Long-run water demand estimation: Habits, adjustment dynamics and structural breaks

Musolesi, Antonio; Nosvelli, Mario

Long-run water demand estimation. Habits, adjustment dynamics and structural breaks

Journal:	*Applied Economics*
Manuscript ID:	APE-08-0334.R1
Journal Selection:	Applied Economics
Date Submitted by the Author:	18-Mar-2009
Complete List of Authors:	musolesi, antonio; INRA, CESAER Nosvelli, Mario; CNR, CERIS
JEL Code:	C22 - Time-Series Models < C2 - Econometric Methods: Single Equation Models < C - Mathematical and Quantitative Methods, Q25 - Water < Q2 - Renewable Resources and Conservation|Environmental Management < Q - Agricultural and Natural Resource Economics
Keywords:	BOUNDS TESTING, ARDL APPROACH TO COINTEGRATION, LONG RUN WATER CONSUMPTION

Long-run water demand estimation. Habits, adjustment dynamics and structural breaks

ABSTRACT

This paper examines a water demand equation for Milan for the second half of the twentieth century: 1950-2001. We focus mainly on the effects of price and habits, but also account for other factors in the demand for water such as climate, income and productive activity. Allowing for trend break stationarity or non-linear trend stationarity we find evidence against the unit root hypothesis for many time series. Based on this result, standard cointegration analysis would not be appropriate; therefore we adopt an alternative estimation and testing procedure. We focus in particular on the so called bounds testing approach which can be applied irrespective of the level of integration of the variables and which can be a useful modelling strategy given that dynamics are important when estimating a water demand equation. The main results are that long run price elasticity is higher than short run elasticity, and that consumption habits are relevant. We also find that both climate, sectoral and technological modifications affect water consumption, while income is not significant. Finally, the changes to pricing schemes in the mid seventies provoked reactions of different magnitudes among households and firms.

Keywords: BOUNDS TESTING; ARDL APPROACH TO COINTEGRATION; LONG RUN WATER CONSUMPTION

JEL codes: C22; Q25

1 Introduction

While there is a large body of literature on econometric estimation of a water demand equation, which use data on highly disaggregated spatial units over short time periods, only a few studies have focused on European cities, and over the long run.

Many recent works generally converge in concluding that there is a significant, but small, negative effect of price on water demand in many European countries (Nauges and Thomas, 2003; Martínez Espiñeira, 2002, 2007). Furthermore, instantaneous price elasticities obtained in a static framework, in general are smaller than long-run elasticities based on dynamic models. This could be due either to the use of durable equipment, such as washing machines, dish washers, showers (Arbuès et al., 2003) or on slow adaptations in consumer behaviour (Martínez Espiñeira and Nauges, 2004).

This paper examines a water demand equation for Milan, for the second half of the twentieth century (1950-2001) using annual data and assesses the impact of price on consumption – in both the short- and the long-term - and measures consumption habits. While the use of cross section or (short) panel data for highly disaggregated spatial units allows many degrees of freedom, low collinearity, building complex behavioural models, the use of a long time series has the comparative advantage to be more appropriate for: i) analysing how structural changes occurred in the economy during some decades have affected consumption, ii) comparing long-run and short-run consumption's behaviours, iii) measuring long-run (persistent) habits.

Analysis of consumption habits is one of the main objectives of this paper and has been a central topic in the theoretical and empirical debates over the economics of water demand. Theoretical findings (Tversky and Kahneman, 1991; Naimzada and Tramontana, 2008) indicate that ignoring habits in the estimation of a consumption function can lead to mis-specification problems and to consequent poor explanatory power.

Although the focus is mostly on the effect of price and habits, other factors affecting demand for water, such as climate, income and production activity have been also accounted for. Focusing on this last determinant of water consumption allows us to take account of crucial sectoral and technological modifications over the second half of the twentieth century.

This paper presents some degree of novelty with respect to the the econometric modelling, estimation and testing procedures. An approach commonly adopted for the estimation of time series (water) demand is to use a cointegrated framework. We argue that this approach could produce misleading results. Firstly, cointegration is sometimes applied to short and high-frequency water consumption series. However, cointegration is a long run concept and - as pointed out by Hakkio and Rush (1991) - requires a long data span rather than a merely large number of observations.

Secondly, and of main concern to this work, cointegration analysis requires that all the variables that enter the water demand equation should be first order integrated. Using non-standard unit root testing procedures, which allow for breaks or nonlinearity in the trend function (Zivot and Andrews, 1992; Clemente et al., 1998; Bierens, 1997), many time series are found to be stationary, making cointegration analysis inappropriate. Therefore, this paper adopts the so called bounds testing approach (Pesaran and Shin, 1999; Pesaran et al., 2001). The main advantage of this approach with respect to standard cointegration techniques is that it can be applied irrespective of whether the variables in the underlying VAR (Vector Auto Regression) are I(0), I(1) or mutually cointegrated. This approach allows the estimation of both long-run and short-run elasticities and, after a re-parameterisation in the ECM (Error Correction Mechanism) form of the original ARDL (Auto Regressive Distributed Lags) model, determination of the speed of adjustment to equilibrium.

Finally, this study provides some implications for urban policy, since some of Milan's economic, demographic and climatic trends could be generalised to other big European cities.

The structure of the paper is as follows. The econometric model is presented in section 2. Section 3 describes the data and a section 4 provides the preliminary results about the order of integration of the variables and the existence of a long-run water demand. The estimation results are in section 5 and some brief conclusions and policy implications are in section 6.

2 Econometric modelling

The procedure developed by Pesaran and Shin (1999) and Pesaran et al. (2001) is adopted. The focus is on testing and estimating a long-run level relationship starting from a dynamic ARDL model which also allows for short run dynamics. Before presenting the results of estimation, the methodology is briefly recalled and adapted to the context of a water demand equation.

Let $\{\mathbf{z}_t\}_{t=1}^{\infty}$ denote a $(k+1)$-vector random process partitioned as $\mathbf{z}_t = (c_t, \mathbf{x}_t')'$ where c is a scalar variable representing the natural logarithm of water consumption in year t, $\mathbf{x}$ is a $k-$vector of explanatory variables (also expressed in logarithmic form) like the price of water and some other variables such as income, the productive structure, demographic and climatic factors, etc. The data-generating process for $\{\mathbf{z}_t\}_{t=1}^{\infty}$ is assumed to be a VAR model of order p:

$$\Phi(L)(\mathbf{z}_t - \boldsymbol{\mu}) = \boldsymbol{\varepsilon}_t,\ t = 1, 2, ... \tag{1}$$

where L is the lag operator, $\boldsymbol{\mu}$ is an unknown $(k+1)$-vector of intercept coefficients, the $(k+1, k+1)$ matrix lag polynomial $\Phi(L) = \mathbf{I}_{k+1} - \sum_{i=1}^{p} \Phi_i L^i$ with $\{\Phi_i\}_{i=1}^{p}$ $(k+1, k+1)$ matrices of unknown coefficients. Under few standard assumptions, Pesaran *et al.*(2001)[1] obtain an ARDL conditional ECM model defined as:

[1] Pesaran et al. (2001)make a non-standard assumption that: The roots of $\left|\mathbf{I}_{k+1} - \sum_{i=1}^{p} \Phi_i L^i\right| = 0$ are either outside the unit circle $|z| = 1$ or satisfy $z = 1$. This enables the elements of $\mathbf{z}_t$ to be I(1),

$$\Delta c_t = a_0 + \phi c_{t-1} + \boldsymbol{\delta}'\mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \psi_i \Delta c_{t-i} + \sum_{i=0}^{p} \varphi_i \Delta \mathbf{x}_{t-i} + \epsilon_t \tag{2}$$

It follows from (2) that if $\phi \neq 0$ and $\boldsymbol{\delta} \neq \mathbf{0}$ there exists a conditional level relationship between c_t and $\mathbf{x}_t-$ the long run domestic water demand equation - defined by:

$$c_t = \theta_0 + \boldsymbol{\theta}'\mathbf{x}_t + v_t \tag{3}$$

where $\theta_0 \equiv -a_0/\phi$, $\boldsymbol{\theta} \equiv -\boldsymbol{\delta}/\phi$ is the k-vector of long run response parameters and$\{v_t\}$ is a zero mean stationary process. Equation (2) can be reformulated in ECM form as:

$$\Delta c_t = a_0 + \varsigma\left(c_{t-1} - \boldsymbol{\theta}'\mathbf{x}_{t-1}\right) + \sum_{i=1}^{p-1} \psi_i \Delta c_{t-i} + \sum_{i=0}^{p} \varphi_i \Delta \mathbf{x}_{t-i} + \epsilon_t \tag{4}$$

where ς represents the speed of adjustment to the long run equilibrium. For $\varsigma = 0$ there is no adjustment to the long run equilibrium. If $\varsigma < 0$ then the long-run relationship between the levels of c_t and $\mathbf{x}_t$ is stable since there is adjustment to the long-run equilibrium. For example, for $\varsigma = -1$ the disequilibrium in the previous year is completely corrected in the current year.

