An Estimation of Residential Water Demand Using co-integration and Error Correction Techniques[1]

Roberto MARTÍNEZ-ESPIÑEIRA[2]

Department of Economics, St Francis Xavier University

Abstract

The purpose of this paper is to measure the short- and long-run effect of the price of water on residential water use. Unit root tests reveal that water use series and series of other variables affecting use are non-stationary. However, a long-run co-integrating relationship is found in the demand model, which makes possible to obtain a partial correction term and to estimate an error correction model. The empirical application uses monthly time-series observations from Seville (Spain). The price-elasticity of demand is estimated as around -0.1 in the short run and -0.5 in the long run. These results are robust to the use of different specifications.

JEL Classification: C22, D12, Q25.

Keywords: seasonal unit roots, residential water demand, price elasticity, time-series, co-integration, Error Correction Model.

[1]The useful cooperation of Céline Nauges with previous drafts of the paper and of EMASESA in Seville for supplying data is gratefully acknowledged. All remaining errors remain the sole responsibility of the author.

[2]Economics, St Francis Xavier University PO Box 5000, Antigonish, B2G2W5, Nova Scotia, Canada. Tel: +1 902 867 5443 Fax: +1 902 867 3610 E-mail: rmespi@stfx.ca

Introduction

While it is generally agreed that there might be substantial differences between short-run and long-run reactions of residential water users to price changes, long-run water demand elasticities in European cities have been rarely estimated. The main purpose of this paper is to estimate short-run and long-run price elasticities of residential water demand using data from Seville (Spain). Monthly time-series data on price and aggregate residential consumption over a ten-year period are matched with climatic data, data on non-price demand policies, and average income. The availability of monthly data allows not only for the use of much more accurate measures of consumption but also to test for seasonal effects in consumption and the peculiarities of dynamic effects that cannot be captured when using yearly data. The analysis is based on the techniques of co-integration (see Engle and Granger, 1987; Johansen, 1988, among others) and error correction (Hendry, Pagan, and Sargan, 1984). To the author's knowledge, no previous published work has applied this econometric methodology to the study of water demand, while it has proved very useful to estimate the demand for other types of transformed natural resources, such as gasoline and electricity. The analysis presented in this paper is similar to these previous studies in that it focuses on a resource whose price could induce adaptations in the purchase patterns of the capital stock (water-consuming durable goods and equipment) and whose consumption might respond partly to habit.

The econometric estimation proceeds in two steps. First, unit root tests are conducted to determine the degree of integration of the main variables. Then, a long-run equilibrium relationship between price and consumption is estimated. The stationarity of the different time-series involved is investigated using seasonal unit root tests. The long-run equilibrium relationship is then used as an error correction term in an Error Correction Model (ECM henceforth). These techniques provide measures of the short- and long-run elasticities as well as the speed of adjustment towards long-run values. The elasticities estimated suggest, as it has been found in the literature, that household water demand is inelastic with respect to its own price but not perfectly so. The results show remarkable consistency between the different techniques used to analyze the dynamics of the relationships.

This paper is organized as follows. Section 1, lists some of the existing studies dealing with the estimation of residential water demand and some applied works that use the techniques of co-integration and error correction. The general characteristics of water demand in Seville are described in Section 2 and the data set is described in Section 3. The econometric methods and the results are presented in Sections 4 and 5 respectively. Section 6 concludes.

1 Background

Residential water demand has been extensively analyzed during the last decades. Most applied studies focus on areas of the USA (e.g. Schefter and David, 1985; Chicoine and Ramamurthy, 1986; Nieswiadomy and Molina, 1989; Renwick and Green, 2000). Some exceptions that use European data are Hansen (1996), Höglund (1999), Nauges and Thomas (2000), and Martínez-Espiñeira (2002). The main objective of this research is to estimate price elasticities of water demand from water demand functions where either individual or aggregate residential water use is made dependent on water price and other variables such as income, climatic conditions and type of residence. Water demand appears as inelastic but not perfectly inelastic. Arbués et al. (2003) and Dalhuisen et al. (2003) provide detailed reviews of the literature.

A number of previous studies have analyzed short-run versus long-run water demand elasticities, finding that short-run elasticities are smaller than their long-run counterparts. This suggests that consumers might need time to adjust their water-using capital stock (durable goods and equipment) and to learn about the effects of their use on their bills (Carver and Boland, 1980). These studies use some type of flow-adjustment model, where lagged consumption is included as one of the explanatory variables. The latter assumes that the actual adjustment to consumption is a fixed ratio of the total *desired* or equilibrium adjustment. The short-run elasticity is then given by a choice of utilization rate of the water-using capital stock while the long-run is defined as the choice of both the size of this capital stock and the intensity of its use. Past consumption is introduced in the model with lags of different length and shape. Carver and Boland (1980), Agthe et al. (1986), Moncur (1987), Lyman (1992), Dandy et al. (1997) are examples of this type of approach. More sophisticated econometric techniques have recently been applied to the dynamic analysis of water demand, including the use of dynamic panel data methods (Nauges and Thomas, 2001) . Normally, lack of data on water-using capital stock prevents the use of stock-adjustment models, although in some cases (e.g. Agthe et al., 1986) a time variable has been used as a crude proxy for the evolution of the capital stock. Renwick and Archibald (1998), using individual-household data, have available information on water related technology and introduce them in a model that explicitly analyzes endogenous technical change. All these studies find that, in agreement with economic theory, short-run responses to price changes are weaker than long-run ones. However, some surprisingly high values for short-run effects have been found. Agthe and Billings (1980), using a linear flow adjustment model, find that the short-run elasticity value (-2.226) is much higher than the long-run value (-0.672). They obtain more *reasonable* results with other methods (such as linear and logarithmic Koyck distributed lag models) but suggest that, with monthly data, there could be an *overreaction* to price changes (a *shock effect* in the short run) and also that other techniques of time series analysis are needed to solve the inconsistency.

None of these studies has used co-integration and/or error correction techniques to estimate the short-run and long-run price effects. These methods have been used in numerous applied studies since the seminal paper of Engle and Granger in 1987. Electricity demand forecasting is among the earliest applications of co-integration (Engle et al., 1989). More recently, co-integration and error correction have been applied to the estimation of energy and gasoline demand. For example, Bentzen (1994), Eltony and Al-Mutairi (1995), and Ramanathan (1999) study the behavior of gasoline consumption in Denmark, Kuwait, and India respectively. Fouquet (1995) investigates the impact of VAT introduction on residential fuel (coal, petrol, gas, and electricity) demand in the United Kingdom, while Beenstock et al. (1999) addresses the issue of seasonality in electricity consumption.

The use of co-integration analysis when estimating demand functions avoids problems of spurious relationships that bias the results and provides a convenient and rigorous way to discern between short-run and long-run effects of pricing policies. One important drawback of this methodology is the lack of power of the unit root tests needed to construct the co-integrating regressions.

2 Water demand in Seville

Residential water use represents about 74% of the demand for drinking water in the Seville and its metropolitan area. The proportion of domestic water use relative to commercial-industrial and institutional use has remained fairly constant during the nineties, with the exception of 1992-93, when the Universal Exposition increased the share of institutional use (*EMASESA*, 2000, pp. 2-3).

The total number of families living in Seville city in 1998 was 226,692 and the water supplier, *EMASESA*, had a total of 190,759 domestic customers at the end of 1998. The number of customers has increased significantly since 1997. This is because the water supplier implemented a campaign (*Plan Cinco*) of replacement of collective meters by individual meters, causing an increase in the average yearly growth of the number of domestic customers from 7% to %10-11 (*EMASESA*, 2000 , p. 4).

According to company's estimates, Sevillan households use 53% of the water in the toilet, in the kitchen, and for washing clothes. These components could be significantly affected by the efficiency of water-using equipment and the frequency of its renewal. An extra 39% is used in showers, which could be determined by both habits and the characteristics of water-use equipment. Outdoor use is minimal (*EMASESA*, 2000, p. 7).

Seville suffered a serious draught during the years 1992-1995, during which important savings were achieved through several measures, such as media campaigns, municipal edicts and the ban of certain uses, water restrictions, and consumption control inspections. At the height of the drought, savings of around 25% with respect to previous years were achieved.

