The Australian Journal of
Agricultural and
Resource Economics

Journal of the Australian Agricultural and Resource Economics Society

The Australian Journal of Agricultural and Resource Economics, 54, pp. 61–80

Household size and residential water demand: an empirical approach*

Fernando Arbués, Inmaculada Villanúa and Ramón Barberán†

The effectiveness of pricing policies depends on the price elasticity of consumption. It is well documented that residential demand for water is influenced by heterogeneity associated with differences in the size of the household and socioeconomic characteristics. In this paper, we focus on household size. Our initial hypothesis is that users' sensitivity to changes in price is different depending on the number of household members. To this end, we carry out an empirical estimation of urban water demand in Zaragoza (Spain) distinguishing between households with different sizes using data at the individual level. As far as we are aware, this approach to urban residential water demand is new in the literature. The analysis suggests that all households are sensitive to prices regardless of size. A more relevant finding is that small households are more sensitive to price changes.

Key words: demand analysis, household size, price elasticity, residential water demand.

1. Introduction

Increased efforts to improve urban water management are focused on demand side policies, seeking to affect the behaviour of users so that a 'reasonable' use of water resources is reached. Pricing reforms are among several measures implemented to encourage this 'reasonable' use of water. The effectiveness of these policies in engaging water consumption depends on the price elasticity of consumption. The larger the price elasticity, the more effective these policies are at reducing water consumption.

In this framework, the accurate characterization of water demand play a major role in obtaining sufficient knowledge about this behavioural response to changes in price. For urban water demand, given that domestic consumption is metered at the level of the dwelling and not individually this accurate

* The authors are very grateful to M.T. Aparicio and two anonymous reviewers for their valuable comments and suggestions, which have helped improve the paper considerably. Authors also acknowledge the support provided by the Zaragoza City Council through the Collaboration Agreement for Water Research made with the University of Zaragoza. Financial support received from the Spanish Ministry of Education and Science (projects SEJ2005-06069/ECON and SEJ2006-09961) is gratefully acknowledged.

† Fernando Arbués (e-mail: farbues@unizar.es) is at the Department of Applied Economics and Institute of Environmental Sciences, University of Zaragoza, Zaragoza, Spain. Inmaculada Villanúa (e-mail: villanua@unizar.es) is at the Department of Economic Analysis, University of Zaragoza, Zaragoza, Spain. Ramón Barberán is at the Department of Applied Economics and Institute of Environmental Sciences, University of Zaragoza, Zaragoza, Spain.

doi: 10.1111/j.1467-8489.2009.00479.x

characterization requires precise water demand models that account for differences in water consumption behaviour across households. The sources of this heterogeneity in the water demand patterns of different households are directly related with its composition both quantitatively (the size of the household) and qualitatively (the socioeconomic characteristics of household members).

Most empirical studies have found residential demand for water is influenced by heterogeneity associated with differences in the size of the household and socioeconomic characteristics such as income, age distribution and household preferences towards water use and conservation (Arbués *et al.* 2003; Worthington and Hoffmann 2008), like energy residential demand (Madlener 1996; Espey and Espey 2004).

It is well documented that household size affects demand for water positively (Arbués *et al.* 2004; Arbués and Villanúa 2006; Hoffmann *et al.* 2006; Schleich and Hillenbrand 2009) and energy (Nesbakken 1999; Yoo *et al.* 2007; Navajas 2009). Income level is also positively related to water consumption in most cases (Dandy *et al.* 1997; Arbués *et al.* 2004; Arbués and Villanúa 2006; Gaudin 2006; Hoffmann *et al.* 2006; Frondel and Messner 2008). This positive relation is also demonstrated in most papers focused on energy consumption analysis (Dubin and McFadden 1984; Narayan and Smyth 2005; Yoo *et al.* 2007; Navajas 2009). Other works find age distribution within the household to be relevant (Lyman 1992; Musolesi and Nosvelli 2007; Kenney *et al.* 2008; Dale *et al.* 2009; Schleich and Hillenbrand 2009, for water demand; Nesbakken 1999; Labandeira *et al.* 2006, for energy demand). Several works on residential demand for water (Nieswiadomy 1992; Musolesi and Nosvelli 2007; Frondel and Messner 2008), and energy (Dubin and McFadden 1984; Reiss and White 2005; Yoo *et al.* 2007; Dale *et al.* 2009) include variables such as lagged consumption, educational level, durable goods and equipments, etc. that reflect the relationship between the preferences towards use and conservation and demand decisions.

However, other variables than price are rarely central in studies on residential water demand. A substantial body of literature uses an aggregate approach to analyse the sensitivity of demand to changes in these variables and so cannot be used to test for heterogeneity across consumer groups (Blundell 1988). Only a few studies have carried out estimations of water demand distinguishing between users with different characteristics. The most common differentiation is according to household income (Agthe and Billings 1987; Renwick and Archibald 1998; Frondel and Messner 2008; Kenney *et al.* 2008; Schleich and Hillenbrand 2009), like energy demand studies (Nesbakken 1999; Reiss and White 2005). Other differentiations used are the availability of certain durable equipments and installations (Renwick and Archibald 1998; Arbués *et al.* 2004, for water; Nesbakken 1999, for energy) and household preferences towards water use and conservation (Krause *et al.* 2003; Gaudin 2006; Hoffmann *et al.* 2006, for water; Labandeira *et al.* 2006, for energy).

Given the above points, if the analysis of urban water demand is intended as a basis for water management measures, regulators will benefit from knowing how household size might affect demand for water. This is the aim of this paper. Our initial hypothesis is that users' sensitivity to changes in price is different depending on the number of household members. To this end, we carry out an empirical estimation of urban water demand in Zaragoza (Spain) distinguishing between households with different sizes using data at the individual level. As far as we are aware, this approach to urban residential water demand is new in the literature.

The interest of this approach lies in two essential points: first, and the most relevant issue, is that a demand analysis that takes into account the heterogeneity of price response across households provides very useful information in order to obtain a more accurate estimation of the impact of water pricing reforms on households. Therefore, water utility managers might implement more effective water demand management programmes. Second, more accurate comparisons between geographical areas with different household structures can be made based on demand models that reflect how household size changes affect demand for water. If household size is explicitly introduced in demand analysis, the risk of error diminishes when results obtained from analysing water demand in a geographical area are used to evaluate the potential effectiveness of water demand management programs in other areas.

The paper is organized as follows. First, in section 2 we describe the variables included in the proposed demand model. The main characteristics of the data collected are reported in section 3. Section 4 contains an intuitive approach to the relationship between household size and water consumption. The method applied to estimate water demand is presented in section 5. The main results of the empirical study are described in section 6. Finally, section 7 closes the paper with considerations of the results obtained.

