Pricing in Public Water Supply

By
Márta Bisztray

Submitted to
Central European University
Department of Economics

In partial fulfillment of the requirements for the degree of Masters of Arts

Supervisor: Professor Gábor Kézdi

Budapest, Hungary
2010

Abstract

The sector of public water supply gains more and more attention these days. The presence of exhaustible aquifers and the declining state of infrastructure are frequently mentioned reasons why current pricing practices should be reconsidered. However, welfare effects of a price change can only be evaluated by anticipating demand side adjustments. In my thesis I focus on the estimation of price elasticity concerning public water supply for residential customers. Estimations are done based on a panel database of about 700 firms from 21 countries of Central- and South-Eastern Europe and the post-Soviet region, using the method of first differences. Water demand is found to be inelastic with an elasticity value of -0.099. I also provide estimates regarding elasticity of variable costs with respect to output. Using a Cobb-Douglas specification increasing returns to scale is found. Finally I give a numerical example in order to illustrate how these estimates can be used for evaluating welfare effect of price changes. For this purpose I propose a dynamic model which also takes into consideration the effect of investments into infrastructural quality.

Acknowledgement

Hereby I would like to thank the valuable comments and support of my supervisor, Professor Gábor Kézdi. I am also grateful for Alexander Danilenko, who gave permission to use the database of IBNET for the purpose of my thesis.

Table of contents

List of tables

1. *Introduction*

Having an access to drinking water is vital for staying alive. In developed countries this is mainly ensured by public water supply. As it is an essential service provided to the vast majority of the population, pricing of supplied water is a question of high importance. The relevance of this issue is further increased by the fact, that the Water Framework Directive of the European Union[1] suggests the reconsideration of current pricing practices. The aim is securing sustainable and efficient water use by imposing prices which cover all costs including externalities.[2] However, potential effects of a change in pricing policy should be carefully analyzed in order to make good decisions. A key part of this analysis is the estimation of demand side responsiveness to price changes.

The main purpose of my thesis is to provide an estimation concerning price elasticity of residential water demand, focusing on Central- and South-Eastern Europe and the post-Soviet region, and to show an example for its application in welfare analysis. Though there exists a broad literature on this topic, the majority of the estimates I found were done for America (especially for the USA) or for Western Europe. Consequently I consider as an interesting question to estimate price elasticity of water demand in the region defined above, containing my home country, Hungary as well. A good overview of the existing literature is provided by Arbués et al. (2003) and Worthington and Hoffman (2006). These show the variety of ways this estimation problem was approached, using aggregate or household level data of different frequency and various estimation methods. I work with a panel database of more than 700

[1] Directive 2000/60/EC
[2] A good description of this topic can be found at the webpage of the European Comission: http://ec.europa.eu/environment/water/participation/pdf/waternotes/water_note5_economics.pdf (accessed: June 2, 2010)

companies, aggregated on the firm level. Using the method of first differences I find residential water demand in the region to be strongly inelastic with respect to price changes, which corresponds to the findings in the literature, where estimated value of elasticity typically lies between 0.25 and 0.75 or even below in absolute value (Worthington and Hoffman 2006).

Beyond estimation of price elasticity my aim is to provide an illustration how this estimate can be used in welfare analysis. For that purpose using again the method of first differences I estimate a Cobb-Douglas cost function in order to characterize the supply side. In the literature several examples of cost function estimates can be found concerning public water supply, with different aims. Some of them estimate returns to scale in water supply (Nauges and van den Berg 2007), others measure cost inefficiencies in the sector (Filippini et al. 2008) or evaluate pricing practice comparing it to estimated marginal costs (Renzetti 1999) and measure welfare effects of price changes (Garcia and Reynaud 2004). The last purpose corresponds to my aims. Accordingly, I provide estimates of cost elasticity with respect to output, which are somewhat lower than those found in the literature, but refer to increasing returns to scale.

As a next step I propose a modeling framework which makes possible to analyze the welfare effect of changes in price setting. Many papers provide such welfare analysis, like Garcia and Reynaud (2004) or García-Valiñas (2005) among others. However, most of them apply a static framework, which does not allow the investigation of over the time effects of decisions. Nevertheless, there are important aspects of water supply connected to pricing decisions, which can only be analyzed in a dynamic framework. These include the usage of aquifer resources or maintenance of infrastructural quality. The first problem is examined by Timmins (2002), who provides a dynamic model for that. Relying on this model, I use the proposed framework for the second question, quality of infrastructure. Then I provide a numerical example, how welfare effects of pricing and infrastructural investment decisions can be analyzed.

The structure of the thesis is as follows: first I give a short description of the water supply sector focusing on price regulation, which is necessary for the understanding of the estimations. In chapter 3 I provide theoretical funding for the following estimations. The dataset used is presented in the next chapter, while methodology and results of demand and cost estimations can be found in chapter 5. In chapter 6 I propose a model for analyzing the welfare effect of different pricing schemes using the estimated elasticity values, and I also present a numerical example. Finally I give a summary of the main results and make some concluding remarks.

2. Description of the sector

Before turning to demand estimation it is necessary to provide a brief description of the public water supply sector. A good overview of this topic can be found in Garcia and Thomas (2001). Water supply for residential, industrial or institutional customers is provided by water utilities, which in several cases are responsible for the collection and treatment of wastewater as well. To some extent these are separate activities, but there are strong complementarities between the two. However, in the followings I focus on water supply only. Typically the process of water service consists of the following steps: water is extracted from a water source, usually it is treated, then delivered to the consumers via a network of pipes, using pumping if necessary. Various kinds of water sources (surface or underground water) and diverse quality of raw water impose different amount of costs on a cubic meter of water via different pumping needs and necessary treatment processes.

Due to the huge investments needed for establishing the network water utilities can be regarded as local natural monopolies (see Nauges and van den Berg 2007). According to theory of industrial organizations, natural monopolies occur when the considerable economies of scale in the technology does not make it possible for more than one firm to produce the same good or provide the same service efficiently. In this case there can be many water supply providers in a country (for example there are about 400 in Hungary[3]), but they all have distinct service area. There is no sense for a potential competitor to build another network at the same place as its enormous sunk costs could never be refunded. Although, on the borders of the service area there can be some threat coming from the possible expansion of the

[3] The webpage of the Hungarian Water Utility Association gives more information on the topic: http://www.maviz.org/?q=english (accessed May 11, 2010)

neighboring provider, this pressure is expected to be limited. In case of a natural monopoly regulation may be needed in order to decrease the inefficiencies and the high burden imposed on customers through the high prices. In order to justify this some further investigations have to be made. It has to be examined whether there are any other goods or services which could serve as substitutes. Candidates for that could be bottled mineral water or an own well. However, the first is not an alternative for water needs different from drinking. Wells may be a real option regarding big industrial customers or households with a garden. Still, water needs to be of sufficient quality and a license is also necessary. These, together with the costs of building such establishments put a limit to substitutability and justify the needs of regulation.

As my thesis is about pricing, in the followings I focus on price regulation. There are several goals, sometimes contradictory as well, which have to be taken into account when designing a pricing scheme. First, safe operation has to be secured by prices covering operating costs and preferably costs of investments as well. This aim is also emphasized by the Water Framework Directive of the European Union.[4] On the other hand lower prices are desirable in order to lower the burden on consumers, especially on households. While setting lower prices can serve as an incentive for the firm to operate more efficiently, these may provide insufficient incentive for the consumers to use water economically which can lead to exhausting water resources (see Garcia and Reynaud 2004, and Montginoul 2007).

The operation of a water utility can be done by a private investor or a public firm. In any case the role of the regulator is played either by the local municipality or a state authority. For

[4] 2000/60/EC

example in Hungary the price authority for the state-owned utilities is the Minister of Environment and Water, while that of the other utilities are the municipalities of the service area.[5] In Romania the National Regulatory Authority for Municipal Services has the responsibility to survey prices of water supply, while the final approval is done by the local public administration.[6] In Poland tariffs are determined by the executives of local communes.[7]

Price regulation in the water supply sector is typically based on price cap regulation. This means that an upper limit of the price is set for a certain period, based on incurring costs, including a path of price change, where cost increases and potentially efficiency improvements are also taken into account (on this topic see Viscusi et al. 1995). Tariff decisions are based on the proposition of the utility and typically are in the form of maximum applicable prices, determined separately for the different consumer groups, i.e. for residential and nonresidential (industrial and institutional) consumers. Sometimes prices are formed using an exact price formula[8], in other cases they are a result of a bargaining process between the regulator and the regulated firm (see Nauges and Thomas 2000). Prices are set in the beginning of each year and are valid for the whole year.

Operating costs, which are the base for regulated prices depend on several characteristics of water utilities, including size, type of aquifer, morphology and population density of the service area (see section 3.2.). Moreover, cost elements which can be taken into consideration in the calculation of prices can also vary among countries. There are differences regarding the

[5] Act LXXXVII /1990
[6] Law no. 241 (2006)
[7] Act on collective water supply and collective sewage disposal of 7 June 2001
[8] Examples for that can be found in Hungary, among others this is the case in the city of Dunaújváros (http://www.dunaujvaros.hu/docfile.php?id=91 – in Hungarian language, accessed June 04, 2010)

question, whether investment costs are included or not (Bobadilla and Camrova 2006). In the sector of public water supply the value of fixed assets is very high compared to variable costs. Network, pumps and machinery represent a considerable amount of sunk costs. Consequently huge expenditures are needed to keep infrastructure in a good quality; otherwise proper quality of supply cannot be secured (Garcia and Thomas 2001). It is a frequent problem, that prices not cover all the costs, especially not investment costs (see Bobadilla and Camrova 2006). As a result necessary infrastructural investments are abandoned or have to be financed by state or municipal subsidies (see for example Kopańska 2009).

