Estimating water demand in urban Indonesia: A maximum likelihood approach to block rate pricing data

1997

Piet Rietveld - Vrije Universiteit Amsterdam
Jan Rouwendal - Landbouwuniversiteit Wageningen
Bert Zwart - Vrije Universiteit Amsterdam

Abstract:

In this paper the Burtless and Hausman model is used to estimate water demand in Salatiga, Indonesia. Other statistical models, as OLS and IV, are found to be inappropiate. A topic, which does not seem to appear in previous studies, is the fact that the density function of the loglikelihood can be made arbitrary high if observations are located exactly on a kink of the budget constraint. To avoid this problem, a discretization technique is used to work with genuine probabilities. The unconditional distribution of water demand is explored with parametric and semiparametric techniques. An important conclusion is that the distribution of water demand is not unimodal and that data are clustered around kinks. Main estimation results are a price elasticity of approximately -1.2 and an income elasticity of 0.05. Price and income elasticities are mutually dependent. The estimated model is finally used to investigate consequences for social welfare when a uniform price level is chosen. It is argued that without loss of total welfare, the complex rate structure can be replaced by a uniform marginal price.

Key words: Nonlinear budget constraints, maximum likelihood estimation, kernel estimation, consumer surplus measure, block rate pricing, welfare effects, compensating variation, Vartia's method

Contents

1. Introduction

Understanding the nature of demand for goods or supply of hours of work can be complicated by nonlinear tax systems or varying marginal prices. In this class of problems the value of the combination of micro-economics and econometrics is widely recognized in the literature. Several authors prefer a statistical model, introduced by Burtless and Hausman (1978) which is a direct extension of some micro-economic analysis in the presence of piecewise linear budget constraints. During the last decades, several authors have followed the same approach. A review can be found in Moffitt (1986, 1990). In these reviews and also in Blomquist (1996) it is argued that standard estimation models as OLS and IV are not appropiate.

Hewitt and Hanemann (1995) review the problem of dealing with systems of block prices in water demand. Such systems lead to a divergence between the average and marginal price paid. From a theoretical viewpoint one would prefer the use of the marginal price when estimating demand functions; however, this entails the use of a 'difference term' to account for the impact on income available for other goods. Examples of such studies of urban water demand are Chichoine and Ramamurthy (1986), Billings (1987), and Nieswiadomy and Molina (1989). However, a limitation of these studies is that they only model water demand directly; they leave unmodelled the choice of the block in which to locate water consumption. As indicated by Hewitt and Hanemann (1995), the Burtless and Hausman model can be used to overcome this limitation.

In this paper the demand of water in Salatiga, Indonesia will be estimated. The nonlinear budget constraint is due to the rate structure. For the first ten cubic meters the marginal price is twice as cheap as the marginal price for more than thirty cubic meters. The marginal price rises when the size of demand crosses the limits of ten, twenty and thirty cubic meters, so there are four different marginal prices.

The organisation of this paper is as follows: In section 2 a short overview of the underlying micro-economic theory will be given, followed by some statistical models that can be used to estimate the demand function. Here special attention will be given to the Burtless and

Hausman model, which will be used for estimation. Section 3 gives a description of the dataset. Also, the empirical distribution of water demand will be studied. The loglikelihood for our purposes will be derived in section 4. A modification of the two-error Burtless and Hausman model is used. The main reason is the unboundedness of the loglikelihood of the standard Burtless and Hausman model. As far as we know, this property has not been met before and is due to the presence of data lying exactly at a kink. To avoid this problem a discretization technique is used. Estimation results and some tests can be found in section 5. Section 6 applies the estimated models to study welfare effects by changing the rate structure. Section 7 concludes.

2. Micro-economic and econometric theory

In this section, the micro-economic theory in the case of a convex budget set will be briefly summarized. More about this subject (for instance, other forms of the budget set) can be found in Moffit (1986,1990). Consider an utility-maximizing consumer with a budget set as in figure 1. For simplicity it will be assumed that the consumer's utility will be based on the amount of the good of our interest (for instance: water) and of the total amount of other goods, and that there is only one kink. The ideas discussed below can be simply generalized for more goods and more kinks.

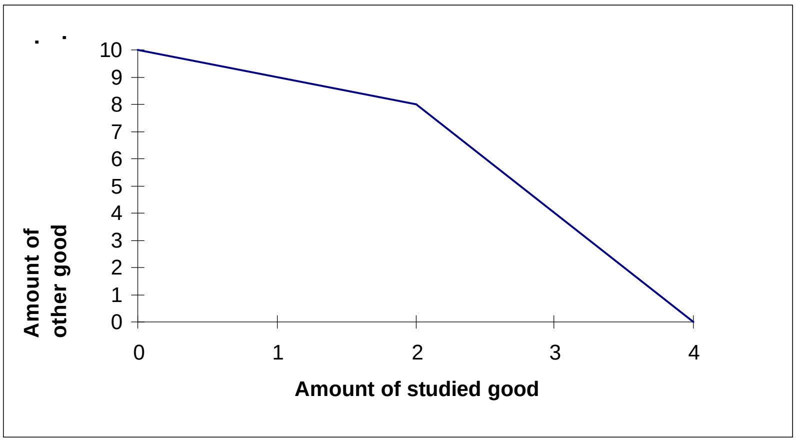

figure 1: Budget set of a consumer

The consumer's decision problem can be stated as follows: The consumer's utility function U is a function of the amount of the good of our interest, denoted by q_1 and of other goods, generically denoted by q_2. Assume that good q_1 is normal and that prices of good q_1 are given:

$P(q_1) = p_1 \quad (q_1 \leq k)$
$P(q_1) = p_2 \quad (q_1 > k)$

Here is k the value of q_1 at the kink point (k = 2 in figure 1). p_1 and p_2 are the prices of q_1, with $p_2 > p_1$. The price of q_2 is normalized at 1. The consumer wishes to maximize U(.) with respect to its budget restriction. The consumer's budget constraint is nonlinear in q_1

(1) $\min(k,q_1)\, p_1 + \max(q_1-k,0)p_2 + q_2 \leq y$

y is the consumer's income. Expression (1) can be rewritten into

(2) $q_1 p_1 + q_2 \leq y \qquad (q_1 \leq k)$
$q_1 p_2 + q_2 \leq y + k(p_2 - p_1) = y_v \qquad (q_1 > k)$

y_v is called the virtual income of the consumer on the region $\{q_1 > k\}$ and corresponds with the intersection of the extrapolation of the steepest line in figure 1 with the Y-axis (remember that the price of q_2 is normalized at 1). The indirect utility function of the consumer is denoted by $V=V(y,P(q_1)) = U(g(P(q_1),y)\ ,\ y-g(P(q_1),y))$, with g the standard demand function. Now the demand of good q_1 (simply denoted by q_1) is equal to:

(3)

$g(p_1,y)$	if $g(p_1,y) \leq k$	and	$V(y,p_1) > V(y_v,p_2)$
$g(p_2,y_v)$	if $g(p_2,y_v) > k$	and	$V(y_v,p_2) > V(y,p_1)$
k	else		

It can be verified that the latter conditions in the first two cases are redundant. Thus, the demand of q_1 can be written as (Moffitt, 1986):

(4) $q_1 = d_1\, g(p_1,y) + d_2\, g(p_2,y_v) + (1-d_1-d_2)\, k$

$$d_1 = I\ \{k > g(p_1,y\ )\}$$

$$d_2 = I\ \{k < g(p_2,y_v)\}$$

with I{.} the indicator function of the expression between brackets. The tedious expression (4) has a simple graphical interpretation, see figures 2 and 3. Both figures show two graphs. One graph is a combination of two standard inverse demand functions[1], the other graph represents the marginal price as a function of demand. The intersection between these two graphs takes place at the amount q as in expression (4). The situation in both figure 2 and 3 is a marginal price of 200 for the first ten units and a marginal price of 400 for a demand of more than ten units. In figure 2 demand lies in the first segment, while the demand in figure 3 sticks at the kink.

