CHAPTER 4: A REVIEW OF DEMAND THEORY AND ITS APPLICATION TO WATER

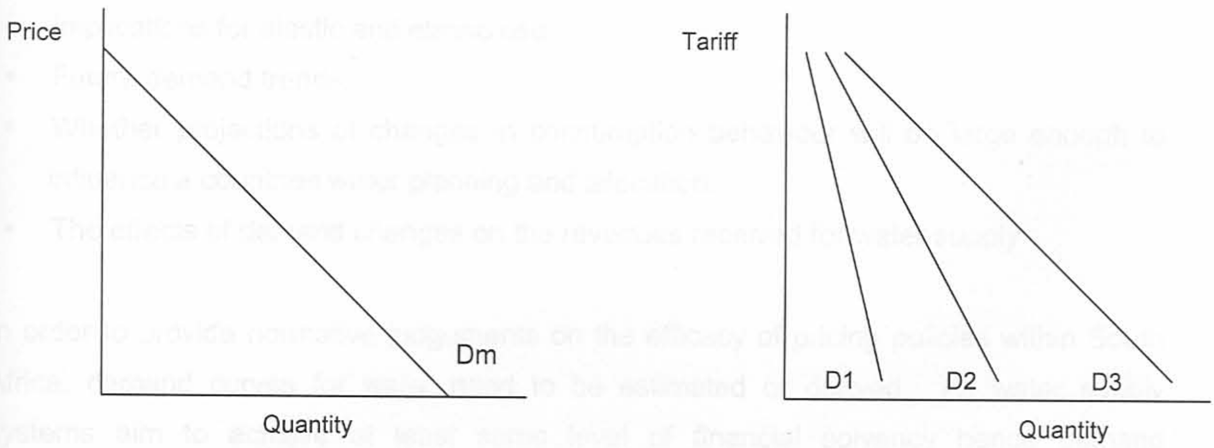
4.1 Introduction

This chapter reviews several methods of demand analysis. The theoretical grounds for various analytical approaches are outlined and compared. Advantages and disadvantages of each approach are discussed and gaps are identified. A limited range of analytical approaches to deriving and estimating demand functions for water have been employed in the literature to map the impacts of price on the quantities demanded of water by various users. Various analytical tools such as elasticities and consumption shares have been used to evaluate the impacts of policy measures on these relationships and the efficacy of water pricing on demand and consumption adjustments. Available models are classified into positive and normative economic approaches. The first approach, positive economics "deals with the objective or scientific explanations of the working of the economy". It is used to provide detailed implications of making a particular choice over another and relies upon the economic mechanisms such as market equilibrium, technology structure and sensitivity analysis. The second approach, normative economics, "offers prescriptions or recommendations based on personal value judgements". It relies on measures of welfare and decision-making based on pareto-optimal criteria (Begg et al, 1991). Conditional normative approaches to economic analysis build on the positive approach by explicitly including optimisation rules and is commonly used in agricultural analysis (Hassan, 1999).

Section 2 of this chapter introduces the concept of demand and its relationship to elasticity. Section 3 outlines the various theoretical approaches to positive demand analysis using market methodologies. Section 4 explains the role of duality in demand theory. Section 5 outlines the pragmatic approaches to econometric estimation of demand for water and section 6 reviews normative approaches to the demand of water. Section 7 outlines the non-economic approaches to demand analysis.

4.2 The water demand curve

Microeconomics is concerned with the explanation and prediction of the economic behaviour of individual units such as consumers, firms and resource owners. It provides an analysis of the relationships between various role players in the market and the factors that influence decision-making. Intrinsic to this process is the understanding of the market and the related workings of its price system (Mansfield, 1994). The demand side of the market may be represented by a market demand schedule that translates into a market demand curve. It represents the quantity of a good that would be purchased at each price. (The market demand curve or aggregate demand curve is the aggregation of a number of different individual demand curves each showing the amount of a particular good that a consumer would purchase at each particular price of that good). The market demand curve represents the level of demand by the market as a whole at each particular market price. Generally, the demand curve represents a negative relationship between price and quantity and the demand curve slopes downwards to the right (as an increase in the income of a particular good leads to a decline in the demand for that particular good). The demand curve also pertains to a particular period in time (a factor which affects the shape and position of the curve), (Mansfield, 1994; Varian, 1996).



Market demand curve

Individual demand curves

Figure 4-1: Hypothetical individual demand curves and market demand curve for water

Source: Own compilation

The distinction between the market demand curve and the individual demand curve is particularly, pertinent when deriving demand curves for water. The actual price of the commodity, water, pertains to the market price whereas the 'tariff' or user charge paid by the consumer pertains to the individual demand curve (figure 4-1).

When relating the price of a commodity or group of commodities to the quantity demanded by consumers or purchasers, the economic concept of demand hinges on a strict set of assumptions implying that the following factors are held constant (Mansfield, 1991, Varian, 1996, Dockel et al, 1979, Begg et al, 1991):

- The prices of other commodities;
- Consumers' incomes;
- Consumption preferences and tastes;
- Consumers' future expectations;
- The number of consumers.

Resource demand such as that for water is also dependent upon other variables such as the nature of the demand for the product, the supply of other inputs, the degree of substitutability between inputs, the time-period of adjustment and the existing market structure (Shumway, 1973). Demand curves play an intrinsic role in economic theory and provide various applications for practical analysis (Dockel et al, 1979). Knowledge of the demand for water by various user groups can be used to determine some of the following:

- The influence of gradual price increases on the level of demand.
- The levels of responsiveness of different user groups to the changes in price.
- Implications for elastic and elastic use.
- Future demand trends.
- Whether projections of changes in consumption behaviour will be large enough to influence a countries water planning and allocation.
- The effects of demand changes on the revenues received for water supply

In order to provide normative judgements on the efficacy of pricing policies within South Africa, demand curves for water need to be estimated or derived. All water supply systems aim to achieve at least some level of financial solvency hence demand projections are necessary. In turn, the determining factors of this demand must be measured. One measure important to projection analysis is that of elasticity (both income and price). Indicating the percentage by which demand changes with changes in income and price respectively, elasticities for 'necessities', such as water, are expected to lie between zero and one (Katzman, 1977).

Theoretical demand represents the relationship between the price and quantity of a particular good while holding all other things constant. It is derived from the underlying preferences of a consumer and maps the optimal amounts of each of the chosen goods as a function of prices and income faced by the consumer during the decision-making process, such that,

$$x_i = x_i\left(p_1, p_2, m\right) \tag{4-1}$$

$$x_2 = x_2\left(p_1, p_2, m\right) \tag{4-2}$$

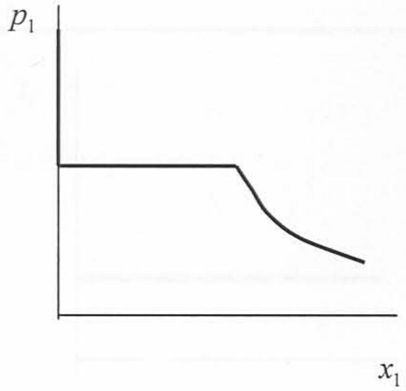
where, x_1 and x_2 are the quantity demanded of good 1 and good 2 respectively; p_1 and p_2 are the prices of good 1 and good 2 respectively; m is household income. The demand curve for x_1 assumes that the price of good 2 (p_2) and income (m) are held constant and the demand curve for x_2 assumes that the price of good 1 (p_1) and income (m) are held constant.