2. Economic model

As it is mentioned above, the central objective of this paper is to carry out diverse urban water demand estimations in Zaragoza considering different household sizes. From available empirical papers for estimating urban water demands in Zaragoza (Arbués *et al.* 2004; Arbués and Villanúa 2006), we propose the following demand model:

$$q_{it} = f(de_{it-2}, DC_{it}, W_{it}, CHW_{it}, DN_{it}, AG20_{it}, AG60_{it}), \qquad (1)$$

where $q_{it} \equiv$ *Daily water consumption measured at the household level.* Although Zaragoza's municipal regulations provide for quarterly meter readings, the standard period is not generally observed and readings are taken at a range of different intervals. For this reason, we have converted consumption data for each billing cycle to a daily basis to obtain temporally homogeneous information for water consumption.

de_{it-2} ≡ *Two-lagged daily expenses*. This lagged average price specification is defined as the amount of the water bill divided by the number of days included in the billing cycle. As Arbués *et al.* (2004) show, this is the most appropriate price specification for Zaragoza residential water consumers given the billing cycle (users receive their water bill 6 months after the meter reading, that is to say, two readings later) and the price structure established in the municipal regulations.

In Zaragoza, water bills include a fixed part and a volume charge. The fixed part is a charge which does not entitle users to use a free number of units of water. The volume-based charge is a variable single charge applied according to a progressive linear tariff of 205 prices. This tariff means that all units recorded on the meter are paid at the same price, which increases progressively as the daily average consumption of a dwelling rises, resulting in an increasing straight line. Because all units are charged at the same price, the average price paid by customers is the price taken from the official tariff. Table 1 shows the linear volume-based charge, for 1996, 1997 and 1998, contained in the Municipal Regulations of Zaragoza City Council.

Figure 1 illustrates the water tariff structure used for residential customers in Zaragoza shown in Table 1. If a household consumes q_A cubic meters of water, according to the official tariff, each unit of water will be charged at the same price p_A, so the household will pay a volume charge equal to $p_A \times q_A$ (area Oq_AAp_A in Figure 1). If water consumption changes to q_B cubic meters, now each unit of water will be charged at the same price p_B, so the household will pay a volume charge equal to $p_B \times q_B$ (area Oq_BBp_B in Figure 1). As we can see, the new quantity consumed implies a volume charge supplement equal to $[(q_B - q_A) \times p_B] + [q_A \times (p_B - p_A)]$ (the shaded area), that is, the volume charge corresponding to additional cubic meters consumed (q_B−q_A), plus the supplement for the first q_A cubic meters consumed (area p_Ap_BCA).

HD_{it} ≡ *Hot days*. This is a variable which indicates the percentage of days with a maximum temperature above 18°C in each meter reading period. The temperature used to define this variable has been selected according to the Koeppen climatic classification system.

W_{it} ≡ *Wealth index*. This variable measures the relationship between consumption and users' lifestyles. Following the approach used in many empirical studies (Dandy *et al.* 1997; Arbués *et al.* 2004), wealth is defined in terms of the fiscal value (in real terms) of the dwelling as recorded in the Urban Property Register.

CHW_{it} ≡ *Collective hot water service*. This variable reflects the existence of individual water consumption registered on a collective water meter, distinct from the private meter. This collective consumption is paid to water utility by the residents' association, which subsequently distributes this water bill among its members. This is a dummy variable which takes a value of one if a collective hot water service is provided and zero otherwise.

Table 1 Zaragoza official water tariff, volume charges (1996, 1997 and 1998)

Daily consumption (m^3)	€/m^3 1996	€/m^3 1997	€/m^3 1998	Daily consumption (m^3)	€/m^3 1996	€/m^3 1997	€/m^3 1998	Daily consumption (m^3)	€/m^3 1996	€/m^3 1997	€/m^3 1998	Daily consumption (m^3)	€/m^3 1996	€/m^3 1997	€/m^3 1998
0.030	0.174	0.168	0.172	0.464	0.349	0.336	0.342	0.656	0.523	0.505	0.516	1.120	0.697	0.673	0.684
0.205	0.180	0.174	0.180	0.468	0.355	0.342	0.348	0.665	0.529	0.511	0.521	1.148	0.703	0.679	0.689
0.210	0.186	0.180	0.185	0.473	0.361	0.348	0.353	0.675	0.535	0.516	0.527	1.170	0.709	0.685	0.695
0.216	0.192	0.186	0.191	0.478	0.367	0.354	0.359	0.685	0.541	0.522	0.533	1.180	0.715	0.690	0.701
0.222	0.198	0.191	0.197	0.483	0.373	0.360	0.365	0.695	0.547	0.528	0.539	1.191	0.721	0.696	0.713
0.228	0.204	0.197	0.203	0.488	0.379	0.365	0.371	0.706	0.553	0.534	0.545	1.201	0.727	0.702	0.718
0.234	0.210	0.203	0.209	0.494	0.385	0.371	0.377	0.717	0.559	0.540	0.550	1.212	0.733	0.708	0.724
0.241	0.216	0.209	0.214	0.499	0.391	0.377	0.382	0.729	0.565	0.545	0.556	1.223	0.739	0.714	0.730
0.248	0.222	0.215	0.220	0.504	0.397	0.383	0.388	0.740	0.571	0.551	0.562	1.234	0.745	0.719	0.736
0.256	0.228	0.220	0.226	0.510	0.403	0.389	0.394	0.752	0.577	0.557	0.568	1.245	0.751	0.725	0.742
0.265	0.234	0.226	0.232	0.516	0.409	0.394	0.400	0.765	0.583	0.563	0.574	1.257	0.757	0.731	0.747
0.273	0.240	0.232	0.238	0.522	0.415	0.400	0.406	0.778	0.589	0.569	0.579	1.269	0.763	0.737	0.753
0.283	0.246	0.238	0.243	0.528	0.421	0.406	0.411	0.791	0.595	0.574	0.585	1.281	0.769	0.743	0.759
0.293	0.252	0.244	0.249	0.534	0.427	0.412	0.417	0.805	0.601	0.580	0.591	1.293	0.775	0.748	0.765
0.304	0.258	0.249	0.255	0.540	0.433	0.418	0.429	0.820	0.607	0.586	0.597	1.305	0.781	0.754	0.771
0.315	0.264	0.255	0.261	0.546	0.439	0.423	0.435	0.835	0.613	0.592	0.603	1.318	0.787	0.760	0.776
0.328	0.270	0.261	0.267	0.553	0.445	0.429	0.440	0.850	0.619	0.598	0.608	1.331	0.793	0.766	0.782
0.342	0.276	0.267	0.272	0.560	0.451	0.435	0.446	0.866	0.625	0.603	0.614	1.344	0.799	0.772	0.788
0.357	0.282	0.273	0.278	0.567	0.457	0.441	0.452	0.883	0.631	0.609	0.620	1.357	0.805	0.777	0.794
0.373	0.288	0.278	0.284	0.574	0.463	0.447	0.458	0.900	0.637	0.615	0.626	1.371	0.811	0.783	0.800
0.390	0.294	0.284	0.290	0.581	0.469	0.452	0.464	0.918	0.643	0.621	0.632	1.385	0.817	0.789	0.805
0.410	0.301	0.290	0.295	0.588	0.475	0.458	0.469	0.937	0.649	0.627	0.637	1.399	0.823	0.795	0.811
0.432	0.307	0.296	0.301	0.596	0.481	0.464	0.475	0.956	0.655	0.632	0.643	1.414	0.829	0.801	0.817
0.437	0.313	0.302	0.307	0.604	0.487	0.470	0.481	0.977	0.661	0.638	0.649	1.429	0.835	0.806	0.823
0.441	0.319	0.307	0.313	0.612	0.493	0.476	0.487	0.998	0.667	0.644	0.655	1.444	0.841	0.812	0.829
0.446	0.325	0.313	0.319	0.620	0.499	0.482	0.492	1.020	0.673	0.650	0.660	1.459	0.847	0.818	0.834
0.450	0.331	0.319	0.324	0.629	0.505	0.487	0.498	1.043	0.679	0.656	0.666	1.475	0.853	0.824	0.840
0.454	0.337	0.325	0.330	0.638	0.511	0.493	0.504	1.067	0.685	0.661	0.672	1.492	0.859	0.830	0.846
0.459	0.343	0.331	0.336	0.646	0.517	0.499	0.510	1.093	0.691	0.667	0.678	1.508	0.865	0.835	0.852