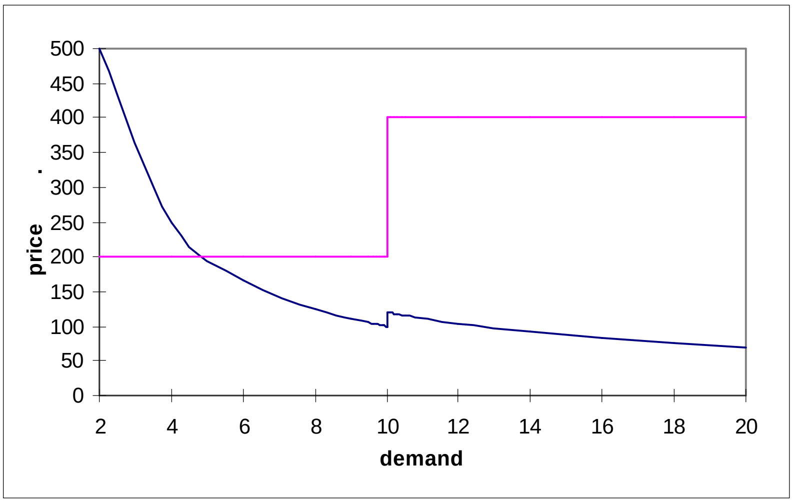

figure 2: Illustration of demand in a segment

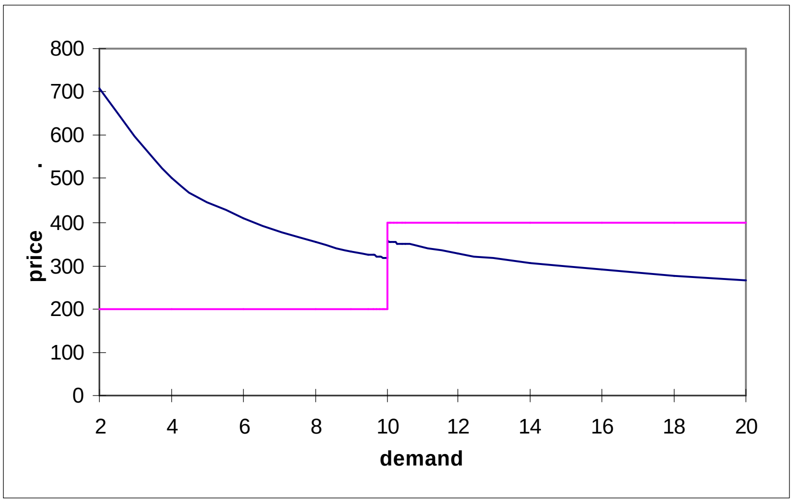

figure 3: Illustration of demand sticking at a kink

[1]one for y (the left part) and one for y_v

The behaviour of q_1 when it is equal to k follows from the comparative statics of q_1, which can be obtained by taking the first derivatives of q_1 with respect tot p_1, p_2, and y. It can easily be shown[2] that both dq_1/dp_1 and dq_1/dp_2 are nonpositive and dq_1/dy is nonnegative. New is the possibility of zero effects in the case $q_1=q^*$. When both d_1 and d_2 are zero, it is of course possible that their value remains unchanged after a perturbation of (p_1,p_2,y). Therefore, the choice of a consumer may stick at a kink. Consequently, the output value $q_1=k$ corresponds to some region of input variables. This can have important implications for a distribution of water demand: a kink will generally correspond with a mass point (this will be shown below). Note that the value of q_1, when it is on segment 1, is completely unaffected by any change in p_2 and vice versa. Note that all these observations can also be interpreted graphically. For instance: a small perturbation of the demand function in figure 3 does not change the fact that demand sticks at a kink. Finally: the jump of the standard demand function is caused by a rise in virtual income of $q^*(p_2-p_1)$.

Statistical models

When we take the micro-economic model and its properties into account, one can expect a nonlinear relation between the demand for water and other input variables as income and marginal price. Furthermore, one might expect a concentration of data around a kink.

Consider the following specification for the demand equation, ignoring the micro-economics leading to expression (4):

$$(5) \qquad q = \beta X + \gamma p_m + \delta y_v + \varepsilon$$

In (5), X is some vector of other explanatory variables, The variables β, γ and δ are unknown parameters. ε is the error term. Equation (5) seems linear in both p_m and y_v , but it is not since both p_m and y_v are functions of q[3]. For a set of realisations of the variables in (5), ordinary least squares (OLS) could be used, ignoring the nonlinearity. The advantage of this method is its simplicity. The disadvantages are however:

- p_m and y_v are functions of q and are therefore correlated with the disturbance term ε. Additionally, in our case, p_m is globally positively correlated with q, so OLS would lead

[2]Under the assumption that good 1 is normal

[3]For example, take p_m in figure 3: p_m = 200 if $q \leq 10$ and p_m = 400 if $q > 10$.

to a positive price elasticity. However, it makes more sense that the price elasticity is negative.

- An implicit assumption is that the observed demand lies on the same segment or kink as the true demand. Therefore, OLS is not a flexible setting
- The OLS setting ignores the possibility of concentration of data around a kink.

The first disadvantage could be removed by applying an instrumental variable technique (see Hanemann (1995) and Blomquist (1996) for a discussion), but the second and the third argument cannot be solved in this way. Results of these estimation methods can be found in the appendix.

The Burtless and Hausman model (Maximum likelihood)

The Burtless and Hausman model is an extension of the micro-economic theory as presented above. Starting point of the model is equation (4), which will be extended with two error terms:

(6) $$q = d_1\,(g(p_1, y) + \alpha) + d_2\,(g(p_2, y_v) + \alpha) + (1-d_1-d_2)\,k + \varepsilon$$

$$d_1 = I\,\{k > g(p_1, y) + \alpha\}$$

$$d_2 = I\,\{k < g(p_2, y_v) + \alpha\}$$

The error term ε is a common measurement error term. To take account of heterogeneity of preferences an extra error term α is added[4]. An important consequence of this setting is the notion that d_1 and d_2 are now unobserved, hence it is allowed that the location of the observed demand is not the same as the location of the utility-maximizing kink or segment[5]. When the (joint) distribution of α and ε is specified, the likelihood for one observation can be derived, which will be done for our (special) case in section 4. The formulation stated above will be referred to as the two-error model. When α is omitted (6) reduces to the one-error model and leaving ε outside (6) is called the heterogeneity-only model.

---

[4]More general, a stochastic demand function $g'(P,M,\alpha)$ is introduced. In our case $g'(P,M,\alpha) = g(P,M)+\alpha$.

[5]This idea is presented in Burtless and Hausman (1978)

What happens with the concentration of data around k ? The equation q = k (ignoring ε), holds when both d_1 and d_2 are zero. This is equivalent to:

$$k - g(p_2,y_v) < \alpha < k - g(p_1, y)$$

So, for a given realisation of ε, demand sticks at k for an interval of realisations of α. The same insight can be archieved by looking at figure 3. A small perturbation of the demand function does not change the x-coordinate of intersection, while it does in figure 2. This property of the Burtless and Hausman model gives some flexibility for the case of a concentration of data around a kink.