Furthermore prices also differ in the sense that various price structures can be applied. In case of a flat rate, which is rarely used, the monthly price of service is independent of the quantity consumed. Prices can be purely proportional to quantity consumed (variable price), or may be combined with a fixed element in a two-part tariff. A more complicated version is block pricing. In this case intervals are formed regarding the quantity of water consumed. Either the amount of fixed charge or that of the volumetric part differs in each interval (see Montginoul 2007).

As a result prices show a considerable variation among utilities. They depend on the given characteristics of the firm and its service area, the decisions of different regulatory authorities including the possible bargaining process with the utility. Additionally, they might also vary within a utility depending on the type of customers and the different municipalities within the service area.

With the knowledge of the sector's characteristics described above now it is possible to turn to the econometrical analysis of the demand and supply side.

3. *Theoretical background*

3.1. Demand estimation

In order to estimate the price elasticity of residential water demand first a proper demand function has to be specified. It is necessary to consider the main factors influencing the quantity of water demanded. Arbués et al. (2003) and Worthington and Hoffman (2006) provide a good starting point by giving a review of papers which focus on this question. According to these studies it can be concluded, that income, composition of households, housing characteristics, weather, characteristics of billing and water restrictions are the ones which have a considerable effect on individual water demand, besides the main variable of interest, which is price. Let us take a closer look at these factors one by one. In the followings I largely rely on the two papers mentioned above.

From microeconomic theory it can be expected, that price has a negative effect on demand, assuming that water is not a Giffen-good. However, it is not easy decide how to capture the measure of price. The main problem lies in the different tariff structures applied by water utilities. In case of two-part tariffs, where the volumetric fee is accompanied by a fixed component, price measure can either be the marginal price, the average price or some combination of these taking into account the fixed tariff element as well. Starting from basic microeconomic theory I would suggest marginal prices to be used, as the fixed part is assumed not to influence quantity choice there, it only affects consumer surplus. This approach is frequently chosen in the literature as well, taking Taylor et al. (2004) as an example. On the other hand advocates of average price (used by Nauges and Thomas (2000) among others) claim that in reality consumers react only to the average price they pay for one cubic meter of water, the true price structure does not matter for them. Shin (1985) suggests a way in the middle by introducing a so-called price-perception parameter which is defined as

$$P_p = P_m \left(\frac{P_a}{P_m}\right)^k \tag{1}$$

where P_p denotes price perceived by the consumers, P_m and P_a are marginal and average prices respectively, finally k is the price-perception parameter. A fourth solution is used mostly in case of complicated block pricing schemes. Price is captured by two variables: the one is marginal price and the other is a difference variable, introduced by Taylor (1975) and modified by Nordin (1976). The latter can be calculated by using the following formula:

$$B - P_m(Q) \cdot Q \tag{2}$$

where B is the total bill paid by the consumer and P_m is the marginal price at quantity Q demanded. This difference variable measures the negative income effect coming from the price structure, corresponding to the given level of consumption. Demand estimation in the presence of increasing or decreasing block pricing raises several further problems, which are not to be considered here, as block pricing does not emerge in the sample used for the following estimations. However, the approach with a difference variable can also be used with two part tariffs where marginal price is constant. In this case difference variable is equal to the fixed part of the tariff. Finally it is worth mentioning, that some authors (e.g. Arbués et al. 2004) suggest that lagged prices should be used in the demand function. Their argument is based on the assumptions that new prices are learnt by the customers only ex post, when they receive the water bill. In the current work it does not seem reasonable, as prices are changed yearly, while billing happens usually monthly. Consequently using annual data it can be expected, that consumers are familiar with current prices.

Turning to the next factor, income, microeconomic theory can be used as a starting point again. If water is a normal good, quantity demand increases when income is higher. Moreover, people with higher income tend to spend a lower fraction of income on paying the water bill. Thus it can be expected, that richer people are less sensitive to price changes, price elasticity may depend negatively on income. On the other hand income can be positively correlated with education, which may exert an opposite effect on demand, if it means higher environmental consciousness and more concern about water savings (see Worthington and Hoffman 2006).

Material environment should be considered as well. The presence of a garden is a demand increasing factor in itself. If there is a huge lawn or a swimming pool, demand will be even higher. These two are positively correlated with income. However, it is expected to be also true for water conserving appliances, which lower demand. Consequently the precise effect of income is not clear if one cannot control for factors like presence of a pool or water saving appliances (see Worthington and Hoffman 2006).

Personal composition of household matters in water demand from two aspects: size and age. It is showed (Arbués et al. 2004) that there exists an economies of scale in water consumption, i.e. per capita consumption in households with more members tends to be lower. Age is also an influential factor, as some water consuming activities are more likely in certain ages. Younger people have typically higher demand, especially with small children. On the other hand elderly people work on average more in the garden, which is a highly water demanding activity (see Arbués et al. 2003).

Weather characteristics, like temperature and rainfall also play an important role in water demand. I would expect that rainfall influences mainly outdoor water use, so it really matters for those who have a garden, while temperature influences demand of people living in all kinds of households. More rain decreases water demand and higher temperature has the opposite effect. Rain is captured either as the number of rainy days (e.g. Maidment and Miaou 1986) or the amount of rainfall (e.g. Nauges and Thomas 2000). Temperature is claimed to have a nonlinear effect on water demand, increasing demand only above a threshold level (Maidment and Miaou 1986).

Finally it is claimed by Stevens et al. (1992 cited Arbués et al. 2003) that more frequent billing results in higher price awareness, which can lead to higher price elasticity. Metering at the household level can have similar consequences (Yepes and Dianderas 1996 cited Worthington and Hoffman 2006). It is debated whether price structure affects price awareness and elasticity, but the results of Nieswiadomy and Molina (1989) support this assumption. If there are non-price water restrictions, like non-availability of supply after a given quantity consumed, these should also be considered as demand lowering factors.

In addition to the enumerated factors, which were emphasized in the review papers of Arbués et al. (2003) and Worthington and Hoffman (2006), two other features of water supply can be thought of which might have an influence on demand. The first is supply quality and the second is the presence of wastewater collection and treatment together with the price charged for it. Product quality or in this case service quality is considered to have an important effect on demand according to the microeconomic theory as well. There are several models in the field of industrial organizations analyzing the choice of quality taking into consideration its effect on demand. (e.g. Buehler et al. 2004) Consumers are ready to pay more for the same quantity if the quality is higher. In this way increasing quality shifts out the demand curve thus it increases consumer surplus. The same effect is expected to be valid at water supply as well. Consequently it is advisable to include some measure of service quality in the demand function, though it is not a usual practice, according to the papers on demand estimation in public water supply mentioned above.

The other factor is the availability of sewage treatment together with the corresponding price. It is reasonable to include these, if wastewater collection and treatment is thought of as a complementary service to water supply. From the households' point of view wastewater collection and treatment may even be regarded as an additional service within water supply. This view is supported by the fact that in most cases the wastewater bill of households is based on the quantity of water consumed, so customers can only react to wastewater tariff changes by adjusting their water consumption. In this way some studies, like Billings and Agthe (1980) suggest that wastewater price should also be included in the demand function. However, based on all the papers referred to in this section, it can be said that in most cases the effect of wastewater price is not taken into consideration, water demand is estimated separately.

Summarizing what is written above the demand function is assumed to be of the following form:

$$Q_{ijt} = f(p_{jt}, y_{ijt}, h_{ijt}, w_{jt}, b_{ijt}, q_{jt}, s_{jt}) \tag{3}$$

Lower indices *i*, *j* and *t* refer to consumer, firm and time period respectively. All kind of price variables are included in *p*, *y* is personal income, *h* captures all household characteristics, personal and material as well, *w* consists of weather variables, *b* are behavioral features referring to mostly water saving consciousness, *q* is supply quality and *s* contains all other important features of supply, like metering, non-price restrictions or the availability of sewage collection and treatment. As the database used consists of data aggregated at the level of water utilities, individual demand functions should be aggregated to firm level demand, using average values as an approximation.

3.2. Cost estimation

After specifying the determinants of demand let us turn to the supply side. Supply can be characterized by the cost function. It is important to calculate marginal costs in order to evaluate the effect of quantity changes on costs, which is a crucial part of estimating the welfare effect of price changes. First, determinants of the cost function have to be considered, like in case of demand. According to Renzetti (1999) a general cost function can be written in the following form:

$$C = f(p^i, Q, Z) \tag{4}$$

where p^i stands for input prices, Q is the quantity of output and Z is a vector of controls including technological characteristics and features of the service area. C can either be variable costs (e.g. Antonioli and Filippini 2001) or total costs (e.g. Filippini et al. 2008) depending on the purpose of the estimation. Variable costs can be defined as the sum of labor, electricity, material and other operating expenses, while total costs include additionally expenditures on capital, including investments and renewal expenses (Garcia and Thomas 2003). According to García-Valiñas (2005) operating costs are proper for estimating short-term relationships, while total costs, also including expenditure on capital serve for the estimation of a long-run cost function. In the following I will focus on the estimation of the short-run cost function. In this case it can be assumed, that the firms minimize operating costs subject to given conditions, represented by the vector Z. (Garcia and Reynaud 2004)

Cost function is obviously non-decreasing in input prices. Most frequently unit cost of labor, energy, capital and average material costs or a subset of these is included among them (see Garcia and Thomas 2001 and García-Valiñas 2005 among others), as these are the most important inputs used. Additionally Nauges and van den Berg (2007) include unit cost of services being contracted out. Reference to their work is essential, as they used the same database I do, only for different countries and years. Their aim was to estimate a cost function and analyze economies of density and scale. Though their suggestions were useful especially regarding the limitations of the database, I did not completely follow the estimation method used in their paper.