Table 1 (*Continued*)

Daily consumption (m^3)	€/m^3 1996	€/m^3 1997	€/m^3 1998
1.525	0.871	0.841	0.857
1.542	0.877	0.847	0.863
1.560	0.883	0.853	0.869
1.578	0.889	0.859	0.875
1.597	0.896	0.864	0.881
1.616	0.902	0.870	0.886
1.635	0.908	0.876	0.892
1.655	0.914	0.882	0.898
1.676	0.920	0.888	0.904
1.697	0.926	0.893	0.910
1.718	0.932	0.899	0.915
1.740	0.938	0.905	0.921
1.763	0.944	0.911	0.927
1.786	0.950	0.917	0.933
1.810	0.956	0.922	0.939
1.834	0.962	0.928	0.944
1.859	0.968	0.934	0.950
1.885	0.974	0.940	0.956
1.912	0.980	0.946	0.962
1.939	0.986	0.951	0.968
1.967	0.992	0.957	0.973
1.996	0.998	0.963	0.979

Daily consumption (m^3)	€/m^3 1996	€/m^3 1997	€/m^3 1998
2.026	1.004	0.969	0.991
2.057	1.010	0.975	0.997
2.088	1.016	0.980	1.002
2.121	1.022	0.986	1.008
2.154	1.028	0.992	1.014
2.189	1.034	0.998	1.020
2.225	1.040	1.004	1.025
2.262	1.046	1.009	1.031
2.301	1.052	1.015	1.037
2.340	1.058	1.021	1.043
2.381	1.064	1.027	1.049
2.424	1.070	1.033	1.054
2.468	1.076	1.038	1.060
2.514	1.082	1.044	1.066
2.561	1.088	1.050	1.072
2.610	1.094	1.056	1.078
2.661	1.100	1.062	1.083
2.715	1.106	1.067	1.089
2.770	1.112	1.073	1.095
2.828	1.118	1.079	1.101
2.888	1.124	1.085	1.107
2.951	1.130	1.091	1.112

Daily consumption (m^3)	€/m^3 1996	€/m^3 1997	€/m^3 1998
3.016	1.136	1.096	1.118
3.085	1.142	1.102	1.124
3.157	1.148	1.108	1.130
3.232	1.154	1.114	1.136
3.311	1.160	1.120	1.141
3.393	1.166	1.125	1.147
3.480	1.172	1.131	1.153
3.572	1.178	1.137	1.159
3.668	1.184	1.143	1.165
3.770	1.190	1.149	1.170
3.878	1.196	1.154	1.176
3.992	1.202	1.160	1.182
4.113	1.208	1.166	1.188
4.242	1.214	1.172	1.194
4.378	1.220	1.178	1.199
4.524	1.226	1.183	1.205
4.680	1.232	1.189	1.211
4.848	1.238	1.195	1.217
5.027	1.244	1.201	1.222
5.221	1.250	1.207	1.228
5.429	1.256	1.212	1.234
5.656	1.262	1.218	1.240

Daily consumption (m^3)	€/m^3 1996	€/m^3 1997	€/m^3 1998
5.901	1.268	1.224	1.246
6.170	1.274	1.230	1.251
6.463	1.280	1.236	1.257
6.787	1.286	1.241	1.263
7.144	1.292	1.247	1.275
7.541	1.298	1.253	1.280
7.984	1.304	1.259	1.286
8.483	1.310	1.265	1.292
9.049	1.316	1.270	1.298
9.695	1.322	1.276	1.304
10.441	1.328	1.282	1.309
11.311	1.334	1.288	1.315
12.339	1.340	1.294	1.321
13.573	1.346	1.299	1.327
15.081	1.352	1.305	1.333
16.967	1.358	1.311	1.338
19.390	1.364	1.317	1.344
22.622	1.370	1.323	1.350
27.147	1.376	1.328	1.356
33.933	1.382	1.334	1.362
45.244	1.388	1.340	1.367
67.867	1.394	1.346	1.373
135.733	1.400	1.352	1.379

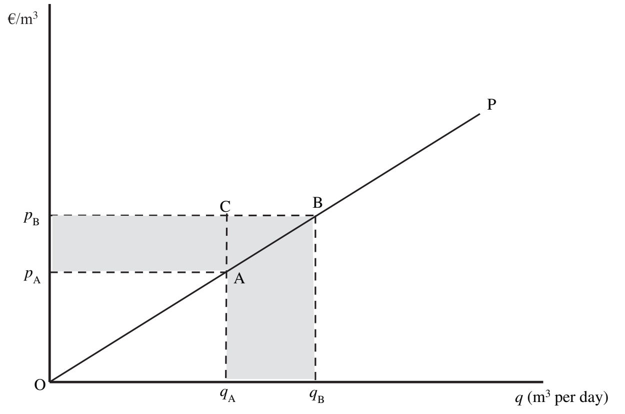

Figure 1 Zaragoza, linear increasing tariff structure.

$AG20_{it}$ ≡ *Young household members*. This is a variable control for possible differences in water consumption due to age. This is a dummy variable which takes a value of one if there are in the household members 20 years old or less and zero otherwise.

$AG60_{it}$ ≡ *Elder household members*. This is a variable control for possible differences in water consumption due to age. This is a dummy variable which takes a value of one if there are household members 60 years old and over and zero otherwise.