I(0) or cointegrated.

According to Pesaran and Shin (1999) and Pesaran et al. (2001), the ARDL modelling approach to cointegration requires the following two steps. In the first step, the absence of any long run relationship between c_t and $\mathbf{x}_t$ is tested through exclusion of the lagged level variables c_{t-1} and $\mathbf{x}_{t-1}$ in (2).

Critical values of the test statistics are tabulated in Pesaran et al. (2001). In the second step of the analysis – following estimation of the ARDL model defined by (2), which gives the short run impacts – we obtain the long run parameters of (3) from the coefficients of the lagged variables of (2); finally, by re-parameterising (2), the ECM model (4) is estimated and the speed of adjustment to the equilibrium is determined.

This econometric framework is very useful for our research objectives. Firstly, it provides statistical information on consumption habits and on their modification over-time: i) in the ARDL specification, the parameter associated to the auto-regressive component indicates the persistence of the series and thus provides a measure of habits; ii) the parameter associated to the ECM term gives useful information on adjustment dynamics; iii) long-run estimated parameters indicate how a modification in the price, personal income, etc modify long-run habits in consumption. Second, it allows comparing long-run and short-run agents' behaviours by comparing long-run and short-run estimated parameters. Finally, it is a suitable framework for estimating a model irrespective of the degree of integration of the variables of interest.

3 Data and variables

We use an annual time series data set which covers the period 1950-2001 for the city of Milan. All variables have been transformed in logarithmic form.

3.1 Variables and sources

WATER CONSUMPTION. The annual data on daily average water consumption per-user - which is the sum of residential, commercial and industrial consumptions[2] - was collected by the municipal water network Acquedotto del Comune di Milano (c_t, in litres). The explanatory variables used to estimate the demand equation are:

[2] Probably, the main limitation of our data is that they do not allow the decomposition of total water consumption with respect to different typologies of users (residential, industrial, commercial) except for few years at the end of the period. In a micro-econometric setting with disaggregated data, water demand for water of each typology of user is estimated in a specific way. While residential water demand is obtained by maximising consumers' utility, industrial water is introduced as an input of production (generally it is considered as a substitute of labour and complementary to capital; Renzetti, 1992) and it is often estimated in a cost function framework. In spite of such a limitation, we believe that our data allow us to provide interesting new information about long run structural dynamics of total water consumption in a relevant policy scenario, which is that of a big European city. Moreover, we did some non reported robustness checks using an alternative definition of the dependent variable relative to consumption since we add total employment to total population within the denominator. This alternative definition could be relevant for studying industrial water consumption. Non reported results confirm our main findings.

$$\mathbf{x}_t' = (price_t, rain_t, prod_t, m_t) \, .$$

PRICE. $price_t$ is the price of water in real terms. It is defined as the nominal water price (Acquedotto del Comune di Milano, 2003) deflated by the average consumer price(ISTAT, 2007). Since the foundation of Milan's water network, the price of water in Milan has followed a nonlinear tariff structure. The choice of which price to include in the demand function is controversial. Economic theory suggests the use of marginal prices, while some studies on water consumption point out that users react to average prices rather than marginal prices (CREDOC, 1997). An approach commonly adopted in the face of a multiple-block price structure is to include in the demand function both the marginal price and a difference variable in order to take account of the intramarginal structure (Nordin, 1976).

In our study, it is not possible to calculate actual average and marginal prices, given the lack of data relative to disaggregated water consumption over the period considered. Indeed, the use of ex post average prices produces a problem of simultaneity, and predetermined information is required in order to identify the demand equation (Taylor, 1975). Despite the continuing debate among researchers on the most appropriate price variable to adopt, recent surveys show that results in terms of sign and intensity of price elasticity do not depend hugely on the type of price variable selected (Arbues et al., 2003).

In a first stage, we employ the value (in real terms) of the first block tariff as a proxy for water price. This block accounts for the largest share of consumption[3] and represents the core of the entire tariff plan, which applies to all typologies of users. For these reasons, the majority of users should perceive the effective water price to be the price of the lower block, rather than the ex-post average or marginal prices that also account for the other blocks, representing more specific typologies of consumers such as small and large firms. In a later stage, our analysis focuses on the price structure and its evolution.

RAINFALL. The variable $rain_t$ represents rainfall (per year), measured in millimetres per year (Osservatorio Metereologico di Milano Duomo, 2002).

PRODUCTION ACTIVITY. Production activity $prod_t$ is included in the model to take account of non-domestic consumption. Here, we adopt different proxies for the productive structure:

[3] In the year 2001, the most relevant part of the water bill regards the first block price rate which considers per capita consumption within 350 litres per day. The average water supplied for domestic consumption by Milan's water network is around 353 litres per day, obtained by 172,7 million m3 supplied in the year to the 1.339.933 citizens. Statistics about water consumption relative to the fifty years considered in this analysis and depicted in figure 1 concern total water consumption and not only residential/domestic water consumption. A recent comparative report on water consumption in Europe shows that Milan, with around 350 litres per day, presents a very high per capita domestic consumption compared to other European cities (Ambiente Italia, 2006).

- tot_emp≡ log (total employment/population) ;
- ind_emp≡ log (employment in manufacturing industry/population) ;
- serv_emp≡ log (employment in services/population) ;

which measure respectively, the intensity of total employment, manufacturing employment and services employment (Crenos, 2007).

INCOME. Another variable suggested by economic theory is personal income, even if in empirical studies on water consumption, it is often non-significant. Time series of personal income have not been gathered in Italy at the urban level. In a log-linear specification, the parameters that are of interest are elasticities, which are invariant to a change of scale. Under the quite reasonable assumption that during the period covered by our analysis, the evolution of personal income has been similar for Milan and Italy, then per-capita income for Milan can be proxied by per capita income for Italy, m_t (Heston et al., 2006)[4].

3.2 Description of the time series

The time series plots of the five variables included in the VAR are given in figure 1.

[4]Obviously there are many other variables involved in water demand, such as the demographic structure of the city. Unfortunately, yearly data on this and other variables are not available; many time series that would have provided interesting data are collected by the national census only every ten years.

FIGURE 1 ABOUT HERE

WATER CONSUMPTION. Daily per user water consumption increased up to the 1960s, when it stabilised at around 500 litres. There are some exogenous events, not accounted for here, that might have had some short run effects on consumption. These are primarily supply side factors, mostly linked to modifications to the structure and management of water systems. The early 1960s also saw the emergence of pollution problems in Milan's water system. This led to the closure in 1963 of 37 out of 55 wells. In the 1974 another discovery of serious water pollution, mainly due to the intensive industrialisation, was accompanied by a lowering of the water table due to over-exploitation (Bianchi, 1989). The Milan water network solved the pollution problems in the 1980s by temporary closure of the polluted wells. These closures did not affect water supplies, which were guaranteed by more intensive utilisation of the unpolluted wells. However, wells closures could reasonably impact on operating costs and, indirectly, on water price.

PRICE (first block). Up to 1974 the nominal price of water was set by a government committee, with no reference to management costs. In 1984, a law established that any increase in the price of water should not be higher than the forecast inflation rate. The real price of water shows a generally slowly decreasing trend, with dramatic shocks in the 1970s and 1980s primarily due to the rate of inflation in Italy, which in these periods went up to 20%.

RAINFALL. The rainfall series shows a very slightly decreasing trend with relatively constant variability over time: about 80% of the data fall within the interval 750-1,250 mm. There are some particular climatic anomalies in three years in the period: 1587 mm in 1950, 1583 mm in 1958 and 643 mm in 1990.

INCOME. Real per-capita income for Italy shows a general increase around a logarithmic trend with some shocks, corresponding to the oil crises.