Turning to the left-hand side variable, it can be noted, that either variable costs or total costs are used, short-run marginal costs can be obtained by taking the derivative of the cost function with respect to output. Consequently the quantity of output is the variable of main interest, if the reaction of costs to changes in output is to be analyzed. Output can be defined as the quantity of water produced (Nauges and van den Berg 2007) or the quantity sold (like Garcia and Reynaud 2004). The difference between the two can be captured by a loss variable, which includes own usage of the firm, leakages in the system and water being stolen. In case of using quantity sold as the measure of output, it is good to include the amount of losses as well, since the presence of these results in increased costs (see Garcia and Thomas 2001). It is reasonable to choose the latter solution if it is assumed, that water sold and water lost impose different amount of costs.

Water losses are an obvious link to the broad set of controls included in the cost function. These are features of the technology used and characteristics of the service area. Several factors can belong to this category. One subgroup is technology used, which clearly influences the evolution of costs. It can be described by factors like type of raw water treatment if necessary at all (Renzetti 1999), or the type of water source (Filippini et al. 2008). As a rule of thumb higher level of treatment needed results in higher costs in the first case, though in reality treatment levels are not so easy to determine. Type of water source, i.e. surface water or different kinds of underground water is partly correlated with the necessary treatment, partly influence costs through the pumping needs. Size and quality of capital stock have to be included as well, since these have an influence on operating costs. More pumps or cleaning machines need more electricity, while a lower quality of the infrastructure results higher losses of output and makes more maintenance necessary, which requires more labor and material. Length of distribution network can serve as a proxy for capital stock (Garcia and Reynaud 2004). The state of the infrastructure can be characterized by a measure of water losses (Garcia and Thomas 2003) or the number of pipe breaks during the year (Nauges and van den Berg 2007). A further suggestion of Nauges and van den Berg (2007) is to include average daily duration of water supply in case it is different from 24 hours, as it can refer to an obviously lower service quality and has a cost decreasing effect. Finally characteristics of the service area, like its size (Filippini et al. 2008) and morphology (Antonioli and Filippini 2001) are also to be included among the controls. Both are in connection with pumping needs, thus with the necessary volume of electricity used. Size can be approximated by the number of towns served (Garcia and Thomas 2001). Finally Antonioli and Filippini (2001) claim that operating costs also depend on the number and type of customers served. The latter can be quantified as the ratio of residential consumption (Nauges and van den Berg 2007 or Renzetti 1999), assuming that nonresidential output has a different effect on costs.

4. Data used

The dataset I worked with was provided by the International Benchmarking Network for Water and Sanitation Utilities (IBNET). This is an initiative of World Bank, collecting data about firms providing water supply and sewage treatment all over the world. Participation in data provision is voluntary and it is supported by local coordinators in each country.[9]

For the estimations I used a subset of this huge database, containing data of 702 water and sewage utilities coming from 21 countries of Central- and South-Eastern Europe and the post-Soviet region. The list of countries included and the yearly number of firms providing data can be found in Appendix A. The structure of the database is an unbalanced panel of annual frequency, containing data aggregated on the firm level, coming from the period of 2000-2008. The same database was used in the analysis of the Regional Centre for Energy Policy Research investigating the factors which influence unit costs in the sector. (Bisztray et al. 2009)

In the IBNET database nominal data including prices, GDP, costs and revenues were given in local currency units. In order to make them comparable, I converted them into US dollars, making adjustments based on purchasing power parity. For this purpose I used the purchasing power parity conversion factor given by World Bank.[10] To enable over the time comparisons

[9] More information about IBNET can be found on its webpage: http://www.ib-net.org/
[10] PPP conversion factor data are available at the official United Nations site for the Millennium Development Goals Indicators: http://mdgs.un.org/unsd/mdg/Metadata.aspx?IndicatorId=0&SeriesId=699/ (accessed May 11, 2010)

as well, I calculated real values from all nominal measures using 2000 as the base year. For this transformation I used consumer price index data provided by Word Bank. [11]

Additionally, weather data were also included in demand estimation, coming from sources other than IBNET. Daily average temperature data being representative for a country were used from the database of the National Climatic Data Center of the United States.[12] The source of daily precipitation data is the European Climate Assessment[13]. In case of more than five weather stations reporting data in a country a random sample of five stations was used to represent yearly tendencies within the country.

Definitions and descriptive statistics of the data used are to be found in Appendix A.

[11] Consumer price index data are available at the webpage of World Bank: http://data.worldbank.org/indicator/FP.CPI.TOTL.ZG (accessed May 11, 2010)
[12] Temperature data are available at the webpage of the University of Dayton: http://www.engr.udayton.edu/weather/ (accessed April 13, 2010)
[13] Precipitation data are available at the webpage of the European Climate Assessment: http://ecad.knmi.nl/dailydata/customquery.php (accessed May 11, 2010)

5. *Estimation results*

5.1. Demand estimation

In the previous section about the theoretical background of demand estimation concerning water supply a demand function was established. As the next step the method of estimation should be decided, given the data available. Based on the theoretical arguments the true model can be formulated as follows:

$$Q_{jt} = \beta_0 + \beta_1 \cdot p_{jt} + \beta_2 \cdot y_{jt} + \beta_3 \cdot h_{jt} + \beta_4 \cdot w_{jt} + \beta_5 \cdot h_{jt} \cdot w_{jt} + \beta_6 \cdot b_{jt} + \beta_7 \cdot q_{jt} + \beta_8 \cdot s_{jt} + D_t + a_j + u_{jt} \tag{5}$$

All values being previously individual are now averaged at the level of the firm. Variable D_t is a composition of year dummies, a_j denotes firm fixed effects not measured by other variables and u_{jt} is the usual error term.

However, using pooled cross section estimation as indicated above is likely to result in biased and inconsistent estimates for the price coefficient, as there are several variables on the right hand side which are not measurable. These include such influential factors of demand like average income or household characteristics. Thanks to the panel structure of the database the problem of endogeneity coming from time-fixed omitted variables can be solved by using panel estimation methods. From econometrics it is known that three such methods are to be chosen from: first differencing, fixed effects or random effects. Random effects can be ruled out, as the necessary assumption for consistency, i.e. $\mathrm{cov}(a_j, x_{kjt}) = 0$ with x_k being the k-th right hand side variable is not likely to hold having omitted variables like household characteristics. Using either of the other two methods eliminates time-invariant factors from the equation. Although variables like housing characteristics (h) or water saving behavior (b) have a time index in the model, they can be assumed to remain constant over time in the short-run. The same is assumed for metering level, billing frequency or non-price water restrictions. Then the first-differenced demand equation will be:

$$\Delta Q_{jt} = \gamma_0 + \beta_1 \cdot \Delta p_{jt} + \beta_2 \cdot \Delta y_{jt} + \beta_4 \cdot \Delta w_{jt} + \beta_5 \cdot h_j \cdot \Delta w_{jt} + \beta_7 \cdot \Delta q_{jt} + D_t + \Delta u_{jt} \tag{6}$$

Using fixed effect estimation, the equation is similar, only the time-demeaned version of $\ddot{x}_{jt} = x_{jt} - \overline{x}_j$ is used instead of the first differences (Δx_{jt}). The main deciding factor between these two methods is the presence of serial correlation in the original error term. Having a strong serial correlation the method of first differences is preferable, while using fixed effects is better if serial correlation is negligible.

Before turning to other questions regarding specification it is useful to take a look at the variables used in the estimations. The left-hand side variable is average yearly per capita water consumption. The price variable is defined as average price of water, making the assumption that consumers react to that instead of regarding only marginal price. Using average price is also supported by the fact that marginal price could not be estimated reliably based on the available database. Moreover, the majority of firms did not report to apply two-part tariffs. However, different price measures are also used for testing robustness. Calculation method of marginal price and the difference variable can be found in Appendix A. Income data were unfortunately not available on the firm-level. As a remedy per capita GNI was used as a proxy for over the time changes in income. This variable cannot measure within country differences, but can capture differences across countries, which can be more relevant. Weather conditions could also be considered only at the country level, so they are an imprecise measure of the true factors. Several specifications were tested based on the literature (Maidment and Miaou 1986, Martinez-Espiñeira 2002, and Nauges and Thomas 2000). The list and description of these variables can be found in Appendix A. In spite of the fact that household characteristics are assumed to be constant over time in the short-run, they remain in the differenced equation as part of the interaction term with weather variables. The most important feature of households in this respect is whether there is a garden. Though it is not included in the database, type of service area can be used as a proxy. Three categories are defined: urban, rural and both. I assumed that there are more gardens if the service area can be characterized as rural. The third category lies somewhere in between, regarding the presence of gardens. A proxy was used for supply quality as well, defined as the loss ratio, which corresponds to the fraction of water production being not billed. This includes water loss through leakages in the network and own water usage of the firm, but also the amount of water thefts. Assuming that the last factor is negligible compared to the first two, loss ratio can expected to be correlated with the quality of network and appliances and also with the overall quality of service. Using this measure is supported by the paper of Garcia and Thomas (2003), who use a somewhat similar proxy for the physical state of the distribution network. This variable is defined as the volume of water lost per customer divided by the length of the network. Precise definition of the variables used in the estimations, together with notations and descriptive statistics can be found in Appendix A.

