. The researcher is free to decide whether she wishes to estimate willingness to pay or willingness to accept and her decision will determine the direction of error. This decision may also influence how certain comparisons are made between resources, once demand curves have been generated, in particular when ascertaining the quantity of one resource which is worth the same as a given quantity of another. In this context, willingness to pay is a measure of the number of units of an alternative resource that the subject would be prepared to give up to secure a given quantity of the resource of interest. It corresponds to the number of 'lost' units of the alternative good, counting downwards, possibly from a satiating quantity, that are worth the same as a given quantity of the resource of interest, counting upwards from zero (e.g. Kirkden 2000). Hence, when using the consumer surplus as an estimate of willingness to pay, it makes sense to count units of the alternative good downwards. In contrast, willingness to accept specifies the number of units of the alternative resource that the subject would have to be given to compensate her for not receiving a given quantity of the resource of interest. It corresponds to the number of units of the alternative good, counting upwards, possibly from zero, that are worth the same as a given quantity of the resource of interest, counting upwards from zero. Thus, when treating the consumer surplus as an estimate of willingness to accept, units of the alternative good should be counted upwards.

There is one circumstance under which it makes more sense to use the consumer surplus as an estimate of willingness to pay than as an estimate of willingness to accept. This is when the first unit of the alternative good, counting upwards from zero, is worth much more than the quantity of the resource of interest to be evaluated. The trouble with using willingness to accept is that it may be difficult to identify a suitable consumption level of the alternative good from which to count upwards. A level of zero is inappropriate, because a small fraction of a unit cannot be counted accurately, whilst a larger level will often be arbitrary. In this situation, it makes more sense to use the willingness to pay. Because the first 'lost' unit of the alternative resource, counting downwards from a satiating quantity, is generally worth much less than the first unit, counting upwards from zero, it may not be necessary to count down small fractions of units.

The problem associated with the income effect could in principle be avoided entirely by abandoning the consumer surplus and using either the equivalent surplus or the compensating surplus instead. These indices are not influenced by changes in the purchasing power of the subject's income. They are therefore more valid than any of the other indices that we have discussed. However, their use has several practical drawbacks for applied ethologists. First, they require the systematic manipulation of both price and income, which would substantially increase the amount of experimental time required. Second, the calculations that are necessary to estimate them are relatively complex.

DISCUSSION

It is necessary to consider how tractable the shortcomings of the elasticity of demand and consumer surplus indices might be, given modifications to the design of preference tests. The elasticity of demand indices are considered first. Except where stated, shortcomings pertain to internal validity.

The first shortcoming (affecting external validity) is that ethologists assume that a single elasticity value can be assigned to each resource. This is an auxiliary assumption that does not follow inevitably from Dawkins' (1983a, 1988, 1990) rationale. In principle, a series of point elasticity estimates (Mansfield & Yohe 2000, page 27) might conceivably be used instead, corresponding to particular consumption levels.

The second shortcoming is that these indices confound the defence of a preferred consumption level with satiation. This shortcoming is intractable. A readiness to maintain consumption in the face of increasing costs and a tendency to become rapidly satiated are perfectly confounded with one another.

The third shortcoming is that the elasticity of demand indices fail to take into account the amount that a subject is required to pay. It compromises the validity of these indices in all circumstances except where the resources to be compared are consumed in the same quantity when cost is low. However, a solution may be to factor out the effects of initial consumption level by treating it as a covariate (Warburton & Mason, in press).

The fourth shortcoming (affecting external validity) is that the price and income elasticity of demand indices are specific to the level of income or price at which the demand curve or Engel curve was generated, respectively. This can be resolved by designing preference tests in such a way that their results are relevant to the environments in which they are applied. For the price elasticity of demand, if the currency is one of time, it is sufficient to ensure that the chosen income level is similar to the period of access that animals would have to such resources in a real-life environment. If time is not a major factor, it should be ensured that the subject's income permits similar quantities of alternative resources to be consumed, when the price of the resource of interest is at its lowest level, as would be the case in the real-life environment. For the income elasticity of demand, in an operant test, the price should be kept as low as possible. In a choice test, the price of a resource may be defined in terms of the procurement time required per unit return, or in terms of the identity or importance of alternative resources. Provided that the experimental environment resembles the real-life environment in the way the resource of interest is procured, as well as in the kinds and

available quantities of alternatives, the price of the resource of interest should be comparable to its price in this environment.