DN_{it} ≡ *Household size*. Five household sizes are considered with the aim of analysing if sensitivity to changes in price is different across households with different sizes. Five dummy variables are defined to capture size-specific effects:

$$D1_{it} = \begin{cases} 1 & \text{if household } i \text{ has one member} \\ 0 & \text{otherwise} \end{cases}$$

$$D2_{it} = \begin{cases} 1 & \text{if household } i \text{ has two members} \\ 0 & \text{otherwise} \end{cases}$$

$$D3_{it} = \begin{cases} 1 & \text{if household } i \text{ has three members} \\ 0 & \text{otherwise} \end{cases}$$

$$D4_{it} = \begin{cases} 1 & \text{if household } i \text{ has four members} \\ 0 & \text{otherwise} \end{cases}$$

$$D5_{it} = \begin{cases} 1 & \text{if household } i \text{ has five or more members} \\ 0 & \text{otherwise} \end{cases}$$

3. Description of the sample

As Blundell *et al.* (1992) show, individual level data have numerous advantages in the analysis of consumer demand decisions because they better reflect the heterogeneity of preferences, avoid aggregation biases and produce robust estimations of parameters in the demand function. Given these advantages, this study uses a panel of individual data from a random sample of 1507 households connected to the Zaragoza public water network. The sample contains the following information: metered consumption of dwellings (10 time observations covering the period 1996–1998); the fiscal value of dwellings; the number of individuals living in each dwelling according to the Municipal Census; and the availability of a common hot water facility in condominiums. This information was provided by the City Council.

The data set was further augmented by the unit price paid in real terms for the water consumed, and weather data. The water price was calculated by applying the official tariffs established in the municipal regulations. Weather data (daily maximum temperature over the period 1996–1998) were obtained from the Spanish Meteorological Institute.

Summary statistics for the variables included in the demand model (1) are presented in Table 2.

4. Household size and water consumption: an intuitive approach

Before proceeding with the econometric estimation of water demand, it may be useful to describe the relationship between water consumption and the number of residents in a household. Table 3 focuses on this issue from two perspectives: of total and per capita water consumption.

From a total water consumption perspective, Table 3 shows a direct relationship between total water consumption and household size, ranging from 0.1845 m^3 per day ($n = 1$) to 0.4831 m^3 per day ($n \geq 5$). When dealing with

Table 2 Summary statistics

	Mean	SD	Minimum	Maximum
All consumption (m^3) ($n = 1315$)	0.3351	0.1425	0.0300	0.9647
Size 1 consumption (m^3) ($n = 105$)	0.1845	0.0752	0.0300	0.4302
Size 2 consumption (m^3) ($n = 398$)	0.2640	0.1249	0.0306	0.9473
Size 3 consumption (m^3) ($n = 313$)	0.3326	0.1001	0.0626	0.7386
Size 4 consumption (m^3) ($n = 348$)	0.3998	0.1127	0.0937	0.8842
Size ≥ 5 consumption (m^3) ($n = 151$)	0.4831	0.1452	0.1428	0.9647
de_{it-2} (Euros)	0.1820	0.0813	0.0309	0.6888
W_{it} (Euros)	27 874.5542	16 421.7870	5058.8398	279 381.2761
HD_{it} (%)	67.9651	32.7224	2.0618	100.0000

Table 3 Size and water consumption

Size of household	Total consumption (m^3/day)	Per capita consumption (m^3/day)	Variation in per capita consumption (m^3/day)
1	0.1845	0.1845	–
2	0.2640	0.1320	−0.0525
3	0.3326	0.1109	−0.0211
4	0.3998	0.0999	−0.0110
≥ 5†	0.4831	0.0909	−0.0090

†The average size of households with a size ≥ 5 is 5.31 people. This size was obtained by means of the expression: $\Sigma_{n=5}^{8}(n \times h_n)/\Sigma_{n=5}^{8}h_n$; where n = household size, and h_n = number of households of size n.

per capita water consumption, the relationship observed in Table 3 between household size and per capita water consumption have an inverse relationship, pointing to economies of scale in water use, associated with uses connected more to a set of indivisible basic forms of consumption allocated to common household uses (i.e. domestic cleaning) than to the number of household residents.

5. Estimation of the model

The demand equation (1) is specified in semi-logarithmic form:

$$\begin{aligned}\ln q_{it} = \beta_0 + \delta_1 de_{it-2} + \delta_2(D1_{it} \times de_{it-2}) + \delta_3(D2_{it} \times de_{it-2}) \\ + \delta_4(D3_{it} \times de_{it-2}) + \delta_5(D4_{it} \times de_{it-2}) + \delta_6 HD_{it} + \beta_1 W_{it} \\ + \beta_2 CHW_{it} + \beta_3 AG20_{it} + \beta_4 AG60_{it} + u_{it}\end{aligned} \quad (2)$$

having introduced the household size dummies in multiplicative form with the price variable. This specification allows us to capture differences in the reaction to changes in price across households with different sizes.

Following Baltagi (2005), the error term is $u_{it} = \mu_i + v_{it}$ with $\mu_i \sim IID(0, \sigma_\mu^2)$, and $v_{it} \sim IID(0, \sigma_v^2)$, where μ_i is time-invariant and it accounts for any individual specific effect that is not included in the regression and v_{it} varies with individuals and time and can be thought of as the usual disturbance in the regression.

Given the price structure in Zaragoza, where the price is linked to the amount consumed, it may be observed that:

$$de_{it-2} = f(q_{it-2}) \quad (3)$$

in such a way that model (2) is indirectly dynamic through (3) (Blundell *et al.* 1992).

The relationship given in (3) implies that the variable de_{it-2} is correlated with the error term u_{it} by means of μ_i. Within this framework, the Ordinary Least Squares estimation of the parameters is biased and inconsistent, and even v_{it} is not autocorrelated.

Because of this, the estimation of model (2) requires a two-step procedure. In the first step, the first difference transformation of model (2) is estimated by Generalized Least Squares. Thus, a first estimation of the parameters of price, wealth and climate variables is obtained, which will be used to calculate the residuals ($\Delta\hat{v}_{it}$). The second step consists of employing Generalized Least Squares once again to estimate the first difference transformation of model (2), but now using the residuals obtained in the previous step. Although this procedure is similar to that proposed by Arellano (1989) and Arellano and Bond (1991) several differences exist between these two estimation techniques. Whereas in the original procedure there is a problem of correlation between the error term and the explanatory variable (the lagged endogenous variable) in the first differenced model, which makes the use of instrumental variables necessary, in our procedure this problem does not occur: because the price variable is two period lagged, the explanatory variable and the error term are not related in the first differenced model. There is, like Arellano and Bond (1991) autocorrelation between the different components of the error term. Specifically:

$$E\left(\Delta v_i \Delta v_i'\right) = \sigma_v^2 (I_N \otimes GS)$$

where $\Delta v_i = \begin{pmatrix} \Delta v_{i4} & \Delta v_{i5} & \cdots & \Delta v_{i10} \end{pmatrix}'$, being each Δv_{ij} the error term of first differenced model:

$$\begin{aligned}\Delta \ln q_{it} = {} & \delta_1 \Delta de_{it-2} + \delta_2 \Delta(D1_{it} \times de_{it-2}) + \delta_3 \Delta(D2_{it} \times de_{it-2}) \\ & + \delta_4 \Delta(D3_{it} \times de_{it-2}) + \delta_5 \Delta(D4_{it} \times de_{it-2}) + \delta_6 \Delta HD_{it} + \Delta v_{it}\end{aligned} \tag{4}$$

and GS is a $(T-3) \times (T-3)$ matrix with a form:

$$\begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 \end{pmatrix}$$

Expression (4) will be correct provided that problems of autocorrelation do not appear in the initial model. However in our model, it seems logical that water consumption in a period of time depends on consumption during a preceding period. This fact could lead to other autocorrelation problems in the initial model (2). Given that the data correspond to water meter reading (four times a year), a relationship between the error term in the period t and the error term in the period $t + 4$ (the same water meter reading a year before) is very probable.