In mid-1992, imbalances between supply and demand started to arise. Media campaigns were launched to ask for voluntary water conservation. Then this was made compulsory, since from September water supply was reduced to 20 hours daily, inducing savings of 15%. Daily water supply was reduced to 16 hours and at the end of 1992 consumption began to reflect a 25% reduction. At the beginning of 1993 the company had to resort to the emergency intakes as the only source of supply. During the first half of 1995, a 28% reduction with respect to the consumption previous to the drought was achieved. Restrictions increased to 10 hours a day. Eventually, the rain came at the end of 1995 and the drought was overcome thanks to the savings achieved in that period. The awareness campaign continued (in spite of the reservoirs having enough water) to maintain the population's saving habits (*EMASESA*, 2000, pp. 6-7). A more detailed description of the measures implemented to reduce demand can be found in the Appendix. See also García-Valiñas (2002).

3 Dataset description

The main data used for the estimations were provided by *EMASESA*, the private company in charge of supplying water and sewage collection services in Seville. They include information for the period 1991-1999 on tariffs, number of domestic accounts, and total domestic use.

The tariff consists of a fixed quota and an increasing three-block rate. Table 1 shows the evolution of the block sizes. The price for the first seven-unit block applies only to those users who use a total of less than seven cubic meters. If the consumer exceeds this level of use, the price of the second block applies also to these first seven cubic meters. This type of *step-rate* structure is in this case explicitly aimed at rewarding water conservation efforts. The rest of the tariff is based on conventional increasing blocks. The tariff includes a water supply fee, a sewage collection fee, and a treatment fee, and, from 1994, a waste-water infrastructure fee (*canon*) was collected on behalf of the Andalusian government. Finally, from 1993 to 1997, a temporary extra fee was charged for the company's finances to recover from the impact of the drought. The value of the fixed quota depends on the size of the meter, but the most common one for domestic users (13 mm) was adopted. The evolution of the prices in each block between 1991 and 1999, including all the elements of the water and sewage bill, is detailed in Table 2. All prices are expressed in constant pesetas (ESP) of 1992, translated into EURO equivalents (1 EURO = 166.386 ESP).

The original data were manipulated into the following variables (where the subscript t refers to Month t):

- Q_t (m^3/capita month) is average per capita domestic water use. The raw data consist of 108 monthly values for total use. The company reads meters quarterly and estimates monthly use in the following manner. The average daily use during the reading period is calculated, then this average use is allocated to each month according to the number of days corresponding to that month in that particular reading period.[3] Annual data on the number of accounts were also collected. However, instead of using this variable to calculate average water use per account, values of total population in Seville were used to calculate monthly average use per capita. The reason is that during the study period the water company substantially increased the

[3] For example, if the reading period goes from 28-04-00 to 03-08-00 and the reading is 91 m^3, since the length of the period is 97 days, average daily use is 0.93 m^3. This average daily use would be multiplied by 2 to obtain April's consumption, by 31 for May, by 30 for June, by 31 for July and by 3 for August.

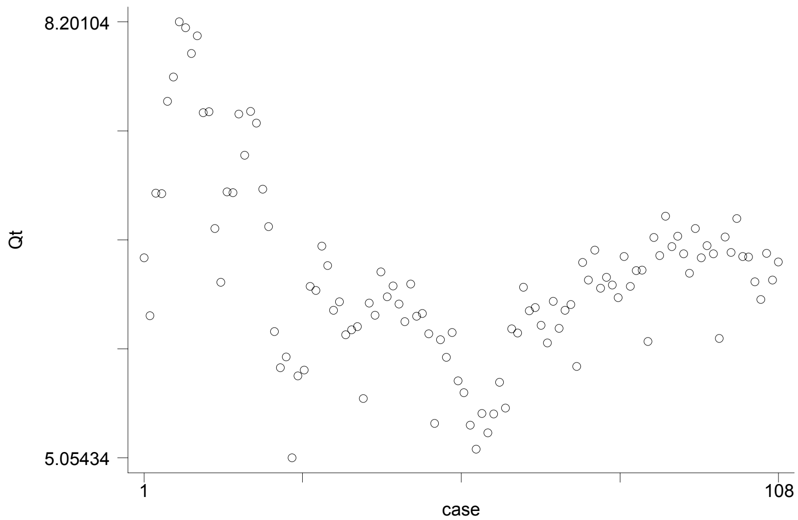

Figure 1: Evolution of water use per head (cubic metres per month)

number of individual meters. The evolution of the values of both consumption per account and inhabitants per account strikingly show this effect of the introduction of individual meters described in Section 2. Figure 1 shows the evolution of the values of Q_t, including the effect of the drought during the first half of the decade. Conservation efforts persisted after the end of the drought, as described in Section 2 and water use levels did not fully return to pre-drought levels

- P_t (1992 EURO equivalents/m^3) is the marginal price of water. It corresponds to the Taylor-Nordin specification (Taylor, 1975 (Taylor 1975) ; Nordin, 1976 (Nordin 1976)) for multipart tariff structures. It is an instrumental marginal price derived from a linear regression of the theoretical water bills associated with each and every one of the integer values of potential monthly water use per account between 1 m^3 and 25 m^3 on these integer values (see Billings, 1982(Billings 1982)). This instrumental marginal price is the slope of the estimated function. This formulation avoids problems of price endogeneity and also reflects the fact that consumers have only an imperfect knowledge of the tariff structure and the block they are consuming in at each point in time. Monetary values are deflated using the official provincial-wise retail price index. No single available series of the price-index would be long enough to cover the whole price series, so the published series with base 1983 was adapted to merge with the series with base 1992.

- VI_t (1992 EURO equivalents) is *virtual income*. It is the difference between the average salaries (W_t) and D_t, the instrument for the *Nordin-difference* (Nordin, 1976) variable. It is the intercept of the estimated linear function used to derive P. The average salaries series (available from the *Instituto Nacional de Estadística*) is used as a proxy for household income. It had originally a quarterly frequency, so it was linearly interpolated to get monthly values. The values for P_t and D_t were calculated using the tariff schedules applied in each period.

- $RAIN_t$ is the current level of precipitation. Unit:mm/month.

- $TEMP_t$ is the average of the daily maximum temperatures in Month t. Unit: °C/10.

- $REST_t$ (hours/day) refers to the number of daily hours of supply restrictions applied as part of the emergency control measures during the worst drought periods. The number of hours of restriction a day is weighted by the number of days in the month to which that number applied. This variable has been calculated directly from the relevant city council drought-emergency decrees *EMASESA*, 1997(EMASESA 1997).

- BAN_t is a binary variable with value 1 when temporary outdoor-use bans were applied during the drought.

- $INFOR_t$ is a binary variable with value 1 if water conservation information campaigns were being applied during the drought.

- SUM_t is a binary variable with value 1 for the months of May, June, July, and August

Summary statistics for all variables are provided in Table 3.

4 Econometric methods

The techniques of co-integration (see Engle and Granger, 1987) and error correction (see Hendry et al., 1984, among others) are used to investigate the dynamics of household water consumption and to measure the short-run and long-run effects of the price of water on household demand.

Let us consider the simple form of a dynamic model:

$$y_t = \mu + \gamma_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t, \quad t = 1, \ldots, T, \tag{1}$$

where y_t and x_t could represent respectively consumption and price at time t. The error term (ε_t) is assumed independently and identically distributed. We will assume in the following that x_t is a one-dimensional vector for ease of exposition. $\mu, \gamma_1, \beta_0, \beta_1$ are unknown parameters. It is well known that in Model 1 the short-run and long-run effects of x on y are measured respectively by β_0 and $(\beta_0 + \beta_1)/(1 - \gamma_1)$.

Re-arranging terms in Model 1, we obtain the usual ECM:

$$\Delta y_t = \mu + \beta_0 \Delta x_t - (1 - \gamma_1)(y_{t-1} - \theta x_{t-1}) + \varepsilon_t, \quad t = 1, \ldots, T, \qquad (2)$$

where Δ represents the difference operator (e.g. $\Delta y_t = y_t - y_{t-1}$) and $\theta = (\beta_0 + \beta_1)/(1 - \gamma_1)$. So, the estimation of the ECM model gives directly a measure of the short-run and long-run effects of x on y through the estimates of β_0 and θ. The second term in Model 2 $(y_{t-1} - \theta x_{t-1})$ can be seen as a *partial correction* for the extent to which y_{t-1} deviated from the equilibrium value corresponding to x_{t-1}. In other words, this representation assumes that any short-run shock to y that pushes it off the long-run equilibrium growth rate will gradually be corrected, and an equilibrium rate will be restored. The expression $(y_{t-1} - \theta x_{t-1})$ corresponds to the *residual* of the long-run equilibrium relationship between x and y.[4] Therefore, this error correction term will be included in the model as long as there exists a long-run equilibrium relationship between x and y or, in other words, if both series are co-integrated in the sense of Granger (see Engle and Granger,

[4]For this reason, it is commonly said that $(1 - \gamma_1)$ provides a measure of the *speed of adjustment* towards the long run values.