Typically, as the price of a good increases, the demand for that good decreases, provided other influencing factors such as income remain constant, hence, price and quantity tend to move in opposite directions and the demand curve slopes downward such that,

$$\frac{\Delta x_1}{\Delta p_1} < 0 \tag{4-3}$$

Only in the exceptional case of giffon goods will demand increase as price increases (Begg, 1991; Varian, 1996).

For goods that have a high degree of substitutability and there are alternatives available, the price of these alternatives will impact the demand for the original good. If the alternative good is a substitute then the demand for the original good, say (x_1) will increase if $p_1 < p_2$ to $^{m}/_{p_1}$; remain the same or change to any random quantity if $p_1 = p_2$ and decrease towards zero when $p_1 > p_2$. For perfect complements there exists a combination of goods that a consumer is encouraged to consume together, such that the quantities demanded of both goods will change in the same way depending on the price changes, irrespective of whether the prices are the same or different (Varian, 1996). Figure 4-2 depicts the demand curves for substitutes and for compliments.



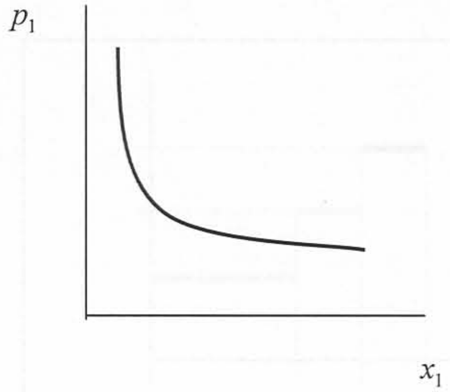


Figure 4-2: Demand curves dependent on the price of relative goods

Source: Varian, 1996

The second factor that impacts the level of demand is consumer income. Assuming prices remain fixed, one can compare optimal choices at different levels of income. Normal goods are defined as goods, the demand for which increases as income rises, conversely inferior goods are goods for which demand decreases as income rises. Normal goods are further classified into luxury and necessary items, dependent on the proportionate change to income. Should the demand for a good increase by more than the increase in income, it is a luxury but where the increase in demand is less than the increase in income the good is a necessity (Begg, 1991; Varian, 1996).

As mentioned in previous chapters, water is regarded as a normal good and a necessity. The degree of substitutability of water with other products is relatively negligible for most uses, unless technology is greatly revised, for example a shift in electricity generation from hydroelectric power to other forms. The purpose of this study is to determine the demand for domestic water use and this incorporates life support uses for which there are no substitutes. Hence, measures such as the cross-price elasticity of demand will not be determined. Figure 4-3 depicts various demand curves that may result from the different tariff structures for water outlined in chapter 6.

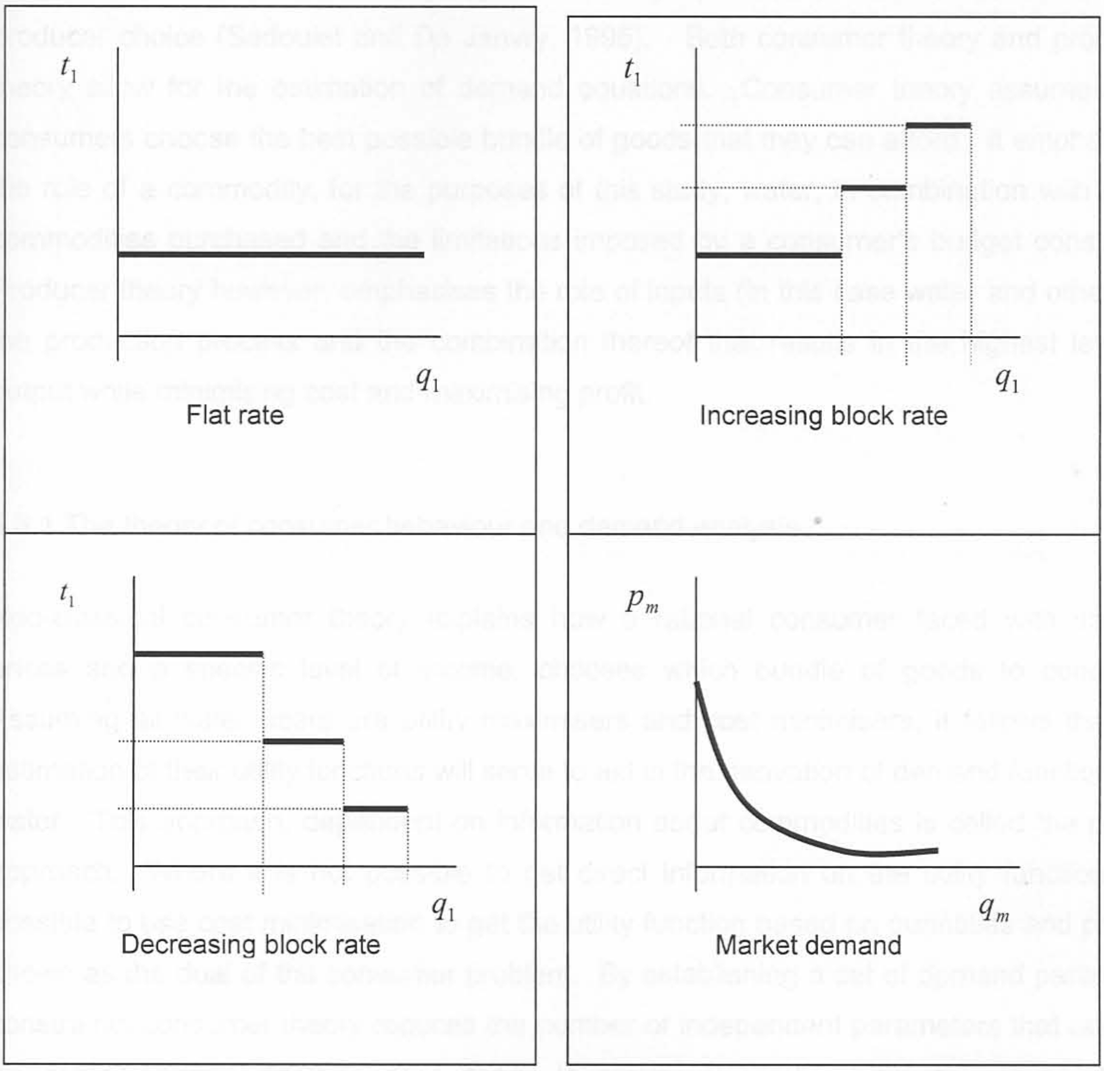


Figure 4-3: Hypothetical demand curves based on water tariff structures

Source: Own graphical representation

For a further discussion on these pricing structures, see section 3.4.2.2 on delivery cost pricing. The market demand curve depicted here is a hypothetical aggregate of demand for the various structures based on the flat rate and decreasing block rate scenarios.