PRODUCTION ACTIVITY. The variables for productive activity, IND_EMP, SERV_EMP and TOT_EMP were plotted. The process of heavy industrialisation (IND is the number of workers in industry) that occurred after the Second World War was coupled with a significant increase in population (POPU) up to the end of the 1960s. From the early 1970s a process of de-industrialisation was accompanied by a decrease in population, which also occurred in many other Italian cities mainly do to the declining birth rates and increased emigration from downtown to peripheries outside of the municipal territory.

Consequently, the share of industrial workers in the population (IND_EMP) remained fairly stable over the entire period. However, we can see that during the 1960s this share shows a generally decreasing trend, while in the 1990s it increases due to the stabilisation of employee numbers in the industrial sector which was coupled with a decrease in total population.

The number of employees in services (SERV) was fairly stable until the 1970s when a strong tertiarisation process occurred, resulting in a major increase in service employees that has continued. Thus, the share of service employment in the population (SERV_EMP) shows a U shaped trend: it decreased up to the beginning of the 1970s and increased thereafter.

Total employment (TOT=SERV+IND) shows fairly linear growth over the entire time period. Up to the 1970s it was driven by industry and after that time received a boost from services employment. The share of total employment in the population (TOT_EMP) follows a U shaped trend.

4 Preliminary results

4.1 Detecting the order of integration: are the time series really non-stationary?

Before testing for the existence of a long run relationship and estimating the model, we conducted a preliminary statistical analysis in order to detect the order of integration of the variables. Since the Augmented Dickey-Fuller (ADF) test was introduced, a number of alternative tests have been proposed that improve the size and power of the ADF. Elliott et al.'s (1996) quasi-differencing variant through local GLS detrending of the data (DFGLS) yields the greatest advantages with respect to alternative approaches.

The ADF and the DFGLS tests (table 1) both provide strong evidence favouring the unit root hypothesis for many time series. The ADF test indicates that all the variables, except RAIN, are non-stationary, while the DFGLS test shows both RAIN and IND_EMP to be stationary.

TABLE 1 ABOUT HERE

However, since Neslon and Plosser (1982) seminal work, many econometricians have shown that, for many economic time series, standard unit root tests fail to reject the null hypothesis of a unit root. For example, Kwiatkowski *et al.* (1992) argue that the standard unit root tests are not very powerful against relevant alternatives. They propose the so called KPSS test in which the unit root is the null hypothesis to be tested, and the way in which classical hypothesis testing is carried out ensures that the null hypothesis is accepted unless there is strong evidence against it.

When we apply the KPSS test (column iv, table 1) we cannot cross-validate the outcomes of ADF and DFGLS unit-root tests since for many of the time series under consideration the KPSS test strongly supports the stationarity null hypothesis.

The failure of the ADF and DFGLS tests to reject the unit root null hypothesis could be due to breaks or non-linearities in the trend function. Based on figure 1 there is reason to suspect that there may be a break in the trend functions for C, TOT_EMP and SERV_EMP. We therefore applied the Zivot and Andrews (1992) test, which allows for a break in the trend, with the result that two (TOT_EMP and SERV_EMP) out of three of the series are trend break stationary. Similarly, we can apply the Clemente et al. (1998) Innovative Outlier test to the price (PRICE) series. This test allows for two structural breaks -which occur gradually - in the mean of the series. Under the innovative outlier hypothesis, the price of water can now be considered stationary around a mean that changes in 1974 and 1985. (Results are presented in table 2 and figure 2).

TABLE 2 ABOUT HERE

FIGURE 2 ABOUT HERE

The idea that a time series can be stationary around a breaking deterministic linear trend, as in Zivot and Andrews (1992) and Perron (1989), was generalised by Bierens (1997), who adopted the notion of integration around a deterministic non-linear trend. The notion of non-linear trend stationarity can usefully be applied to the C, PRICE, TOT_EMP, SERV_EMP series, which present some breaks. It can also be applied to the income (M) series, which is clearly non-linear.

Here, we focus on the Bierens (1997) revised non-linear Dickey-Fuller test. Let denote the Chebishev polynomial as $P_{0,t}, ..., P_{m,t}$ where $P_{0,t}$ equals to 1, $P_{1,t}$ is a linear trend and $P_{2,t}, ..., P_{m,t}$ are cosine functions. The Bierens (1997) augmented Dickey Fuller test is based on the following auxiliary regression model:

$$\Delta z_t = \alpha z_{t-1} + \sum_{j=1}^{p} \phi_j \Delta z_{t-j} + \theta^T P_{t,n}^m + \varepsilon_t. \tag{5}$$

Bierens (1997) considers the null unit root hypothesis with drift against the alternative of non-linear trend stationarity, and develops several test statistics:

$\widehat{t}(m)$ is the t-statistic of the estimated coefficient $\widehat{\alpha}$,

$\widehat{A}(m) = \frac{n\widehat{\alpha}}{\left|1-\sum_{j=1}^{p}\widehat{\phi}_j\right|}$.

Where under the null $\alpha = 0$, while under the alternative $\alpha < 0$. The most important fractiles of the null distribution of $\widehat{t}(m)$ and $\widehat{A}(m)$ for $m = 1, \ldots, 20$ are tabulated in Bierens (1997) on the basis of 10000 replications of a Gaussian random walk with sample size 500. Since these two tests do not take account of all the available information, Bierens proposes two other tests $\widehat{F}(m)$ and $\widehat{T}(m)$ for the joint hypothesis that α and the last m components of the parameter vector θ are zero. The $\widehat{F}(m)$ is a conventional F-test whereas $\widehat{T}(m)$ is a χ^2 test based on non-conventional testing principles.

We conducted the non-linear ADF test on the basis of the auxiliary regression (1) with a lag length p chosen using the AIC and the Chebishev time polynomial order m, settled in accordance with the degree of non-linearity of the series.[5]. In order to

[5]In choosing m we are faced with a tradeoff: if m is too low – i.e. the nonlinerar trend is more non-linear than the Chebishev polynomial approximation – this might reduce the power of the test.

correct for the sise distortion we simulate the critical values using wild bootstrap[6] based on 2000 replications.

TABLE 3 ABOUT HERE

Looking at the asymptotic critical values we can reject the null unit root hypothesis for all variables excepting consumption – C. The results also show evidence of substantial sise distortion and the simulations ($F(m)$ and $T(m)$) still indicate the rejection of the unit root hypothesis for PRICE, TOT_EMP, SERV_EMP and M at the 5% significance level.

To sum up, based on the evidence from the different unit root testing procedures, we can conclude that P, TOT_EMP, SERV_EMP, M and RAIN are stationary time series, and C is first order integrated.

4.2 Testing for the existence of a long-run relationship

As in the cointegration analysis, after checking the order of integration of the variables and before estimating the model, we need to test for the existence of a long-run

On the other hand, if m is too high it may also reduce the power of the test by estimating superfluous parameters. We conducted the test using different values for m; the results were similar and are available upon request.

[6] Draw the model errors from normal distribution with zero mean and variance from the squared OLS residuals.

relationship (long-run water demand). The results from the previous section render cointegration tests and estimation inappropriate. We therefore test for the existence of a level relation between water consumption and its determinants by computing the F-statistic to test the significance of the lagged levels of the variables in the error correction specification of the underlying "conditional ECM" model defined by equation 2, as suggested by Pesaran et al. (2001). However, the asymptotic distribution of the F-statistic is not standard, irrespective of whether the regressors are I(0) or I(1). Pesaran et al. (2001) give two sets of critical values: one assumes that all regressors are purely I(1), and the other is computed assuming that all regressors are purely I(0), thereby providing for all possible classifications of the regressors as purely I(0), purely I(1) or mutually cointegrated.

To determine whether a deterministic trend is required we estimated the ARDL model with and without a linear trend, for p=1,2,3 and p chosen using the AIC. We found the time trend to be insignificant in both equations. We therefore chose not to introduce a linear trend. It should be noted also that a visual inspection of the time series reveals the absence of common trends in the variables.

With respect to the intercept, we concentrate on the unrestricted intercept model (model III) which does not incorporate the constraint $a_0 = -(\phi, \boldsymbol{\delta}')\,\boldsymbol{\mu}$. It should be noted that the Data Generating Processes for the models with restricted and unrestricted intercept are identical while estimation and hypothesis testing for model

III do not incorporate such restriction[7].

TABLE 4 ABOUT HERE

Table 4 presents the F-statistics for testing the existence of a level water demand equation. In all cases the F-statistics fall outside the critical bounds tabulated in Pesaran et al. (2001) at the conventional significance level: the hypothesis that there is no level water demand equation is rejected irrespective of the lag order and the proxy used for the productive structure.