Another specification, which appears to be useful is

$$X = (d_1\, g(p_1,y)\, e^{\alpha} + d_2\, g(p_2,y_v)\, e^{\alpha} + (1-d_1-d_2)k)e^{\varepsilon}$$
$$d_1 = I\,\{k > g(p_1, y)\, e^{\alpha}\}$$
$$d_2 = I\,\{k < g(p_2, y_v)e^{\alpha}\}$$

This specification can be simplified by taking logarithms:

$$\ln(q) = \ln(d_1\, g(p_1,y)\, e^{\alpha} + d_2\, g(p_2,y_v)\, e^{\alpha} + (1-d_1-d_2)k) + \varepsilon$$

Now observe that exactly one of d_1 , d_2 and $1-d_1-d_2$ is equal to one (and both others are zero).
This allows us to write

(7) $$\ln(q) = d_1\,(\ln(g(p_1, y\,)) + \alpha) + d_2\,(\ln(g(p_2,y_v\,)) + \alpha) + (1-d_1-d_2)\ln(k) + \varepsilon$$
$$d_1 = I\{\ln(k) > \ln(g(p_1, y\,)) + \alpha\}$$
$$d_2 = I\{\ln(k) < \ln(g(p_2, y_v)) + \alpha\}$$

Note that after replacing in (6) q, k , and g() by their logarithms, one obtains (7). This alternative specification appears to be useful in the case when g(p,y) is of the form: $a\,p^b y^c$.

This reduces to a linear specification for ln(g(.)): ln(g(p,y)) = ln(a) + b ln(p) + c ln(y).[6] The derivation of the likelihood is similar for both expressions.

Before applying the Burtless and Hausman model to our observations in section 4, we discuss some properties of the data in section 3.

3. Description of the data

The dataset that is used in this study is part of a survey carried out in 1994 about demand for urban facilities in Indonesian medium sized cities (Supramono and Wijayanto, 1995). In a survey among 951 households in the Central Java city of Salatiga (population about 100.000 persons) it is found that about 50% has a private connection to the piped water system of the regional water company. In addition, about 6% gets water from a neighbour who has such a connection, and about 8% gets water from a public terminal serviced by the same company. In the present piece of research we only focus on those households having a private water connection. The number of useful observations is 220.

This section presents some descriptive statistics of relevant variables, and examines some relationships between these variables. The empirical distribution of water demand is examined. An important related question is the presence of local maxima in the distribution at the kinks. This would give an extra justification of using the Burtless and Hausman model.

Some descriptive statistics of the 220 observations and mutual correlations of three important variables can be found in tables 1 and 2.

---

[6]Note the similarity with ordinary least squares: Obtaining nonlinear least squares estimators for a, b and c when $g(p,y) = a\, p^b\, y^c$ is rarely done because this specification is nonlinear in the parameters b and c. After taking logarithms the least squares problem becomes trivial (however, the original problem is not solved). An interesting question is which of the specifications (6) and (7) works best (in a computational point of view) in the case $g(p,y) = a\, p^b\, y^c$.

	demand/month (m^3)	income/month (rupiah[7])	size of household
mean:	18.70	166872	4.827
standard deviation:	10.69	105998	2.121
coefficient of variation:	0.571	0.635	0.439
median:	17	150000	5.0
0.25 quantile:	10.0	100000	3.0
0.75 quantile:	22.0	200000	6.0
minimum:	10.0	30000	1.0
maximum:	100	650000	15.0
skewness:	3.411	1.738	0.914

Table 1: Descriptive statistics of demand related variables

	demand	**income**	**size of household**
demand	1.000	0.060	0.438
income	0.060	1.000	0.114
size of household	0.438	0.114	1.000

Table 2: Correlations between demand related variables

One can see that the distributions of water demand and income are highly skewed to the right. The correlation between demand and income is low.

Some results about water demand that cannot be found in the table are of interest. About 37 percent of the data (82 observations) is equal to ten, twenty of thirty, and are thus lying on one of the kinks. This fact may be due to rounding errors, but gives some extra support for use of the Burtless and Hausman model. To get more insight into the distribution of water demand, a kernel estimation of the density will be performed. A kernel density estimator, for observations $x_1,\ldots,x_n$ has the following form:

$$z(t) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{t - x_i}{h}\right)$$

The function K, a density function with zero expectation and variance equal to one, is called the kernel and h is a scale parameter called the bandwidth. For large values of h, one

[7]One dollar is about 2000 rupiah

gets a smoothed estimator. Figure 4 shows the kernel estimator of the density of water demand. The normal kernel (that is, for K the standard-normal density function) is used with h=2 [8].

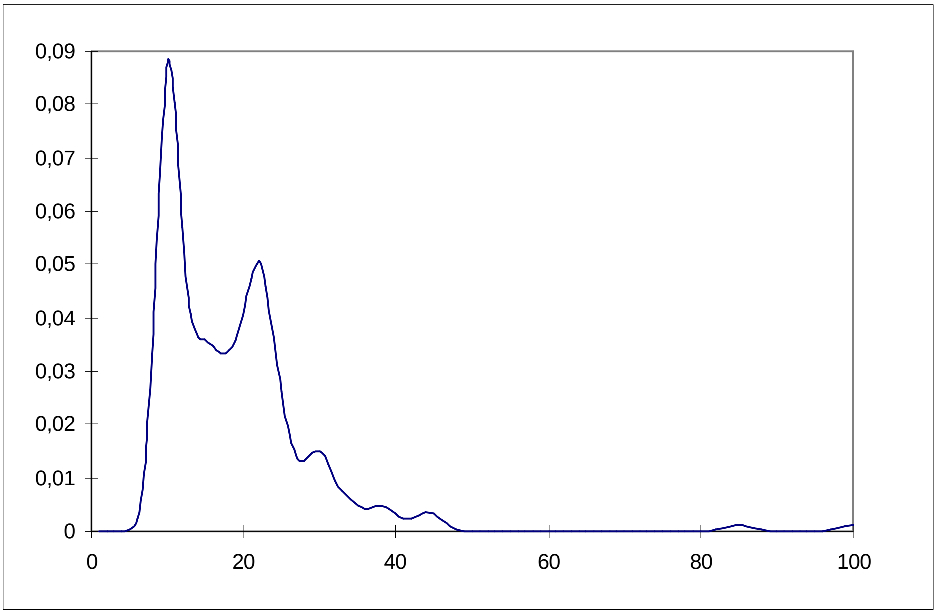

figure 4:Kernel estimator of density water demand

Clearly the distribution is not unimodal. Local modes are clearly around ten and twenty and maybe also around thirty. This observation is consistent with the micro-economic theory as in section 2. A parametric estimator of the density can be obtained by estimation of the parameters of a lognormal distribution or a pareto distribution (of course, other distributions are also possible). The pareto distribution provides the best loglikelihood with a place parameter of 10 and scale parameter 1.98. To take the multimodality into account, a mixture of standard density functions can be used. A convex combination of two pareto distributions has the following density:

$$f(x) = I\{x \geq x_a\}\ p\ \theta\ x_a^{\theta}\ x^{-\theta-1} + I\{x \geq x_b\}\ (1-p)\ \mu\ x_b^{\mu}\ x^{-\mu-1} \qquad \theta, \mu > 0 \quad 0 \leq p \leq 1$$

[8]Under the assumption that the true distribution is normal, the (according to the minimum mean integrated square error) optimal bandwidth is about 3.3, however the true distribution is clearly not normal and a band-width of 3.3 was found too big. For a discussion about various types and properties of Kernel estimators, see Härdle (1993).

Taking x_a = 10 and x_b = 20 and applying maximum likelihood estimation gives the estimation result p = 0.605, θ = 4.42, μ = 3.80. The gain in loglikelihood compared with a single pareto distribution is 36 points, enough to reject the standard pareto distribution in a Likelihood Ratio test. In particular, we can reject unimodality since the density function is not unimodal.

4. A discretized modification of the Burtless and Hausman model

The results in the previous section give some extra support in applying the Burtless and Hausman model, more particular the incorporation of an error term representing heterogeneity of preferenes. The two error model will be applied[9] with expression (6) in section 2 as a starting point. We assume that both the heterogeneity error α and measurement error ε are normally distributed and that their correlation is zero[10] For this case, derivation of the likelihood for one observation is straightforward.

In our case the likelihood consists out of nine components, which can be divided into two groups. First, we extend the model in section 2 to the presence of three kinks. While the demand is bounded from below by zero and bounded from above by the budget constraint, the distribution is censored[11], which is equivalent to adding two kinks at zero and at the maximum possible consumption of water (denoted by max). This results in four segments and five kinks. Table 3 shows those segments and kinks and the corresponding marginal prices:

---

[9]Since so many observations are lying at a kink, the heterogeneity-only model can also be chosen. For a discussion see Moffitt (1986). However, we will see later dat measurement error is certainly not degenerated.