A somewhat different specification is the one including wastewater price. In this case correct specification has to be made with care as not all residential customers obtaining water supply are connected to the wastewater network as well. Consequently individual demand can be written as follows:

$$Q_{ijt} = \boldsymbol{\beta} \cdot \boldsymbol{x}_{ijt} + \alpha \cdot p_{jt}^{ww} \cdot d_{ijt}^{ww} + \delta \cdot d_{ijt}^{ww} + \epsilon_{ijt} \tag{7}$$

All previously mentioned variables are included in the vector of $\boldsymbol{x}$, wastewater price is denoted by p^{ww} and d^{ww} is a dummy being one if the given person gets the service of wastewater connection and treatment. Getting this service influences water demand through both wastewater charges and the enlargement of service. This latter effect is captured by including the wastewater dummy in itself. Taking the averages at the firm level the following equation is obtained:

$$Q_{jt} = \boldsymbol{\beta} \cdot \boldsymbol{x}_{jt} + \alpha \cdot p_{jt}^{ww} \cdot r_{jt}^{ww} + \delta \cdot r_{jt}^{ww} + \epsilon_{jt} \tag{8}$$

Instead of dummies the ratio of customers obtaining both services are included. As this ratio may change over time, both new variables are to be included in the differenced or time-demeaned versions as well. However, there are areas, where water supply and sewage treatment services are provided by different firms. An example for that can be Budapest. In this case the above suggested process would assign a zero value to the variable *r*, although wastewater price can still have an effect on water demand, only it is not known from the data. Consequently for the estimations including wastewater price I only used that subset of the database, where the condition $0 < r \leq 1$ holds.

The next question to decide is the functional form used in the estimation. Three main versions are used in the literature (see Arbués et al. 2003, p.90-93 and Worthington and Hoffman 2006, p.16-17): linear, logarithmic and log-lin, all with different characteristics. The linear version can be formulated as $Q = max\{0;a - bP\}$, with *a* and *b* being parameters or functions of variables different from price. Its advantage is that price elasticity is lower in absolute value at lower prices ($(dQ/dP)*(P/Q) = 1 - a/(a-bP)$) and there is a satiation level at price zero. However, assuming this functional form price changes result in the same amount of change in quantity at every level of consumption ($dQ/dP = -b$). Moreover, above a threshold price demand becomes zero, though water is an essential good. The latter two characteristics claimed to be unrealistic, but this form is still used in some papers (e.g. Martinez-Espiñeira 2002). The log-log specification ($log(Q)=a-b*log(P)$) can be derived from a demand equation of the form $Q = c/P^b$, where *c* and *b* are again parameters independent of price. Assuming this form price elasticity will be constant (-*b*), which can be debated again. Still, many papers work with this specification. (e.g. Nauges and Thomas 2000) Lastly log-lin form (*logQ = a – b*P*) has the advantage that a given subsistence level of water will be consumed even if price is very high. Timmins (2002) can be an example for application of this form. However, I find this assumption debatable, as I expect that after a really high threshold consumers tend to find other ways to get water, including digging own wells, going to the river, or stealing water from the network. All in all I chose the logarithmic functional form in the estimations, partly because it is widely used, and estimation of price elasticity is straightforward.

Before estimating the demand function there are still some questions to be clarified, which are well-known problems from econometrics. Outliers which I considered highly atypical or erroneous were left out of the sample not to distort the results. It affected only a few cases. Multicollinearity can be a problem when both water and wastewater prices are included as they can be expected to move together. However, these two variables proved to be only moderately correlated ($\rho = 0.37$) in the dataset. Heteroskedacticity was controlled for, and White period standard errors were used throughout the estimation process, as remedy of serial correlation.

Finally three possible sources of potential endogeneity, i.e. self-selection, measurement error and simultaneity have to be considered as well. The database consists of a nonrandom sample of water utilities from certain countries of Central- and South-Eastern Europe and the post-Soviet region. In some countries coverage of the full population is quite high, while it is not the case in others. (The table of coverage ratios by country can be found in Appendix A.) Taking part in the IBNET data collection is basically voluntary, though there are countries, like Albania, where it is a regulatory prescription for all water utilities. Still, in most countries it depends on the local coordinator, how many and which firms send their data to IBNET. Consequently it can be suspected, that bigger or more efficient firms are more likely to provide data.[14] This fact is especially problematic at cost estimation, but it may also cause a selection bias at the demand estimation. However, from demand point of view it may be an exogenous selection, being independent of per capita consumption and rather correlated with right-hand side variables as income or service quality measured by the loss ratio. Having this assumption selection does not change the expected value of per capita consumption conditional on the right-hand side variables, and it does not cause any bias. A potential bias can further be mitigated by including only those countries in the estimation where coverage ratio is high enough.

A next source of bias can be measurement error. This problem occurs at several right-hand side variables used, including proxies for income, presence of gardens and supply quality, weather data and estimations of the marginal price. Though, these do not cause any bias in the estimated coefficient of price if true values of these variables are uncorrelated with price, which can be assumed.

Lastly the potential problem of simultaneity has to be taken into account as well. Water price is typically set based on incurring costs. This suggests that prices can be regarded exogenous. However, prices can also be modified based on consumption quantity forecasted for the following year. In this way per capita quantity consumed may have an influence on price, assuming that forecasts are close to real consumption and there are no unexpected changes in the number of consumers throughout the year. This simultaneity problem could be handled by using an instrument for the price variable and applying the estimation method of two-stage least squares. A candidate for this role is unit costs of water supply, or a combination of the

[14] source: András Kis, IBNET coordinator of Hungary (personal communication, May 13, 2010)

elements of unit cost, like cost of labor, electricity or services contracted out. All these are expected to be exogenous regarding per capita consumption but being correlated with average price. However, these measures were found to be not good instruments of price. In the further estimations I assumed that prices are exogenous. It can be argued, that price changes are expected to be made mostly based on cost changes. Quantity forecasts have a negligible role if any concerning per capita consumption, they can be of higher importance regarding consumption changes coming from the changing number of customers.

Results of demand estimations are presented in Table 1. Left-hand side variable is yearly per capita residential water consumption in all cases, right-hand side variables used can be found in the first column. White period standard errors are presented in parentheses below the coefficient estimates. One, two and three stars denote variables being significant at a significance level of 10%, 5% and 1% respectively. Estimations are numbered in the first raw of the table. The last two rows contain the number of observations (N) included and the R^2 value.

Table 1: Results of demand function estimations

	(1)	(2a)	(2b)	(3)	(4)
log(APW)	-0.129***	-0.099**	-0.1**		-0.097*
	(0.0421)	(0.0504)	(0.0476)		(0.0503)
log(MPW)				-0.082**	
				(0.0372)	
RWW*					-0.152*
log(APWW)					(0.0794)
log(LOSS)		-0.046*		-0.047*	-0.039
		(0.0264)		(0.0282)	(0.026)
log(GNI)		-0.487***	-0.505***	-0.554***	-0.612***
		(0.146)	(0.1434)	(0.1501)	(0.1615)
RUR*		-0.0016***	-0.002***	-0.002***	
DRAIN13		(0.0004)	(0.0004)	(0.0005)	
RUR*		0.0288**	0.036***	0.025*	
log(RAINY)		(0.0134)	(0.0135)	(0.014)	
log(RAINY)					-0.019**
					(0.0081)
log(ATEMPY)		0.234***	0.242***	0.24***	
		(0.0783)	(0.0794)	(0.0809)	
constant	-0.071***	-0.016	-0.017	-0.015	-0.011
	(0.012)	(0.0151)	(0.0154)	(0.0157)	(0.0168)
N	1655	1182	1182	1121	942
R^2	0.087	0.129	0.12	0.136	0.152

As fixed effect estimations suffered from considerable serial correlation, while serial correlation was much lower in case of first differences, I used the latter estimation method in all cases presented above. Both left and right-hand side variables are in logarithmic form, with the exception of the rural dummy and the number of rainy days. Year dummies are also included in all equations, but the coefficient estimates are not presented here, though they were jointly significant in all cases. Specification (1) is a simple equation without any controls, including only year dummies. Specification (2) includes additional controls, all being significant, though not always of the expected sign. The coefficient of price, which is the variable of main interest decreased to some extent in absolute value, but the change is not large. Loss ratio, being the inverse measure of quality is also significant, though only at the

10% significance level and it has the expected sign. Specification (2b) is estimated without including loss ratio, which was found to be significant only on a level of 10%. It can be seen that there is no considerable change concerning the coefficient of price. The same specification was used in Specification (3) with marginal price included instead of average price. The results are not changed much and they remain robust if the sample used contains only those firms who did not report any fixed price components. This result is not surprising, as these water utilities represent the vast majority of the database. Wastewater price is included in Specification (4), using only that subsample, where both services are provided within the same utility and the population getting water supply is not lower than the number of people obtaining sewage collection and treatment services. Both water and wastewater prices are significant, though only on 10%, but the water price remains robust compared to previous estimations. As a further robustness check I also repeated estimations excluding post-Soviet countries, but the coefficient estimates of price remained unaffected, which supports the reliability of estimation results.