The fifth shortcoming applies to the price elasticity index. The problem is that the price elasticity of demand is influenced by the income effects of price changes. This problem is tractable, to a limited extent. In principle, any modification that reduces the currency value of the change in real income, Δm, in relation to the subject's income level, should be beneficial. Because $\Delta m \approx x\Delta p = xp^1 - xp^0$, where p^0 and p^1 are the initial (approaching zero) and final price levels and x is the (near satiating) quantity consumed when price=p^0, modifications should aim to achieve a large difference between income and what the cost would be if the quantity, x, were consumed at p^1. However, the effectiveness of this procedure will be limited, for many resources, by the large magnitude of the price change, Δp, over which elasticity is measured, since this makes for a large p^1.

The consumer surplus does not suffer from the first three shortcomings attributed to the elasticity of demand indices. The first shortcoming of the consumer surplus index (affecting external validity) is that it is specific to the income level at which the demand curve was generated. As with the price elasticity index, this can be resolved by selecting an income level at which the experimental environment resembles the environment in which the results are to be applied.

The second shortcoming is that the consumer surplus is influenced by the income effects of price changes. This problem is tractable. The key is to ensure that the currency value of a given quantity of the resource of interest amounts to a small proportion of the subject's income, thereby minimizing the $|CS|/m$ term in equation (10). There are several ways in which this might be achieved in a preference test. The subject's time income might be maximized, or her motivation to consume other resources might be increased by enriching the experimental environment with stimuli and restricting their availability to experimental sessions. However, care must be taken to avoid enriching the experimental environment with resources that are unavailable in a real-life environment, because this may compromise the external validity of the findings. The subject will also attach less value to a small quantity of the resource of interest than to a large quantity. Hence, the smaller the quantity that is evaluated, the more accurate a consumer surplus estimate of its value will be. In many cases, the value of the first unit of a resource will be of more interest than the value of a larger quantity, because the value of the first unit reflects the importance of having the resource available at all. The values of larger quantities reflect increasingly the pleasure that can be gained from consumption, above and beyond the relief of deprivation.

A more radical modification to the standard operant test, based on a procedure used by Rachlin et al. (1976) for a different purpose, is expected to improve substantially the validity of the consumer surplus index. Instead of making the resource of interest available on a single fixed ratio schedule, the resource of interest and an alternative resource, such as food, would be made available simultaneously on a concurrent fixed ratio schedule. The animal would be permitted to perform a fixed number of operant responses each session, defined as her income, and would be free to allocate these between the available resources as she chose. The price of the resource of interest, defined as the number of responses required for access to a fixed quantity, would be gradually increased from zero whilst holding the price of the alternative resource constant, to generate a demand curve. The magnitude of the subject's income should be just sufficient to permit a satiating quantity, or some other specific quantity, of the alternative to be purchased, ensuring that consumption of the resource of interest always entailed giving up some quantity of the alternative, except when its price was zero. The procedure could be repeated for a series of different resources of interest, using the same alternative each time, in order to rank or otherwise compare them.

This procedure is effectively a hybrid between an operant test and a choice test. The subject expends relatively little time or effort in acquiring a resource, limited by the size of her income, but the principal sacrifice made is the quantity of the alternative resource that is rendered unaffordable by doing so. As long as a substantial quantity of the alternative resource is highly valued by the subject, this procedure ensures that the currency value of a given quantity of the resource of interest, measured in terms of lost units of the alternative resource, will be relatively low and will constitute a relatively small proportion of the subject's income, thereby minimizing the income effects of price changes. To ensure that the resulting value estimates are relevant to the environment in which they are applied, the number of responses that constitute the subject's income should be sufficient to allow a similar quantity of the alternative resource to be consumed, as would be the case in the real-life environment. The subject should also receive free access to similar quantities of other resources as would be available in such an environment. The principal limitation of this procedure is that care must be taken to ensure that the resource of interest and the alternative are not substitutes.