4.3 Theoretical approaches to positive demand analysis

The theoretical approach to the estimation of demand equation parameters relies on the theory of demand to guide the choice of functional form and the inclusion of the relevant variables. This approach proves to be valuable in its ability to impose constraints on the demand parameters thereby reducing the need to estimate unnecessary independent parameters where limitations on available data exist. It also ensures that the estimated demand equations are based on mathematically specified models of consumer and producer choice (Sadoulet and De Janvry, 1995). Both consumer theory and producer theory allow for the estimation of demand equations. Consumer theory assumes that consumers choose the best possible bundle of goods that they can afford. It emphasises the role of a commodity, for the purposes of this study, water, in combination with other commodities purchased and the limitations imposed by a consumer's budget constraint. Producer theory however, emphasises the role of inputs (in this case water and others) in the production process and the combination thereof that results in the highest level of output while minimising cost and maximising profit.

4.3.1 The theory of consumer behaviour and demand analysis

Neo-classical consumer theory explains how a rational consumer faced with various prices and a specific level of income, chooses which bundle of goods to consume. Assuming all water users are utility maximisers and cost minimisers, it follows that the estimation of their utility functions will serve to aid in the derivation of demand functions for water. This approach, dependent on information about commodities is called the primal approach. Where it is not possible to get direct information on the utility function it is possible to use cost minimisation to get the utility function based on quantities and prices, known as the dual of the consumer problem. By establishing a set of demand parameter constraints consumer theory reduces the number of independent parameters that need to be estimated and ensures consistency in results (Sadoulet and De Janvry, 1995; Mansfield, 1991; Varian, 1996). Intrinsic to the estimation of demand curves based on consumer theory are utility, engel curves and consumer cost minimisation.

4.3.1.1 Utility

Economics has defined utility as a measure of a person's well-being or happiness, this unit is however extremely difficult to measure or compare conceptually and has therefore been reformulated in terms of a consumers preferences, described by the concept of utility, figure 4-4.

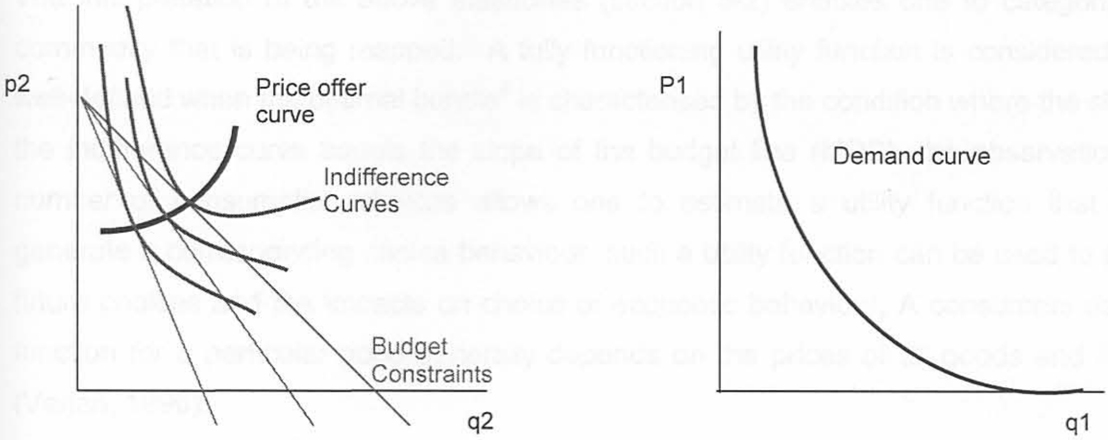


Figure 4-4: Utility and demand

Source: Sadoulet and De Janvry, 1995

The utility maximization problem, assumes that an individual consumer is faced with an utility function subject to some budget constraint,

$$U = v(q, z) \tag{4-4}$$

subject to

$$m = p * q \tag{4-5}$$

using the Lagrange multiplier, the constrained optimisation problem is,

$$U_{\max} = v(q, z) - \lambda(m - p * q) \tag{4-6}$$

solving the above equation using first order conditions, generates a set of Marshallian demand functions,

$$qi = qi(m, p, z), \quad for\ all\ i = 1, \ldots, n. \tag{4-7}$$

From the above demand function, income and price elasticities of demand can be obtained as follows,

$$\eta_i = \frac{\partial q_i}{\partial m} * \frac{m}{q_i} \quad \text{- income elasticity} \tag{4-8}$$

$$e_{ii} = \frac{\partial q_i}{\partial p_i} * \frac{p_i}{q_i} = \frac{MU}{AU} \quad \text{- price elasticity} \tag{4-9}$$

The interpretation of the above elasticities (section 5-2) enables one to categorise the commodity that is being mapped. A fully functioning utility function is considered to be well-defined when the optimal bundle[6] is characterised by the condition where the slope of the indifference curve equals the slope of the budget line (MRS), the observation of a number of consumption choices allows one to estimate a utility function that would generate a corresponding choice behaviour, such a utility function can be used to predict future choices and the impacts on choice of economic behaviour, A consumers demand function for a particular good generally depends on the prices of all goods and income (Varian, 1996).

For the purposes of empirical estimation, time series water data is required for the utility function. Clear observations of price changes over a reasonable time span allow for the estimation of the respective price elasticities.

4.3.1.2 Engel curves

An optimal choice of goods exists at each level of income. For example, the optimal choice of good (x1) can be written as x_1 (p_1, p_2, m), the demand function for good 1. If the prices of the goods, p_1 and p_2 are held constant and the impacts of demand are observed due to changes in income, m, then an Engel curve is generated. An Engel curve represents demand for water as a function of income and constant prices (Varian, 1996).

The Engel curve takes the mathematical form,

$$qi = qi(m, z), \qquad for\ all\ i = 1,...,n. \tag{4-10}$$

where, m denotes income and z denotes characteristics that vary across households, including family size, geographical location, age and education. The Engel curve may take any functional form[7] but the semi-logarithmic tends to perform the best under empirical analysis (Sadoulet and De Janvry, 1995). A fully functioning engel curve exists when the budget constraint is satisfied so that total expenditure and predicted expenditure for each commodity should be equal; all forms of goods such as luxuries, necessities and inferior goods may be represented as well as variable income elasticities; as income

[6] the optimal choice of the consumer is that bundle in the consumer's budget set that lies on the highest indifference curve.