Based on all these I chose Specification (2b) as the best estimation of price elasticity. The obtained value of price elasticity is somewhat lower than those observed in the literature. Price elasticity is claimed to be typically within the range of 0.25-0.75 in absolute value or even smaller than that (Worthington and Hoffman 2006). My result falls into the lower part of this interval, and as an example it is close to the estimations of Martinez-Espiñeira (2002), who found short-run price elasticity in Spain to fall into the range of 0.12-0.17 in absolute value.

5.2. Cost estimation

Estimating a variable cost function requires the same steps which were done in case of the demand estimation. Starting with the functional form it can be noted that two main approaches are used in the literature: either a Cobb-Douglas or a translog specification is used. An example for the former can be found in the papers of Antonioli and Filippini (2001) or García-Valiñas (2005), while among others Nauges and van den Berg (2007) used the second functional form. Based on García-Valiñas (2005) but using the notations introduced in Section 3.2 a Cobb-Douglas cost function can be written as:

$$C = \delta \cdot Q^{\alpha} \cdot \prod_k {p_k^i}^{\beta_k} \prod_l {Z_l}^{\gamma_l} \tag{9}$$

where α, β_k, γ_l and δ are parameters, p_k^i denotes the price of the k^{th} input and Z_l refers to the l^{th} component of technology included. Taking logs leads to the following estimateable functional form:

$$\ln(C_{jt}) = \ln(\delta) + \alpha \cdot \ln(Q_{jt}) + \sum_k \beta_k \cdot \ln(p_{kjt}^i) + \sum_l \gamma_l \cdot \ln(Z_{ljt}) + D_t + a_j + u_{jt} \tag{10}$$

Following the paper of Garcia and Thomas (2001) the starting point of the translog specification is a Cobb-Douglas cost function, where parameters may also depend on the value of the controls included in Z:

$$C = \delta \cdot Q^{\alpha(Z)} \cdot \prod_k {p_k^i}^{\beta_k(Z)} \prod_l {Z_l}^{\gamma_l(Z)} \tag{11}$$

and the parameter functions are specified as

$$\eta(Z) = \eta_0 + \sum_l \eta_l \cdot \ln(Z_l) \quad \text{with } \eta = \{\alpha; \beta_k;\ \gamma_l\} \tag{12}$$

Taking the logarithm and including additional squares and interaction terms of the variables gives the translog version of the cost function:

$$
\begin{aligned}
\ln(C_{jt}) = \mu_0 + \mu_1 \cdot \ln(Q_{jt}) + \sum_k \mu_{2k} \cdot \ln(p^i_{kjt}) + \sum_l \mu_{3l} \cdot \ln(Z_{ljt}) + \mu_4 \cdot (\ln(Q_{jt}))^2 \\
+ \frac{1}{2} \sum_k \sum_m \mu_{5km} \cdot \ln(p^i_{kjt}) \cdot \ln(p^i_{mjt}) + \frac{1}{2} \sum_l \sum_n \mu_{6ln} \cdot \ln(Z_{ljt}) \cdot \ln(Z_{njt}) \\
+ \sum_k \mu_{7k} \cdot \ln(Q_{jt}) \cdot \ln(p^i_{kjt}) + \sum_l \mu_{8l} \cdot \ln(Q_{jt}) \cdot \ln(Z_{ljt}) \\
+ \sum_k \sum_l \mu_{9kl} \cdot \ln(p^i_{kjt}) \cdot \ln(Z_{njt}) + D_t + a_j + u_{jt}
\end{aligned}
\tag{13}
$$

Though this form is more flexible, due to the squares and interaction terms many variables have to be included, which may lead to multicollinearity. Consequently I followed the solution suggested by Antonioli and Filippini (2001) and chose the simpler Cobb-Douglas functional form.

Therefore the true model is assumed to be described by Equation (10). However, it is useful to take the advantage provided by the panel structure of the database and use panel methods in the estimation. From similar reasons as in case of the demand function, I chose the method of first differencing. This means that cost influencing variables being constant over time can be omitted without making parameter estimates inconsistent. Among the factors listed in Section 3.2 type of water source and raw water treatment needed, average daily duration of supply, size and morphology of the service area can be regarded constant over time in the short-run. Neither the length of network, as a proxy for capital stock does show much variation over time, though it changes in a number of cases. As overall capital stock cannot be assumed to be constant over time, another measure should be found instead of network length. Thus I decided to use the gross book value of fixed assets connected to water supply.

After considering the factors being constant over time the following equation is to be estimated:

$$
\Delta(\ln(C_{jt})) = \alpha_0 + \alpha_1 \cdot \Delta(\ln(Q_{jt})) + \sum_k \beta_k \cdot \Delta(\ln(p^i_{kjt})) + \sum_l \gamma_l \cdot \Delta(\ln(Z_{ljt})) + D_t + \Delta u_{jt} \tag{14}
$$

The left-hand side variable, C is defined as total yearly operating expenses of water supply. Q is total quantity of water produced. Input prices, i.e. price of labor, electricity and services being contracted out together with other expenses falling to one cubic meter produced are incorporated in the vector p^i. Price of labor is calculated as total labor expenses over the number of staff. The other three input prices are projected to one unit of output, as no quantity measures of inputs were available. Though pure price effects cannot be distinguished in this way, the lack of data made this solution necessary also proposed by Nauges and van den Berg (2007). Z is a vector of controls including ratio of water losses to total quantity produced, number of pipe breaks falling to one kilometer of the distribution network, gross book value of fixed assets, number of customers served and the ratio of water consumed by residential customers. Definitions and descriptive statistics of the variables used together with the method of calculation if necessary can be found in Appendix A.

Outliers considered mistaken or highly atypical data were excluded before the estimations. Presence of multicollinearity among the right-hand side variables was also checked. I found a strong correlation between total quantity produced and the number of residential customers, therefore I left out the latter from the estimations.

Exogeneity should be ensured in order to get consistent estimates for the parameters of interest, which are the coefficient of quantity produced and that of the loss ratio in the second row. Selection bias, mentioned in Section 5.1 may be problematic if firms not providing data are different in the sense that they tend to have higher costs even if conditioning on the controls included. In this case estimations will not be consistent for all the water utilities of the countries included in the estimation. Nevertheless, it can be argued that selection bias is expected to be mostly based on the size of the utility, which is included among the right-hand side variables. This kind of selection bias does not cause any inconsistency of the estimates.

Still, the presence of selection based on the left-hand side variable cannot be entirely closed out. Measurement error is presumed in case of capital stock and input prices except for the price of labor. Again, if true values of incorrectly measured variables can assumed not to be correlated with the variables of interest, coefficient estimates of the latter remain consistent. This can be the true for quantity produced. However, loss ratio may be correlated with capital stock, as a larger network can be connected with more lost water through leakages, given the same amount produced.

Finally a problem of simultaneity may occur between total operating costs and quantity produced. Since prices are set based on costs, the latter can have an influence on quantity demanded, thus also on quantity produced, being these two closely related. If that is the case, quantity has to be instrumented in the estimation and the method of two-stage least squares is suggested. However, total quantity is mostly based on population served and the number and size of nonresidential consumers. Moreover, water losses which do not depend on prices give a significant part of total quantity produced. All in all, the volume produced is considered to be exogenous in the following estimations.

Table 2 shows the results of two different specifications, estimating the effect of quantity on costs with and without controls. The left-hand side variable is total yearly operating expenses of water supply. When it was necessary White period standard errors were used as a remedy for serial correlation, as indicated in the table. Otherwise the table has the same form as in case of the demand estimation.

Table 2: Results of cost function estimations

	(1)	(2)	(3)
log(Q)	0.442***	0.451***	0.423***
	(0.0901)	(0.0546)	(0.1132)
log(CUST)			0.106
			(0.2007)
log(NETW)			0.221***
			(0.0785)
log(LPRICE)		0.003	0.005
		(0.0161)	0.0215
log(EPRICE)		0.072***	0.064
		(0.0193)	(0.0397)
log(SPRICE)		0.027**	0.024
		(0.0103)	(0.018)
log(OPRICE)		0.059***	0.057
		(0.0085)	(0.0444)
log(FA)		0.012	0.008
		(0.0129)	(0.0103)
log(LOSS)		-0.031**	-0.029
		(0.0152)	(0.0203)
constant	-0.023*	-0.01	-0.01
	(0.014)	(0.0215)	(0.0209)
standard errors	White period	ordinary	White period
N	1493	412	409
R^2	0.171	0.243	0.264

Coefficient estimates of quantity produced prove to be robust, though using a smaller subset of controls gives somewhat larger or smaller estimates. However, I chose Specification (2) as the best one. Although including the ratio of residential consumption and the number of pipe breaks does not change the coefficient estimates of interest, they are omitted in the chosen specification, as they proved to be highly insignificant and including them is not fully supported by the papers which I referred to in Section 3.2.

The coefficient of interest is the one corresponding to output. This measures the cost elasticity with respect to output. The estimated value of 0.45 is in line with results observable in the literature, although it is somewhat lower. Using a similar Cobb-Douglas cost function

Antonioli and Filippini (2001) estimated a value of 0.69 for Italian water utilities. With a translog functional form applied for Brazil, Colombia, Moldova and Vietnam Nauges and van den Berg (2007) found even higher values, situated between 0.61 and 0.96.