5 Estimation results

5.1 Benchmark results

The set of specifications presented below takes account of the inclusion of different variables relative to the productive structure, which we suppose to be a crucial determinant of water consumption over the long-run. In a first specification, we include a variable that represents the average intensity of productive activity and is equal to the share of total workers in the population (TOT_EMP). In a second estimation we focus on services intensity - measured as number of service workers in total population (SERV_EMP) - since the service sector historically has increased in Milan as in other

[7] See Pesaran and Shin (1999) and Pesaran et al. (2001) for a more detailed discussion of the different specifications.

modern economies. In the third and fourth estimations we analyse and control for the role of the prevailing sector in different periods during the 50 years considered. This last approach is aimed at studying the role of structural change in the productive structure and its effect on long-run water demand. In the third specification we introduce two variables relative to productive activity. The first is built by multiplying the industry intensity (IND_EMP) with a step dummy DU5074 that takes the value 1 if year$<$ 1975; the second is obtained by multiplying the service intensity variable (SERV_EMP) by a step dummy DU7501 that takes the value 1 if year$\geq$1975. We label these two variables IND_EMP5074 and SERV_EMP7501 respectively. The idea behind this modelling is that, given the trends in relative sise observed for the different macro-sectors, industrial production will push non-domestic water consumption in the first half of the sample, while in the second half it will be services that have the main influence. The fourth specification cross-validates the third specification by estimating the impact of services in the first half of the time period (SERV_EMP5074), and the impact of manufacturing in the second half (IND_EMP7501). We expect the values and significance of the estimated coefficients to be lower.

The specification defined in eq. 2 is based on the assumption that the disturbances are serially uncorrelated. It is therefore important that the lag order p is selected appropriately. There is of course a delicate balance between choosing a p that is sufficiently large to allow serially uncorrelated residuals and, at the same time, sufficiently small such that the "unrestricted ECM" is not over-parameterised. The orders of the ARDL (p1, p2, p3, p4, p5) were chosen using the Akaike Information Criterion (AIC) subject to a maximum lag of 3, which seems a reasonable point of departure giving the (annual) frequency of the data and the number of observations available for estimation. For example, for both the specifications with TOT_EMP and with SERV_EMP as proxies for productive activity, the AIC results in an ARDL (1,0,3,2,0) specification. The "basic" ARDL then becomes:

$$
\begin{aligned}
c_t \;=\;& a_0 + \phi c_{t-1} + \alpha_1 price_t + \alpha_2 rain_t + \alpha_3 rain_{t-1} + \alpha_4 rain_{t-2} + \\
& + \alpha_5 rain_{t-3} + \alpha_6 tot\_emp_t + \alpha_7 tot\_emp_{t-1} + \alpha_7 tot\_emp_{t-3} + \\
& + \alpha_8 m_t + \epsilon_t
\end{aligned} \tag{6}
$$

while the conditional ECM can be written as:

$$
\begin{aligned}
\Delta c_t \;=\;& a_0 + \varsigma \left(c_{t-1} - \boldsymbol{\theta}'\mathbf{x}_{t-1}\right) + \varphi_1 \Delta price_t + \varphi_2 \Delta rain_t + \varphi_3 \Delta rain_{t-1} + \\
& + \varphi_4 \Delta rain_{t-2} + \varphi_5 \Delta tot\_emp_t + \varphi_6 \Delta tot\_emp_{t-1} + \varphi_7 \Delta m_t + \epsilon_t
\end{aligned} \tag{7}
$$

and long run level water demand is:

$$
c_t = \theta_0 + \theta_1 price_t + \theta_2 rain_t + \theta_3 tot\_emp_t + \theta_4 m_t + \upsilon_t \tag{8}
$$

The results from estimating the demand equation are presented in tables 5 and 6. All specifications pass the set of diagnostic checks for serial correlation, non-normal errors, functional form and heteroskedasticity.

The estimated coefficients and standard errors obtained from estimating the model with all explanatory variables, are presented in columns vi to viii, while the results obtained excluding income from the list of the explanatory variables as suggested by a general-to-specific modelling strategy[8] are in columns ii to v.

SHORT RUN ESTIMATES. Among the results obtained from the different econometric specifications, four are relevant.

The first, very clear result, concerns the variable for lagged water consumption. In each econometric specification it presents a positive coefficient around 0.7, always highly significant. It clearly reveals that persistent habits are crucial for water consumption.

The theoretical bases for estimating a dynamic water demand are described in the literature. Tversky and Khaneman (1991) point out that current consumer utility depends, not only on present consumption levels, but also on past consumption. Nauges and Thomas (2003) show that a dynamic model of residential water consumption can

[8] Note, that since price and income present some degree of collinearity (see also Martinez Espiñeira, 2007) we also estimated the model including income, but excluding price as a regressor. The results (which are not reported here) confirm that income is never significant and presents an estimated coefficient still very close to zero.

be derived from a structural optimisation programme solved by the communities.

TABLE 5 ABOUT HERE

These theoretical findings imply that ignoring habits in the estimation of a consumption function could lead to problems of mis-specification and biased results.

Empirical studies show that water consumption habits depend either on the use of durable equipment, such as washing machines, dishwashers, showers (Martínez Espiñeira, 2002; Arbuès et al., 2003) or on slow adaptations in consumer behaviour (Martínez Espiñeira and Nauges, 2004).

Among the studies that focus on the role of habits[9], we would draw attention to work that estimates autoregressive models, for example, Nauges and Thomas (2003) who find lower values for the estimated autoregressive component. However, direct comparisons are not easy, given the panel data structure of these studies.

A second and related point concerns the parameter associated with the ECM, which, as previously discussed, represents the speed of adjustment following a disturbance, for the restoration of equilibrium in the dynamic model. This is always significant and is further evidence of a stable long-term relationship. In our estimates, the estimated ECM coefficient is about - 0.2, which corresponds to a correction of

[9]See, among others, Carver and Boland (1980), Aghte and Billings (1980), Dandy et al. (1997), Renzetti (2002), Rossi (2005), Nauges and Thomas (2000; 2003).

20% in the next period and is in line with Martinez Espiñeira (2007) who analysed a monthly time series data set for the city of Seville. This means that adjustment takes place relatively slowly.

Third, there are short-run price elasticities showing the expected negative sign, significant but very low in absolute values, ranging from -0.037 to -0.018 and about four times lower than their long-run counterparts.

Fourth, the short-run parameter estimates of the other explanatory variables reveal some complex short-run dynamics. The estimated parameters associated with the current and past rainfall values seem to indicate that water users do not react quickly to climatic changes, since only the lagged coefficients are significant. The short-run dynamics relative to the variables representing productive activity crucially depend on which variable is introduced. For both TOT_EMP and SERV_EMP the AIC chooses two lags and the estimated parameters reveal some rather complex behaviour, producing fluctuations in the signs. In the specification that simultaneously incorporates IND_EMP5074 and SERV_EMP7501, the picture is clearer. In fact, the AIC chooses a model without lagged values for either variable and the estimated short-run contemporaneous elasticities are estimated at 0.05 and 0.06 and are significant.

LONG RUN ESTIMATES. Let us first discussing the results obtained from estimating the model with all explanatory variables (columns vi to viii) and focusing on the role of income. The long-run estimated coefficients for income are close to zero and never significant. This result for water consumption is quite usual; many previous studies show that water consumption is either inelastic or presents little elasticity with respect to income. Martinez Espiñeira and Nauges (2004) state explicitly that income is not *"a determinant factor to explain water use"* (p. 1702) since water represents a very limited component of European households' budgets (Arbues et al., 2004). There are also other explanations, which can be summarised as follows. First, consumers' lack of awareness of the cost of water may cause income elasticity to be close to zero (Dalhuisen et al., 2003). Second, interesting results regarding income effects typically emerge from estimating different levels of demand for different levels of income/wealth using micro data (Hanke and de Maré, 1982; Agthe and Billings, 1987). Third, income effects on water demand can only be indirect since they depend, at least partially, on the characteristics of the water-using capital stock (Martínez Espiñeira, 2007).

We now turn of our preferred estimates excluding income and presented in columns ii to v. Long-run price coefficients are significant in every specification, and with a negative sign, which confirms previous findings. The clearest result in terms of magnitude and significance, emerges from the specification where we introduce employment in the service sector (SERV_EMP) as a proxy for productive activity (the estimated coefficient of price is -0.145 with a p value equal to 0.02). In general, in our estimates, the intensity of long-run price elasticities is lower than in other European studies using dynamic models in the range (-0.50 to -0.30).