[10]See Hewitt and Hanemann(1993) for the generalized case. Moffitt (1986) notes that there is no a priori rationale for the two different kinds of error to be correlated - and we agree. Correlation between the two errors is omitted in most of the previous studies (Moffitt 1986)

[11]A truncated distribution is also possible, for a discussion see Pudney (1989)

segment	marginal price (rupiah)	interval	kink	value
			1	0
1	215	(0,10)	2	10
2	280	(10,20)	3	20
3	370	(20,30)	4	30
4	430	(30,max)	5	max[12]

Table 3: Kinks, segments and marginal prices

The loglikelihood can be written as the sum of densities of nine disjoint events:

(8) $$L(\theta \mid q,X,y) = \sum_{i=1}^{4} S_i(\theta. q, X, y) + \sum_{j=1}^{5} K_j(\theta, q, X, y)$$

S_i and K_j correspond to the contribution to the likelihood of the event that the true demand lies at segment i or kink j given the observed data and the parameter values, summarized with θ. Expressions for S_i and K_i can be derived in the same way as expression (7) in section 2. Denote the density functions of α and ε with f_α and f_ε , the value of the demand in the k-th kink with q_k and the j-th marginal price with $p_{m,j}$. Then:

(9) $$S_i = \int_{a_i}^{b_i} f_\alpha(\alpha)\, f_\varepsilon(q - g(p_{m,i}, y_{v,i}, X, \theta) - \alpha)\, d\alpha \qquad (i=1,...,4)$$

$$K_j = f_\varepsilon(q - q_i) \int_{b_{j-1}}^{a_j} f_\alpha(x)dx \qquad (j=1,...,5)$$

$$a_i = q_{i-1} + g(p_{m,i}, y_{v,i}, X, \theta) \qquad (i=1,...,5)$$

$$b_i = q_i + g(p_{m,i}, y_{v,i}, X, \theta) \qquad (j=0,...,4)$$

With b_0 and a_5 we mean minus infinity and infinity, respectively. q_i is equal to the consumption level in kink i. Both types of expressions S and K can be rewritten, by some manipulation with normal densities, such that for practical purposes only the cumulative

[12]max is equal to $y_{v,4}$ / $p_{m,4}$

normal distribution is required. The final step is specifying the demand function g and the parameter set θ. In the case of a linear demand function (in p_m and y_v) we would have:

$$(10)\quad g(p_m,y_v,X,\theta) = \beta X + \gamma p_m + \delta y_v \qquad \theta = (\beta, \gamma, \delta, \sigma_\alpha, \sigma_\varepsilon)$$

By applying expressions (8), (9) and (10), a maximum likelihood estimator can be obtained. Since the log-likelihood is not globally concave, the choice of a good set of starting values is important.

In estimating the two-error model we encounter another problem: the value of the (log)likelihood goes towards infinity. Analysis of this phenomenon reveals that this can be established by taking σ_ε arbitrarilly small. Numerical evidence can be found in table 4.

value σ_ε	value loglikelihood
0.5	-737.4
0.5^{4}	-588.3
0.5^{16}	91.9
0.5^{100}	4866

Table 4: Unboundedness of the loglikelihood

An explanation of this unexpected property can easily be found: consider an observation lying exacty on a kink, for instance the kink at 10. Then the value of the loglikelihood for that observation can be made arbitrarily high by sending σ_e to zero. The value of the loglikelihood, more precisely the values of S_i, i=1,...,4 , for other observations can be bounded from below by holding $\sigma^2_\alpha + \sigma^2_\varepsilon$ constant. Under some weak assumptions (which are valid in our case), a more formal statement can be made by the following theorem, which states that even when there is only one observation lying exactly on a kink the maximization problem is unbounded.

Theorem 1: *Let* $\mathbf{q} = (q_1,...,q_n)$ *be a sample of n observed demands and let* $\mathbf{X} = (X_1,...,X_n)$ *and* $\mathbf{y}=(y_1,...,y_n)$ *be samples of observed explanatory variables and incomes. Let*

$\ell(\ ) = \prod_{i=1}^{n} L(\ \mid q_i, X_i, y_i)$ *be the loglikelihood corresponding to the sample* ***(q, X, y)*** *with the function L the loglikelihood for one observation defined as above.   , shorthand for (,   ,   ,   ,   ) is a parameter vector lying in the parameter space   =* R^n R_{++} R_{++}*[13]. Assume the demand function g(, , ,*X_i*,*y_i*,p) is strictly decreasing in p for some   and all observations[14] and that there is an observation i with the observed demand* q_i *lying exactly on a kink. Then the maximization problem*

$$\max \{ \ell(\theta) \mid \theta \in \Theta \}$$

is unbounded.

The proof of this theorem for the case of normally distributed error terms and one kink can be found in Appendix A. The theorem can be generalized for more kinks and even for other continuous distribution functions for α and ε when σ_α and σ_ε are scale parameters.

Of course, one wonders why the result of theorem 1 has escaped the attention of the many researchers who have used maximum likelihood methods to deal with kinked budget constraints and who have investigated its econometric properties[15]. The probable answer is that the large majority of these studies refer to labour supply. In that context the probability of observing an individual who is located *exactly* at a kink is extremely small. It would require the taxable income to be exactly on the point where the marginal rate shifts, which may only be possible if the number of working hours equals a specific number that will only by accident be an integer. However, in the present context the kinks refer to demand that equals a multiple of 10, and since the water meters indicate only integers, it is highly probable that at least some observations are exactly at a kink point.

---

[13] $R_{++} = (0,\infty)$

[14] In our case this means taking $\gamma<0$ in (10)

[15] In addition to the work of Burtless and Hausman [1978], and Moffitt [1986] mentioned above, we refer to Blundell [1993] for a review of structural labor supply analysis and MaCurdy, Green and Paarsch [1988], van Soest, Kapteyn and Kooreman [1993] and Blomquist [1995] for some debate on the appropriate econometric technique for dealing with kinked budget constraints.

To avoid this problem, we assume that observations, now denoted by q*, follow a discrete distribution. More in particular we assume that the observed demand q* is the output of a rounding procedure: a value q is rounded to the nearest integer q*. The new likelihood becomes:

$$(11) \quad L^{*}(\theta \mid q^{*},X,y) = \int_{q^{*}-0.5}^{q^{*}+0.5} L(q,\theta,X,y)dq$$

The fine that has to be paid for this is the inevitable need for a numerical integration procedure. The results presented in the next section are derived by applying the trapezium rule. Note that this discretization technique is not just a technical solution to deal with the unboundedness of the loglikelihood. The observations are really discrete since the water-meters only yield water consumption in integers.

5. Estimation results

In this section, estimation results of the two error discretisized Burtless and Hausman model are presented. Both a linear (corresponding with (6) in section 2) and a logarithmic (with (7) as starting point) specification are considered. After evaluating both loglikelihoods and taking the Jacobian into account the logarithmic specification[16] performed 30 points better, so the starting point is the following equation for the logarithm of water demand:

$$(12) \quad \ln q = \beta X + \gamma \ln p_m + \delta \ln y_v + \alpha$$

Table 5 gives results (standard errors between brackets). The number of members of a household will be denoted by H. D is a dummy variable with value one if a household has an extra water source (usually the river or a private well) and value zero elsewhere. Unfortunately, no data are available on whether or not neighbours are co-consumers of the water.