Specifically, there could be fourth order error serial correlation. Nevertheless, it is necessary to carry out a test in order to corroborate if such a problem exists, and then to estimate the model appropriately.

To test the existence of serial correlation, once the model (5) is estimated we carry out the following test:

$$H_0 : E(\Delta v_{it}\Delta v_{it+s}) = 0 \text{ jointly for } s = 2, 3, \cdots, p(\leq T-4)$$
$$H_A : E(\Delta v_{it}\Delta v_{it+s}) \neq 0 \text{ for some } s$$

As we have previously indicated, we focus on $p = 4$ (although if other order serial correlation ($p < 4$) exists, it will be collected). In order to carry out this hypothesis testing there are some statistics in the econometric literature. Arellano and Bond (1991) propose the so-called m_2 test and the Sargan's difference test. Nevertheless, Yamagata (2008) shows some faults in these tests (the first one because of its lower power against some alternative hypotheses, and the second one because it cannot be used for some kinds of correlation) and he proposes a new test denoted as $m^2_{2,p}$, which can be expressed as:

$$m^2_{2,p} = i'_N\hat{H}\left(\hat{G}'\hat{G}\right)^{-1}\hat{H}'i_N$$

where i_N is a $N \times 1$ vector of ones, $\hat{H} = \begin{pmatrix}\hat{\eta}_1 & \hat{\eta}_2 & \cdots & \hat{\eta}_N\end{pmatrix}'$ being $\hat{\eta}_i = \begin{pmatrix}\hat{\eta}_{i2} & \hat{\eta}_{i3} & \cdots & \hat{\eta}_{ip}\end{pmatrix}'$, $\hat{\eta}_{is} = \Sigma_{t=4}^{T-s}\Delta\hat{v}_{it}\Delta\hat{v}_{it+s}$, with

$$\begin{aligned}\Delta\hat{v}_{it} = \Delta\ln q_{it} - \hat{\delta}_1\Delta de_{it-2} - \hat{\delta}_2\Delta(D1_{it} \times de_{it-2}) - \hat{\delta}_3\Delta(D2_{it} \times de_{it-2}) \\ - \hat{\delta}_4\Delta(D3_{it} \times de_{it-2}) - \hat{\delta}_5\Delta(D4_{it} \times de_{it-2}) - \hat{\delta}_6\Delta HD_{it}\end{aligned}$$

and $\hat{G} = (\hat{g}_1, \hat{g}_2, \cdots, \hat{g}_N)'$, being $\hat{g}_i = (g_{i2}, \cdots, g_{ip})'$, with $\hat{g}_{is} = \hat{\eta}_{is} - \hat{\omega}'_{Ns}\left(\Delta X'\hat{\Omega}^{-1}\Delta X\right)^{-1}\Delta X'\hat{\Omega}^{-1}\Delta\hat{\hat{v}}_i$, $\hat{\Omega} = \Sigma_{i=1}^{N}\Delta\hat{\hat{v}}_i\Delta\hat{\hat{v}}'_i$ and being $\Delta\hat{\hat{v}}_i$ the $(T-3) \times 1$ residual vector corresponding to the first stage of the estimation procedure described above. As $N \rightarrow \infty$ and T fixed:

$$m^2_{2,p} \xrightarrow{d} \chi^2_{p^{-1}}$$

When this statistic is calculated for model (4), we obtain $m^2_{2,p} = 122.85$ (higher than $\chi^2_3(0.05)$), so at 5 per cent of statistical significance, H_0 is rejected, and there is evidence about serial correlation in model (4).

With the aim of correcting the problem of serial correlation, the fourth lag of the endogenous variable (q_{it-4}) is added to (4) as a regressor. Thus, model (4) becomes:

$$\begin{aligned}\ln q_{it} = {} & \beta_0 + \delta_1 de_{it-2} + \delta_2(D1_{it} \times de_{it-2}) + \delta_3(D2_{it} \times de_{it-2}) \\ & + \delta_4(D3_{it} \times de_{it-2}) + \delta_5(D4_{it} \times de_{it-2}) + \delta_6 HD_{it} \\ & + \delta_7\ln q_{it-4} + \beta_1 W_{it} + \beta_2 CHW_{it} + \beta_3 AG20_{it} + \beta_4 AG60_{it} + u_{it}\end{aligned} \quad (5)$$

and the first differenced model can be written as:

$$\begin{aligned}\Delta \ln q_{it} &= \delta_1 \Delta de_{it-2} + \delta_2 \Delta(D1_{it} \times de_{it-2}) + \delta_3 \Delta(D2_{it} \times de_{it-2}) \\ &\quad + \delta_4 \Delta(D3_{it} \times de_{it-2}) + \delta_5 \Delta(D4_{it} \times de_{it-2}) \\ &\quad + \delta_6 \Delta HD_{it} + \delta_7 \Delta q_{it-4} + \Delta v_{it}\end{aligned} \tag{6}$$

We now estimate (6) following the same two-stage procedure described at the beginning of this section, and we again test serial correlation. The value of $m^2_{2,p}$ is 2.91, so we conclude that there is not serial correlation.

Finally, a method similar to that described in Hsiao (1986) is used to obtain the parameter of W_{it}, CHW_{it}, $AG20_{it}$ and $AG60_{it}$, four variables that do not appear in the first difference transformation of model (2). This method substitutes δ_j for $\hat{\delta}_j$ in model (2) and estimates the β_1, β_2, β_3 and β_4 parameters as a static panel by means of the random effects technique.

6. Results

Table 4 reports the results obtained from the empirical estimation of demand equation (6). The signs of most coefficients in Table 4 agree with our *a priori* expectations.