1987). If the series are co-integrated they will, in the long run, tend to grow at similar rates, because their data generating processes may be following the same stochastic trend, or may share an underlying common factor.

The econometric analysis will proceed in two steps. In the first step, we test if x and y are co-integrated series. If this proves to be the case, the estimation of the Granger co-integration relationship will give a measure of the long-run effect of x on y. In a second step, the co-integration residuals are used as an error correction term in the ECM model above and the short-run effect and the speed of adjustment can be estimated. This two-step procedure is now described in more detail.

The test for co-integration requires a test for the stationarity of the series. If the series are integrated of the same order, a co-integrating vector might be then found such that a linear combination of the non-stationary variables obtained with that vector is itself stationary.

4.1 Tests for order of integration

A time series is said to be $I(i)$ (integrated of order i) if it becomes stationary after differencing it i-times. Since a non-stationary series can be represented by an autoregressive process of order p, the most widely used unit-root tests for a variable y_t rely on transformed equations of the form:

$$\Delta y_t = \mu + \lambda t + (\gamma - 1) y_{t-1} + \sum_{i=1}^{p-1} \gamma_j \Delta y_{t-i} + \varepsilon_t \tag{3}$$

This test, known as the Augmented Dickey Fuller, ADF, (Dickey and

Fuller 1981) allows for an $AR(p)$ process that may include a nonzero overall mean for the series and a trend variable (t). The inclusion of the term $\sum_{i=1}^{p-1} \gamma_j \Delta y_{t-i}$ simply allows for the consideration of a $p > 1$ in $AR(p)$. The special case where $p = 1$ corresponds to the Dickey-Fuller (DF) test. Its test statistics would be invalidated if the residuals of the reduced form equation

$$\Delta y_t = \mu + \lambda t + (\gamma - 1) y_{t-1} + \varepsilon_t$$

were autocorrelated.

To test the null hypothesis of nonstationarity, the t-statistic of the estimate of $(\gamma - 1)$ is compared with the corresponding critical values, calculated by Dickey and Fuller (Dickey and Fuller, 1979 and 1981) . A key consideration is how many lags of variable y to include in Equation 3 and whether to include a constant and a trend variable. The best model can be selected on the basis of the $\overline{R}^2$, the Akaike Information Criterion (AIC) the Schwartz (1978) Bayesian Information Criterion and the Schwert (1989) criterion. These criteria might lead to conflicting recommendations. Therefore, for consistency, the sequential-t test proposed by Ng and Perron (1995) was used.

If the null of a unit root cannot be rejected, a second test is conducted to check whether the series are integrated of order one, or whether the other of integration is more than one. The ADF test serves this purpose. It consists of testing for the null hypothesis of a unit root in the residual series of a regression in which the series has been differenced once. If the null of unit root is now rejected, the series is deemed $I(1)$ or integrated of order one.

4.1.1 Seasonal case

The tests described above for the stationarity of the series are not sufficient when the data exhibit a seasonal character, since seasonal unit roots must be investigated. A number of seasonal unit root tests have been proposed for the case of monthly data (Franses, 1991; Beaulieu and Miron, 1993) as an extension to the one suggested by Hylleberg et al. (1990).

A characteristic of seasonal unit root tests is that they exhibit poor power performance in small samples[5] and that the power deteriorates as the number of unit roots under examination increases. For example, in a simple test regression with no deterministic variables, the HEGY Hylleberg et al., 1990 test procedure in the quarterly context requires the estimation of four parameters, whereas in a monthly context this number increases to twelve. In addition, the algebra underlying monthly seasonal unit root tests is more involved than in the quarterly case and the associated computational burden non-negligible. To circumvent these problems, the analysis of seasonal unit roots draws on the results found by Rodrigues and Franses (2003). These authors find out which unit roots affecting monthly data can also be detected by applying tests on quarterly data and, in particular, they show that 'with regard to the zero frequency unit root, there is a direct relationship between the monthly and quarterly root'. This means that the problem

[5] Rodrigues and Osborn (1999) provide Monte Carlo evidence on the monthly seasonal unit root tests.

of non-stationarity of the series can be highly simplified by collapsing the monthly data into quarterly data (obtaining $N/3$ quarterly observations on all relevant variables by summing the monthly values or averaging them, depending on the nature of the variable) and then using the original HEGY test. If all the null hypotheses of any type of seasonal roots can be rejected based on the quarterly test, the monthly series can be also deemed free of seasonal unit roots.

To test for a seasonal unit root in the $\{y_t, t = 1, \ldots, T\}$ series, HEGY propose to apply OLS on the following model:

$$y_t - y_{t-4} = \pi_0 + \pi_1 z_{1,t-1} + \pi_2 z_{2,t-1} + \pi_3 z_{3,t-2} + \pi_4 z_{3,t-1} + \varepsilon_t, \qquad (4)$$

$$\begin{aligned} \text{where} \quad z_{1t} &= (1 + L + L^2 + L^3) y_t, \\ z_{2t} &= -(1 - L + L^2 - L^3) y_t, \\ z_{3t} &= -(1 - L^2) y_t, \end{aligned}$$

with L, the lag operator. To find that y_t has no unit root at all and is therefore stationary, we must establish that each of the $\pi_i (i = 1, \ldots 4)$ is different from zero. Moreover, we will reject the hypothesis of a seasonal unit root if π_2 and either π_3 or π_4 are different from zero, which therefore requires the rejection of both a test for π_2 and a joint test for π_3 and π_4. Hylleberg et al. (1990) derive critical values for the tests corresponding to each of the following null hypothesis:

$$H_{01} : \pi_1 = 0,$$

$$H_{02} : \pi_2 = 0,$$

$$H_{03} : \pi_3 = 0,$$

$$H_{04} : \pi_4 = 0,$$

$$H_{03+04} : \pi_3 = 0 \text{ and } \pi_4 = 0.$$

The tests statistics are based on $Student$-statistics (t-stat) for the first four tests and on a $Fisher$-statistic (F-stat) for the last one.

4.2 Co-integration

If, on the basis of these unit root tests, the series are found integrated of the same order, their long-run relationship is then investigated applying OLS on the simple model:

$$y_t = \theta x_t + \varepsilon_t \tag{5}$$

Series x and y are said to be co-integrated if there exists a linear combination of those non-stationary variables that is itself stationary. This means that their linear combination yields a stationary deviation (the residuals series is stationary). As suggested by Engle and Granger (1987) the stationarity of the estimated co-integration residuals ($\widehat{\varepsilon}_t$) from this regression is analyzed. A unit root test[6] is applied whereby the resulting t-statistic is compared with

[6] The ADF test applied in this instance does not contain neither a trend nor a constant term, since the OLS residuals will be mean zero with a constant included in the

the critical values provided by Engle and Yoo (1987).[7] The null hypothesis, in this case, is that of non-cointegration. Therefore, rejecting a unit root in the residuals in a Dickey-Fuller type of test will constitute evidence of a co-integrating relationship among the variables. The OLS estimates have the desirable property of *superconsistency* (Stock, 1987). This means they are not only consistent estimates of the underlying parameters of the data generation process, but they converge on the population values more quickly than OLS estimates in the context of stationary regressors.

If the series are proved to be co-integrated, $\hat{\theta}$ in Equation 5 provides a measure of the long-run effect of x on y. Therefore, the long-run estimates of the price-elasticities are calculated using the estimated coefficients of the price variables in this equation. Additionally $\hat{u}_t$ can be used as an error correction term in the ECM model:

$$\triangle y_t = \mu + \beta_0 \triangle x_t - (1 - \gamma_1)\hat{u}_{t-1} + \varepsilon_t, \quad t = 1, \ldots, T, \tag{6}$$

where $\widehat{\beta}_0$ and $(1 - \hat{\gamma}_1)$ represent the short-run effect and the speed of adjustment towards the long-run values respectively. Short-run price elasticities are then derived from the estimates of price variables in this model.

cointegration regression.