[7] See appendix 3 for a description of the functional forms

increases a saturation point in consumption of commodities should occur. From the above equation, it is possible to determine income elasticities. The price elasticities are however more difficult to determine as they intrinsically require some observation of demand changes for water under various price conditions. One approach to estimating price elasticities of demand from cross-sectional household expenditure survey data with spatial variations in prices is to 'cluster' households with access to the same markets, thereby isolating price changes to location specific variables or transactions. Using expenditure and quantity data it is then possible to calculate the unit values and expenditure shares for each household. From these, price elasticities of demand may be directly estimated,

$$\frac{y_i}{q_i} = v_i, \qquad \frac{y_i}{y} = w_i \tag{4-11}$$

The shortcoming of the above approach is that it tends to overestimate the true price elasticity of demand (Sadoulet and De Janvry, 1995).

Where no information on the observations of price variations is available, but household budget surveys yield cross-sectional data, it is possible to estimate an Engel curve for water instead of a utility function, in order to get demand. Data on income, other household variables and levels of demand are required (Sadoulet and De Janvry, 1995).

4.3.1.3 Cost minimisation

In (section 4.3.1) the consumers' problem was formulated as maximising utility for a given outlay or cost, giving a solution of some utility level, u. This problem can however be reformulated to find the amount of goods that would minimise the cost of producing some level of utility, u. This approach is referred to as the dual of consumer choice (Deaton et al., 1980).

Consider the original problem,

$$Maximise\ U = v(q) \tag{4-12}$$

Subject to,

$$x = \sum p_i \, q_i \tag{4-13}$$

Which generates a solution set of Marshallian demands,

$$q_i = g_i \left(x , p\right) \tag{4-14}$$

The dual of the cost problem however considers some expenditure function,

$$Minimise\ x = \sum p_i \, q_i \tag{4-15}$$

Subject to some level of utility,

$$v\left(q\right) = U \tag{4-16}$$

that in turn yields a set of cost-minimising demand functions dependent not on prices and costs, but on prices and a fixed level of utility, commonly referred to as the Hicksian or compensated demand functions, as they show us how quantity demanded changes when prices change and utility is held constant,

$$q_i = h_i \left(u , p\right) \tag{4-17}$$

From the above set of Marshallian demand functions and Hicksian demand functions, partial derivatives may be obtained,

$$f'\left(q_i\right) = \frac{\partial q_i}{\partial p_i} = MP \tag{4-18}$$

$$f'\left(qi\right) * AP = \frac{\partial qi}{\partial pi} * \frac{pi}{qi} = \frac{MP}{AP} \tag{4-19}$$

As the Marshallian and Hicksian demand functions coincide, they may be substituted back into their respective original functions to determine the maximum utility and the minimum cost that can possibly be obtained,

$$
\begin{aligned}
U &= v(q_1, q_2, \ldots, q_n) = v[g_i(x,p), \ldots, g_n(x,p)] \\
&= \psi(x,p) \qquad (4\text{-}20) \\
&= \text{maximum utility}
\end{aligned}
$$

$$x = \sum p_x h_x(u,p) = c(u,p) = \text{minimum cost} \qquad (4\text{-}21)$$

The above formulations may be rewritten as follows,

$$\psi(x,p) = \max_q [v(q); \sum p * q = x] \quad - \textit{indirect utility function} \qquad (4\text{-}22)$$

$$c(u,p) = \min q [\sum p\, q; v(q) = U] \quad - \cos t \textit{ function} \qquad (4\text{-}23)$$

Interestingly, the dual of the utility and cost functions can be used to formulate different procedures that ultimately determine the underlying Hicksian and Marshallian demand functions. First, it is possible to begin with a cost function specified in terms of utility and prices. Taking the derivative thereof, according to Shephard's Lemma, yields the Hicksian demand, $h_i(u, p) = q_i$. Second, by substituting the maximising level of utility back into the Hicksian demand function, it is possible to derive Marshallian demands,
$q_i [u(x, p); p] = g_i(x, p)$. Third, taking the derivative of the indirect utility function and applying Shephard's lemma, will yield Marshallian demands,

$$q_i = q_i(x,p) = \frac{-\partial\psi / \partial p_i}{\partial\psi / \partial x} \qquad (4\text{-}24)$$

In a similar way, revealed and stated preferences can be used to determine hicksian and marshallian demand functions, discussed in section 4.7, where data on prices and quantities demanded for water are not specifically observable. These can be further used to determine compensating variation, equivalent variation and consumer surplus welfare measures.

The utility-based cost function is fully functioning when it is homogenous of degree one in prices and for a scalar $\phi > 0$. Therefore as prices double, the expenditure required to remain on the same indifference curve also doubles, $c(u, \phi p) = \phi c(u, p)$; increasing in utility, u; non-decreasing in prices, p and increasing in at least one price; concave in prices for all forms of indifference curves, implying that costs will rise linearly as prices increase. For a review of this proof see (Deaton et al., 1980); continuous in prices, p, and is twice differentiable; and taking the first partial derivative of the utility based cost function yields the Hicksian demand functions, otherwise known as Shephard's Lemma (Deaton et al, 1980; Sadoulet et al, 1989; Hassan, 1999):

$$\frac{\partial c(u, p)}{\partial p_i} \equiv h_i(u, p) = q_i$$

Estimations of the utility-based cost function require data on utility and prices. The former is however particularly difficult to obtain as it is based in the theoretical bounds of economics and is limited in its practical empirical application. The derived demand based on the utility-cost function can however be explained through theoretical derivations and requires data on the prices and quantities of the respective good.

4.3.1.4 Multi-output demand equations

Individual demand modelling discussed above allows for the variation of functional form and for the addition or removal of explanatory variables at any researcher's discretion. Individual demand modelling is highly flexible but adheres more closely to the pragmatic approach, as the only theoretical restriction that really provides value is that of homogeneity. Theory however, becomes far more relevant for complete demand estimation or demand systems estimation and a greater range of restrictions may be applied (Deaton et al, 1980).

The choice of demand system is important as this choice holds certain consequences for the estimated parameters. Evidently, demand systems derived from additive utility functions such as the Linear Expenditure System (LES) or the Generalised Linear Expenditure System (GLES) tend to have own-price elasticities proportional to the income elasticities, while the Almost Ideal Demand System (AIDS) does not (Mergos and Donatos, 1972).

The LES is based on the Stone-Geary utility function of the form,

$$U = \Sigma_i \, b_i \ln (q_i - c_i) \tag{4-25}$$

Where ci is the "committed" or minimum subsistence of consumption. Through utility maximisation the following demand curves are obtained,

$$p_i\, q_i \;=\; c_i\, p_i \;+\; b_i \left(Y - \Sigma_j\, c_j\, p_j\right) \qquad \textit{for all } i, j = 1, 2, ..., n \tag{4-26}$$

$\Sigma c_i p_i$ is the subsistence expenditure and the term (y - $\Sigma c_j p_j$) is generally interpreted as "uncommitted" or "supernumerary" income which is spent in fixed proportions b_i between the commodities.