As equation (1) is semi-logarithmic, estimated coefficients can be easily interpreted as semielasticities. This measure gives the percentage change in q_{it} in response to a one unit change of the explanatory variable. As Hatznikolau

Table 4 Estimation results

	Estimated value	*t*-ratios
Constant	−0.7026	−31.0658
$de_{it-2}(\hat{\delta}_1)$	−0.2645	−45.4835
$de_{it-2} \times D1_{it}$ $(\hat{\delta}_2)$	−1.0525	−6.6062
$de_{it-2} \times D2_{it}$ $(\hat{\delta}_3)$	−0.9509	−29.4903
$de_{it-2} \times D3_{it}$ $(\hat{\delta}_4)$	−0.1983	−5.9545
$de_{it-2} \times D4_{it}$ $(\hat{\delta}_5)$	−0.0078	−1.2646
HD_{it} $(\hat{\delta}_6)$	−0.0409	−30.2796
q_{it-4} $(\hat{\delta}_7)$	0.3228	265.5147
W_{it} $(\hat{\beta}_1)$	0.2941×10^{-5}	5.4393
CHW_{it} $(\hat{\beta}_1)$	−0.1087	−6.0003
$AG20_{it}$ $(\hat{\beta}_3)$	0.0684	3.0547
$AG60_{it}$ $(\hat{\beta}_4)$	−0.0692	−3.4079

and Ferentinos (2004) show, semielasticities are more appropriate when only small changes in explanatory variables are possible.

Given that household size dummies interacting with the price variable have been used, the estimated coefficients $\hat{\delta}_1$ to $\hat{\delta}_5$ must be interpreted in a special way. Only the estimated coefficient $\hat{\delta}_1$ represents the price semielasticity for households with five or more members (control size). The estimated coefficients $\hat{\delta}_2$ to $\hat{\delta}_5$ measure the difference in the slope for sizes 1 to 4 versus control size (5 or more). So, the corresponding price semielasticities will be $(\delta_1 + \delta_2)$ for households with one member and $(\delta_1 + \delta_3)$, $(\delta_1 + \delta_4)$ and $(\delta_1 + \delta_5)$ for households with two, three and four members respectively.

Furthermore, the coefficients obtained from (6) for the dummy variables (CHW_{it}, $AG20_{it}$ and $AG60_{it}$) cannot be directly interpreted as semielasticities either. This issue can be resolved following the transformation suggested by Kennedy (1981):

$$\hat{p} = 100 \times \left(\exp(\hat{\psi} - \frac{1}{2}\hat{V}(\hat{\psi})) - 1 \right) \tag{7}$$

where $\hat{\psi} = \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4$.

The semielasticities obtained are presented in Table 5. Each reported price semielasticity estimate represents the percentage change in water consumption when the water price increases by one cent of a euro. The wealth semielasticity represents the percentage change in water consumption given a 1000€ change in wealth variable.

As Table 5 shows, the price semielasticity is less than 0 for all sizes of household. An interesting result is that price semielasticity decreases as household size increases, ranging from −1.317 (one member) to −0.2645 (five or more members). Moreover, two significant decreases in price semielasticities are found when household size changes from two to three members and when household size changes from three to four members, pointing to different patterns of water consumption. To test these differences, we can carry out several hypothesis tests by using the *t*-ratios of δ_2 to δ_5 included in Table 4.

Table 5 Semielasticities of variables

Price (n = 1)	−1.3170
Price (n = 2)	−1.2154
Price (n = 3)	−0.4628
Price (n = 4)	−0.2723
Price ($n \geq 5$)	−0.2645
$HD_{it}t$	−4.09
q_{it-4}	32.28
W_{it}	0.2941×10^{-3}
CHW_{it}	−11.7615
$AG20_{it}$	4.4282
$AG60_{it}$	−8.5914

Moreover, the t-ratio of δ_1 shows the statistical significance of price semielasticity estimates for the control size (five or more members). Therefore, the hypothesis $H_0 : \delta_i = 0$ (with $i = 2, 3, 4, 5$) tests if there are statistically significant differences in price semielasticities between the control household size and every other household sizes considered. As we can see in Table 4, only the t-ratio corresponding to δ_5 is not significant at 5 per cent, which means that price semielasticity of households with four members is not statistically different from that of household with five or more members.

To test other differences in price semielasticities between households with different sizes, we estimate model (6) again, changing the control size (see Appendix I). By carrying out these tests, we found that there are not significant differences between the semielasticities of households with one and two members, and between the semielasticities of households with four and five or more members. Therefore, the results obtained indicate three patterns of water consumption, one for small households (one and two members), another for medium households (three members) and another for large ones (four, and five or more members). This means that small households are better able to adjust to changes in the price of water than large households.

There are two issues which explain these different water consumption patterns:

1. The existence of endogenous transaction costs related with the introduction and spread of new practices to improve the efficiency in the use of water appliances like taps, tanks, bathtubs and showers, and white goods like washing machines and dishwashers. These transaction costs arise due to lack of information, misplaced incentives, decisions influenced by custom and habits, heterogeneity of household members, and difficulties associated with the constant monitoring and verification of water consumption in the household, among other factors.
   Household size will be positively related to endogenous transaction costs. It is clear that the organization and supervision of household activities is more complex in larger households. Thus endogenous transaction costs represent a constraint to the decision-making process and place significant barriers in the way of water saving strategies, reducing incentives for large households.
2. It seems likely that household size will affect the capacity to improve efficiency in water consumption by changing the way people use their white goods like washing machines and dishwashers. Usually, white goods utilization is less efficient in small households than larger ones due to the fact that they are not able to fully exploit economies of scale related to their use. Therefore small households will be better able to obtain efficiency improvements in water consumption in response to exogenous incentives. For example, in a household with one or two people, it is common to run washing machines and dishwashers below full capacity. Therefore it is

possible to reduce water consumption by making better use of the machine's capacity, putting more clothes or dishes in each wash and running it less often. When household size increases, however, these machines run more loads, because the volume of dirty clothes and dishes increases (especially, if children and/or teenagers live in the dwelling), making it more difficult to cut water consumption.

Based on the results shown in Table 5, wealth semielasticity is positive, and tends to be close to zero. Therefore, its relative participation in water consumption decisions is reduced. This result needs to be interpreted in terms of the short run perspective adopted, where a change in wealth does not imply an immediate change of the users' lifestyle.

The semielasticity of the climate variable included in Table 5 reveals a seasonal factor in water consumption, which means that water consumption decreases by 4.09 per cent when the hot days recorded during the consumption period increase by 1 per cent point. Although a positive relation could have been expected in line with other studies (Kenney *et al.* 2008; Schleich and Hillenbrand 2009), this result is consistent with two observed facts: in the first place, outdoor uses of water are not relevant in Zaragoza; and, in the second, in the summer, when it is very hot, many people go away on holiday, or go to sports clubs for the day, or to towns on the outskirts of Zaragoza (to spend the day, or to live during these hot months, returning to the city only to work).

The percentage effect of the dummy indicating the availability of a collective hot water service is (−11.7615). This means that the existence of such supply leads to a fall of 11.7615 per cent in the consumption registered on the individual meters, that is, about 12 per cent of the water used is obtained from the collective supply system installed in the building.