[7] The conventional critical values calculated by Dickey and Fuller are not appropriate, since the distribution of the t-statistic is affected by the number of variables in the cointegration regression (Engle and Yoo, 1987).

5 Results

All the econometric analysis was conducted using STATA 7.0 (Stata 2001). The reader is directed to the manuals and online documentation for details on calculations.[8]

5.1 Unit root tests

First, the order of integration of all relevant series was investigated, using a test of seasonal integration (see Section 4.1.1) and the ADF test 4.1. Table 4 in the Appendix summarizes the seasonal tests applied on the series collapsed into quarterly data. The non-rejection of H_{01}, together with the rejection of both H_{02} and the joint hypothesis H_{03+04}, suggests the presence of a unit root at the zero frequency and no seasonal unit roots. Since there is a correspondence between the quarterly and the monthly root at the zero frequency, not detecting seasonal unit roots at the quarterly level is enough to consider that the series is affected only by unit roots at the zero level and no testing at the monthly level is necessary. The table shows that the null of seasonal unit roots can be rejected for all series,[9] with the not very surprising exception of $TEMP$.

After detecting with the seasonal approach the presence of only unit roots at the zero frequency, the order of integration of the series was further

[8] Details about the specific procedures employed are available form the author upon request.

[9] The test also permitted to reject the null hypothesis of seasonal unit roots in the $INFOR$ series, although the estimates are not shown, since this variable is not used in most of the main final water demand models.

tested using Dickey-Fuller-type tests. Two auxiliary DF regressions with a without a trend were used, and the optimal lag was chosen by an automatic sequential t-test. The results, shown in Table 5 in the Appendix, reveal that the trend component is not relevant in most cases. Table 5 shows that most variables proved to be $I(1)$.[10] The hypothesis tests permit the rejection of the null of non-stationarity of the differenced series at the 99% level of confidence. Once again, there is some doubts about the climate variables. $TEMP$ appears to be stationary, but the seasonal unit root tests did not reject the hypothesis of seasonal roots, so this variable should be considered with caution, since it might be $I(0,1)$. In the case of $RAIN$, we also see that the series might actually be stationary in levels also at all frequencies, $I(0)$. Since the possibility of seasonal unit roots was rejected, there is no problem with introducing this variable in a co-integration regression, whether it is $I(0,0)$ or $I(1,0)$.

The augmentation of the basic DF regression with extra lags described in Section 4.1 above was motivated by the need to generate iid errors. An alternative solution is the Phillips–Perron (PP) test (Phillips and Perron, 1988). This test uses the same models as DF but, instead of lagged variables, it employs a non-parametric correction (Newey and West, 1987) for serial correlation. The critical for both the Dickey Fuller and Phillips Perron tests have the same distributions. Critical levels are reproduced in Hamilton (1994). In principle, the PP tests should be more powerful than the ADF

[10] The test also permitted to reject the null hypothesis of seasonal unit roots in the INFOR series, although the estimates are not shown, since this variable is not used in most of the main final water demand models.

alternative, so the unit root tests have been conducted using both the ADF and PP tests. Since the ADF tests suffice to prove that the desirable dynamic properties of the variables, the results of the PP test are not reported but available upon request.

5.2 Co-integration regression analysis

Since all the series in first-differences are stationary, the next step is to check that there exists a long-run equilibrium relationship between the variables (that the series are co-integrated in the sense of Granger). This requires an extension of the linear relationship between water consumption and a series of variables that the economic theory suggest appropriate. The model given by Equation (5) was extended into two alternative models (time subscripts have been dropped to simplify the exposition):

$$Q = \alpha + P + P^2 + REST + VI + BAN + SUM + \varepsilon \tag{7}$$

which includes the binary variable SUM instead of the climatic variables (see Section 3) and:

$$Q = \alpha' + P + P^2 + REST + VI + BAN + TEMP + RAIN + \varepsilon' \tag{8}$$

Tables 6 and 7 show the OLS estimated coefficients of each of the variables and their t-statistics in these estimations.

The ADF test shows that the hypothesis that the residuals in Regression

7 are non-stationary can be rejected. The relevant t-ratio is -5.625[11] in the usual test of a unit root and must be compared with the critical values provided by (Engle and Yoo 1987), which depend on the dimension of the time-series and on the number of variables included in the model. The DW statistic is also higher than the R^2, which suggests the existence of the co-integration relationship.[12]

The long run price-elasticity calculated at the means of price and quantity according to Model 7 is -0.491. All the variables present the expected signs and are highly significant.

The ADF test shows that the hypothesis that the residuals in Regression 8 are non-stationary can be rejected. The relevant t-ratio is -5.364[13] in the usual test of a unit root and must be compared with the critical values provided by (Engle and Yoo 1987). The DW statistic is higher than the R^2. The long run price-elasticity calculated at the means of price and quantity according to Model 7 is once again -0.494, which is basically the same obtained with Model 7. Once again, all the variables present the expected signs and are highly significant. The exception is $RAIN$, which presents a

[11] This value permits the rejection of the null of no cointegration at a 99% confidence level, but it is achieved when the auxiliary regression includes no lags. Five lags are selected by Ng Perron's sequential t-ratio and the Akaike Information Criterion test, yielding a t-statistic of -3.053 and one lag is selected by the Hannan-Quinn (Hannan and Quinn 1979) criterion, yielding a t-statistic of -3.804.

[12] This is based on an alternative cointegrating regression test developed by (Sargan and Bhargava 1983). This uses the DW statistic from the cointegrating regression. If the residuals are non-stationary, DW will approach zero as the sample size increases. This means that large values of DW suggest that a cointegrating relationship exists.

[13] This value permits the rejection of the null of no cointegration at a 99% confidence level, but it is achieved when the auxiliary regression includes no lags. If the auxiliary regression is run with the optimal number of lags (three) the t-ratio is -4.192.

positive sign while we would normally expect more precipitation to reduce water use, but it cannot be rejected that its coefficient is null.

According to the ADF tests, the null of no co-integration can only be rejected if the lag length of the auxiliary regression is not optimally chosen. However, the value of the DW test and economic intuition suggest that a long run relationship would govern the variables concerned.

In the presence of persistent roots, the Engle-Granger tests tend to lack power to detect a co-integrating relationship in the data, even when one is present. It is difficult to discern whether the inability to reject the null hypothesis actually reflects a non-cointegrated system or simply the weak power of these co-integration tests. Additionally, there could exist more than one co-integrating relationship.

To obtain more definite evidence on the existence of a co-integrating regression, the Johansen and Juselius maximum likelihood method for co-integration (see Johansen, 1988; Johansen and Juselius, 1990 and Osterwald-Lenum, 1992 for details) was used to determine the number of co-integrating relationships. The summarized results are shown in Tables 10 and 11. The eigenvalues and the maximal eigenvalue and trace statistics for the VAR matrix are shown as well as the relevant critical values. The null hypothesis of more than one co-integrating relationship was rejected at the 1% level of significance in all cases, except in the case of the the trace test for Model 7, which rejects the null of no-cointegration only at about the 15%. Likelihood-ratio and Wald test statistics for the exclusion of variables from that co-integrating relationship were also conducted, and all variables included in the co-integration tests were found relevant. Therefore, the Johansen tests support the assumption of co-integration for both models.

5.3 ECM models

Since most of the evidence points towards the stationarity of the residuals of the co-integrating regressions, their residuals can be introduced as error correction terms in two ECM models. The x_t variables in Equation 6 are substituted by first differences and lagged differences[14] of the co-integrating variables. The first error correction specification, $ECM7$ includes a summer variable, whereas the second model, $ECM8$ includes $TEMP$ and $RAIN$ (although $TEMP$ could well suffer problems of seasonal unit roots, so this second model should be considered with caution). Tables 8 and 9 report the results of these OLS estimations. These include lagged values of the differences of some variables. VI was left out of the ECM models, since it showed problems of multicollinearity with the price variables and its introduction made them non-significant. It is reasonable to assume that changes in income tend to affect water use only in the long run, most likely through impacts on the composition of the capital stock. BAN was found non-significant too and it was removed from the ECM models.