These n equations generate the following price and income elasticities

$$E_{ii} = -1 + (1 - b_i)\frac{c_i}{q_i}, \tag{4-27}$$

$$E_{ij} = -\frac{b_i c_j p_j}{p_i q_i}, \tag{4-28}$$

$$\eta_i = \frac{b_i}{w_i} \tag{4-29}$$

From the above demand curves it is evident that data on the prices of the goods consumed as well as expenditure for a household is required for estimation. The almost ideal demand system (AIDS) is however, based on the budget share and its role in utility maximisation and hence requires data on the shares of a households budget that are spent on consuming certain goods, along with the prices of these goods and the quantities consumed. Both may be used to estimate the demand for household water use, although the estimations would require extensive surveying as much of this information is not readily available, hence this approach is not followed in this study.

4.3.2 The theory of producer behaviour and demand analysis

Another approach to estimating the demand for water is based on the theory of producer behaviour. This is based on the premise that water may be considered as one of the inputs into the production process that in turn generates some level of output and one can determine the factor demand of the input in terms of quantities of inputs and outputs or in terms of prices and quantities of inputs. This approach is not applicable to the estimation of residential water demand as households are utility maximisers and not profit maximisers. The approach is however, reviewed here as it defines an invaluable approach to the estimation of demand theory. Furthermore, the Tshwane study includes the estimation of demand for industrial users using the pragmatic approach, and it is interesting to provide the theoretical framework that could otherwise be applied to this kind of demand estimation.

4.3.2.1 Production functions and the related profit and cost functions

The production function is based on the assumption that a relationship exists between inputs and outputs, it measures the maximum level of output that can be obtained from a given vector of inputs (Chambers, 1988; Coelli et al, 1998; Varian, 1996). The function is single valued so that for any combination of inputs, only one value of output may be obtained, such that,

$$Y = f(x_1, x_2) \tag{4-30}$$

This function merely represents a technical relationship that aims to exclude any technical inefficiencies. For the modelling of water as an input into the production process, quantities of water are used as one of the inputs on the right-hand side of the equation above.

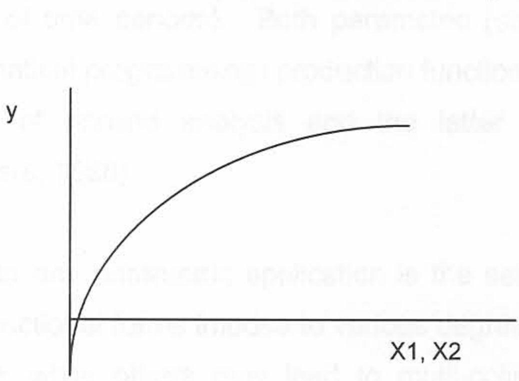


Figure 4-5: A simple production function

A fully-functioning production function is assumed to be twice continuously differentiable, therefore the first and second derivatives are continuous and can both be differentiated. By implication the second cross partials are symmetric and equal each other,

$\frac{\partial^2 Y}{\partial X_1 \partial X_2} = \frac{\partial^2 Y}{\partial X_2 \partial X_1}$; strictly increasing in x, implying monotonicity[3], positive marginal products and concavity; inputs are non-zero for any level of output; the production of

[3] Monotonicity implies that the function does not decrease.

inputs is feasible so their sets must be closed and non-empty; and the function is finite, non-negative and real valued for all values of x and y.

Although the above properties of the production function are adequate for theoretical production function analysis, they are not restrictive enough for strictly applied quantitative analysis or empirical tractability. As most entrepreneurial bodies remain relatively unmoved by the theoretical dictates that price rises should lead to greater output supplied, they will however be more interested in the specifics as to the quantitative amounts by which this output should change in accordance to various price level changes. Consequently, the more restrictive assumptions such as homotheticity and seperability impose structural constraints on the production function that facilitate applied economic analysis.

Production functions may be estimated provided information on input and output quantities is available. This may be done using sample data (Chambers, 1988). The sample data may be in the form of cross-sectional data (involving a number of observations within a particular time period), time-series-data (aggregate data observed over a number of time periods), or panel data (a number of observations made over a number of time periods). Both parametric (econometric methods) and non-parametric (mathematical programming) production functions have been estimated, the former for the purpose of applied analysis and the latter for the purpose of efficiency analysis (Chambers, 1988).

Implicit to any parametric application is the selection of an appropriate functional form. Some functional forms impose to various degrees, restrictive properties on the production structure, while others may lead to multi-collinearity and limited degrees of freedom problems (Chambers, 1988). The most commonly used functional forms for the estimation of production functions are the Leontief and Cobb-Douglas functions.

Water plays a pivotal role in modern agriculture therefore there exists a need to determine the optimal use of water and land resources including the improvement of water management systems through the understanding of the efficiency of production systems (De Juan et al, 1996). In order to solve the problem of optimal water management on farms, information and knowledge on the water consumption of each crop, crop responses to irrigation, agro-climatic data and their effects on crops, and maximum crop evapotranspiration, all of which can be combined in crop production functions for water. From the above it is evident that a production function for water may be applied to any relationship that characterises a crops response to different input combinations such as water, fertilisers, energy, crop-yield and the seasonal amount of applied water (Letey et al, 1985).

Water production functions are usually only valid for a single crop at a single location under conditions of "optimal deficit sequence". They often exclude information on the environmental impacts of irrigation and energy saving. Rao et al (1988) developed a two-stage model that incorporated the above variables and estimated a functional relationship between crop yield and water use. A dated water-production function model was also derived from crop growth stage yield response factors, deducing that a simple heuristic form is applicable over a multitude of stress conditions that can be used where data on crop yield and water use is poor. Letey et al (1985) also recognised that profit maximising conditions were useful as a means of determining the efficient allocation of resources such as water and compared the yields of tall fescue to variations in water salinity and the quantity of irrigation water. The results showed that as water was applied the yield increased to a point beyond which it plateaued, the results were however limited by the specification of the original parameters, see figure 4-6 below.