TABLE 6 ABOUT HERE

The underlying and relevant evidence arising from comparison of the short and long-term results, is that water users are more likely to adjust their consumption following a permanent increase in tariffs than in response to a one year shift. This could derive from the specific typology of the good we are considering, since water is essential for life and a large component of its consumption is not reducible at least in the short-run (Arbues et al., 2003; Haneman, 2006). One major effect on long-run water consumption is due to the renewal of water using capital stock, which refers to both domestic and industrial durable equipment (Nauges and Thomas, 2003; Martinez-Espineira, 2007). The slow rate of change in the stock of appliances, and the role of well entrenched habits, will contribute to a process of slow adaptation to price changes.

The next result concerns rainfall, which is significant in almost all cases. The estimated coefficients have negative signs, in line with the literature (Arbués et al., 2003; Martinez Espiñeira, 2002, 2007). This mostly reflects the effect of climate on water consumption due to increases in temperature and evapotranspiration rates (Renzetti, 2002). In this respect, outside water use, mainly for irrigation and street cleaning, plays a significant role, which is included in our dependent variable.

A third result relates to the group of variables used to analyse the impact of the productive structure. The variable that measures the average intensity of total productive activity (TOT_EMP) shows a positive and significant long-run elasticity, which is estimated at 0.16. The role played by water in industrial processes is crucial, despite the increasing efficiency of water consumption induced by technological innovation (Renzetti, 2002) and the growing relative importance of services. This latter aspect is highlighted in the model that includes intensity of the service sector (SERV_EMP) as an explanatory variable: the estimated coefficient is significant and positive (0.11), but is lower than that for total productive activity referred to above.

Our analysis estimates the impact on water demand of the currently prevailing sector, through the inclusion of the variables IND_EMP5074 and SERV_EMP7501. The estimated coefficients are highly significant and their magnitudes are higher with respect to the estimated parameters obtained through the inclusion of total productive sectors, and the service sector alone. This shows very clearly that changes in the productive structure are a relevant explanatory factor in the evolution of long-term water consumption.

The cross validation exercise provides confirmation of these results. It estimates the impact of services in the first period, and of manufacturing in the second period. The estimated coefficients are close to zero and highly non-significant.

5.2 Modification of the price structure over time and its impact on water consumption

In the previous section we focused on the price relative to the first block tariff, which includes all typologies of consumers: households, and small and large firms.

In this section we focus on the modification to the price structure over time, and its impact on water consumption. When the Milan water network was first established, prices were based on a three-block structure. In the 1970s, this structure was substantially modified, and up to 1975 water users experienced decreasing block rates; however, since 1976, the price structure has been based on increasing block rates (figure 3).

FIGURE 3 ABOUT HERE

This change towards an increasing block structure should be interpreted as a policy instrument designed to benefit households and lower level consumers, such as small firms. Prior to 1976, the decreasing block tariff structure benefited industry by providing lower water costs for this sector.

The evolution of water prices in Milan shows two distinct phases. First, in the first and third blocks, prices are significantly lower than in other European cities at the time and the current project to progressively increase water prices to the level of 1 euro per m^3 by the end of 2009, is designed to bring Milan's water prices within the "European range"[10]. Second, the inter-block price differences are higher in the second half of the period because of the sharp increase of the price of the third block following the 1976 changes to the price structure.

In order to investigate how these changes in price structure affected consumption, we re-estimated the demand equation and simultaneously include the two prices variables relative to the two periods. We built the first price variable by making the price interact with a step dummy DU5075 that takes the value 1 if year$<$ 1975; we built the second by multiplying the price by a step dummy DU7601 that takes the value 1 if year$\geq$1976. In a first set of estimations we include the prices for the first block (P1) for the two periods 1950-1975 and 1975-2001 (P15075 and P17601). In the next estimations we replace these variables with the prices for the third block (P3) for the same two periods (P35075 and P37601). As before, we estimate three different specifications according to the three proxies for productive structure.

The results of our estimations show that both short and long-run price elasticities, relative to the first block tariff (table 7, columns ii, iii and iv), are higher (about 1.35 times) for the period 1950-75, which is characterised by a higher first block price, than 1976-2001 under a decreasing block structure. These results provide clear evidence that water users reacted to the mid-1970s' modifications to the water price structure.

[10] E.g., in 2001, the price of the third block was 0.41 euros per m3, while in many other European cities it was 1-2 euros (UNEP, 2004) .

We now turn to discussion of the estimates obtained using the price variables relative to the third block (P35075 and P37601, in columns v, vi and vii of table 7). As expected, the results are reversed: consumption elasticities with respect to the price corresponding to the third block are higher (about 1.7 times) in the last period 1976-2001, under an increasing block tariff, than in the first period. The effect on price elasticities of changing the price structure, is greater for the third than the first block. Moreover, the third block short and long-run price elasticities for the period 1950-75, under decreasing price rates, are often not significant.

TABLE 7 ABOUT HERE

To better understand these results we need to consider that the first block accounts for all types of users, both domestic and firms, whereas the third block involves only large consumers, typically industrial firms.

According to our findings, even though the price change stimulated reactions from both small and big consumers, its most incontrovertible effect was on firm behaviour, since variation in price elasticity is most evident for large scale consumers. This could be due to the fact that demand for industrial water also depends both on the degree of substitution from other inputs and on the level of other prices. We can assume that in the first period, 1950-75, water was a very convenient input with respect to some other factors of production. Finally, higher price elasticity for large scale consumption in the second half of the time period, might derive from technological innovation, which should increase water utilisation efficiency and the capacity to save water.

6 Concluding remarks and some comments on water policy

This paper estimated a water demand function on a 50 year sample for the city of Milan. Allowing for trend break stationarity or non-linear trend stationarity, we find that the evidence does not support the unit root hypothesis for many time series. Accordingly, the standard cointegration analysis is not appropriate and therefore we adopted an alternative estimation and testing procedure. We focused especially on the so-called bounds testing approach proposed by Pesaran et al. (2001) which can be applied irrespective of the level of integration of the variables and which provides a useful modelling strategy given that dynamics are important in the estimation of a water demand equation (Nauges and Thomas, 2003).

We estimated the effect on water consumption per user, of different factors: price, climate, income and productive structure, over the period 1950-2001.

The first relevant result is related to the estimated parameter of lagged water consumption, which is positive and highly significant, showing the existence of well entrenched habits. Alternatively, we can say that after a disturbance, water users adjust their levels of consumption towards a long run equilibrium, only slowly.

A second important finding is evidence that long run price elasticity is higher than short run elasticity. This could be due to the intrinsic economic characteristics of water - since it is essential for life and in part, not substitutable - and to the long time needed to change water using equipment. We find that long run price elasticities are significant even though they are smaller than those obtained in other studies on dynamic water demand.

The tariff scheme introduced in the mid 1970s promoted substantial changes in the pricing structure, from a decreasing to an increasing block structure. These changes have led to a decrease in the price elasticity for the first block and an increase in the elasticity for the third block. Moreover, the third block price elasticity was not significantly different from zero before the change and become significant thereafter. The third block is related to large scale consumption typically attributable to firms, which seems to show a significant reaction to changing price schemes.

These results on price elasticities present quite clear implications for policy. Price policies can have an effect on water consumption over the long term. Water price policies could be fundamental to changes in behaviour since, for all types of consumers, households and firms, they signal the need to adopt new consumption habits in the face of an increasingly scarce and essential good such as water.

However, since our results show that habits play an important role in water consumption, other policies will need to be combined with price policies in order for consumption levels to be moderated. Such policies could include information campaigns. As in other European countries, these campaigns should be aimed at informing households about both the prices/costs of water supply and about the most efficient ways to use and save water.

References

Acquedotto di Milano. 2003. Banca dati. Milano.

Agthe DE, Billings B. 1980. Dynamic Models of Residential Water Demands. Water Resources Research, 16 : 476-480. DOI: 10.1080/00420980601038255

Agthe DE, Billings RB. 1987. Equity, price elasticity, and household income under increasing block rates for water. American Journal of Economics and Sociology. 46 : 273-286. DOI: 10.1111/j.1536-7150.1987.tb01966.x

Ambiente Italia. 2006. Urban ecosystem Europe, www.ambienteitalia.it

Arbués Garcia F, Angeles Garcìa Valiñas M, Martìnez-Espiñeira R. 2003.Estimation of residential water demand: a state-of-the-art review. Journal of Socio-Economics. 32 : 81-102. DOI: 10.1016/S1053-5357(03)00005-2

Bianchi R. 1989. L'acquedotto di Milano: sviluppo e prospettive. Rivista Milanese di Economia, 32: 84-97.