The speed of adjustment towards equilibrium is given by -0.218 in $ECM7$ and -0.249 in $ECM8$. It can be seen that these error correction terms are both significant and have the expected negative sign. The signi-

[14] The significance of lagged values was also tested.

ficance of the coefficient associated with the error correction term further supports the acceptance of the co-integration hypothesis.

The Ramsey RESET-test (using powers of the fitted values of ΔQ_t) shows that the null hypothesis that Models $ECM7$ and $ECM8$ have no omitted variables cannot be rejected. Tables 8 and 9 include a battery of diagnostic tests used to check that the residuals are normally distributed and are neither autocorrelated nor heteroskedastic. These include a Jarque-Bera (1980) test for normality of the residuals; White's (1980) general test statistic and Cook-Weisberg (1983) test[15] which uses fitted values of ΔQ_t) tests for heteroskedasticity a Lagrange multiplier test for autoregressive conditional heteroskedasticity (ARCH), based on Engle (1982); and a Breusch (1978)-Godfrey (1978) LM statistic. They all present acceptable values, with the exception of the Breusch-Godfrey LM test, which leads to the rejection of the null of non-autocorrelation in $ECM7$. An alternative model with extra lagged values of the price variables solves this problem and yields a short-run elasticity of -0.073, as reported below. The results of this additional augmented regression do not differ significantly from the ones reported and are available upon request.

5.4 Price elasticities

The computation of short-run price elasticities (e_{SR}) using the average price and water consumption, yields the following results. Using $ECM7$ and the co-integration regression in Model 7, $e_{SR} = -0.159$ (while the augmented

[15] Also known as Breusch-Pagan (1979) test for heteroskedasticity.

model used to correct for autocorrelation would yield $e_{SR} = -0.073$) and the $e_{LR} = -0.494$. Similarly, $ECM7$ and Model 8 yield $e_{SR} = -0.101$ and $e_{LR} = -0.491$.

These estimates of price-elasticities confirm that residential water demand is inelastic to its price, but not perfectly so. Almost all the papers published on residential water demand agree on this result. Additionally these results confirm the intuition that long-run elasticities are higher (in absolute values) short-run ones (Dandy, et al., 1997; Nauges and Thomas, 2003, Martínez-Espiñeira and Nauges, 2004) and also than most of the measures that have been obtained in other European countries.[16] The use of the co-integration approach to model the demand for water yields rather sensible results and help to distinguish between the short-run effects and the long-run effects of pricing policies.

5.5 Wickens-Breusch one-step approach

The Engle-Granger procedure described above enjoys important attractive asymptotic properties but it also suffers weaknesses. In finite samples, the parameter estimates are biased. The extent of this bias will depend on omitted dynamics and failure of weak exogeneity among other things. This bias can be extremely severe. The reasonable size of the sample and the fact that the estimates agree with economic theory and previous empirical research suggest that this might be a minor problem in this case. Another

[16] See Arbués et al. (2003) for a review of water demand studies with a special focus on European cases.

problem is that there is no possibility to test the long run parameters. The limiting distributions of the β parameters are non-normal and non-standard. Standard hypothesis testing is invalid as t and F statistics do not have t or F distributions in the context of the co-integrating regression.

For this reason, an additional regression was run using the one-step Wickens-Breusch approach.[17] The results are reported in Tables 12 and 13. The associated price-elasticities, calculated at the means of price and quantity are $e_{SR} = -0.08$ and $e_{LR} = -0.405$ in the model that uses SUM and $e_{SR} = -0.113$ and $e_{LR} = -0.514$ in the model that uses $TEMP$ and $RAIN$. The estimates are very close to the ones calculated with the Engle-Granger approach, which suggests that they can be accepted with more confidence.

6 Conclusions and suggestions for further research

This study is innovative in two aspects. This is the first time that co-integration and error correction techniques are employed in the field of water consumption. Moreover, the estimation of residential water demand using time-series monthly data is still rather uncommon in Europe. The application of these techniques to monthly data to the case of Seville leads to satisfactory results. The fit of the Granger co-integration relationship between water use and the variables that should be expected to influence it

[17] See (Wickens and Breusch 1988) for details on the algebra of this one-step approach..

in the long run and of the Error Correction Models is quite good. The dynamic properties of the series have been analyzed using different approaches and two alternative specifications for the water demand functions have been used. However, the results in terms of price elasticities, most of all in the short run, are remarkably close. This robustness to specification and testing procedures leads to confidently accept the main results.

The estimates of the price effects obtained are less than one in absolute value, which confirms the inelasticity of household demand with respect to the price of water. As predicted by the theory, the long-run price elasticities are greater, in absolute value, than their short-run counterparts.

The measure of the impact of pricing policies on the behaviour of households depending on the changes that these policies introduce in the tariff structure is still an open research area. The long-run effects of water pricing on water use should be investigated using other datasets, involving different regions, and if possible longer time-series or panel data. Ideally, studies should be conducted at the individual level, with observations linked to the ownership and frequency of renewal of capital stock.

References

Agthe, D. E. and R. B. Billings (1980). Dynamic models of residential water demand. *Water Resources Research 16*(3), 476–480.

Agthe, D. E., R. B. Billings, J. L. Dobra, and K. Rafiee (1986). A simultaneous equation demand model for block rates. *Water Resources Research 22*(1), 1–4.

Arbués, F., M. A. García-Valiñas, and R. Martínez-Espiñeira (2003). Estimation of residential water demand: A state of the art review. *Journal of Socio-Economics 32*(1), 81–102.

Beaulieu, J. J. and J. A. Miron (1993). Seasonal unit roots in aggregate U.S. data. *Journal of Econometrics 55*(1), 305–328.

Beenstock, M., E. Goldin, and D. Nabot (1999). The demand for electricity in Israel. *Energy Economics 21*(2), 168–183.

Bentzen, J. (1994). Empirical analysis of gasoline demand in Denmark using cointegration techniques. *Journal of Energy Economics 16*(2), 139–143.

Billings, R. B. (1982). Specification of block rate price variables in demand models. *Land Economics 58*(3), 386–393.

Breusch, T. S. (1978). Testing for autocorrelation in dynamic linear models. *Australian Economic Papers 17*, 334–355.

Breusch, T. V. and A. R. Pagan (1979). A simple test for heteroskedasticity and random coefficient variation. *Econometrica 47*(5), 1287–1294.

Carver, P. H. and J. J. Boland (1980). Short-run and long-run effects of price on municipal water use. *Water Resources Research 16*(4), 609–616.

Chicoine, D. L. and G. Ramamurthy (1986). Evidence on the specification of price in the study of domestic water demand. *Land Economics 62*(1), 26–32.

Cook, R. D. and S. Weisberg (1983). Diagnostics for heteroscedasticity in regression. *Biometrika 70*(1), 1–10.

Dalhuisen, J. M., R. Florax, H. L. F. de Groot, and P. Nijkamp (2003). Price and income elasticities of residential water demand: Why empirical estimates differ. *Land Economics 79*(2), 292–308.

Dandy, G., T. Nguyen, and C. Davies (1997). Estimating residential water demand in the presence of free allowances. *Land Economics 73*(1), 125–139.

Dickey, D. A. and W. A. Fuller (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association 74*(366), 427–431.

Dickey, D. A. and W. A. Fuller (1981). Likelihood ratio statistics for autoregressive processes. *Econometrica 49*(4), 1057–1072.

Eltony, M. N. and N. H. Al-Mutairi (1995). Demand for gasoline in Kuwait. an empirical analysis using cointegration techniques. *Energy Economics 17*(3), 249–253.

EMASESA (1997). *Crónica de una Sequía*. Sevilla: EMASESA.

EMASESA (2000). Water demand management: The perspective of EMASESA. Technical report, EMASESA.

Engle, R. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of united Kingdom inflation. *Econometrica 50*(4), 987–1007.

Engle, R. and C. W. J. Granger (1987). Cointegration and error correction: Representation, estimation and testing. *Econometrica 55*(2), 251–76.

Engle, R. F., C. W. J. Granger, and J. S. Hallman (1989). Merging short and long run forecasts: An application of seasonal cointegration to monthly electricity sales forecasting. *Journal of Econometrics 40*(1), 45–62.