The internal and external validity of the reservation price index is the same as that of the consumer surplus of the first unit of a resource, or of the quantity consumed in the first visit if the resource is unrationed. The smaller this quantity is, the less that the income effect will compromise the internal validity of the index.

Two experimental studies have attempted to compare the price elasticity index with the consumer surplus. This comparison is not straightforward, since the two indices are designed to evaluate different things: a resource and a given quantity of a resource, respectively. The approach that has been taken is to obtain consumer surplus estimates of the value of comparable quantities of the resources, such as the first unit, or the quantity required for satiation, and to use these to rank the resources. These measures do not, in reality, rank resources, but they do at least compare like with like. In one study, Mason et al. (2001) compared the motivational strength of mink, *Mustela vison*, to consume seven different resources. The average price elasticity of demand and the consumer

surplus of a satiating quantity, corresponding to the entire area under the demand curve, ranked the resources in the same way. However, in a second run of the experiment, using a different group of mink, considerably less agreement was found between the rank orders generated by the two indices (G. Mason, personal communication). Kirkden (2000) compared the motivational strength of growing pigs, *Sus scrofa*, to interact with novel and familiar objects. Consumer surplus estimates of the value of the first unit and of the value of a satiating quantity both ranked the novel objects above the familiar objects for all five subjects, but the average price elasticity of demand ranked the familiar objects above the novel objects for three of the five subjects.

There is also some experimental evidence to support the argument we have advanced that the elasticity of demand indices are biased by the costliness of maintaining consumption at its initial level, tending to underestimate the relative value of goods that are consumed in large quantities when cost is low. Cooper & Mason (2001) measured an analogue of the income elasticity of demand in mink for seven resources and reported that the goods that the subjects spent the most time consuming when they had continual access to them showed the greatest proportional reduction in consumption time when the daily time budget was reduced.

CONCLUSIONS

It is not possible to establish, on the basis of a theoretical analysis, how serious each of the problems of validity will be for the various indices of motivational strength, because they depend in part upon the properties of the resources in question. Only an empirical enquiry can shed light upon this. However, it is possible to draw some general conclusions concerning the overall validity of the competing indices. The main conclusion is that, under most circumstances, the consumer surplus is a more internally valid index of motivational strength than the price elasticity of demand. This follows from two facts.

(1) The price elasticity index has two shortcomings that compromise its internal validity under most circumstances and from which the consumer surplus does not suffer.

(2) The price elasticity index shares a third shortcoming (the income effect) with the consumer surplus. Two factors act, via the income effect, to compromise the validity of the consumer surplus. One, the income elasticity of demand, is similar to a factor affecting the price elasticity index, the rate at which demand changes with income. The other, the value that the subject attaches to the quantity evaluated as a proportion of income, is tractable by modifying the experimental procedure. Hence, there are few circumstances under which the consumer surplus will be influenced more by the income effect than the price elasticity of demand.

A general conclusion cannot be drawn concerning the internal validity of the income elasticity and consumer surplus indices, because the shortcomings of the consumer surplus are not a subset of the problems affecting the income elasticity index. Although the income elasticity index has more shortcomings than the consumer surplus, including one that is intractable, in the absence of empirical information concerning their severity it is unclear, on balance, which index is the more internally valid.

We can also conclude that the price and income elasticity indices, as they are used at present, are less externally valid and considerably less versatile than the consumer surplus, because they attempt to assign a single value to a resource. The consumer surplus is capable of addressing a range of practical questions that the price and income elasticities of demand cannot.

Furthermore, the reservation price index may be a useful shortcut to estimating the consumer surplus when only the value of the first unit or visit is of interest. It may be particularly useful for evaluating resources that cannot be divided into rations without altering their value to the subject (Cooper & Mason 1997b, 2001; Mason et al. 1998; Olsson & Keeling 2002; Olsson et al. 2002). By increasing the range of resources that can be evaluated, as well as the speed with which value estimates can be obtained, the reservation price index further increases the versatility of the consumer surplus.