The obtained coefficient of output indicates increasing returns to scale. However, in the literature the measure of returns to scale is considered to be different from the value of cost elasticity with respect to output in itself. (Antonioli and Filippini 2001 or Nauges and van den Berg 2007) A change in the size of the firm is defined as a parallel change in output, number of customers and service area. Economies of scale are captured as the percentage change in costs caused by a one percent increase in each of these three factors. Using Specification (3), where both population and network length as a proxy for the size of service area are included allows the estimation of returns to scale, defined as described above. It can be calculated by taking the sum of the coefficients of these two variables and that of output, then taking its inverse (Filippini et al. 2008). Economies of scale are found if this value is greater than one, which is the case here. However, the reliability of this estimate is constrained by the fact, that the number of customers is an insignificant variable in the estimation.

Turning back to the chosen Specification (2), signs of the estimated coefficients correspond to what was expected. Only the negative sign of the loss ratio may seem unexpected. Still, taking into consideration that coefficients have to be interpreted in a ceteris paribus way, also this result becomes reasonable. This means, that given a fixed level of quantity produced, having a higher loss ratio results in lower costs. A considerable part of costs comes from the distribution of water to consumers. As a large part of water lost (excluding thefts) does not go through the full process of distribution, it is obvious that lower costs are imposed by this kind of water.

6. *Welfare analysis – an application*

6.1. Modeling framework

In the previous section I presented estimations regarding price elasticity of demand and response of variable costs to changes in quantity produced. These can serve as a starting point for further applications. In the following section I suggest a model, which makes it possible to evaluate the welfare impacts of present pricing practices applied by water utilities and compare it to different pricing schemes. Calculations are based on previous estimation results. Although the model to be set up can serve as a framework for such evaluations, reliability of the exact calculations is unfortunately constrained by the availability of data. Still, their sign and magnitude can call the attention to problems connected to present pricing practices.

The majority of papers analyzing welfare effects of pricing in public water supply evaluate the welfare losses caused by deviations from first best pricing (e.g. Renzetti 1999). It is well-known from theory that prices being equal to marginal costs provide the efficient or first-best solution (on this issue see Garcia and Reynaud 2004). Additionally fixed prices can be imposed, which only leads to redistribution of surplus between producer and consumers in order to cover losses suffered by the firm, but does not change the overall welfare. However, this framework is a static one, which does not take into account the effect of current decisions on future outcomes. Yet, there are very important aspects of water supply which can only be examined using a dynamic model, where also time dimension is included. One such question is in connection with water extraction. Though water is a renewable resource, a certain aquifer can be exhausted if the amount used from it is steadily higher than the sum of natural and artificial recharge. Moreover, water quality can also easily deteriorate. The effect of increased consumption on the volume of water available in the aquifer is regarded in a paper of Timmins (2002). He also considers the increased pumping costs caused by the diminishing level of water height in the aquifer. Using a dynamic framework Timmins (2002) suggests that prices should be made equal to long-run marginal costs, which also take into account this cost increasing effect of increased consumption. The model I propose here builds on the work of Timmins (2002), but I employ the suggested concept for a different problem.

In the following model I focus on the quality of infrastructure and its evolution over time. This is a question of both theoretical interest and strong practical motivation. It is frequently claimed (e.g. Bobadilla and Camrova 2006), that revenue does not cover all expenditures which would be necessary to maintain the quality of infrastructure, do the necessary repairs and replacements. Consequently low prices not only result in high consumption and lower profits due to inelastic demand, but also lead to a diminishing quality of infrastructure if necessary investments are postponed and not financed from other sources provided by the state or from bank loans. In the following model I disregard the last possibility.

Besides the maintenance of current infrastructure new investments can also be necessary or even inevitable in the longer run. These include implementing modern technology, applying higher level of raw water treatment or expanding the network to areas not yet reached. Although the problem of such new investments is of high importance, it will not be taken into consideration in the future analysis, as it would make the model much more complicated and evaluating the needs of new investments would require a far richer database than what was available for me. By disregarding the necessity of new infrastructure enlarging investments, it can be said that the role of investment and its potential effect on pricing will be underestimated, the true overall effect is expected to be even higher than the here estimated value.

Let us start from the basic model of static welfare estimation then make it dynamic by including the quality of infrastructure. A textbook version of the static welfare maximization problem, i.e. the problem of the regulator could be written in the following form:

$$max_p\{\nu \cdot CS(p) + (1-\nu) \cdot \pi(p)\} \tag{15}$$

where the variable of choice is the price (p), and the weight corresponding to the consumer surplus (CS) relative to the profit (π) in the objective function is ν.

Including the quality of infrastructure makes it necessary to think over which elements of the welfare function specified above are influenced by this factor. First it has a direct effect on costs thus on profit as well. Lower infrastructural quality is expected to raise operating cost due to higher value of lost output or more frequent repairing needed. Second, infrastructural quality is likely to influence service quality perceived by the customers, as it may be in connection with the number of interruptions of supply or the length and frequency of periods when perceptibly lower quality (e.g. discolored) of water is supplied because of repairs on the network. Based on microeconomic theory demand is expected to be an increasing function of service quality if prices are kept constant. Therefore it is also an increasing function of infrastructural quality being correlated with service quality. Hence the quality of infrastructure has an indirect impact on both revenues and costs, thus also on profits via changing the quantity consumed. Third, maintaining the quality of infrastructure requires expenditures or investments to infrastructural quality, which directly lowers profits.

Additionally to all these time dimension needs to be also considered as future quality is connected to current quality and current investments into quality. Obviously less investments result ceteris paribus in lower quality, and lower quality today leads to lower quality tomorrow if investments remain the same.

In order to formalize what was written above, based on Timmins (2002) a dynamic problem can be written in the following form, using a Bellman equation:

$$V(q_t) = max_{p_t,i_t} E\{\nu \cdot CS(p_t, q_t) + (1-\nu) \cdot \pi(p_t, q_t, i_t) + \beta \cdot V(q_{t+1})\} \tag{16}$$

Expectation operator is used, because true demand and costs are not perfectly known ex ante for the decision maker, as they also depend on random factors. *V(.)* is the value function, measuring the discounted present value of welfare over time. The state variable, q is the quality of infrastructure. Besides the price the amount of investments to the quality of infrastructure (i) is a new variable of choice. Finally β is the discount factor. This general optimization problem is constrained by the law of motion of the state variable:

$$q_{t+1} = g(q_t, i_t) \quad \text{with the first derivatives} \quad g_q > 0 \quad \text{and} \quad g_i > 0 \tag{17}$$

The amount of spending on infrastructure which maintains the current quality level will be denoted by i^*:

$$q_t = g(q_t, i_t^*) \tag{18}$$

It can also be expected, that the quality maintaining amount of investment, i^* changes with the level of quality, i.e. a network in a worse state may need more spending not to deteriorate more. However, this effect of quality will be disregarded in the followings. If the assumption of $\partial i^*/\partial q > 0$ were true, ignoring this relationship would result in underestimating the benefits of higher quality.

6.2. A numerical example

Using the previously presented model and the estimated elasticities the effect of different pricing policies can be evaluated, taking into consideration investment decisions to quality of infrastructure. Though the model is simplified in many aspects and crucial parameters could only be roughly estimated, the sign and magnitude of changes may be of interest. Furthermore this numerical example can demonstrate how the estimation results of Chapter 5 can be used for evaluations of pricing schemes. The dynamic model presented in Section 6.1 is an approach of such welfare evaluations I found especially interesting.

In order to evaluate the welfare effect of price changes, functional forms and parameter values have to be specified. By starting from Equation (16) a value has to be assigned to the discount factor and to the weight of consumer surplus in the welfare function. Values of $\beta = 0.95$ and $v = 0.5$ are set based on Timmins (2002). Consumer surplus can be written as the sum of residential and nonresidential surpluses:

$$CS(p_t, q_t) = \int_{p_t}^{\infty} Q(s, q_t) ds \cdot N + \int_{p_t^n}^{\infty} TQ^n(s, q_t) ds \tag{19}$$

Q(.) is the previously estimated per capita demand function of residential customers. *N* denotes the number of residential customers and assumed to be constant over time in the following calculations. Though loss ratio was found to be insignificant on a significance level of 5% as specified in the previous estimations, it can be expected that quality still has an effect on quantity demanded and on consumer surplus. Both measures are likely to increase due to a higher quality level if prices remain unchanged. However, estimation results do not help to find out the true form of this relationship. Hence in the following calculations I will not take into connsideration the demand and consumer surplus enhancing property of infrastructural quality and focus only on its effect on future costs. As a result I am most likely to underestimate the benefits coming from investment to infrastructural quality.

Turning back to the equation above, $TQ^n(.)$ denotes total demand of the nonresidential customer group. I assumed that residential and nonresidential surplus have the same weight in the welfare function. Although nonresidential demand function cannot be estimated reliably based on the data available, it has to be included in the calculations, as costs are not separated in residential and nonresidential output.

The next component of the welfare function is profit. I use this term for producer surplus (i.e. total revenue minus operating costs) less the amount of expenditures on infrastructural quality. Profit can be written in the following way:

$$\pi(p_t, q_t, i_t) = Q(p_t, q_t) \cdot N \cdot p_t + TQ^n(p_t^n, q_t) \cdot p_t^n - C(p_t, q_t) - i_t \tag{20}$$

where *C(.)* is the variable cost function estimated in Chapter 5.