Engle, R. F. and B. S. Yoo (1987). Forecasting and testing in co-integrated systems. *Journal of Econometrics 35*(1), 143–159.

Fouquet, R. (1995). The impact of VAT introduction on UK residential energy demand. an investigation using the cointegration approach. *Energy Economics 17*(3), 237–247.

Franses, P. (1991). Seasonality, nonstationarity and forecasting of monthly time series. *International Journal of Forecasting 7*, 199–208.

García-Valiñas, M. Á. (2002). Residential water demand: The impact of management procedures during shortage periods. *Water Intelligence Online 1*(May).

Godfrey, L. G. (1978). Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. *Econometrica 46*(6), 1293–1301.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton NJ: Princeton University Press.

Hannan, E. and B. G. Quinn (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society Series B 41*(2), 190–95.

Hansen, L. G. (1996). Water and energy price impacts on residential water demand in Copenhagen. *Land Economics 72*(1), 66–79.

Hendry, D. F., A. R. Pagan, and J. D. Sargan (1984). Dynamic specification. In Z. Grilliches and M. D. Intrilligator (Eds.), *Handbook of Econometrics*, Volume III. Amsterdam: North-Holland.

Höglund, L. (1999). Household demand for water in Sweden with implications of a potential tax on water use. *Water Resources Research 35*(12), 3853–3863.

Hylleberg, S., R. F. Engle, C. W. J. Granger, and B. S. Yoo (1990). Seasonal integration and cointegration. *Journal of Econometrics 44*(1), 215–238.

Jarque, C. M. and A. K. Bera (1980). Efficient tests for normality, homoskedasticity, and serial independence of regression residuals. *Economics Letters 6*, 255–259.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control 12*, 231–254.

Johansen, S. and K. Juselius (1990). Maximum likelihood estimation and inference on cointegration - with applications to the demand for money. *Oxford Bulletin of Economics and Statistics 52*, 169–210.

Lyman, R. A. (1992). Peak and off-peak residential water demand. *Water Resources Research 28*(9), 2159–67.

Martínez-Espiñeira, R. (2002). Residential water demand in the Northwest of Spain. *Environmental and Resource Economics 21*(2), 161–187.

Martínez-Espiñeira, R. and C. Nauges (2004). Is really all domestic water consumption sensitive to price control? *Applied Economics forthcoming.*

Moncur, J. (1987). Urban water pricing and drought management. *Water Resources Research 23*(3), 393–398.

Nauges, C. and A. Thomas (2000). Privately-operated water utilities, municipal price negotiation, and estimation of residential water demand: The case of France. *Land Economics 76*(1), 68–85.

Nauges, C. and A. Thomas (2001). Dynamique de la consommation d'Eau potable des ménages: Une étude sur un panel de communes françaises. *Economie et Prevision 143-144*, 175–184. numéro spécial: 'Economie de l'Environnement et des Resources Naturelles'.

Nauges, C. and A. Thomas (2003). Long-run study of residential water consumption. *Environmental and Resource Economics 26*(1), 25–43.

Newey, W. K. and K. D. West (1987). A simple positive semi-definite heteroskedasticity and autocorrelation consistent covariance matrix.

Econometrica 55(3), 1029–1054.

Ng, S. and P. Perron (1995). Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag. *Journal of the American Statistical Association 90*(268-81), 268–81.

Nieswiadomy, M. L. and D. J. Molina (1989). Comparing residential water estimates under decreasing and increasing block rates using household data. *Land Economics 65*(3), 280–289.

Nordin, J. A. (1976). A proposed modification on Taylor's demand-supply analysis: Comment. *The Bell Journal of Economics 7*(2), 719–721.

Osterwald-Lenum, M. (1992). A note with fractiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics: Four cases. *Oxford Bulletin of Economics and Statistics 54*(4), 361–78.

Phillips, P. C. B. and P. Perron (1988). Testing for a unit root in time series regression. *Biometrika 75*(2), 335–346.

Ramanathan, R. (1999). Short- and long-run elasticities of gasoline demand in India: An empirical analysis using cointegration techniques. *Energy Economics 21*(4), 321–330.

Renwick, M. E. and S. O. Archibald (1998). Demand side management policies for residential water use: Who bears the conservation burden? *Land Economics 74*(3), 343–359.

Renwick, M. E. and R. D.Green (2000). Do residential water demand side management policies measure up? An analysis of eight California water agencies. *Journal of Environmental Economics and Management 40*(1), 37–55.

Rodrigues, P. M. M. and P. H. Franses (2003). A sequential approach to testing seasonal unit roots in high frequency data. Institute Reports, Econometrics Department, Erasmus University Rotterdam. Available at http://www.eur.nl/WebDOC/doc/econometrie/feweco20030410140109.pdf.

Rodrigues, P. M. M. and D. R. Osborn (1999). Performance of seasonal unit root tests for monthly data. *Journal of Applied Statistics 26*(8), 985–1004.

Sargan, J. D. and A. Bhargava (1983). Testing residuals from least squares estimators for being generated by the Gaussian Random Walk. *Econometrica 51*(1), 153–174.

Schefter, J. E. and E. L. David (1985). Estimating residential water demand under multi-tariffs using aggregate data. *Land Economics 61*(3), 272–80.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics 6*(2), 461–464.

Schwert, G. W. (1989). Tests for unit roots: A Monte Carlo investigation. *Journal of Business and Economic Statistics 7*(2), 147–159.

StataCorp (2001). *Statistical Software: Release 7.0.* College Station, TX: Stata Corporation.

Stock, J. H. (1987). Unit roots, structural breaks and trends. In R. F. Engle and D. McFadden (Eds.), *Handbook of Econometrics*, Chapter 46, pp. 2740–2831. Amsterdam: Elsevier.

Taylor, L. D. (1975). The demand for electricity: A survey. *The Bell Journal of Economics 6*(1), 74–110.

White, H. (1980). A heteroskedasticity-consistent covariance estimator and a direct test for heteroskedasticity. *Econometrica 48*(4), 817–838.

Wickens, M. R. and T. S. Breusch (1988). Dynamic specification, the long run and the estimation of the transformed regression models. *The Economic Journal 98*(390), 189–205.

Appendix: Summary of demand measures

Measures taken to reduce demand were of three types:

1. Changes in tariff structure to promote savings (see Tables 1 and 2).

2. Meter Replacement Campaign (*Plan Cinco*) to increase the reliability of consumption readings. In the year 2000, meters in Seville and its metropolitan area were on average less than four years old.

3. Promotion of the replacement of collective meters in blocks of dwellings by individual ones. A total of 18,822 supplies corresponding to 226,034 buildings, 87% of them located in the city of Seville, was to be included in the project. 50% of these buildings have between two and eight dwellings, 28.4% have between 9 and 16 dwellings, 10.8% are buildings with between 17 and 24 dwellings and the same percentage corresponds to buildings with more than 25 dwellings. The supply company has provided a series of measures to facilitate the replacement, taking into account the problems and disadvantages encountered (EMASESA 2000, pp. 9-10) .

Table 1: Evolution of pricing-block sizes

	1991-1995	1996-1999
Block 1	0-7 m^3	0-7 m^3
Block 2	0-20 m^3	0-17 m^3
Block 3	$>$20 m^3	$>$ 17 m^3

Table 2: Tariff evolution (1992 EURO equivalents, excluding VAT)

	Water						Sewage and Treatment			
Year	Fixed	PBL_1^*	PBL_2	PBL_3	Canon	TEC**	Fixed	Sewage	Treat.	Canon
1991	1.063	0.139	0.214	0.386	0.000	0.000	0.000	0.075	0.130	0.000
1992	1.010	0.138	0.212	0.384	0.000	0.000	0.000	0.082	0.123	0.000
1993	1.133	0.132	0.206	0.378	0.000	0.020	0.000	0.089	0.120	0.000
1994	1.187	0.126	0.204	0.398	0.016	0.019	0.270	0.088	0.118	0.016
1995	1.246	0.125	0.213	0.421	0.016	0.021	0.285	0.092	0.124	0.016
1996	1.443	0.126	0.252	0.505	0.015	0.093	0.505	0.111	0.131	0.015
1997	1.524	0.130	0.260	0.609	0.015	0.093	0.550	0.125	0.140	0.015
1998	1.540	0.131	0.263	0.616	0.050	0.000	0.555	0.126	0.141	0.040
1999	1.533	0.131	0.262	0.614	0.048	0.000	0.553	0.126	0.141	0.039

* PBL_i = water price in block i.