Bierens HJ. 1997. Testing the unit root with drift hypothesis against nonlinear trend stationarity, with an application to the US price level and interest rate Journal of Econometrics, 81 : 29-64. DOI: 10.1016/S0304-4076(97)00033-X

Carver P, Boland J. 1980. Short run and long run effects of price on municipal water use. Water Resources Research. 16 : 609.616.

Clemente J. Montañés A. Reyes M. 1998. Testing for unit roots in variables with a double change in the mean. Economics Letters 59 : 175–182. doi:10.1016/S0165-1765(99)00131-7

CREDOC. 1997. L'Eau et les Usages Domestiques. Comportements de consommation de l'eau dans les ménages, Cahier de Recherche. 104.

CRENOS. 2007. Regional Accounts, CRENOS Databanks, http://www.crenos.it/databanks/ ital_reg/italianregions.php#Reg_Acc, 15 june 2007.

Dalhuisen J, Florax RJGM, De Groot HLF, Nijkamp P. 2003. Price and income elasticities of residential water demand: a meta analysis. Land Economics. 79. 292-308. DOI: 10.3368/le.79.2.292

Domene E, Saurì D. 2006. Urbanisation and water consumption: influencing factors of the metropolitan region of Barcelona. Urban Studies. 43 : 1605-1623. DOI: 10.1080/00420980600749969

Dupont D, Renzetti S. 2001. Water's Role in Manufacturing. Environmental & Resource Economics. 18 : 411-432. DOI: 10.1023/A:1011117319932

Elliott G, Rothenberg TJ, Stock JH. Efficient tests for an autoregressive unit root, Econometrica, 64 : 813–836. DOI: 10.2307/2171846

Hakkio CS. Rush M. 1991. Cointegration: how short is the long-run? Journal of International Money Finance 10 : 571–581. DOI: 10.1016/S0261-5606(99)00028-5

Haneman WM. 2006. The economic conception of water, in Water Crisis: myth or reality? (Eds.) P.P. Rogers, M.R. Llamas, L. Martinez-Cortina, Taylor & Francis plc., London.

Hanke SH, De Maré L. 1982. Residential water demand: a pooled time-series cross-section study of Malmö, Sweden. Journal of the American Water Resources Association. 18. 621-625. DOI: 10.1111/j.1752-1688.1982.tb00044.x

Heston I, Summers R, Aten B. 2006. Penn World Table Version 6.2. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, September, http://pwt.econ.upenn.edu/

ISTAT. 2007. Il valore della moneta dal 1861 al 2006. ISTAT. 18 november 200

Klein B, Kenney D, Lowrey J, Goemans C. 2007. Factors influencing residential water demand: a review of the literature, University of Colorado, Western Water Assessment, Working Paper, 29 february 2008, http://wwa.colorado.edu/resources/ water_demand_and_conservation/ literature_review_version_1_12_07.pd

Kwiatkowski D, Phillips PCB, Schmidt P, Shin Y, 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?. Journal of Econometrics. 54 : 159–178. DOI: 10.1016/0304-4076(92)90104-Y

Martínez Espiñeira R, Nauges C. 2004. Is all domestic water consumption sensitive to price control?. Applied Economics. 36 : 1697-1703. DOI: 10.1023/A:1014547616408

Martínez Espiñeira R. 2002. Residential Water Demand in the Northwest of Spain. Environmental and Resource Economics. 21 : 161-187. DOI: 10.1023/A:1014547616408

Martínez Espiñeira R. 2007. An Estimation of Residential Water Demand Using Co-Integration and Error Correction Techniques. Journal of Applied Economics. 10 : 161-184

Naimzada A.K., Tramontana F. 2008. Un modello dinamico del consumatore con razionalità limitata e analisi globale. Economia Politica. April. 1 : 59-94. DOI: 10.1428/26456

Nauges C, Thomas A. 2003. Long-run study of residential water consumption. Environmental and Resource Economics. 26 : 25-43. DOI: 10.1023/A:1025673318692

Nelson CR, Plosser CI. 1982. Trends and random walks in macro-economic time series. Journal of Monetary Economics 10 : 139–162. DOI: 10.1016/0304-3932(82)90012-5

Newey W, West K. 1994. Automatic lag selection in covariance matrix estimation. Review of Economic Studies 61: 631-653. DOI: 10.2307/2297912

Nordin J. 1976. A Proposed Modification of Taylor's Demand Analysis: Comment. The Bell Journal of Economics. 7 : 719-721. DOI: 10.2307/3003285

Osservatorio Meteorologico di Milano Duomo. 2002. Serie storica delle precipitazioni a Milano 1889-2001, Milano, mimeo.

Pesaran MH, Shin Y. 1999. An autoregressive distributed lag modelling approach to cointegration analysis, in Strom S (eds). Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium, Cambridge University Press, Cambridge.

Perron P. 1989. The great crash, the oil price shock, and the unit root hypothesis. Econometrica. 57 : 1361-1401. DOI: 10.2307/1913712

Pesaran, MH, Shin Y, Smith RJ. 2001. Bounds testing approaches to the analysis of level relationships. Journal of Applied Econometrics. 16 : 289-326. DOI: 10.1002/jae.616

Renzetti, S. 1992. Estimating the Structure of Industrial Water Demands: The Case of Canadian Manufacturing. Land Economics 68(4), 396–404.

Renzetti S. 2002. The economics of water demand. Kluwer Academic Publisher, Boston.

Taylor R. 1975. The Demand for Electricity: A Survey, The Bell Journal of Economics, 6 : 74-110. DOI: 10.2307/3003205

Tversky A, Kahneman D. 1991. Loss Aversion in Riskless Choice: A Reference-Dependent Model, Quarterly Journal of Economics. 106 : 1039-1061. DOI: 10.2307/2937956

UNEP. 2004. Freshwater in Europe.

http://www.grid.unep.ch/product/publication/freshwater_europe.php

Zivot E, Andrews D. 1992. Further evidence of the great crash, the oil-price shock and the unit-root hypothesis. Journal of Business and Economic Statistics. 10 : 251–270.