Our overall conclusion is that applied ethologists stand to benefit from using the consumer surplus as an index of motivational strength, both in terms of the range of questions it permits them to ask and in terms of the validity of their findings. The reservation price index, which is a special case of the consumer surplus, is also recommended, especially where it is not practical or convenient to generate a demand curve.

Acknowledgments

We thank Georgia Mason for valuable comments on the manuscript and Tony Dickinson, Alan Grafen, Patrick Green, Rowena Guzy, Kristin Hagen, Suzanne Held, Georgia Mason and Tim Roper for useful comments made at various stages in the development of these ideas.

References

Baldwin, B. A. & Start, I. B. 1981. Sensory reinforcement and illumination preference in sheep and calves. *Proceedings of the Royal Society of London, Series B*, **211**, 513–526.

Baldwin, B. A. & Start, I. B. 1985. Illumination preferences of pigs. *Applied Animal Behaviour Science*, **14**, 233–243.

Blom, H. J. M., van Tintelen, G., Baumans, V., van den Broek, J. & Beynen, A. C. 1995. Development and application of a preference test system to evaluate housing conditions for laboratory rats. *Applied Animal Behaviour Science*, **43**, 279–290.

Broom, D. M. & Johnson, K. G. 1993. *Stress and Animal Welfare.* Dordrecht: Kluwer.

Bubier, N. E. 1996a. The behavioural priorities of laying hens: the effect of cost/no cost multi-choice tests on time budgets. *Behavioural Processes*, **37**, 225–238.

Bubier, N. E. 1996b. The behavioural priorities of laying hens: the effects of two methods of environmental enrichment on time budgets. *Behavioural Processes*, **37**, 239–249.

Chiang, A. C. 1984. *Fundamental Methods of Mathematical Economics*, 3rd edn. New York: McGraw-Hill.

Cooper, J. J. & Appleby, M. C. 1995. Nesting behaviour of hens: effects of experience on motivation. *Applied Animal Behaviour Science*, **42**, 283–295.

Cooper, J. J. & Appleby, M. C. 1996. Demand for nest boxes in laying hens. *Behavioural Processes*, **36**, 171–182.

Cooper, J. J. & Appleby, M. C. 1997. Motivational aspects of individual variation in response to nestboxes by laying hens. *Animal Behaviour*, 54, 1245–1253.

Cooper, J. J. & Mason, G. J. 1997a. Income elasticity as an indicator of behavioural priorities in mink (*Mustela vison*). In: *Proceedings of the 31st International Congress of the International Society for Applied Ethology, 13–16 August 1997, Prague, Czech Republic* (Ed. by P. H. Hemsworth, M. Špinka & L'. Košt'ál), page 64. Prague-Uhříněves, Czech Republic: Research Institute of Animal Production.

Cooper, J. J. & Mason, G. J. 1997b. The effect of cost of access on consumption of environmental resources in mink. In: *Animal Choices* (Ed. by J. M. Forbes, T. L. J. Lawrence, R. G. Rodway & M. A. Varley), pp. 129–130. Penicuik, Midlothian: British Society for Animal Science.

Cooper, J. J. & Mason, G. J. 1999. Assessing the behavioural needs of mink (*Mustela vison*) using three methodologies from human microeconomics. In: *Proceedings of the 33rd International Congress of the International Society for Applied Ethology, 17–21 August 1999, Lillehammer, Norway* (Ed. by K. E. Bøe, M. Bakken & B. O. Braasted), page 82. Ås: Agricultural University of Norway.

Cooper, J. J. & Mason, G. J. 2000. Increasing costs of access to a resource cause re-scheduling of behaviour in American mink (*Mustela vison*): implications for the assessment of behavioural priorities. *Applied Animal Behaviour Science*, **66**, 135–151.