**TEC = Temporary Extra Charge

Table 3: Summary Statistics

Variable	N	Min	Max	Mean	Std. Dev.
ABONS	108	109,082	201,385	148,310	27,671
POP	108	683,028	719,588	702,529	9684
Q	108	5.054	8.201	6.352	0.648
P	108	0.472	0.700	0.571	0.080
W	108	1128	1410	1235	71,379
D	108	0.522	1.413	0.999	0.369
RES	108	0	12.00	1.40	2.99
BAN	108	0	1	0.273	0.445
SUM	108	0	1	0.333	0.474
$TEMP$	108	152	385	255.259	68.965
RAIN	108	0	3105	421.926	605.559
INFOR	108	0	1	0.319	0.466

Table 4: Quarterly seasonal unit root test

Variable	Test	HEGY model specification$^{(a)}$							
		SEAS		TREND		STREND		CONST	
		t-stat$^{(b)}$	Lags$^{(c)}$	t-stat	Lags	t-stat	Lags	t-stat	Lags
Q	H_{01}	-1.947	8	-2.379	0	-1.199	8	-2.929*	0
	H_{02}	-3.998**		-3.684**		-3.867**		-3.848**	
	H_{03+04}	17.857**		6.176**		16.407**		6.485**	
P	H_{01}	-0.033	1,4,5	-1.136	0	-1.528	0	-1.108	0
	H_{02}	-4.116**		-2.708**		-2.795*		-2.708**	
	H_{03+04}	12.143**		11.666		13.407**		11.701**	
P^2	H_{01}	0.044	1,4,5	-1.237	0	-4.016**	4	-1.124	0
	H_{02}	-4.455**		-2.750**		-3.475**		-2.743**	
	H_{03+04}	14.219**		12.233**		20.187**		12.194**	
VI	H_{01}	1.181	0	-0.787	1	0.311	0	-0.791	1,5
	H_{02}	-3.172**		-0.418		-3.051		-0.829	
	H_{03+04}	6.711**		0.332		6.604		0.744	
$REST$	H_{01}	-2.970*	6	-2.643	3,5,7	-3.549*	0	-2.732*	5,6
	H_{02}	-3.177**		-4.419**		-4.299**		-3.291**	
	H_{03+04}	15.982**		2.785*		7.027**		13.050**	
BAN	H_{01}	-1.757	1	-5.310**	1,2	-3.543*	1	-2.153	1
	H_{02}	-3.788**		-2.779**		-3.585**		-3.547**	
	H_{03+04}	11.913**		9.464**		12.272**		10.323**	
$TEMP$	H_{01}	-2.419	0	-3.401*	2,3,8	-2.399	0	-3.800**	2,3,6,7,8
	H_{02}	-2.133		-1.281		-2.111		-1.057	
	H_{03+04}	8.117**		0.040		7.850**		0.280	
$RAIN$	H_{01}	-2.062	4	-1.888	0	-2.775	2,4,8	-1.779	0
	H_{02}	-1.674*		-1.674*		-3.193**		-1.677*	
	H_{03+04}	2.437*		2.437*		7.753**		2.400*	

(a) Test specifications: SEAS (Seasonal dummies + constant) TREND (Constant + trend) STREND (Seasonal dummies + constant + trend) CONST (constant only)

(b) HEGY estimates, ** and * denote a t-ratio significant at the 5% and 10%

(c) lag length and lags of the fourth difference of the time-series to be included in the auxiliary regression

Table 5: Unit root tests to determine the order of integration of the series

Variable	Test	No trend		With trend		
		t-stat[(a)]	Lags	t-stat[(b)]	Lags	Trend t-ratio
Q	ADF in levels	-2.527	4	-2.384	4	-0.06
	ADF in differences	-6.287***	5	-6.966***	8	2.48**
P	ADF in levels	-1.165	0	-1.888**	0	1.50
	ADF in differences	-3.801***	7	-10.550***	0	-0.26
P^2	ADF in levels	-1.202	0	-1.847	0	1.42
	ADF in differences	-10.569***	0	-10.525***	0	-0.28
VI	ADF in levels	0.333	8	-0.524	8	1.76*
	ADF in differences	-12.709***	8	-13.800***	8	3.35***
$REST$	ADF in levels	-2.547	1	-3.369*	2	-1.23
	ADF in differences	-5.472***	1	-5.454***	1	-0.28
BAN	ADF in levels	-2.130	0	-2.404	0	-1.22
	ADF in differences	-9.494***	0	-9.470***	0	-0.47
$TEMP$	ADF in levels	-8.921***	6	-9.016***	6	1.17
	ADF in differences	-12.124 ***	6	-12.041***	6	-0.22
$RAIN$	ADF in levels	-6.849***	0	-6.834***	4	0.43
	ADF in differences	-12.334***	0	-12.273***	0	0.09

(a) t-ratio of estimates *** ,** and * denote a t-ratio significant at the 1%, 5% and 10%

(b) The number of lags (with a maximum of 8) to be included was selected using the Ng-Perron sequential-t test

Table 6: Cointegration regression results, Model 7

Variable	Coef.	Std. Err.	t	P/t/	[95% Conf.	Interval]
P	-78.62897	11.63832	-6.76	0.000	-101.7163	-55.54167
P^2	64.06969	10.0208	6.39	0.000	44.19111	83.94827
$REST$	-.065878	.0198825	-3.31	0.001	-.1053196	-.0264364
VI	.0028973	.0005331	5.44	0.000	.0018398	.0039548
BAN	-.4509199	.1399911	-3.22	0.002	-.7286246	-.1732152
SUM	.3165457	.0797838	3.97	0.000	.158276	.4748153
$CONS$	26.48916	3.335331	7.94	0.000	19.87276	33.10556

$\overline{R}^2 = 0.6438$ F(6,101)=33.23, , Prob > F = 0.0000

N=108 Durbin-Watson d-statistic= 0.919

Table 7: Cointegration regression results, Model 8

Variable	Coef.	Std. Err.	t	P>/t/	[95% Conf. Interval]	
P	-76.9496	12.07195	-6.37	0.000	-100.9	-52.99919
P^2	62.56866	10.40075	6.02	0.000	41.93387	83.20345
$REST$	-.0669561	.0206834	-3.24	0.002	-.1079914	-.0259208
VI	.0024044	.0005813	4.14	0.000	.0012512	.0035576
BAN	-.4732595	.146367	-3.23	0.002	-.7636475	-.1828715
$RAIN$	.000081	.0000733	1.11	0.271	-.0000643	.0002264
$TEMP$	.0020339	.0006714	3.03	0.003	.0007017	.003366
$CONS$	26.1972	3.457589	7.58	0.000	19.33744	33.05696
$\overline{R}^2 = 0.6194$			F(7,100)=25.88 , Prob > F = 0.0000			
N=108			Durbin-Watson d-statistic= 0.8559396			

Table 8: OLS results of Model $ECM7$

Variable	Coef.	Std. Err.	t	P>/t/	[95% Conf. Interval]	
ΔP_t	-37.43696	21.09048	-1.78	0.079	-79.29572	4.421804
$\widehat{\varepsilon}_{t-1}$	-.218386	.0853624	-2.56	0.012	-.3878068	-.0489652
$\Delta {P^2}_t$	31.26435	17.50574	1.79	0.077	-3.479696	66.0084
$\Delta REST_t$	-.0764792	.024869	-3.08	0.003	-.1258373	-.027121
ΔSUM_t	.316758	.0667966	4.74	0.000	.1841852	.4493308
ΔP_{t-1}	-71.00558	20.68853	-3.43	0.001	-112.0666	-29.94457
$\Delta {P^2}_{t-1}$	54.72853	17.2249	3.18	0.002	20.54189	88.91518
ΔQ_{t-1}	-.27307	.0880438	-3.10	0.003	-.4478126	-.0983274
$CONS$	.0190848	.0266035	0.72	0.475	-.0337157	.0718853
$\overline{R}^2 = 0.3864$	Jarque-Bera normality test: 2.044 $\kappa(2) = 0.3598$					
$AIC = 0.306$	ARCH-LM test statistic, order(1): 2.913434 $\kappa^2(1)$ P-value = 0.0878					
RESET= 1.48	Breusch-Godfrey LM-statistic:16.91713 $\kappa^2(1)$ P-value = 0.000					
p value = 0.23	White's general test statistic : 44.02462 $\kappa^2(1)(44)$ P-value = 0.4706					
	Cook-Weisberg test $\kappa^2(1) = 3.26$, Prob > $\kappa^2 = 0.0711$					