[16]In the logarithmic case, the extra kink for zero demand can be omitted, because the logarithm of zero is equal to minus infinity.

$$\ln(q) = 8.764 + 0.526 \ln(H) - 0.220\, D - 1.280 \ln(p_m) + 0.501\, 10^{-6} \ln(y_v) + 8.604\, 10^{-3} \ln(p_m)\ln(y_v)$$

$$(1.22) \quad (0.099) \quad (0.142) \quad (0.235) \quad (0.348)\, 10^{6} \quad (1.164)\, 10^{-3}$$

$$\alpha \sim N(0\,,\, 0.558^2) \qquad \varepsilon \sim N(0\,,\, 0.020^2)$$

$$(0.036) \qquad\qquad (0.0052)$$

loglikelihood = -681.06

Table 5: Estimation results

The value of the coefficient of D differs significantly from zero when a likelihood ratio test is performed. Water demand becomes 22 percent smaller when an extra water source is available. The only coefficient which does not differ significantly from zero is the coefficient for the virtual income y_v. This is also the conclusion when the product $\ln(p_m)\ln(y_v)$ is left out of the specification. This product term is significant and adds an extra richness to the model in the sense that the price elasticity is not constant anymore. Table 6 shows various values for the price elasticity for a given value of the (virtual) income. The mean price elasticity is -1.176. Thus we arrive at the conclusion that water demand in Salatiga, Indonesia, is strongly sensitive to the price of water. The values of the price elasticities obtained here are higher (in an absolute sense) than the of about -0.3 to -0.9 mentioned by Nieswiadomy and Molina (1989). On the other hand Hewitt and Hanemann (1995) arrive at an elasticity of about -1.6, and mention some other studies leading to elasticities of water demand clearly higher than one (in absolute sense). Most of the studies in this field relate to the US, where indoor water consumption is thought to be rather price independent whereas outdoor consumption (for watering lawns) is much more price dependent. This context cannot be easily transferred to a developing country such as Indonesia where lawn watering is not usual.

income (rupiah)	price elasticity
0	-1.280
30000 (minimum)	-1.191
166872 (mean)	-1.176
650000 (maximum)	-1.164

Table 6: Price elasticity as a function of income

Another advantage that appears when the product term is added is the fact that income effects are clearly positive. The mean income elasticity[17], is equal to 0.049 and its range lies between 0.046 and 0.052.

The ratio between variances of the disturbance terms gives an indication of the amount of clustering of data around a kink (Moffitt, 1986). The ratio of heterogeneity to measurement variance is 778, which is large compared to previous results (see again Moffitt 1986, table 1). This result indicates a clustering of data around kinks and coincides with what we found in section 3. Although the estimator of the measurement variance is small, its standard error is small enough to state that the measurement error term is nondegenerate.

In Appendix B we present the estimations with OLS and instrumental variables. The results are implausible and quite different from those obtained with the maximum likelihood method presented here.

6. Welfare Effects of a Uniformization of the rate structure

The idea behind the rate structure for water users seems clear. Households with small budgets and - presumably - small water consumption are favoured because the marginal price is relatively low for small consumption levels. It appears from our estimation results that income has a significant, but, nevertheless, only small effect on water consumption, whereas household

[17]The mean income elasticity is defined as the income elasticity for the mean marginal price observed in the sample.

size has a more substantial influence. This implies that the main effect of the current rate structure may well be that larger households have to pay a higher marginal price than small households, whereas income is an unimportant determinant of the marginal price. Such an effect is almost surely unintended, and probably undesired, by the authorities that determine the price schedule. It is therefore of some interest to investigate the welfare consequences of an uniformization of the rate structure.

The appropriate tool for doing so is the cost function, or the closely related indirect utility function. Either of these allows us to determine the compensating variation, that is minus the change in income that is required to keep a household at the same indifference curve after a change in the price schedule for water has occurred. However, the cost function and the indirect utility function that are associated with the demand function estimated here is unknown, and the method proposed by Hausman [1981] does not lead to an (easily) solvable differential equation. We will therefore make use of the procedure developed by Vartia [1983], which requires only knowledge of the ordinary (Marshallian) demand functions. In our case, the system of demand functions consists of two equations: one for water, the other for a (Hicksian) composite consumption good. All conditions required by Vartia's method are satisfied if the demand equation for water satisfies the Slutsky-condition:[18]

$$\frac{\partial q}{\partial p_m} + q\frac{\partial q}{\partial y_v} \; < \; 0$$

In order to check the validity of this condition, we computed this Slutsky term on the basis of our estimation results. The computations involved are complicated by the fact that observed demands depend on the random variables ε and α. We dealt with this problem as follows. For the purposes of the present section we assume that all realizations of the ε's are equal to zero. Since the estimated variance of this error term is very small (see Table 5), the associated error is negligible. With respect to α, the random term that refers to taste variation, we can now determine the interval in which it must lie, given the observed demand for water and the presumed rounding procedure. The Slutsky term was then computed for both the lower and upper bound of this interval, taking into account that for demands located on the kinks the

[18]This conclusion is motivated, in the present context, in an appendix that is available upon request from the authors.

marginal prices and virtual incomes are different for the lower and upper bounds. The computed Slutsky terms are all negative. Hence Vartia's procedure can be used.

An illustration of the outcome of the computation is given in figure 5. Under the current rate structure the household is on the second segment of the kinked budget line. A change to the uniform price per unit, while keeping the utility level constant, implies a move along the original indifference curve until the point where the marginal rate of substitution corresponds with the uniform price for water. The straight line that touches the indifference curve at this point is the budget restriction that would lead the utility maximizing individual to choose that point as its optimum. The income implied by this budget line, denoted as $\overset{*}{y}$ is the income that would be needed to make the consumer indifferent between the current rate structure and the uniform price. The difference between actual income y and $\overset{*}{y}$ is the compensating variation. If the compensating variation is negative, a higher income is needed to keep the consumer on the same indifference curve after the uniformization of the rate structure. If it is positive, as is the case in the situation shown in figure 5, a lower income would suffice. Clearly, in the former case the change is to the disadvantage of the household, in the latter to its benefit.

We considered two possibilities for uniformization of the rate structure. In the first the new uniform price was computed at such a level that total demand remains unchanged. This price is equal to Rp 296.10 per m^3. In the second the new price per unit equals the average price paid under the current rate structure by the households in the sample, which is equal to Rp 263.40 per m^3 .[19] The first procedure has the advantage that no assumptions about the cost structure of the water company have to be made. The disadvantage is that we have to assume how the additional profits are used. For the purpose of these computations it is assumed that the benefits were distributed equally over the household connected to the water pipeline. The Indonesian context is that the profits of water companies are used for public provisions that benefit the whole population. This means that a part of these additional profits will be spent in favour of

[19]In all computations we took the α's first to be equal to the lower bounds that are compatible with observed demands, then to the upper bounds, and averaged the results. Remember that all ε's have been put equal to 0.