Finally the infrastructural quality law of motion (Equation 17) needs to be specified. First a measure of infrastructural quality has to be found. As I suggested in previous sections the ratio of water lost will be used for this purpose, as a reverse measure. Though it depends on several factors, I assume that it is also highly correlated with infrastructural quality. More concretely, a better network quality is assumed to be accompanied by a lower loss ratio. A similar loss measure was used by Garcia and Thomas (2003) as a proxy of network quality. Their measure used was water per capita losses over one kilometer of the network. I chose a different normalization (i.e. water losses over total quantity produced) to get rid of the impact which increased production may have on the level of losses (Garcia and Thomas 2001).

Loss ratio will be denoted by l_t with $l_t \in (0; 1)$. There may be a lower bound greater than zero if a certain level of own water use is inevitable for the technology. However, this value is considered negligibly small. Before describing the law of motion for water losses I need to introduce the measure of underinvestment. I define underinvestment as a ratio of investment being necessary to maintain the present quality, i.e. water losses remain constant over time:

$$UI_t = \frac{i_t^* - i_t}{i_t^*} \tag{21}$$

There is no underinvestment ($UI_t = 0$) if $i_t = i_t^*$. A negative value of UI_t refers to overinvestment, which leads to a decreasing loss ratio. Variable UI_t is only bounded from above: $UI_t \in (-\infty; 1]$. This ratio is a more convenient measure than the absolute value of investment in itself, as this kind of normalization makes underinvestment comparable between utilities with different size of infrastructure and different investment needs.

Then the law of motion of loss ratio is proposed to be of the following form:

$$l_{t+1} = l_t \cdot e^{UI_t(1-l_t)} \tag{22}$$

This form has some crucial properties listed below:

1) $l_{t+1} = l_t$ if $UI_t = 0$ Based on the definition of underinvestment ratio it has to maintain loss ratio when it takes the value of zero (i.e. $i_t = i_t^*$).
2) $l_{t+1} \leq 1$ for all values of l_t and UI_t
3) $\partial l_{t+1}/\partial l_t > 0$ Loss ratio is persistent. A higher value today implies a higher value tomorrow given the same UI_t.
4) $\partial l_{t+1}/\partial UI_t > 0$ Loss ratio is a monotone increasing function of the underinvestment ratio. Higher value of investment leads to a lower loss ratio, provided that the initial loss ratio was the same.
5) $\lim_{u \to -\infty} l_{t+1}(l_t, UI_t) = 0$ for all values of l_t

A transformed version of Equation (22) could be estimated with the help of the database used.

$$\Delta(\ln(l_{t+1})) = c + \beta D_t(1 - l_t) + u_t \tag{23}$$

D_t is a vector of year dummies and one for the base year, β is the corresponding vector of coefficients. The value of UI_t can be obtained with the help of estimated values of vector β. They proved to be significant together and the estimated value for year 2008 was approximately 0.41. However, the constant proved to be significantly negative, which refers to a decreasing trend in water losses or a misspecification of the functional form. Still, as an approximation I will use the estimated value of *UI* in the following calculations.

For the calculations it is inevitable to determine the value of i_t^*. Having no other reliable data I assumed that it is equal to the value of depreciation of capital stock: δK_t. This estimate is motivated by papers evaluating the macroeconomic effects of investment in infrastructure (e.g. Rioja 2003). There the capital law of motion is described as usually in macroeconomics: $K_{t+1}=(1-\delta)K_t -I_t$. The value of capital stock is maintained if $I_t = \delta K_t$. In the current case I assume that maintaining the value of capital stock (i.e. mainly infrastructure) is equivalent to maintaining its quality level, as quantity of capital is assumed to be fixed over time. Depreciation rate was assigned a value of 0.1, frequently used in the literature (e.g. Rioja 2003). As for capital stock, having no better option its value is measured as the gross book value of capital connected to water supply. Though it is a very rough measure of quality maintaining investment needed, some tendencies still can be showed using this value.

As a further requirement of the calculations also the parameters of nonresidential demand function have to be estimated. The available database provided not enough information considering nonresidential water consumption, so the estimates of nonresidential demand are not expected to be highly reliable. Nevertheless, using the method of first differences lowers potential biases by eliminating those factors which are constant over time. In this way, I regressed total nonresidential consumption on average nonresidential price and time dummies. The estimated price elasticity was -0.4393.

Before turning to the calculation results a few words have to be said about the assumptions regarding choice variable *i*. In the model there are two possible sources of investment into infrastructural quality: operating profits and state allowances. I assume that losses of the firm and state allowances are both financed by the same taxpayers. Hence expenditures on infrastructure simply decrease profits being either positive or negative, irrespective of the exact source of investments.

In the followings three scenarios will be compared: current pricing practice (Scenario I), efficient, i.e. marginal cost based pricing (Scenario II) and full cost coverage (Scenario III). I created a 'typical' firm by using the average values of firms based on the data from the last year of the estimations, which was 2008. The value of i^* was also determined in this way. Then coefficients of the demand and cost functions were calculated, given quantities, prices and previously estimated elasticities. The exact numbers used can be found in Appendix B. As discussed before, the equation below can be regarded as the starting point of the calculations:

$$V(l_t) = max_{p_t,p_t^n,i_t} E\left\{0.5 \cdot \left[\int_{p_t}^{\infty} Q(s)ds \cdot N + \int_{p_t^n}^{\infty} TQ^n(s)ds\right] + 0.5[Q(p_t) \cdot N \cdot p_t + TQ^n(p_t^n) \cdot p_t^n - C(p_t, l_t) - i_t] + 0.95 \cdot V(l_{t+1})\right\}$$

$$s.t. \;\; l_{t+1} = l_t e^{UI_t(1-l_t)} \tag{24}$$

As the purpose of this numerical example is only to illustrate dynamic welfare effects of price and infrastructural investment decisions, instead of providing full over the time optimization I make some simplifications. I assume that from the second period on first best pricing is used with prices being equal to marginal cost and underinvestment ratio is zero. In this way the ratio of water losses remains constant over time. First-period pricing decisions have an effect on first-period surpluses only. As opposed to that investment decisions also influence future surpluses through changing next period quality of infrastructure, which leads to a shift in marginal costs, thus also in prices. As value of welfare remains the same from second period on, assuming an infinite horizon the problem can be written in the following way:

$$V(l_1) = max_{p_1,p_1^n,i_1} E\left\{0.5 \cdot CS(p_1,p_1^n) + 0.5 \cdot \pi(p_1,p_1^n,l_1,i_1) + \frac{0.95}{1-0.95} \cdot 0.5 \cdot (CS_2(l_2) + \pi_2(l_2))\right\}$$

$$s.t.\ \ l_2 = l_1 e^{Ul_1(1-l_1)} \tag{25}$$

The baseline for evaluating welfare changes is Scenario I with an underinvestment ratio corresponding to the estimated value of 0.41. Next period prices are calculated by maximizing per period surplus with respect to both residential and nonresidential prices. Optimal prices are independent of investments which can be considered as fixed costs, they only depend on the loss ratio influencing marginal costs. As second period values of cost, price and quantity demanded are all influenced by the loss ratio l_2, second period profit also depends on that measure. Efficient residential and nonresidential prices will be equal due to the fact that marginal cost of each is the same.

Prices in Scenario II are obtained in exactly the same way. This corresponds to the first-best case, where total welfare is maximized. However, marginal cost pricing does not lead to cost coverage, which is an objective also formulated in the previously mentioned Water Framework Directive of the European Union. It is proposed that all the costs connected to water supply should be covered by the revenue coming from the prices paid by customers. Under cost coverage now I understand an extended version of operating cost coverage, including also expenditures on infrastructural quality, as it is also suggested by the Water Framework Directive. Considering the different possible pricing schemes described in Chapter 2, marginal cost pricing still can be brought in line with cost coverage by using fixed fees. There are two problems with this solution. First, as it was mentioned in Chapter 3, it can be expected that consumers react to average prices at least to some extent. Hence introducing fixed price elements would lead to adjustments in the quantity demanded. Second, using high fixed fees with low variable prices imposes a disproportionately large burden on households with low consumption. As the papers (e.g. Nieswiadomy and Molina 1989) estimating a positive income elasticity show, that low income is typically coupled with lower consumption, this pricing structure would be especially disadvantageous for the poor. So this kind of policy can be regarded undesirable from the social point of view (on different pricing structures see Garcia and Reynaud 2004). Based on these arguments, in Scenario III I assume that cost coverage is achieved by using only variable prices, keeping the ratio of residential and nonresidential prices constant. In this case the chosen level of expenditures on infrastructural quality has an effect on both first and second period consumer surpluses besides profits.

In all three scenarios two sub-cases are examined. In the first, underinvestment ratio is assumed to be equal to the estimated value of 0.41, while in the second I used its optimal value. This optimum can be obtained by maximizing Equation (25) with respect to the value of i_1. The amount of infrastructural investments is the only choice variable, as prices are determined in each scenario by the different assumptions.

Results are presented in Table 3. Each sub-scenario is compared to the baseline, i.e. to Scenario I with an underinvestment ratio of 0.41, which corresponds to the current practice. The first column contains the underinvestment ratio of period one, while the second shows required percentage changes of first period residential price, compared to the baseline value of P_0. Changes of consumer surplus, profits and overall welfare are denoted by ΔCS, $\Delta\pi$ and ΔW respectively, and are expressed as a percentage of current total revenues of the firm. Time periods are indicated by lower indices. The last three columns contain the discounted present value of these measures, taking an infinite time horizon. Sub-scenarios 1 and 2 refer to current and optimal underinvestment ratios respectively.