Finally, the results obtained for the age dummies indicate an inverse relation between the age of household members and water consumption. The presence in a household of elderly people tends to reduce the water consumption by 8.59 per cent, while the presence of young people tends to increase water consumption by 4.42 per cent. These results may be explained by the different consumption patterns existing between these age groups (elderly people have a greater tendency to use less water for washing and hygiene than young people). Another possible explanatory factor for this result is the fact that the average income of households with elderly members is about 55 per cent of the average income of households with young members.

7. Final considerations

In this paper, we have estimated a residential water demand function which includes multiplicative dummy variables that enables us to analyse if users' sensitivity to changes in price is different depending on the number of household members. This demand function has been estimated by adopting a dynamic panel data approach, using household level data. Therefore, this paper offers a new perspective on urban water demand analysis by estimating different water price semielasticities according to household sizes, a criterion to establish different categories of users not considered in the literature until now.

The empirical analysis suggests that all households, regardless of size, are sensitive to prices (all coefficients estimated are significant and different from zero). Semielasticities are lower than zero for households of all sizes.

A relevant result related to the price semielasticities obtained is that small households are more sensitive to price changes. This implies that pricing measures will have a greater impact on small households (one and two members) than medium (three members) and large ones (four or more members). This finding confirms the initial hypothesis that sensitivity to changes in the price will differ depending on household size.

Hence there is a dependence relation between the demographical framework and the sensitivity of residential water consumption to price changes. This issue is especially relevant when changes in household size are happening, as has been occurring in Zaragoza, where the average occupation density of homes changed from 3.04 people per home in 1991 to 2.23 people per home in 2008. In this context, water utility managers should be aware of demographic changes affecting household size and adjust their information about residential water demand. Likewise this dependence relation should be taken into account when comparing geographical areas with different household structures and when using the results obtained in a specific geographical area to analyse water tariff reforms in other ones.

Furthermore, if we estimate model (2) without the household size multiplicative dummies, we obtain an aggregate price semielasticity of −0.57. If managers consider only aggregate response, although they will come close to the response of households with three members (semielasticity of −0.46), they will overstate the response of households with four or more members (semielasticities ranging from −0.26 to −0.27) and they will especially understate that of small households (semielasticities below −1).

These results suggest that information at size-aggregated level about water demand commonly used to evaluate the effects of water tariff reforms can cause a misinterpretation of the residential users' attitude towards changes in water prices. Only if price changes by the same proportion for all users, regardless of water consumed, will the results obtained at size-aggregated level be right. Nevertheless, water tariff reforms usually mean price changes in different proportions for different water consumptions (as a result of the substitution of a flat-rate tariff by a block tariff, or the modification of the number of price blocks, their width or the level of prices associated with each block in a existing block tariff). Therefore, as a significant positive relation exists between household size and water consumption, it is clear that households will be faced with a price change that will depend on their size. Moreover, an accurate characterization of water demand that enables us to discover the response to changes in price of households with different sizes offers a better framework to analyse the effect of water tariff reforms on water consumption.

Finally, taking household size into account in the design of the urban water rate would be useful in order to improve the management effectiveness of the utility. Therefore, in a standard increasing block tariff design, increasing the applicable prices corresponding to blocks where the average consumption of small households is situated may be more effective in moderating water consumption. However, if the aim is to increase revenues without affecting basic consumption, then managers could increase the price levels in blocks where the average consumption of large households is situated.

Therefore, the results of this residential water demand function estimation which includes multiplicative dummy variables strongly suggest that household size is an important determinant of water use, so policymakers need to include it in the design of demand-side water management measures. These results should be an incentive to extend this research to geographical contexts with different socioeconomic and environmental features to those observed in Zaragoza with the aim of verifying whether there are relevant changes in the behavioural characteristics of residential water demand observed. Furthermore, another relevant comparison needing future research is the separate price response for households with a size larger than five. Also, future research could analyse, how qualitative characteristics of the household (age of members, education level, etc.) affect price responses for each size. Nevertheless, these research lines will be, in most cases, very limited, by the lack of information of the quality necessary to obtain significant results from both a statistical and economic perspective.

References

Agthe, D.E. and Billings, R.B. (1987). Equity, price elasticity, and household income under increasing block rates for water, *American Journal of Economics and Sociology* 46, 273–286.

Arbués, F. and Villanúa, I. (2006). Potential for pricing policies in water resource management: estimation of urban residential water demand in Zaragoza, Spain, *Urban Studies* 43, 2421–2442.

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

Arbués, F., Barberán, R. and Villanúa, I. (2004). Price impact on urban residential water demand: a dynamic panel data approach, *Water Resources Research* 40, W11402, doi: 10.129/2004WR003092.

Arellano, M. (1989). A note on the Anderson-Hsiao estimator for panel data, *Economics Letters* 31, 337–341.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, *Review of Economic Studies* 58, 277–297.

Baltagi, B.H. (2005). *Econometric Analysis of Panel Data*, 3rd edn. John Wiley and Sons, Hoboken.

Blundell, R. (1988). Consumer behaviour: theory and empirical evidence. A survey, *The Economic Journal* 98, 16–65.

Blundell, R., Bond, S., Devereux, M. and Schiantarelli, F. (1992). Investment and Tobin's *Q*: evidence from company panel data, *Journal of Econometrics* 51, 233–257.

Dale, L., Fujita, S., Guerrero, S., Hanemann, M., Lutzenhiser, L., Moezzi, M. and Vásquez, F. (2009). *Price Impact on the Demand for Water and Energy in California Residences*. Public Interest Energy Research (PIER) Program Reports CEC-500-2009-032-D, California Energy Commission, Sacramento, CA.

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

Dubin, J.A. and McFadden, D.L. (1984). An econometric analysis of residential electric appliance holdings and consumption, *Econometrica* 52, 345–362.

Espey, J.A. and Espey, M. (2004). Turning on the lights: a meta-analysis of residential electricity demand elasticities, *Journal of Agricultural and Applied Economics* 36, 65–81.

Frondel, M. and Messner, M. (2008). *Price Perception and Residential Water Demand: Evidence from a German Household Panel*. 16th Annual Conference of the European Association of Environmental and Resource Economists in Gothenburg, 25–28 June. Available from URL: http://www.webmeets.com/files/papers/EARE/2008/362/water.pdf [accessed 14 December 2009].

Gaudin, S. (2006). Effect of price information on residential water demand, *Applied Economics* 38, 383–393.

Hatznikolau, D. and Ferentinos, K. (2004). On the role of semielasticity in statistics, *The American Statistician* 58, 150–151.

Hoffmann, M., Worthington, A. and Higgs, H. (2006). Urban water demand with fixed volumetric charging in a large municipality: the case of Brisbane, Australia, *The Australian Journal of Agricultural and Resource Economics* 50, 347–359.