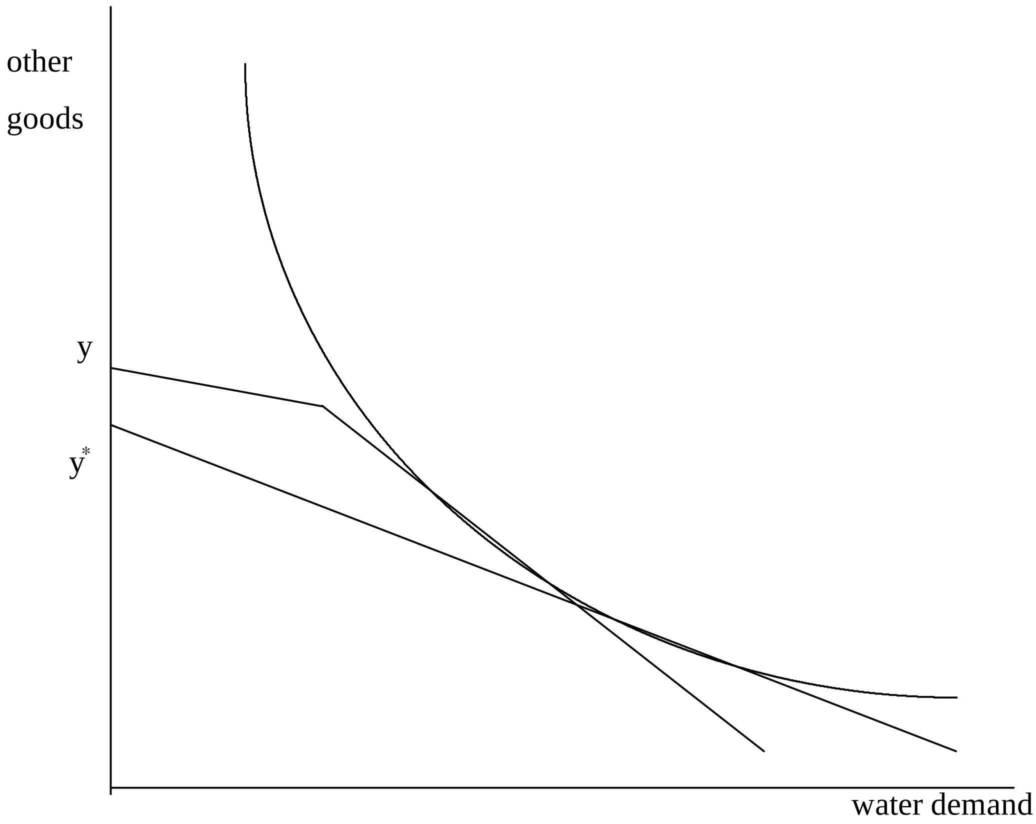

figure 5: Welfare effect of a uniformization of the price of water

households not connected to the water pipeline. However, it is difficult to assess the effects of such expenditures, and we have taken the redistribution of the additional profits among households not connected to the water pipeline as a kind of benchmark. Actual benefits may be lower or higher than the figures we computed, depending on the way the additional profits are spent.

The second procedure is only reasonable if water supply can be increased without a considerable increase in average cost. If the water company operates under constant returns to scale, has no fixed costs and makes no profit, then the compensating variations computed give the complete welfare effect. However, it seems plausible that the water companies have considerable fixed costs, whereas marginal costs are probably low and independent of the volume of production in a wide range. This means that an increase in the production at a constant average price will lead to an increase in the profits of the water companies. The compensating variations that have been computed for this case are therefore most probably underestimates of the total welfare effects.

The averaged compensating variations that have been computed for various subgroups of the population are listed Table 7. For both uniformizations the average compensating variation is positive, and for both uniformizations the group of households with a higher than average income gains more than the group with a lower than average income. However, the decomposition into quartile groups reveals that the welfare effect is not monotonous in income: although the averaged effects for all quartile groups are positive, the second and fourth quartile gain less than both the first and third. The effect of uniformization on households of various size is close to monotonous, and larger households gain most, as we expected.

			% of households with	
Household group	Number	Effect (Rp/month)*	welfare gain (%)	higher demand (%)
uniform price Rp 296.09				
all	220	207.79	.35	.37
*income***				
first quartile	52	235.00	.38	.38
second quartile	51	36.83	.29	.29
third quartile	74	369.51	.36	.41
fourth quartile	43	99.35	.33	.42
household size				
single person	4	-106.38	.25	.25
two persons	25	31.41	.16	.16
three persons	29	- 82.25	.17	.17
four persons	49	- 68.72	.18	.18
five persons	41	146.98	.44	.51
six persons	32	597.02	.50	.53
seven persons	16	655.86	.56	.63
eight or more p.	24	645.12	.58	.67
extra water source available				
yes	20	425.46	.20	.25
no	200	186.02	.36	.39

Table 7: Welfare effects of uniformization of the rate structure for two different price levels

Household group	Number	Effect (Rp/month)*	% of households with welfare gain (%)	% of households with higher demand (%)
uniform price Rp 263.40				
all	220	347.39	.38	.70
*income***				
first quartile	52	381.55	.38	.67
second quartile	51	66.66	.29	.67
third quartile	74	570.74	.41	.69
fourth quartile	43	254.68	.42	.81
household size				
single person	4	-233.17	.25	.50
two persons	25	-101.94	.16	.40
three persons	29	-212.63	.17	.52
four persons	49	-137.09	.18	.59
five persons	41	370.98	.51	.85
six persons	32	1018.27	.53	.91
seven persons	16	1147.02	.63	.81
eight or more p.	24	1110.17	.67	.92
extra water source available				
yes	20	492.84	.25	.45
no	200	332.85	.39	.73

* Average compensating variation for the relevant group.
** Due to a large number of equal (reported) incomes, the number of households per quartile differs.

Table 7 (continued)

The effect of availability of an extra water source is somewhat unexpected: households with an extra source gain more than households without an extra source. The probable explanation is that the households with an extra source are often large households.

It is important to note that for both levels of a uniform price, the average welfare effect is clearly positive. However, in terms of the percentage of households that benefit from a uniformization it is only a minority of households (35-38 %) that are net beneficiaries. Thus, there is a rather large group of households that experience a small welfare loss, and a smaller group of households that experience a relatively large welfare gain. The beneficiaries can mainly be found among the large households. If we would carry out the analysis per individual instead of per household, the share of beneficiaries is considerably higher.

The following observations are relevant for a comparison of the two uniformizations considered. The average welfare gains of a uniformization to the current average price are much larger than those of a uniformization to a price where total demand remains unchanged. It has, moreover, been observed above that the effects of the former are probably an underestimate of the total welfare effect. However, a drawback of the uniformization to the current average price are the stronger negative effects on small households, and this makes it impossible to formulate an unconditional best choice on the basis of the computed welfare effects.

7. Conclusions

In this paper a discretized version of the Burtless and Hausman model is used to estimate water demand. The unconditional distribution of water demand is characterized by a clustering of observations around the kinks. In particular, the distribution of water demand is clearly not unimodal. This property is consistent with the Burtless and Hausman model and related micro-economic theory. The estimation results indicate a large (in absolute value) price elasticity and a small income elasticity. The price elasticity is decreasing (in absolute value) in income. Average price elasticity is about - 1.17.

Section 6 shows that replacing the block rate structure by one marginal price leads on average to an increase in household welfare. When the uniform price level is chosen such that total demand remains unchanged, the water company has higher revenues owing to the higher average price. An equal redistribution of these revenues among households implies an average increase in welfare of Rp 208 per month. The strongest beneficiaries of a uniform price are large households. Differences among income groups are small, however. A similar result is found when a uniform price is introduced at the level of the average price in the block rate system.

We conclude that in the case of Salatiga (Indonesia) , the block rate structure does not reach its aim of helping the poor. In stead, it appears to favour small households, and these include both poor and rich.

Acknowlegdement

The authors Prapto Yuwono and Supramono for useful advice and for making available the data.

References

Billings, R.B.(1987): Alternative Model Estimators for Block Rate Pricing, Water Resources Bulletin, vol. 23, 1987, pp. 341-345.

Blomquist, S. (1996): Estimation Methods for Male Labour Supply Functions; How to take account of Nonlinear Taxes, Journal of Econometrics 70 pp. 383-405.

Blundell, R. (1993) Evaluating Structural Microeconometric Models of Labour Supply, pp. 3-48 in C. Sims (ed.) Advances in Econometrics, Sixth World Congress, Cambridge University Press, Cambridge.

Burtless, G. and Hausman, J. (1978): The Effect of Taxation on Labour Supply: Evaluating the Gary Income Maintenance Experiment, Journal of Political Economy 86. 1103-1130.

Chichoine, D.L. and G. Ramamurthy, Evidence on the Specification of Price in the Study of Domestic Water Demand, Land Economics, vol. 62, 986, pp. 26-32.

Greene W.H. (1993): Econometric Analysis, Macmillan, New York.

Härdle W. (1993): Applied nonparametric regression, Cambridge University Press.

Hausman, J.A. (1981) Exact Consumer's Surplus and Deadweight Loss, American Economic Review, **71**, 662-676.