Table 9: OLS results of $ECM88$

Variable	Coef.	Std. Err.	t	P>/t/	[95% Conf. Interval]	
ΔP_t	-32.6092	20.1055	-1.62	0.108	-72.5183	7.299901
$\widehat{\varepsilon'}_{t-1}$	-.2488099	.0775516	-3.21	0.002	-.4027486	-.0948713
$\Delta P^2{}_t$	27.57213	16.69743	1.65	0.102	-5.572012	60.71628
$\Delta TEMP_t$	.0035901	.0006697	5.36	0.000	.0022609	.0049194
$\Delta REST_t$	-.0981625	.0238574	-4.11	0.000	-.145519	-.050806
ΔP_{t-1}	-79.55463	20.42766	-3.89	0.000	-120.1032	-39.00605
ΔQ_{t-1}	-.2756471	.0836188	-3.30	0.001	-.4416292	-.109665
$\Delta P^2{}_{t-1}$	61.88111	17.07977	3.62	0.000	27.97803	95.78418
$\Delta RAIN_t$	.0000862	.0000444	1.94	0.055	-1.82e-06	.0001743
$CONS$	.018672	.0257402	0.73	0.470	-.032422	.0697659
$\overline{R}^2$ =0.4254	Jarque-Bera normality test: 2.059 $\kappa(2) = 0.3572$					
AIC = 0.249	ARCH-LM test statistic, order(1): 0.0007344 $\kappa^2(1)$ P-value = 0.9784					
RESET= 1.75	Breusch-Godfrey LM-statistic: 1.876757 $\kappa^2(1)$ P-value = 0.1707					
p value = 0.16	White's general test statistic : 75.72078 $\kappa^2(1)(44)$ P-value = 0.0272					
	Cook-Weisberg test $\kappa^2(1) = 0.78$, Prob > κ^2 = 0.3773					

Table 10: MODEL 1 Johansen-Juselius cointegration rank test

Eigenvalues (lambda)	H0: rank<=(r) r	H1: Max-lambda statistics (rank<=(r+1))	Trace statistics (rank<=(p=7))
.40067781	0	54.779285	115.46473
.2222569	1	26.895414	60.685446
.16540982	2	19.347149	33.790032
Osterwald-Lenum Critical values (99% interval):			
Table/Case: 1* (assumption: intercept in co-integrating Equation)			
	H0:	Max-lambda	Trace
	0	51.91	143.09
	1	46.82	111.01
	2	39.79	84.45
Table/Case: 1 (assumption: intercept in VAR)			
	H0:	Max-lambda	Trace
	0	51.57	133.57
	1	45.10	103.18
	2	38.77	76.07
Sample: 1 to 108	N= 107		

Table 11: MODEL 2 Johansen-Juselius cointegration rank test

Eigenvalues (lambda)	H0: rank<=(r) r	H1: Max-lambda statistics (rank<=(r+1))	Trace statistics (rank<=(p=8))
.58686628	0	94.586283	185.98108
.2725507	1	34.048574	91.394798
.23272295	2	28.345084	57.346224
Osterwald-Lenum Critical values (99% interval):			
Table/Case: 1* (assumption: intercept in co-integrating Equation)			
	H0:	Max-lambda	Trace
	0	57.95	177.20
	1	51.91	143.09
	2	46.82	111.01
Table/Case: 1 (assumption: intercept in VAR)			
	H0:	Max-lambda	Trace
	0	57.69	168.36
	1	51.57	133.57
	2	45.10	103.18
Sample: 1 to 108	N= 107		

Table 12: Wickens Breusch one-step cointegration regression, Model 1 (dependent variable: ΔQ_{t-1})

Variable	Coef.	Std. Err.	t	P>/t/	[95% Conf.Interval]	
P_{t-1}	-14.8257	8.765719	-1.69	0.094	-32.23517	2.583775
Q_{t-1}	-.1971793	.0584825	-3.37	0.001	-.3133306	-.0810279
${P^2}_{t-1}$	12.20513	7.459134	1.64	0.105	-2.609352	27.01962
ΔQ_{t-2}	.2262604	.0813084	2.78	0.007	.0647749	.3877459
BAN_{t-1}	-.1857024	.0789577	-2.35	0.021	-.3425193	-.0288856
ΔP_t	-4.904221	1.691726	-2.90	0.005	-8.264135	-1.544306
ΔSUM_t	.2359263	.0761873	3.10	0.003	.0846117	.3872408
ΔSUM_{t-1}	-.2278216	.068657	-3.32	0.001	-.3641804	-.0914628
$\Delta REST_t$	-.0536682	.0226287	-2.37	0.020	-.0986108	-.0087256
ΔP_{t-2}	68.63656	19.74489	3.48	0.001	29.4215	107.8516
$\Delta {P^2}_{t-2}$	-53.91469	16.35266	-3.30	0.001	-86.39248	-21.43691
SUM_{t-1}	.1210409	.0743297	1.63	0.107	-.0265843	.268666
$CONS$	5.663556	2.738408	2.07	0.041	.2248421	11.10227
$\overline{R}^2 = 0.4553$	Jarque-Bera normality test: 21.49 $\kappa(2) = 0.000$					
AIC=0.166	ARCH-LM test statistic, order(1): 3.153843 $\kappa^2(1)$ P-value = 0.757					
RESET = 0.55	Breusch-Godfrey LM-statistic: .8856157 $\kappa^2(1)$ P-value = 0.3467					
(p value= 0.6464)	White's general test statistic : 101.2766 $\kappa^2(1)(44)$ P-value = 0.0732					
	Cook-Weisberg test $\kappa^2(1) = 0.75$, Prob > $\kappa^2 = 0.3851$					

Table 13: Wickens Breusch one-step cointegration regression, Model 2 (dependent variable: ΔQ_{t-1})

Variable	Coef.	Std. Err.	t	P>/t/	[95% Conf.Interval]	
P_{t-1}	-12.81087	9.062847	-1.41	0.161	-30.80788	5.186152
Q_{t-1}	-.2188747	.0571935	-3.83	0.000	-.3324497	-.1052997
${P^2}_{t-1}$	10.12214	7.725367	1.31	0.193	-5.218908	25.46319
BAN_{t-1}	-.2861252	.0822907	-3.48	0.001	-.4495382	-.1227122
ΔQ_{t-2}	-.4171249	.084715	-4.92	0.000	-.5853522	-.2488976
$\Delta TEMP_{t-2}$	.0036692	.0007519	4.88	0.000	.0021761	.0051624
$\Delta REST_t$	-.0711663	.026048	-2.73	0.008	-.1228924	-.0194402
ΔP_{t-2}	-6.273908	1.880937	-3.34	0.001	-10.00908	-2.53874
$RAIN_{t-1}$	.0001814	.0000621	2.92	0.004	.0000581	.0003048
$\Delta INFOR_t$	-.5578359	.2148499	-2.60	0.011	-.9844853	-.1311865
$\Delta RAIN_t$	.0000904	.000051	1.77	0.080	-.0000109	.0001916
ΔBAN_t	.3630526	.2119329	1.71	0.090	-.0578042	.7839094
$CONS$	5.353634	2.813476	1.90	0.060	-.2333728	10.94064
$\overline{R}^2 = 0.4207$	Jarque-Bera normality test: 5.669 $\kappa(2) = 0.0587$					
AIC=0.282	ARCH-LM test statistic, order(1): 0.0265371 $\kappa^2(1)$ P-value = 0.8706					
RESET=1.75	Breusch-Godfrey LM-statistic: 1.717154 $\kappa^2(1)$ P-value = 0.1901					
(p value= 0.1618)	White's general test statistic : 89.59179 $\kappa^2(1)(44)$ P-value = 0.0341					
	Cook-Weisberg test $\kappa^2(1) = 1.95$, Prob > $\kappa^2 = 0.1625$					