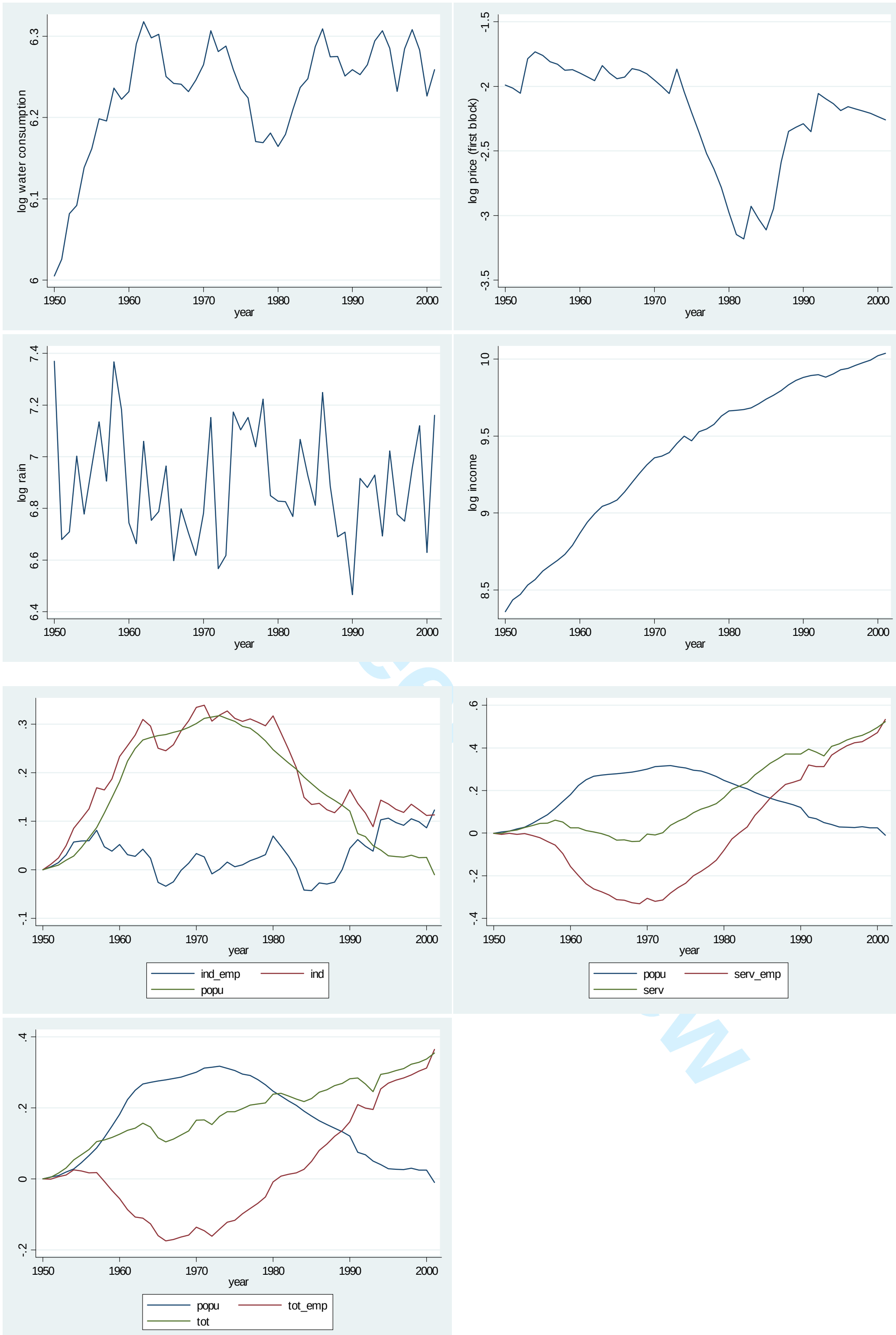

Figure 1. The time series (logarithms)

Figure 2. Unit Root tests with break

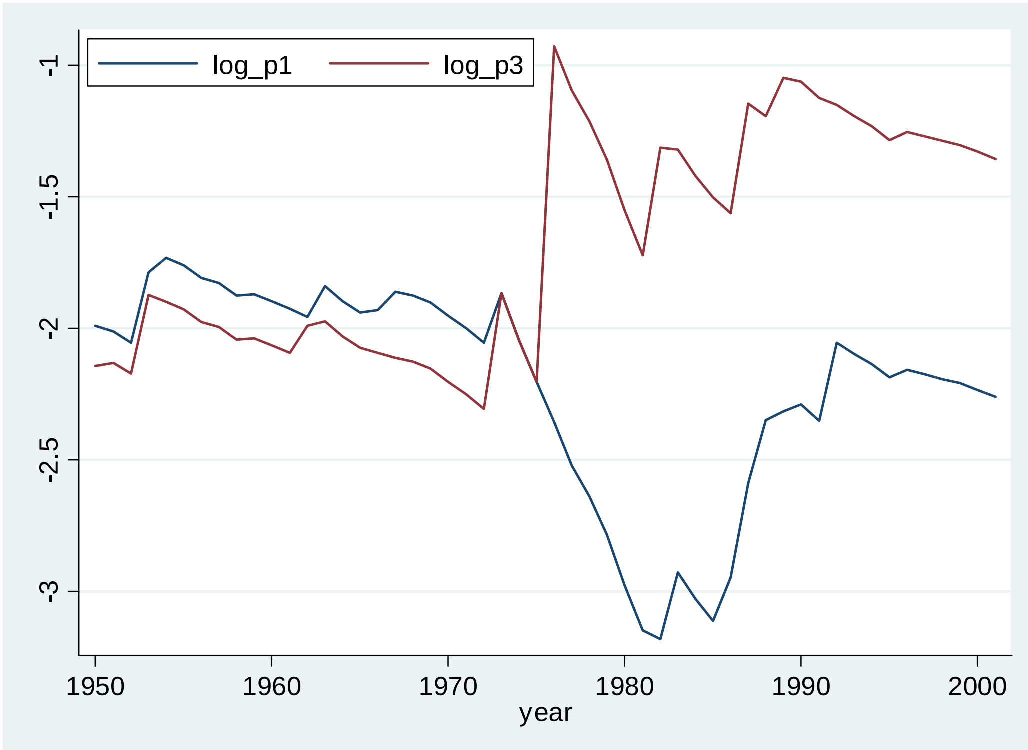

Figure 3. Modification in the price structure (logarithms)

Table 1: Unit root tests

Variable	DC	ADF	DFGLS	KPSS
c	c, t	-3.101	-1.704	0.115
c	c	-3.668*	-0.852	0.39
Δc	c	-6.157 *	-6.096*	0.326
$price$	c, t	-1.105	-1.173	0.131
$\Delta price$	c	-5.143 *	-5.187*	0.109
$rain$	c, t	-6.808 *	-5.538*	0.063
$\Delta rain$	c	-11.717 *	-8.551*	0.083
m	c, t	-1.030	-0.171	0.250
Δm	c, t	-6.780 *	-4.620*	0.146
tot_emp	c,t	-1.355	-1.135	0.238*
Δtot_emp	c, t	-5.329*	-2.813	0.103
ind_emp	c	-1.637	-2.139	0.245
Δind_emp	c	-5.5718*	-5.637*	0.236
$serv\_emp$	c, t	-2.5413	-1.819	0.233*
$\Delta serv\_emp$	c, t	-2.017	-1.840	0.121

Notes.

Δ: first-difference operator.

*: rejection of null hypothesis at 5% level.

DC: deterministic component. c is a constant while t is a linear time trend.

ADF: Augmented Dickey-Fuller test.

DFGLS: quasi-differencing variant the ADF proposed by Elliott et al. (1996) through local GLS detrending of the data.

ADF,DFGLS: the order of the auto-regressive component has been chosen with the AIC.

KPSS uses the automatic bandwidth selection procedure proposed by Newey and West (1994).

Table 2. Testing for unit root with breaks

Variable	DC	ZA	CMR	Type of break	time of the break
c	c, t	-3.058		t	1993
price	c		-5.538*	c1,c2 (IO)	1974, 1985
tot_emp	c,t	-5.015*		t	1967
serv_emp	c,t	-4.866*		t	1963

*: rejection of null hypothesis at 5% level.

DC: deterministic component. c is a constant while t is a linear time trend.

ZA is the Zivot and Andrews (1992) test.

CMR: Clemente et al. (1998) test allowing two changes in the mean.

t indicates that the break occurs in the deterministic linear trend while c means a break in the mean.

IO: Innovative Oulier version of the test allowing gradual changes.

Table 3. Test of the unit root with drift hypothesis against nonlinear trend stationarity (Bierens, 1997).

Variable	p	m	Test type	test statistics	Fractiles		simulated
					0.05	0.10	p-values
c	*0*	10	$\widehat{t}\,(m)$	-5.7862	-6.67	-6.29	0.2780
			$\widehat{A}\,(m)$	-38.9295	-80.30	-73.70	0.6820
			$\widehat{F}\,(m)$	5.6622	2.15	2.36	0.8510
			$\widehat{T}\,(m)$	1654.7126	280.57	359.51	0.9370
price	*1*	10	$\widehat{t}\,(m)$	-6.947	-6.67	-6.29	0.0630
			$\widehat{A}\,(m)$	-135.86	-80.30	-73.70	0.0610
			$\widehat{F}\,(m)$	8.7316	2.15	2.36	0.9530
			$\widehat{T}\,(m)$	2597.6870	280.57	359.51	0.9050
m	*1*	5	$\widehat{t}\,(m)$	-4.2173	-5.16	-4.83	0.2685
			$\widehat{A}\,(m)$	-37.2462	-48.70	-43.40	0.4015
			$\widehat{F}\,(m)$	6.6236	1.83	2.08	0.9680
			$\widehat{T}\,(m)$	1417.7731	18.60	30.15	0.9955
tot_emp	*1*	5	$\widehat{t}\,(m)$	-5.7052	-5.16	-4.83	0.0050
			$\widehat{A}\,(m)$	-71.6307	-48.70	-43.40	0.0587
			$\widehat{F}\,(m)$	10.8567	1.83	2.08	0.9997
			$\widehat{T}\,(m)$	2737.2529	18.60	30.15	0.9997
serv_emp	*1*	5	$\widehat{t}\,(m)$	-4.161	-5.16	-4.83	0.3824
			$\widehat{A}\,(m)$	-54.589	-48.70	-43.40	0.1784
			$\widehat{F}\,(m)$	9.887	1.83	2.08	0.9948
			$\widehat{T}\,(m)$	7165.041	18.60	30.15	0.9999

Nonlinear ADF test with lag length p chosen using the AIC.