Cooper, J. J. & Mason, G. J. 2001. The use of operant technology to measure behavioral priorities in captive animals. *Behavior Research Methods, Instruments and Computers*, **33**, 427–434.

Dawkins, M. S. 1977. Do hens suffer in battery cages? Environmental preference and welfare. *Animal Behaviour*, **25**, 1034–1046.

Dawkins, M. S. 1978. Welfare and the structure of a battery cage: size and cage floor preferences in domestic hens. *British Veterinary Journal*, **134**, 469–475.

Dawkins, M. S. 1980. *Animal Suffering: The Science of Animal Welfare.* London: Chapman & Hall.

Dawkins, M. S. 1981. Priorities in the cage size and flooring preferences of domestic hens. *British Poultry Science*, **22**, 255–263.

Dawkins, M. S. 1983a. Battery hens name their price: consumer demand theory and the measurement of ethological 'needs'. *Animal Behaviour*, **31**, 1195–1205.

Dawkins, M. S. 1983b. Cage size and flooring preferences in litter-reared and cage-reared hens. *British Poultry Science*, **24**, 177–182.

Dawkins, M. S. 1988. Behavioural deprivation: a central problem in animal welfare. *Applied Animal Behaviour Science*, **20**, 209–225.

Dawkins, M. S. 1990. From an animal's point of view: motivation, fitness, and animal welfare. *Behavioral and Brain Sciences*, **13**, 1–61.

Dawkins, M. S. 1997. Suffering, demand curves and welfare: a reply to Houston. *Animal Behaviour*, **53**, 1119–1121.

Dawkins, M. S. & Beardsley, T. 1986. Reinforcing properties of access to litter in hens. *Applied Animal Behaviour Science*, **15**, 351–364.

Duncan, I. J. H. 1978. The interpretation of preference tests in animal behaviour. *Applied Animal Ethology*, **4**, 197–200.

Duncan, I. J. H. 1992. Measuring preferences and the strength of preferences. *Poultry Science*, **71**, 658–663.

Duncan, I. J. H. & Kite, V. G. 1987. Some investigations into motivation in the domestic fowl. *Applied Animal Behaviour Science*, **18**, 387–388.

Faure, J. M. 1986. Operant determination of the cage and feeder size preferences of the laying hen. *Applied Animal Behaviour Science*, **15**, 325–336.

Faure, J. M. & Lagadic, H. 1994. Elasticity of demand for food and sand in laying hens subjected to variable wind speed. *Applied Animal Behaviour Science*, **42**, 49–59.

Findley, J. D. 1962. An experimental outline for building and exploring multi-operant behavior repertoires. *Journal of the Experimental Analysis of Behavior*, **5**, 113–166.

Fraser, D. & Matthews, L. R. 1997. Preference and motivation testing. In: *Animal Welfare* (Ed. by M. C. Appleby & B. O. Hughes), pp. 159–173. Wallingford: CAB International.

Gunnarsson, S., Matthews, L. R., Foster, T. M. & Temple, W. 2000. The demand for straw and feathers as litter substrates by laying hens. *Applied Animal Behaviour Science*, **65**, 321–330.

Hanemann, W. M. 1991. Willingness to pay and willingness to accept: how much can they differ? *American Economic Review*, **81**, 635–647.

Harri, M., Kasanen, S., Mononen, J., Ahola, L. & Sepponen, J. 2001. Trade-off between floor level and floor material in farmed silver foxes. *Behavioural Processes*, **53**, 87–95.

Hodos, W. 1961. Progressive ratio as a measure of reward strength. *Science*, **134**, 943–944.

Hogan, J. A. & Roper, T. J. 1978. A comparison of the properties of different reinforcers. In: *Advances in the Study of Behavior. Vol. 8* (Ed. by J. Rosenblatt, R. Hinde, C. Beer & M.-C. Busnel), pp. 155–255. New York: Academic Press.

Houston, A. I. 1997. Demand curves and welfare. *Animal Behaviour*, **53**, 983–990.