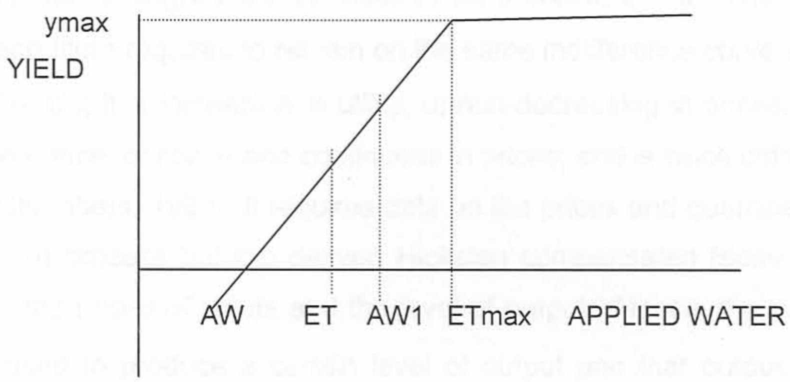


Figure 4-6: Relationship between yield and seasonal applied non-saline water

Source: Letey et al, 1985

Most studies on water production functions emphasise the use of production functions in order to optimise the application of irrigation water (Letey et al, 1985; Sharma and Alonso Neto, 1986; Rajput and Singh, 1986; Russo and Bakker, 1987; Rao et al, 1988; and De Juan et al, 1996). The literature so far, reveals that the production function approach is not used for estimating residential or household water demand and furthermore none of the irrigation based production functions used to derive the demand for water continue further to derive price or income elasticities for water demand.

Three other approaches underlying producer theory are those of profit maximisation and cost minimisation and the revenue function. All three approaches were reviewed for this study but specific studies relating these approaches to water resources were not evident. Only cost functions for service provision could be identified (discussed below). Basically, a well-functioning profit function for water is characterised by positive profits, such that $\Pi(p, w) \geq 0$. Prices are non-decreasing, such that if $p1 \geq p2$, then $\Pi(p1, w) \geq \Pi(p2, w)$ and costs are non-increasing such that if $w1 \geq w2$, then $\Pi(p, w1) \geq \Pi(p, w2)$. The profit function, $\Pi(p, w)$, is convex and continuous in p and w and displays the characteristics of positive linear homogeneity, such that, $\Pi(tp, tw) = t\Pi(p, w)$, where $t > 0$. Estimation of the profit function requires data on the output prices that producers accept as given, the quantity of a good produced, the cost of various inputs used, and the input prices that producers accept as given. All costs of production must be included in the calculation of costs and everything must be measured on a compatible time scale. No studies relating the profit function approach to estimating the demand for water could be found.

The cost function however, represents the "minimum cost of producing a given output level during a given time period expressed as a function of input prices and output", and is homogenous of degree one in prices or for a scalar, $\phi > 0$. Therefore as prices double, the expenditure required to remain on the same indifference curve also doubles, $c(u, \phi p) = \phi c(u, p)$; it is increasing in utility, u; non-decreasing in prices, p and increasing in at least one price; concave and continuous in prices; and is twice differentiable (Binswanger, 1974; Chambers, 1988). It requires data on the prices and quantities of inputs used in the production process but the derived Hicksian compensated factor demand only requires data on the prices of inputs and the level of output. Hence, the price of water and other inputs used to produce a certain level of output and that output $\overline{Q}$. Firestone (1982) developed a statistical cost function on water mains for the Melbourne Metropolitan Board based on economies of scale. The study did not however estimate derived demand functions for water but determined the average and marginal cost options for different supply schemes. The determination of elasticities from cost derived demands does not appear to have been used as an approach for water demand.

Interestingly, the role of duality allows the primal formulation of an economic problem whether cost, profit, revenue, or production optimisation, to be adjusted in terms of dependent variables and rewritten in a different formulation (Sadoulet and De Janvry, 1995). Where a chosen function yields explicit demands, the demand equations are particularly simple to obtain although they require relatively complex algebraic solutions.

However, one can still obtain the derived demand functions for any functional form through the application of five theorems commonly used in duality, namely, Hotelling's Theorem 1 or otherwise known as Shephards Lemma; Shephard's Theorem; Roy's Theorem; World's Theorem and Hotelling's Theorem 2 (Pope, 1982). Duality is measured by a number of successes and failures, they are noted below Pope, 1982).

- Both duals and primals may be used to test structural properties such as: returns to scale, homotheticity, separability, structural change, homogeneity, and other characteristics.
- For some functional forms, solving for first-order conditions does not yield input demands and output supplies, by using Hotelling's theorems, it is possible to obtain the duals from which reduced form input demands and product supplies may be derived.
- Primal systems estimations may be limited by their ability to handle effectively Leontief technologies and estimations when input demands are more co-linear than prices. Duality is able to address these shortfalls of the primal approach.
- Simple estimation methods may be used as duality yields explicit reduced forms where prices are independent.
- The dual approach is easier to use when solving for the welfare impacts of an economic environment.
- Duality is limited in cases where prices are co-linear, such as for intra-seasonal prices of an input (water or fertilisers). In such cases the primal problem will recover the underlying technology more comprehensively.
- Duality is further limited by poorly defined underlying theoretical restrictions or in cases where they are not clearly satisfied, such as where the objective function is non-linear in parameters. For these cases, duality is not an effective approach.

For the purposes of household water demand estimation in this study, the role of duality does not specifically apply as the pragmatic approach is followed. The dual approach based on producer or consumer theory is not applied.

4.4 Pragmatic approaches to positive demand analysis

The pragmatic approach to estimating single equation demand functions relies on the intuition and common sense of the economist facilitating the estimation process. It does not refer to the underlying economic theory for guidance in the selection of the most appropriate functional form or the relevant variables to be included in the model. Although proving to be an attractively simple approach to demand estimation, the pragmatic approach is limited by some serious shortcomings. First, the arbitrary selection of the functional forms for the demand equations and the variables to be included may raise doubt as to the ability of the estimated equation to accurately represent consumer or producer behaviour as well as the underlying behavioural assumptions. Second, the inappropriate choice of functional form may result in the incorrect derivation of the underlying technology structure. For example, for the cobb-douglas functional form, elasticities will be constant over all values of the exogenous variables. They may however vary in reality with long-run price adjustments. Third, the estimated parameters are unlikely to satisfy the restrictions imposed on them by demand theory (Sadoulet and De Janvry, 1995).

Despite these shortcomings, the majority of studies directed at estimating demand functions for water rely on this approach. A pragmatic model usually takes the following form:

$$Q_i = f\left(P_i, P_j, z\right) \qquad (4\text{-}31)$$

Where the quantity of the good (Q_i) is dependent on any selection of independent variables that usually include the price of the good itself (P_i), various other price variables (P_j), and other socio-economic variables (z) that appropriately explain its behaviour.

4.4.1 Application of the pragmatic approach to estimating demand curves for water

The modelling of water demand plays an important role in the planning and management of water resources for urban, industrial and agricultural use (Miaou, 1990a,b). Numerous demand models for urban water use have been proposed based on neo-classical consumer theory and solved by means of linear regression. However, common problems with model specifications have been apparent. First, the length of time for which data is available is usually too short, rendering the degrees of freedom particularly low for any statistical testing. Second, the collected set of explanatory variables must be especially large if it is to be representative. It needs to include socio-economic, environmental, technological and other valid 'input' variables. Third, socio-economic and climatic input variables are often highly correlated, this multicollinearity may adversely affect the statistical validity of results. Fourth, the error terms in models may be autocorrelated or non-stationary (Miaou, 1990a,b).