Hsiao, C. (1986). *Analysis of Panel Data*. Cambridge University Press, New York.

Kennedy, P.E. (1981). Estimation with correctly interpreted dummy variables in semilogarithmic equations, *American Economic Review* 71, 801.

Kenney, D.S., Goemans, C., Klein, R., Lowrey, J. and Reidy, K. (2008). Residential water demand management: lessons from Aurora, Colorado, *Journal of the American Water Resources Association* 44, 192–207.

Krause, K., Chermark, J.M. and Brookshire, D.S. (2003). The demand for water: consumer response to scarcity, *Journal of Regulatory Economics* 23, 167–191.

Labandeira, X., Labeaga, J.M. and Rodríguez, M. (2006). A residential energy demand system for Spain, *Energy Journal* 27, 87–112.

Lyman, A. (1992). Peak and off-peak residential water demand, *Water Resources Research* 28, 2159–2167.

Madlener, R. (1996). Econometric analysis of residential energy demand: a survey, *Journal of Energy Literature* 2, 3–32.

Musolesi, A. and Nosvelli, M. (2007). Dynamics of residential water consumption in a panel of Italian municipalities, *Applied Economics Letters* 14, 441–444.

Narayan, P.K. and Smyth, R. (2005). The residential demand for electricity in Australia: an application of the bounds testing approach to cointegration, *Energy Policy* 33, 467–474.

Navajas, F.H. (2009). Engel curves, household characteristics and low-user tariff schemes in natural gas, *Energy Economics* 31, 162–168.

Nesbakken, R. (1999). Price sensitivity of residential energy consumption in Norway, *Energy Economics* 21, 493–515.

Nieswiadomy, M.L. (1992). Estimating urban residential water demand: effects of price structure, conservation, and education, *Water Resources Research* 28, 609–615.

Reiss, P.C. and White, M.C. (2005). Household electricity demand, revisited, *Review of Economic Studies* 72, 853–883.

Renwick, M.E. and Archibald, S.O. (1998). Demand side management policies for residential water use: who bears the conservation burden? *Land economics* 74, 343–359.
Schleich, J. and Hillenbrand, T. (2009). Determinants of residential water demand in Germany, *Ecological Economics* 68, 1756–1769.
Worthington, A.C. and Hoffmann, M. (2008). An empirical survey of residential water demand modelling, *Journal of Economic Surveys* 28, 842–871.
Yamagata, T. (2008). A joint serial correlation test for linear panel data models, *Journal of Econometrics* 146, 135–145.
Yoo, S.H., Lee, J.S. and Kwak, S.J. (2007). Estimation of residential electricity demand function in Seoul by correction for sample selection bias, *Energy Policy* 35, 5702–5707.

Appendix I

As we can see in Table 6, in Model AI.1 and Model AI.2, only δ_3 is not significant, which means that price semielasticity of households with one member is not statistically different from that of households with two members. Similarly, δ_5 in Model AI.4 is not significant, so price semielasticity of households with four members is not statistically different from that of households with five or more members. The results of model AI.3 show that there are significant differences between the price semielasticities of households with three members and households with any other number of members.

Model AI.1 (control size 1)

$$\begin{aligned}\Delta \ln q_{it} = {} & \delta_1 \Delta de_{it-2} + \delta_2 \Delta(D2_{it} \times de_{it-2}) + \delta_3 \Delta(D3_{it} \times de_{it-2}) \\ & + \delta_4 \Delta(D4_{it} \times de_{it-2}) + \delta_5 \Delta(D5_{it} \times de_{it-2}) \\ & + \delta_6 \Delta HD_{it} + \delta_7 \Delta q_{it-4} + \Delta v_{it}\end{aligned}$$

Table 6 Estimation results of model (6) changing the control size

	Model AI.1			Model AI.2	
	Estimated value	*t*-ratios		Estimated value	*t*-ratios
$\hat{\delta}_1$	−1.3170	−8.2668	$\hat{\delta}_1$	−1.2154	−38.0121
$\hat{\delta}_2$	0.1015	0.6386*	$\hat{\delta}_2$	−0.1015	0.6386*
$\hat{\delta}_3$	0.8548	5.3941	$\hat{\delta}_3$	0.7526	20.0834
$\hat{\delta}_4$	1.0446	6.5529	$\hat{\delta}_4$	0.94309	29.2129
$\hat{\delta}_5$	1.05254	6.6062	$\hat{\delta}_5$	0.95097	29.4903
	Model AI.3			**Model AI.4**	
	Estimated value	*t*-ratios		Estimated value	*t*-ratios
$\hat{\delta}_1$	−0.4628	−13.9742	$\hat{\delta}_1$	−0.2723	−126.1114
$\hat{\delta}_2$	−0.8541	−5.3940	$\hat{\delta}_2$	−1.044	−6.5529
$\hat{\delta}_3$	−0.7526	−20.0834	$\hat{\delta}_3$	−0.9431	−29.2129
$\hat{\delta}_4$	0.1904	5.7007	$\hat{\delta}_4$	−0.1904	−5.7007
$\hat{\delta}_5$	0.1983	5.9545	$\hat{\delta}_5$	0.0078	1.2646*

*Not significant at 5%.

Model AI.2 (control size 2)

$$\begin{aligned}\Delta \ln q_{it} = {} & \delta_1 \Delta de_{it-2} + \delta_2 \Delta(D1_{it} \times de_{it-2}) + \delta_3 \Delta(D3_{it} \times de_{it-2}) \\ & + \delta_4 \Delta(D4_{it} \times de_{it-2}) + \delta_5 \Delta(D5_{it} \times de_{it-2}) + \delta_6 \Delta HD_{it} \\ & + \delta_7 \Delta q_{it-4} + \Delta v_{it}\end{aligned}$$

Model AI.3 (control size 3)

$$\begin{aligned}\Delta \ln q_{it} = {} & \delta_1 \Delta de_{it-2} + \delta_2 \Delta(D1_{it} \times de_{it-2}) + \delta_3 \Delta(D2_{it} \times de_{it-2}) \\ & + \delta_4 \Delta(D4_{it} \times de_{it-2}) + \delta_5 \Delta(D5_{it} \times de_{it-2}) + \delta_6 \Delta HD_{it} \\ & + \delta_7 \Delta q_{it-4} + \Delta v_{it}\end{aligned}$$

Model AI.4 (control size 4)

$$\begin{aligned}\Delta \ln q_{it} = {} & \delta_1 \Delta de_{it-2} + \delta_2 \Delta(D1_{it} \times de_{it-2}) + \delta_3 \Delta(D2_{it} \times de_{it-2}) \\ & + \delta_4 \Delta(D3_{it} \times de_{it-2}) + \delta_5 \Delta(D5_{it} \times de_{it-2}) + \delta_6 \Delta HD_{it} \\ & + \delta_7 \Delta q_{it-4} + \Delta v_{it}\end{aligned}$$