Houston, A. I. & McFarland, D. J. 1980. Behavioral resilience and its relation to demand functions. In: *Limits to Action: The Allocation of Individual Behavior* (Ed. by J. E. R. Staddon), pp. 177–203. New York: Academic Press.

Howell, D. C. 1997. *Statistical Methods for Psychology*, 4th edn. Belmont, California: Duxbury Press.

Hughes, B. O. 1975. Spatial preference in the domestic hen. *British Veterinary Journal*, **131**, 560–564.

Hughes, B. O. & Black, A. J. 1973. The preference of domestic hens for different types of battery cage floor. *British Poultry Science*, **14**, 615–619.

Hursh, S. R. 1984. Behavioral economics. *Journal of the Experimental Analysis of Behavior*, **42**, 435–452.

Hursh, S. R., Raslear, T. G., Bauman, R. & Black, M. 1989. The quantitative analysis of economic behavior with laboratory animals. In: *Understanding Economic Behavior* (Ed. by K. G. Grunert & F. Ölander), pp. 393–407. Dordrecht: Kluwer Academic.

Hutson, G. D. 1988. Do sows need straw for nest-building? *Australian Journal of Experimental Agriculture*, **28**, 187–194.

Hutson, G. D. & Haskell, M. J. 1990. The behaviour of farrowing sows with free and operant access to an earth floor. *Applied Animal Behaviour Science*, **26**, 363–372.

Hutson, G. D. & van Mourik, S. C. 1981. Food preferences of sheep. *Australian Journal of Experimental Agriculture and Animal Husbandry*, **21**, 575–582.

Just, R. E., Hueth, D. L. & Schmitz, A. 1982. *Applied Welfare Economics and Public Policy*. Englewood Cliffs, New Jersey: Prentice-Hall.

Katz, M. L. & Rosen, H. S. 1998. *Microeconomics*. 3rd edn. Boston, Massachusetts: McGraw-Hill.

Kirkden, R. D. 2000. Assessing motivational strength and studies of boredom and enrichment in pigs. Ph.D. thesis, University of Cambridge.

Kline, P. 1998. *The New Psychometrics: Science, Psychology and Measurement*. London: Routledge.

Lagadic, H. & Faure, J. M. 1987. Preferences of domestic hens for cage size and floor types as measured by operant conditioning. *Applied Animal Behaviour Science*, **19**, 147–155.

Lankford, R. H. 1988. Measuring welfare changes in settings with imposed quantities. *Journal of Environmental Economics and Management*, **15**, 45–63.

Lawrence, A. B. & Illius, A. W. 1989. Methodology for measuring hunger and food needs using operant conditioning in the pig. *Applied Animal Behaviour Science*, **24**, 273–285.

Lea, S. E. G. 1978. The psychology and economics of demand. *Psychological Bulletin*, **85**, 441–466.

Manser, C. E., Morris, T. H. & Broom, D. M. 1995. An investigation into the effects of solid or grid cage flooring on the welfare of laboratory rats. *Laboratory Animals*, **29**, 353–363.

Manser, C. E., Elliott, H., Morris, T. H. & Broom, D. M. 1996. The use of a novel operant test to determine the strength of preference for flooring in laboratory rats. *Laboratory Animals*, **30**, 1–6.

Mansfield, E. & Yohe, G. 2000. *Microeconomics: Theory/Applications*, 10th edn. New York: W.W. Norton.

Mason, G. J., Cooper, J. J. & Garner, J. 1997. Models of motivational decision-making and how they affect the experimental assessment of motivational priorities. In: *Animal Choices* (Ed. by J. M. Forbes, T. L. J. Lawrence, R. G. Rodway & M. A. Varley), pp. 9–17. Penicuik, Midlothian: British Society for Animal Science.

Mason, G. J., McFarland, D. & Garner, J. 1998. A demanding task: using economic techniques to assess animal priorities. *Animal Behaviour*, **55**, 1071–1075.

Mason, G. J., Cooper, J. J. & Clarebrough, C. 2001. Frustrations of fur-farmed mink. *Nature*, **410**, 35–36.