In (section 5-3), some of the demand models for residential water demand are reviewed and critically discussed. The analysis emphasises the extent to which demand models based on the pragmatic approach differ and hence re-iterates the need for caution when interpreting the results.

4.4.2 Data requirements for the pragmatic approach

The pragmatic estimation of demand models may include variables based on any available data. Cross-sectional and time-series data may be used that include a wide variety of exogenous variables. For the purposes of water demand estimation any combination of the following exogenous variables may be included: water price (marginal or average); household income; number of persons per household; irrigable lawn area; the market value of the home; household water consumption; age of dwelling unit; average water pressure; sum of water and sewerage charges; number of billing periods; regional price index; temperature; precipitation; number of rainy days; soil moisture content; evapotranspiration; number of water dependent household appliances, such as dishwashers and washing machines; number of water saving devices; swimming pools; and mandatory restrictions. Household estimations, however, rely on household specific data that is often tedious and difficult to collect.

4.5 Normative demand analysis

In contrast to the various aspects of positive economics discussed above that aim to provide objective or scientific explanations on how society and firms make decisions, normative economics ‘offers prescriptions or recommendations based on personal value judgements’, and not distinctions between whether these judgements are right or wrong (Begg et. al, 1991). The normative approach requires the estimation of technology parameters as done in the positive approach. These parameters are then used in mathematical programming in order to find optimal solutions in terms of choices or constraints. The normative approach is valuable in its ability to answer normative questions based on non-historic data, it also becomes valuable when valuing natural resources as the programming solutions can generate marginal or shadow prices for the respective resource in question. The approach is however, limited by its ability to fully account for aggregations, technical changes, dynamics and realistic specifications of the problem (Hassan, 1998). The normative approach allows data to be produced through the techniques of parametric programming. Available data on any of the variables that could impact certain decision-making processes can be included. The models are usually based on time-series data.

4.6 Non-economic approaches to demand analysis

Where an econometric approach to demand analysis is hampered by poor databases for endogenous and exogenous variables including prices and quantities for water demanded, non-economic approaches may be used for estimation. Two approaches exist that may be used to estimate a household or individuals 'willingness to pay' for a particular good or service. The first approach is the contingent valuation method (CVM), from which the demand for water and related price elasticities may be estimated. The CVM approach uses surveys to collect data on household decisions. Households or individuals are asked directly to state how much they would be willing to pay in order to have or maintain a particular good or service. Where the changes in prices are to be evaluated, they may also be asked to state by how much their demand would change should the price of a good or service shift up or down. Although critics of CVM claim that it does not yield reliable results due to strategic bias', gamesmanship, and interviewer influence, it is often used in the field of environmental economics to determine the value of environmental goods and services that do not have market prices, such as rural water (Thomas and Syme, 1988; Whittington et al., 1990; Piper et al., 1997; North and Griffin, 1993).

Thomas and Syme, (1988) used this approach to estimate water demand for the Perth Metropolitan Area, Australia due to a lack of sufficient data for empirical estimations. The questionnaire was administered to 312 households selected from a sample frame of 3300 households. The results indicated that the elasticity of demand with respect to a price increase was –0.2. This low elasticity result was attributed to various aspects such as: households having access to private boreholes or wells, low consumer profiles, high income groups and a few consumers indicating that water was not important to their lifestyles. In-house water use had lower elasticities than outdoor water use, results consistent with economic demand theory. The study suggested that the CVM approach proved to be a valuable tool for elucidating price responses where inadequate time series data was a hindrance. This study did however have the advantage of a substantial database gained from previous surveys.

The user's willingness to pay for facilities and resources often determines the success of water or sanitation supply schemes, hence, Rogerson, (1996), undertook an overview of international literature on 'willingness to pay' research. The findings indicated that relatively few studies have been undertaken on 'willingness to pay' for water services, most of which have been linked to the World Bank and other development agencies. Surveys on water-vending services in developing counties were shown to provide valuable sources of information for water resource planning. Furthermore contingent valuation surveys were shown to disclose useful information on consumer / household preferences despite being criticised as "not truly participatory" and "reliant on external resources to generate information on household's willingness to pay", [Dzikus and Surjadi, (1995), in Rogerson, (1996)]. Consequently, contingent valuation surveys were suggested as useful tools for data collection on willingness to pay especially in developing countries such as South Africa. Another area widely valued by using the CVM approach is that of rural water supply (Whittington et al, 1990; Piper et al, 1997; North and Griffin, 1993).

The willingness to pay problem can be formulated as follows, assuming respondents want to minimise some level of expenditure,

$$\text{Min} \quad p\,\Delta x + q\,z \qquad (4\text{-}32)$$

Subject to some level of utility,

$$u^0 = u(\Delta x, z) \qquad (4\text{-}33)$$

where, p and q are prices given at a fixed level, Δx is the level of expenditure on the service or commodity such as water supply, z is a vector of private goods, and u^0 is the original level of utility. The optimisation problem yields an expenditure function in terms of prices, quantities and utility,

$$E = m(p, q, \Delta x, z, u^0) \qquad (4\text{-}34)$$

Which in term may be used to determine WTP,

$$WTP = m(p, q, \Delta x, z, u^0) - m(p, q, x^*, z, u^0) \qquad (4\text{-}35)$$

This in turn, may be used as the correct measure of welfare (Piper et al, 1997).
The second approach to WTP is the indirect approach that uses various models such as hedonic property valuation, hedonic travel costs methods and varying parameter demand. The indirect methods are applied to data on the observed behaviour of water use including quantities of water used, time spent travelling to the point of collection and user perceptions of water quality, in order to determine the response of consumers to changes in the quantity and quality of the water service provided (Whittington, 1990).

The hedonic model is based on the notion that individuals or households choose a dwelling for rental or purchasing purposes based on the characteristics of the surrounding community, the dwelling itself, and the environment. Comparisons are then made between dwellings observed to have mostly identical characteristics but differ in one or two. The amount paid for the dwellings is then compared and the differences in value are assumed to indicate the amount that the consumer or purchaser is willing to pay for some characteristic evident in the one dwelling unit and not in the other. (North and Griffin, 1993), used hedonic property valuation to determine the value of water source and proximity to a number of rural households in the Phillipines.

The approach by North and Griffin (1993) was based on consumer theory and assumed that "all households were utility maximisers who divided their incomes between housing and non-housing goods".