Hewitt, J.A. and Hanemann, W.M. (1995): A Discrete/Continuous Approach to Residential Water Demand under Block Rate Pricing, Land Economics May 1995 pp. 173-192.

MaCurdy, Th., D. Green and H. Paarsch (1990) Assessing Empirical Approaches for Analyzing Taxes and Labor Supply, Journal of Human Resources, **25**, 414-490.

Moffitt, R.(1986): The Econometrics of Piecewise-Linear Budget Constraints, Journal of Business & Economic Statistics, July 1986 Vol. 4, No. 3, pp. 317-328.

Moffitt, R.(1990): The Econometrics of Kinked Budget Constraints, Journal of Economic Perspectives Vol. 4, nr. 2 spring 1990 pp. 119-139.

Nieswiadomy, M.L. and D.J. Molina (1989): Comparing Residential Water Demand Estimates under Decreasing and Increasing Block Rates using Household Data, Land Economics, vol. 65, 1989, pp. 280-289.

Pudney, S. (1989): Modelling Individual Choice, The Econometrics of Corners, Kinks and Holes, Basil Blackwell, New York .

van Soest, A., A. Kapteyn and P. Kooreman (1993) Coherency and Regularity of Demand Systems with Equality and Inequality Constraints, Journal of Econometrics, **57**, 161-188.

Supramono and B. Wijayanto (1995): "Permintaan riil Prasarana Kota" (Demand for Urban Infrastructure), Universitas Kristen Satya Wacana, Salatiga Indonesia, 1995.

Vartia, Y.O. (1983) Efficient Methods of Measuring Welfare Change and Compensated Income in Terms of Ordinary Demand Functions, Econometrica, **51**, 79-98.

Appendix A Proof of theorem 1

Before we give a proof of theorem 1 we present expressions for the likelihoods of the two-error model and the heterogeneity-only model. We present the latter case first[20]. Let the likelihood function of the heterogeneity-only model be denoted by $\ell_1(\theta)$.

$$\ell_1(\;) = \prod_{i=1}^{n} L_i^1(\;) = \prod_{i:q_i=k} K_i^1(\;) \prod_{i:q_i<k} S_{1,i}^1(\;) \prod_{i:q_i>k} S_{2,i}^1(\;)$$

$$K_i^1(\theta) = \Phi\left(\frac{k-g_{i2}(\theta)}{\sigma_\alpha}\right) - \Phi\left(\frac{k-g_{i1}(\theta)}{\sigma_\alpha}\right)$$

$$S_{1,i}^1(\theta) = \frac{1}{\sqrt{2\pi}\sigma_\alpha}\exp\left(-\frac{1}{2}\left(\frac{q_i-g_{i1}(\theta)}{\sigma_\alpha}\right)^2\right)$$

$$S_{2,i}^1(\theta) = \frac{1}{\sqrt{2\pi}\sigma_\alpha}\exp\left(-\frac{1}{2}\left(\frac{q_i-g_{i2}(\theta)}{\sigma_\alpha}\right)^2\right)$$

The likelihood function for the two error model is $\ell_2(\theta,\sigma_\varepsilon) = \prod_{i=1}^{n} L_i(\;,\;)$, with

$$L_i(\theta,\sigma_\varepsilon) = S_{1,i}(\theta,\sigma_\varepsilon) + S_{2,i}(\theta,\sigma_\varepsilon) + K_i(\theta,\sigma_\varepsilon)$$

$$K_i(\theta,\sigma_\varepsilon) = \frac{1}{\sqrt{2\pi}\sigma_\varepsilon}\exp\left(-\frac{1}{2}\left(\frac{k-q_i}{\sigma_\varepsilon}\right)^2\right)\left(\Phi\left(\frac{k-g_{i2}(\theta)}{\sigma_\alpha}\right) - \Phi\left(\frac{k-g_{i1}(\theta)}{\sigma_\alpha}\right)\right)$$

$$S_{1,i}(\theta,\sigma_\varepsilon) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_\alpha^2+\sigma_\varepsilon^2}}\exp\left(-\frac{1}{2}\frac{(q_i-g_{i,1}(\theta))^2}{\sigma_\alpha^2+\sigma_\varepsilon^2}\right)\int_{-\infty}^{a_{i,1}(\theta,\sigma_\varepsilon)}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}u^2}\,du$$

$$S_{2,i}(\theta,\sigma_\varepsilon) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_\alpha^2+\sigma_\varepsilon^2}}\exp\left(-\frac{1}{2}\frac{(q_i-g_{i,2}(\theta))^2}{\sigma_\alpha^2+\sigma_\varepsilon^2}\right)\int_{a_{i,2}(\theta,\sigma_\varepsilon)}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}u^2}\,du$$

where:

$$a_{i,1}(\theta,\sigma_\varepsilon) = \frac{\sqrt{\sigma_\alpha^2+\sigma_\varepsilon^2}}{\sigma_\alpha\sigma_\varepsilon}\left(k-g_{i,1}(\theta)-\frac{\sigma_\alpha^2}{\sigma_\alpha^2+\sigma_\varepsilon^2}\big(q_i-g_{i,1}(\theta)\big)\right)$$

$$a_{i,2}(\theta,\sigma_\varepsilon) = \frac{\sqrt{\sigma_\alpha^2+\sigma_\varepsilon^2}}{\sigma_\alpha\sigma_\varepsilon}\left(k-g_{i,2}(\theta)-\frac{\sigma_\alpha^2}{\sigma_\alpha^2+\sigma_\varepsilon^2}\big(q_i-g_{i,2}(\theta)\big)\right)$$

---

[20] For compactness, we represent the loglikelihood for observation i $L(\theta|q_i,X_i,y_i)$ with $L_i(\theta)$ and $g(\theta,X_i,y_i,p_j)$ with $g_{ij}(\theta)$. Another change in notation: with θ we mean $(\beta,\gamma,\delta,\sigma_\alpha)$ because σ_ε appears only in the two-error model. Finally: the index i is for observations and not for segments, which was the case in section 4.

The expressions for $S_{1,i}$ and $S_{2,i}$ are derived from (9), section 4. We simplified the expressions by reducing the two quadratic forms behind the integral sign (f_α and f_ε) to one quadratic form. The tedious but rather simple derivations are omitted.

After these preliminaries, we give the main result of this appendix. This result gives some insight in the relationship between both likelihood functions and Theorem 1 turns out to be a trivial by-product.

Proposition A1 Let m be the number of observations i with $q_i = k$. Then

$$\lim \quad \to 0 \ \left(\sqrt{2} \quad \right)^m \ell_2(\ , \quad) = \ell_1(\ )$$

pointwise for all $\theta \in \Theta$.

Corollary (proof of Theorem 1) Note first that $m \geq 1$ in this case. Choose θ such that the demand functions $g_i(\theta,.)$ are strictly decreasing in p for i=-1,...,n. The reader should verify that this gives $\ell_1(\theta) > 0$. Now, proposition A1 guarantees that $\lim \quad \to 0 \ \ \ell_2(\ , \quad) = \infty$. In particular, we conclude the unboundedness of the maximization problem $\max\{\ \ell(\theta) \mid \theta \in \Theta\}$. This completes the proof of theorem 1.

Proof of proposition A1 Assume that q_i=k for i=1,...,m.