Matthews, L. R. & Ladewig, J. 1994. Environmental requirements of pigs measured by behavioural demand functions. *Animal Behaviour*, **47**, 713–719.

Matthews, L. R., Temple, W., Foster, T. M., Walker, J. & McAdie, T. M. 1995. Comparison of the demand for dustbathing substrates by layer hens. In: *Proceedings of the 29th International Congress of the International Society for Applied Ethology, 3–5 August 1995, Exeter, UK* (Ed. by S. M. Rutter, J. Rushen, H. D. Randle & J. C. Eddison), pp. 11–12. Potters Bar: Universities Federation for Animal Welfare.

Nicol, C. J. 1986. Non-exclusive spatial preference in the laying hen. *Applied Animal Behaviour Science*, **15**, 337–350.

Nicol, C. J. 1997. Environmental choices of farm animals. In: *Animal Choices* (Ed. by J. M. Forbes, T. L. J. Lawrence, R. G. Rodway & M. A. Varley), pp. 35–43. Penicuik, Midlothian: British Society for Animal Science.

Ng, Y.-K. 1990. The case for and difficulties in using 'demand areas' to measure changes in well-being. *Behavioral and Brain Sciences*, **13**, 30–31.

Olsson, I. A. S. & Keeling, L. J. 2002. The push-door for measuring motivation in hens: laying hens are motivated to perch at night. *Animal Welfare*, **11**, 11–19.

Olsson, I. A. S., Keeling, L. J. & McAdie, T. M. 2002. The push-door for measuring motivation in hens: an adaptation and a critical discussion of the method. *Animal Welfare*, **11**, 1–10.

Rachlin, H., Green, L., Kagel, J. H. & Battalio, R. C. 1976. Economic demand theory and psychological studies of choice. In: *The Psychology of Learning and Motivation. Vol. 10* (Ed. by G. Bower), pp. 129–154. New York: Academic Press.

Randall, A. & Stoll, J. R. 1980. Consumer's surplus in commodity space. *American Economic Review*, **70**, 449–455.

van Rooijen, J. 1980. Wahlversuche, eine ethologische Methode zum Sammeln von Meßwerten, um Haltungseinflüsse zu erfassen und zu beurteilen. In: *Aktuelle Arbeiten zur artgemäßen Tierhaltung, Vol. 264* (Ed. by K. Zeeb), pp. 165–185. Darmstadt: KTBL-Schrift.

van Rooijen, J. 1982. The value of choice tests in assessing welfare of domestic animals. *Applied Animal Ethology*, **8**, 295–299.

Sherwin, C. M. 1996. Laboratory mice persist in gaining access to resources: a method of assessing the importance of environmental features. *Applied Animal Behaviour Science*, **48**, 203–214.

Sherwin, C. M. 1998. The use and perceived importance of three resources which provide caged laboratory mice with the opportunity for extended locomotion. *Applied Animal Behaviour Science*, **55**, 353–367.

Sherwin, C. M. & Nicol, C. J. 1996. Reorganisation of behaviour in laboratory mice, *Mus musculus*, with varying cost of access to resources. *Animal Behaviour*, **51**, 1087–1093.

Sherwin, C. M. & Nicol, C. J. 1997. Behavioural demand functions of caged laboratory mice for additional space. *Animal Behaviour*, **53**, 67–74.

Siegel, S. & Castellan, N. J. 1988. *Nonparametric Statistics for the Behavioral Sciences*. 2nd edn. New York: McGraw-Hill.

Toates, F. 1986. *Motivational Systems*. Cambridge: Cambridge University Press.

Varian, H. R. 1996. *Intermediate Microeconomics: A Modern Approach*, 4th edn. New York: W.W. Norton.

Warburton, H. J. & Mason, G. J. In press. Is out of sight, out of mind? The effect of resource cuts on motivation in the mink, *Mustela vison*. *Animal Behaviour*.