> A housing unit can be modelled as a composite of several characteristics; for simplicity, limit them to two, proximity to the water source (h_1) and construction materials (h_2). These characteristics are purchased jointly, so the monthly expenditure on rent represents the total cost of the housing unit as a function of proximity to the water supply and materials... Under the usual assumptions of consumer theory, each housing characteristic can be valued in terms of the amount of other consumption that a household gives up to secure another unit of it, keeping in mind the complicated fact that each characteristic is bundled with others.
>
> (North and Griffin, 1993).

Based on these household trade-offs, the value of each dwelling and its related characteristics; bid-rent functions (similar to indifference curves) and hedonic price functions can be estimated. From the hedonic price function, marginal willingness to pay values can be determined by the derivative with respect to each characteristic, from which demand may be found.

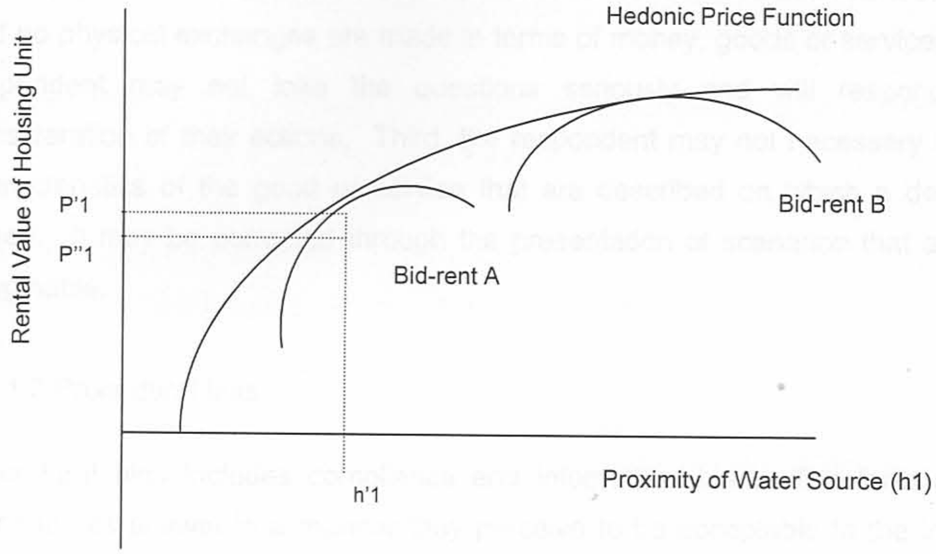


Figure 4-7: Example of a hedonic approach for one characteristic

Source: North and Griffin, 1993

The above figure, (figure 4-7), indicates that the amount that the household is willing to pay (p''1) for the level of the service (h'1) is overstated by the hedonic price function by (p'1). This in turn will overstate the marginal willingness to pay for the service.

4.6.1 Data biases in the willingness to pay approaches

Numerous biases have been identified in the literature, that influence the results obtained from WTP studies of water demand (North and Griffin, 1993; Piper et al, 1997; Rogerson, 1996; Whittington et al, 1990). They are the following:

4.6.1.1 Strategic bias

Strategic bias arises when respondents misrepresent their true positions by answering untruthfully in order to sway investment or policy decisions. Their WTP estimates may be either overstated or understated. This problem may be overcome through transparency of the interview process, where respondents are informed of the reasons for the questions and it is made clear that there are no costs associated with telling the truth and no gains from misrepresentation.

4.6.1.2 Hypothetical bias

Hypothetical bias may result for various reasons. First, the respondent may not understand the full implications of the decision under consideration as it only hypothetical and no physical exchanges are made in terms of money, goods or services. Second, the respondent may not take the questions seriously and will respond without due consideration of their actions. Third, the respondent may not necessary understand the characteristics of the good or service that are described on which a decision is to be based. It may be corrected through the presentation of scenarios that are familiar and imaginable.

4.6.1.3 Procedural bias

Procedural bias includes compliance and information bias. The former occurs when respondents answer in a manner they perceive to be acceptable to the interviewer, and may be corrected when respondents are aware that the costs associated with this type of response are greater than the costs associated with the truth. The latter, refers to the situation where the scenarios presented to the respondents do not accurately reflect reality and may be corrected through the presentation of accurate information.

4.6.1.4 Starting point bias

Starting point bias occurs where a certain price is randomly selected as the first bidding point in the process and respondents believe that this price or value is the one that they should select, and consequently do not give answers that would reflect their actions. Open-ended questions reduce the likelihood of this bias from occurring, as they do not offer an opening bid.

4.6.1.5 Vehicle bias

Vehicle bias occurs where respondents react to the method by which payment is made and not to the level of payment itself. Where this form of payment is familiar and represents a well-defined commodity, vehicle bias is unlikely to occur.

4.7 Summary

In the first part of this chapter, the role of the demand curve for water to better inform decision-makers about the impacts of pricing strategies and policies was discussed. In light of the National Water Act (s36, 1998), better-informed decision-making around pricing is imperative for the implementation of the act in its entirety. The literature indicates that research initiatives around the role of pricing and demand management for developing countries is limited in terms of quantitative measures.

The second part of this chapter, revised the approaches to demand schedule estimation and the studies carried out based on these approaches. From the literature review it became evident that most studies done to estimate water demand curves were based either on the pragmatic approach (Howe and Linaweaver, 1967; Dockel, 1973; Gibbs, 1978; Foster and Beattie, 1979; Howe, 1982; Billings and Agthe, 1980, and Hansen, 1996), or used the production function approach (Letey et al., 1985; Rao et al., 1988 and Juan et al., 1996). Those based on the pragmatic approach tended to estimate water demand for industry and domestic use, while those based on the production function approach tended to estimate water demand schedules for irrigation agriculture.

Most of the industry or domestic water demand studies set as their aim, the determination of price elasticities of demand, hence they continued from the demand schedule to determine the price elasticities of demand and marginal values for water. The irrigation agriculture studies, however, were based on spatially specific micro-economic data and focussed mainly on representing the relationship between price and demand. As a result, once the demand or marginal value schedules had been determined, the studies concluded without calculating the price elasticities of demand or the marginal values for water under irrigation agriculture.

Cost functions were rarely estimated and the studies reviewed here focussed specifically on water supply rather than demand. An example was the study carried out by Firestone (1982), which focussed on the marginal cost of water supply mains. This study was not based on duality theory and it did not attempt to determine price elasticities of demand or values for water supplied.

Non-economic approaches to estimating the demand for water are relatively young in South Africa. One study used contingent valuation to determine the price responsiveness of domestic water users to price adjustments in Gauteng and is discussed further in chapter 5.

Overall, many of these approaches to estimating water demand schedules are highly dependent on data availability, the spatial setting, the nature of the water sector and its demand management strategies. Many research gaps exist for either conducting primary research for water demand schedules particularly in developing countries or for applying and refining studies that have already been done to other regions.

Chapter five continues with the review of demand theory for water, focussing specifically on demand studies for urban household water use and the price elasticity of demand. An extensive literature review of household water studies is given. This is followed by the methodology selected for this study.