Simulated critical values are computed using wild bootstrap based on 2000 replications.

Table 4. F statistics for testing the existence of level water demand equation

Specification p	total employment	service employment	Ind&Serv
1	3.9608**	3.9690**	3.3750**
2	4.6034***	4.7792***	2.6761*
3	3.8464**	3.8808**	4.0127***
AIC	3.9827**	4.1866**	3.3300**

"Total employment": the demand equation is estimated using tot_emp as proxy for the productive activity.
"Service employment": serv_emp is used as a proxy for the productive activity.
"Ind&Serv": both ind_emp5074 and serv_emp7501 are used.
*,**,*** significant respectively at 10% 5% and 1% level.

Table 5. Short run estimates

	Without income				With income		
ARDL	Total	Service	Ind&Serv	Cross val.	Total	Service	Ind&Serv
Regressor							
c(-1)	.726***	.743***	.741***	.696***	.726***	.750***	.792***
	(.010)	(.060)	(.067)	(.093)	(.094)	(.092)	(.098)
price	-.018*	-.037***	-.029*	.005	-.024*	-.037**	-.031**
	(.010)	(.013)	(.015)	(.017)	(.014)	(.014)	(.015)
rain	.015	.020	.017	.023	.127	.020	.013
	(.016)	(.016)	(.017)	(.018)	(.017)	(.017)	(.018)
rain(-1)	-.069***	-.071***	-.040**	-.048**	-.071***	-.071***	-.048**
	(.017)	(.017)	(.017)	(.019)	(.018)	(.017)	(.020)
rain(-2)	.008	.006	.021	-0.004	.008	.006	.019
	(.017)	(.017)	(.017)	(.018)	(.018)	(.017)	(.017)
rain(-3)	-.035*	-.045***	-.034*	-.037**	-.036**	-.046**	-.035**
	(.016)	(.016)	(.017)	(.017)	(.017)	(.017)	(.017)
tot_ emp	-.317				-.263		
	(.212)				(.233)		
tot_ emp(-1)	-.119				–.117		
	(.356)				(.359)		
tot_ emp(-2)	.480**				.436**		
	(.237)				(.251)		
m					-.120	-.002	-.019
					(.020)	(.020)	(.026)
serv_ emp		-.419**				-.412**	
		(.184)				(.200)	
serv_ emp(-1)		.088				.090	
		(.306)				(.310)	
serv_ emp(-2)		.358*				.350	
		(.192)				(.208)	
ind_ emp5074			.051**				.057**
			(.024)				(.025)
serv_ emp7501			.057**				.061**
			(.026)				(.026)
serv_ emp5074				-.015			
				(.173)			
serv_ emp5074(-1)				-.440			
				(.284)			
serv_ emp5074(-2)				.138			
				(.316)			
serv_ emp5074(-3)				.308			
				(.196)			
ind_ emp7501				-016			
				(.176)			
ind_ emp7501(-1)				-.446			
				(.288)			
ind_ emp7501(-2)				.144			
				(.320)			
ind_ emp7501(-3)				.306			
				(.199)			
CONSTANT	2.49***	2.33***	2.15***	2.27***	2.43***	2.32***	2.15***
	(.554)	(.498)	(.546)	(.540)	(.567)	(.518)	.(549)
$ecm(-1)$	-.273***	-.256***	-.258***	–.303***	-.230***	-.249***	-.207**

The estimated coefiicients of c(-1),...,*ind_ emp7501(-3)* are from the "basic" ARDL defined in equation 8

The coefficient of the ecm term is estimated from the conditional ecm defined by equation 9.

“Total”: the demand equation is estimated using tot_emp as proxy for the productive activity.

“Service”: serv_emp is included. “Ind&Serv” means that both ind_emp5074 and serv_emp7501 are used.

“Cross val.”: inclusion of serv_emp5074 and ind_emp7501.

*,**,*** significant respectively at 10% 5% and 1% level.

Standard error in parentheses.

Table 6. Long run estimates

	Without income				With income		
	Total	Service	Ind&Serv	Cross val.	Total	Service	Ind&Serv
price	-.067*	-.145**	-.113*	.019	-.104	-.152*	-.153
	(.040)	(.060)	(.066)	(.054)	(.092)	(.091)	(.113)
rain	-.294**	-.349***	-.137	-.208	-.378*	-.363*	-.242
	(.120)	(.130)	(.123)	(.147)	(.228)	(.194)	(.244)
m					-.052	-.008	-.093
					(.108)	(.082)	(.164)
tot_emp	.159*				.241		
	(.092)				(.203)		
lserv_emp		.106*				.114	
		(.056)				(.096)	
*ind_emp*5074			.198**				.276
			(.095)				(.187)
*serv_emp*7501			.221**				.296
			(.104)				(.191)
*ind_emp*7501				-.037			
				(.207)			
*serv_emp*5074				-.028			
				(.202)			
CONSTANT	9.114***	9.081***	8.342***	7.505***	10.591*	9.294***	10.366***
	(1.022)	(.969)	(.890)	(2.157)	(3.332)	(2.340)	(3.804)

The long run coefficient are estimated from the level relation defined by equation 10.
"Total": the demand equation is estimated using tot_emp as proxy for the productive activity.
"Service": serv_emp is used. "Ind&Serv" means that both ind_emp5074 and serv_emp7501 are used.
"Cross val.": inclusion of serv_emp5074 and ind_emp7501.
*,**,*** significant respectively at 10% 5% and 1% level.
Standard error in parentheses.

Table 7. Estimating price elasticities

ARDL	FIRST BLOCK			THIRD BLOCK		
	Total	Service	Ind&Serv	Total	Service	Ind&Serv
Regressor						
$p15075$	-.0432*	-.060***	-.047*			
	(.024)	(.023)	(.025)			
$p17601$	-.032*	-.048***	-.035**			
	(.016)	(.016)	(.016)			
$p35075$				-.024	-.030*	-.022
				(.016)	(.016)	(.019)
$p37601$				-.046*	-.057**	-.038
				(.024)	(.025)	(.026)
.....						
$Constant$	2.55***	2.34***	2.29***	2.35***	2.07***	2.02***
	(.556)	(.496)	(.570)	(.550)	(.509)	(.573)
Serial corr. χ^2 (1)	.061[.80]	.499[.48]	.282[59]	.510[.47]	.089[.76]	1.302[.25]
Func. form χ^2 (1)	.675[.41]	.238[62]	1.461[.22]	1.207[.27]	.490[.48]	1.920[.16]
Normality χ^2 (2)	.320[.85]	.376[.82]	.040[.98]	1.117[.57]	.717[.69]	.210[.90]
Hetero. χ^2 (1)	.843[.35]	.323[56]	.692[.40]	1.271[.26]	.403[.52]	.641[.42]
Long-run estim.						
$p15075$	-.165	-.244**	-.173*			
	(.1087)	(.117)	(.094)			
$p17601$	-.122*	-.197**	-.127*			
	(.071)	(.084)	(.066)			
$p35075$				-.089	-.153*	-.085
				(.064)	(.086)	(.074)
$p371601$				-.167*	-.235*	-.146
				(.098)	(.137)	(.110)
tot_emp	.306*			.118		
	(.177)			(.116)		
$serv\_emp$		.202*			.409	
		(.108)			(.410)	
$ind\_emp5074$			.251**			.114
			(.109)			(.093)
$serv\_emp7501$			.264**			.123
			(.112)			(.091)
$rain$	-.286**	-.339**	-.111	-.261**	-.204	-.137
	(.125)	(.134)	(.119)	(.117)	(.135)	(.131)
$CONSTANT$	9.77**	9.48***	8.41***	8.58***	7.14***	7.81***
	(1.294)	(1.107)	(.851)	(.965)	(1.005)	(.859)

The ARDL estimated coefiicients of p15075,...,p37501 are from the "basic" ARDL specification
The long run coefficient are estimated from the level relation.
... Non reported ARDL estimates for saving space.
“Total”: the demand equation is estimated using tot_emp as proxy for the productive activity.
“Service” serv_emp is used. “Ind&Serv” means that both ind_emp5074 and serv_emp7501 are used.
“Cross val.”: inclusion of serv_emp5074 and ind_emp7501.
*,**,*** significant respectively at 10% 5% and 1% level.