Warburton, H. J. & Nicol, C. J. 2001. The relationship between behavioural priorities and animal welfare: a test using the laboratory mouse, *Mus musculus*. *Acta Agriculturae Scandinavica, Section A, Animal Science, Supplement*, **30**, 124–130.

van de Weerd, H. A., van Loo, P. L. P., van Zutphen, L. F. M., Koolhaas, J. M. & Baumans, V. 1998. Strength of preferences for nesting material as environmental enrichment for laboratory mice. *Applied Animal Behaviour Science*, **55**, 369–382.

Appendix

In this appendix, we demonstrate from first principles that the rate at which expenditure, expressed as a proportion of income, changes with income is proportional to the income elasticity of demand, but also depends upon the magnitude of expenditure as a proportion of income before the income change.

Let x_i denote the initial level of demand for resource i and let p_i denote the price of this resource. $p_i x_i$ therefore represents the initial level of expenditure on the resource. Furthermore, if m denotes the subject's initial income level, then $S_i=p_i x_i/m$ represents initial expenditure as a proportion of income, or the initial 'expenditure share'. Now, x_i is an endogenous variable and is itself a function of the prices of all n resources in the economy, $\boldsymbol{p}=(p_1, p_2, \ldots, p_n)$, and income: $x_i=x_i(\boldsymbol{p}, m)$. An income change will therefore influence the expenditure share in two ways: directly, in the expression $S_i=p_i x_i/m$; and indirectly, via $x_i(\boldsymbol{p}, m)$.

Total differentiation of the general function, $S_i=S_i(\boldsymbol{p}, x_i, m)$, yields:

$$\mathrm{d}S_i=\frac{\partial S_i}{\partial \boldsymbol{p}}\mathrm{d}\boldsymbol{p}+\frac{\partial S_i}{\partial x_i}\mathrm{d}x_i+\frac{\partial S_i}{\partial m}\mathrm{d}m.$$

Since $\boldsymbol{p}$ is constant, $\mathrm{d}\boldsymbol{p}=0$. Setting $\mathrm{d}\boldsymbol{p}=0$ and dividing through by $\mathrm{d}m$:

$$\frac{\mathrm{d}S_i}{\mathrm{d}m}=\frac{\partial S_i}{\partial x_i}\frac{\mathrm{d}x_i}{\mathrm{d}m}+\frac{\partial S_i}{\partial m}, \qquad \text{(A1)}$$

where: $\frac{\partial S_i}{\partial x_i}\frac{\mathrm{d}x_i}{\mathrm{d}m}$ represents the indirect influence of m upon S_i, via x_i; and $\frac{\partial S_i}{\partial m}$ represents the direct influence of m upon S_i.

Since the income elasticity of demand, $\eta_m \equiv \frac{dx_i}{dm}\frac{m}{x_i}$, it follows that $\frac{dx_i}{dm} = \frac{x_i}{m}\eta_m$. Substituting this term into equation (A1):

$$\frac{dS_i}{dm} = \frac{\partial S_i}{\partial x_i}\frac{x_i}{m}\eta_m + \frac{\partial S_i}{\partial m}. \quad \text{(A2)}$$

Because $S_i = p_i x_i / m$, we can say that:

$$\frac{\partial S_i}{\partial x_i} = \frac{p_i}{m} \quad \text{and} \quad \frac{\partial S_i}{\partial m} = -\frac{p_i x_i}{m^2}. \quad \text{(A3)}$$

Finally, substituting equation (A3) into equation (A2):

$$\frac{dS_i}{dm} = \frac{p_i x_i}{m^2}\eta_m - \frac{p_i x_i}{m^2} = \frac{S_i}{m}(\eta_m - 1).$$

It follows that, when $\eta_m < 1$, $dS_i/dm < 0$; when $\eta_m > 1$, $dS_i/dm > 0$; and when $\eta_m = 1$, $dS_i/dm = 0$. In words, when demand for a resource is income inelastic, income elastic, or has a unitary income elasticity, expenditure as a proportion of income increases, declines or stays the same, respectively, when income is reduced. Furthermore, the rate at which expenditure as a proportion of income changes with income is influenced by the magnitude of expenditure as a proportion of income, S_i, before the income change.