Let first i>m, so $q_i \neq$k. For symmetry reasons we concentrate ourselves on the case q_i< k. Starting point is the relation $L_i(\theta,\sigma_\varepsilon) = S_{1,i}(\theta,\sigma_\varepsilon) + S_{2,i}(\theta,\sigma_\varepsilon) + K_i(\theta,\sigma_\varepsilon)$. Because q_i < k:

$$K_i(\theta,\sigma_\varepsilon) = \frac{1}{\sqrt{2\pi}\sigma_\varepsilon}\exp\left(-\frac{1}{2}\left(\frac{k-q_i}{\sigma_\varepsilon}\right)^2\right)\left(\Phi\left(\frac{k-g_{i2}(\theta)}{\sigma_\alpha}\right)-\Phi\left(\frac{k-g_{i1}(\theta)}{\sigma_\alpha}\right)\right) \to 0 \text{ when } \sigma_\varepsilon \to 0.$$

The derivation of the limit of $S_{1,i}$ is less trivial.

$$\lim_{\sigma_\varepsilon \to 0} S_{1,i}(\theta,\sigma_\varepsilon) = \lim_{\sigma_\varepsilon \to 0} \frac{1}{\sqrt{2\pi}\sqrt{\sigma_\alpha^2+\sigma_\varepsilon^2}}\exp\left(-\frac{1}{2}\frac{(q_i-g_{i,1}(\theta))^2}{\sigma_\alpha^2+\sigma_\varepsilon^2}\right)\int_{-\infty}^{a_{i,1}(\theta,\sigma_\varepsilon)}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}u^2}\,du =$$

$$\lim_{\sigma_\varepsilon \to 0} \frac{1}{\sqrt{2\pi}\sqrt{\sigma_\alpha^2+\sigma_\varepsilon^2}}\exp\left(-\frac{1}{2}\frac{(q_i-g_{i,1}(\theta))^2}{\sigma_\alpha^2+\sigma_\varepsilon^2}\right)\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}u^2}\,1_{\{u\le a_{i,1}(\theta,\sigma_\varepsilon)\}}\,du$$

The term before the integral converges to $\frac{1}{\sqrt{2\pi}\sigma_\alpha}\exp\left(-\frac{1}{2}\frac{(q_i - g_{i,1}(\theta))^2}{\sigma_\alpha^2}\right)$. To derive the limit of the integral, observe that

$$u \le a_1(\ ,\ \ ) = \frac{\sqrt{\ ^2 + \ ^2}}{\ }\left(k - g_{i,1}(\ ) - \frac{\ ^2}{\ ^2 + \ ^2}\left(q_i - g_{i,1}(\ )\right)\right) \Leftrightarrow$$

$$u \le \frac{\sqrt{\ ^2 + \ ^2}}{\ }\left(k - g_{i,1}(\ ) - \frac{\ ^2}{\ ^2 + \ ^2}\left(q_i - g_{i,1}(\ )\right)\right)$$

Taking the limit for $\sigma_\varepsilon \to 0$ in both sides gives the pointwise convergence of the function $1_{\{u \le a_1(\theta,\sigma_\varepsilon)\}}$ to $1_{\{q_i \le k\}} = 1$ (note that in our case $q_i < k$). Interchanging limit and integration is allowed according to the dominated convergence theorem, hence we have

$$\lim_{\sigma_\varepsilon \to 0} S_{1,i}(\theta,\sigma_\varepsilon) = \frac{1}{\sqrt{2\pi}\sigma_\alpha}\exp\left(-\frac{1}{2}\frac{(q_i - g_{i,1}(\theta))^2}{\sigma_\alpha^2}\right)\int_{-\infty}^{\infty}\lim_{\sigma_\varepsilon \to 0}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}u^2}1_{\{u \le a_{i,1}(\theta,\sigma_\varepsilon)\}}du =$$

$$\frac{1}{\sqrt{2\pi}\sigma_\alpha}\exp\left(-\frac{1}{2}\frac{(q_i - g_{i,1}(\theta))^2}{\sigma_\alpha^2}\right)\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}u^2}du = \frac{1}{\sqrt{2\pi}\sigma_\alpha}\exp\left(-\frac{1}{2}\frac{(q_i - g_{i,1}(\theta))^2}{\sigma_\alpha^2}\right).$$

The same arguments show that $S_{2,i}(\theta,\sigma_\varepsilon) \to 0$ when $\sigma_\varepsilon \to 0$. The first conclusion is that $\lim_{\sigma_\varepsilon \to 0} L_i(\theta,\sigma_\varepsilon) = L_i^1(\theta)$

Assume now that $i \le m$, so $q_i = k$. We wish to calculate

$$\lim_{\sigma_\varepsilon \to 0}\sqrt{2\pi}\sigma_\varepsilon L_i(\theta,\sigma_\varepsilon) = \lim_{\sigma_\varepsilon \to 0}\sqrt{2\pi}\sigma_\varepsilon S_{1,i}(\theta,\sigma_\varepsilon) + \lim_{\sigma_\varepsilon \to 0}\sqrt{2\pi}\sigma_\varepsilon S_{2,i}(\theta,\sigma_\varepsilon) + \lim_{\sigma_\varepsilon \to 0}\sqrt{2\pi}\sigma_\varepsilon K_i(\theta,\sigma_\varepsilon)$$

The first two limits are zero. The third limit is equal to $K_i^1(\theta) = L_i^1(\theta)$ (note that $q_i = k$):

$$\lim_{\sigma_\varepsilon \to 0}\sqrt{2\pi}\sigma_\varepsilon K_i(\theta,\sigma_\varepsilon) = \lim_{\sigma_\varepsilon \to 0}\Phi\left(\frac{k - g_{i2}(\theta)}{\sigma_\alpha}\right) - \Phi\left(\frac{k - g_{i1}(\theta)}{\sigma_\alpha}\right) = \Phi\left(\frac{k - g_{i2}(\theta)}{\sigma_\alpha}\right) - \Phi\left(\frac{k - g_{i1}(\theta)}{\sigma_\alpha}\right) = K_i^1($$

Summarized, we have

$$\lim_{\sigma_\varepsilon \to 0}\left(\sqrt{2\pi}\sigma_\varepsilon\right)^m \ell_2(\theta,\sigma_\varepsilon) = \lim_{\sigma_\varepsilon \to 0}\left(\sqrt{2\pi}\sigma_\varepsilon\right)^m \prod_{i=1}^{n} L_i(\theta,\sigma_\varepsilon) =$$

$$\prod_{i=1}^{m}\left(\lim_{\sigma_\varepsilon \to 0}\left(\sqrt{2\pi}\sigma_\varepsilon\right)^m L_i(\theta,\sigma_\varepsilon)\right)\prod_{i:q_i<k}\left(\lim_{\sigma_\varepsilon \to 0} L_i(\theta,\sigma_\varepsilon)\right)\prod_{i:q_i>k}\left(\lim_{\sigma_\varepsilon \to 0} L_i(\theta,\sigma_\varepsilon)\right) =$$

$$\prod_{i=1}^{m}\left(K_i^1(\theta)\right)\prod_{i:q_i<k}\left(S_{1,i}^1(\theta)\right)\prod_{i:q_i>k}\left(S_{2,i}^1(\theta)\right) = \ell_1(\theta)$$

This completes the proof of the proposition.

Appendix B: Estimation results of OLS and IV models

The results of OLS and IV estimations can be found in tables 9 and 10. The choice of instrumental variables for p_m and y_v are the marginal price and virtual income on the segment predicted by the OLS estimator for ln(q), respectively. The results are terrible. In both cases, the price elasticity is positive. Moreover, also other estimations of parameters are having the 'wrong' sign, for instance the income elasticity and the coefficient of the dummy for an extra water source in the OLS case. We note that the main contribution on both R^2 's comes from p_m: the correlation between ln (p_m) and ln(q) is almost 0.9.

$$\ln(q) = -2.014 + 0.039 \ln(S) + 0.043D + 0.817 \ln(p_m) - 0.431 \ln(y_v) + 0.078 \ln(p_m)\ln(y_v)$$

(5.724) (0.027) (0.039) (1.007) (0.348) (0.085)

R^2 = 0.870

table 9: Ordinary Least Squares estimation of demand function

$$\ln(q) = -1.334 + 0.011 \ln(S) - 0.094 D + 0.671 \ln(p_m) - 0.505 \ln(y_v) + 0.094 \ln(p_m)\ln(y_v)$$

(7.066) (0.032) (0.104) (1.240) (0.595) (0.104)

R^2 = 0.826

table 10: Instrumental Variable estimation of demand function