Table 3: Results of the numerical example

	UI_1	ΔP_1	ΔCS_1	$\Delta\pi_1$	ΔW_1	ΔCS_2	$\Delta\pi_2$	ΔW_2	ΔCS	$\Delta\pi$	ΔW
Scenario II/1	0,41	-38,0%	48,1%	-41,4%	6,7%	0%	0%	0%	48,1%	-41,4%	6,7%
Scenario III/1	0,41	4,3%	-4,3%	3,6%	-0,7%	0%	0%	0%	-4,3%	3,6%	-0,7%
Scenario I/2	-1,79	0%	0%	-29,6%	-29,6%	9,5%	5,3%	14,8%	181,0%	70,4%	251,5%
Scenario II/2	-1,79	-38,0%	48,1%	-71,0%	-22,9%	9,5%	5,3%	14,8%	229,1%	29,0%	258,2%
Scenario III/2	-1,73	40,7%	-38,9%	3,6%	-35,2%	9,5%	5,3%	14,7%	141,6%	103,4%	245,0%

These numbers show the estimated effects of different price structures and investment decisions. A clear trade-off can be observed between present and future welfare as well as between consumer surplus and profits. In the first two rows only first period prices are changed, underinvestment ratio remains 0.41. It is not surprising, that efficient prices corresponding to marginal costs lead to an increase in welfare (Scenario II/1). As current revenue does not ensure cost coverage, prices need to be increased to reach this objective. Since marginal cost is estimated to be lower than current prices, this price increase is expected to decrease welfare, as it can be seen in Scenario III/1. However, there are substantial welfare gains in case of all pricing schemes, provided that the value of infrastructural investments is optimal. These gains are of a much higher size than that of efficient pricing alone. This result is caused by the long-lasting effect of infrastructural quality improvements. However, in each case there are welfare losses in the first period, which can be considered high as well, though future gains provide a generous compensation. It can be important to analyze this trade-off between current and future welfare in case of policy evaluations. Another notable result is the effect of investment financing decisions on first period consumer surplus and profit. The magnitude of first period losses is similar in case of all pricing schemes, though it is the highest at full cost coverage. In Scenario III/2 losses are born by the consumers, while in the other two cases they result in falling profits. Welfare gains of efficient pricing can only be reached by a huge decline in profits, corresponding to 71% of current total revenues. On the other hand full cost coverage results in high price increases and significant losses on the consumer side. The trade-off between declining consumer surplus and profit losses financed by taxpayers is also an interesting question when price schemes are evaluated. It is also worth emphasizing, that increasing the quality of infrastructure is highly desirable in any of the three scenarios. Moreover, infrastructural investments should be strongly increased in order to reach an optimal value. These values are close to each other under the different pricing methods, though optimal investments are a little lower in case of full cost coverage. The main message of this table lies in comparing the magnitude of welfare changes caused by different price schemes or different amount of infrastructural investments. The latter seems to have a much more considerable effect, though it creates current losses. Who bears these losses and how much they are is influenced by the chosen pricing method.

As a sensitivity test I also carried out the calculations using a much lower discount value of 0.6. This can be interpreted as a case, where agents have a shorter horizon, i.e. future values are taken into consideration with lower weight. Though total discounted welfare gains of sub-scenarios with optimal underinvestment rate are found to be lower due to the lower discount factor, they are still positive. Similarly, optimal value of investment into infrastructural quality is lower, but optimal rate of underinvestment remains positive. Exact numerical results can be found in Appendix B.

Similar calculations could also be helpful in policy analysis, provided that more reliable data are available concerning investment needs or nonresidential demand. However, a further limitation to these calculations is given by the fact that elasticity estimates provided in Chapter 5 are expected to be valid only in the short-run. Long-run elasticity is supposed to be higher in absolute value than the short-run, so estimating true long-run effects of policy changes would require long-run estimates of elasticities as well. Still, the previously obtained elasticity estimates can be used in welfare evaluations where the focus is on the near future.

7. Conclusions

In my thesis I examined the effects of pricing in the sector of public water supply. I presented estimates concerning price elasticity of demand and elasticity of variable costs with respect to output, using the panel data method of first differencing. Residential demand was found to be inelastic, with a price elasticity of -0.099. This value is in line with estimation results found in the literature, falling into the lower range of estimates. Considering the supply side I estimated a Cobb-Douglas variable cost function, which gave a value of 0.45 for the elasticity of costs with respect to output, and increasing returns to scale were found. Based on these estimates I proposed a model, which provides a framework for estimating dynamic welfare effects of pricing by including long-term effects of investments into infrastructural quality.

A contribution to existing research is twofold. First, estimations were done for the region of Central- and South-Eastern Europe together with post-Soviet countries, using most recent data available. Although there exists a broad literature on the topic of demand estimation concerning public water supply, to my knowledge there are no such estimates based on data from the entire region described above. Second, the suggested dynamic modeling framework including the effects of investment decisions may serve as a good starting point for further investigations.

The model proposed makes possible to evaluate trade-offs between short-run and long-run welfare effects of price and investment decisions. However, the interpretation of exact numerical results is strongly limited by the fact that important data were not available and I could only use rough measures instead. A possible direction of future research could be to find reliable measures for these data, including the quality maintaining amount of infrastructural investments compared to current practice and probably a better measure of infrastructural quality. In order to provide trustworthy estimates one would also need a better specified nonresidential demand function and law of motion for infrastructural quality, which was not possible here due to data constraints. If a more detailed database were available, the model together with the previous elasticity estimates could also be applied for policy purposes, as short and long-term effects of price and investment decisions can be analyzed with its help. It could also support decisions concerning the question how the burdens of fully cost covering pricing should be divided among the different customer groups. In the presence of more detailed data about investment effects and needs the optimal path of infrastructural investment decisions could also be elaborated. This would require additional enrichment of the model by including the effects of supply quality on the demand side.

A further extension could be the inclusion of wastewater services in the model. As these are strongly connected to water supply, spillover effects are expected to be between them. Moreover, utilities which provide both services can be analyzed as multi-product firms, where pricing decisions concerning the two services are linked.

The suggested framework could also be applied in case of other sectors, where a large infrastructure has considerable investment needs. Examples are network industries like electricity or gas supply, district heating and public transport. In the presence of an inelastic demand modeling results are expected to be similar to those obtained in case of public water supply.

Finally, a possible extension of theoretical interest could be to take into consideration the process of regulation including informational asymmetries in the proposed model. A situation could be modeled, where prices are set by the regulator, while investment decisions are made by the regulated firm.

Appendix A – Data used, variable definitions and descriptive statistics

Table 4: Ratio of total country population living in the service area of firms in the database per country

Albania	100%
Armenia	87%
Belarus	32%
Bosnia and Herzegovina	50%
Bulgaria	67%
Croatia	42%
Czech Republic	56%
Georgia	43%
Hungary	60%
Kazakhstan	39%
Kyrgyz Republic	12%
Macedonia, FYR	60%
Moldova	37%
Poland	95%
Romania	25%
Russia	40%
Slovakia	84%
Tajikistan	18%
Turkey	4%
Ukraine	18%
Uzbekistan	16%

Population data calculating coverage ratio are obtained from World Bank database[15]

[15] Population data available at the webpage of World Bank: http://data.worldbank.org/indicator/SP.POP.TOTL (accessed: May 11, 2010)

Table 5: Number of observations by year and country

	2000	2001	2002	2003	2004	2005	2006	2007	2008	total
Albania	0	0	0	0	0	30	64	55	55	204
Armenia	1	1	1	3	3	5	5	5	5	29
Belarus	8	9	10	11	12	13	22	0	0	85
Bosnia and Herzegovina	22	22	22	22	22	21	22	22	0	175
Bulgaria	0	0	0	0	20	20	20	20	20	100
Croatia	21	21	21	21	21	0	0	0	0	105
Czech Republic	20	20	20	20	20	20	0	0	0	120
Georgia	28	28	28	28	28	29	17	17	16	219
Hungary	23	23	23	23	23	24	20	20	0	179
Kazakhstan	9	11	14	15	17	19	22	24	0	131
Kyrgyz Republic	9	9	9	9	9	9	9	0	0	63
Macedonia, FYR	0	0	2	2	12	15	15	15	0	61
Moldova	42	42	42	42	42	41	41	41	39	372
Poland	0	0	0	36	36	36	36	36	0	180
Romania	25	25	25	29	29	25	25	25	0	208
Russia	88	88	81	86	86	86	84	84	80	763
Slovakia	0	0	0	2	4	5	6	7	0	24
Tajikistan	0	9	9	9	9	9	0	0	0	45
Turkey	0	0	0	0	20	20	20	20	20	100
Ukraine	82	83	24	24	24	16	16	16	0	285
Uzbekistan	1	1	1	3	3	5	5	3	0	22
total	379	392	332	385	440	448	449	410	235	3470

Table 6: Definitions and descriptive statistics of the variables used, without outliers

Name	Definition	Unit of measurement	Mean	Std. Dev.	Max	Min
CONSW	average per capita residential water consumption, by firm (yearly residental water consumption at the firm/ number of residential customers)	m3/capita/year	59.50	33.6917	295.17	3.54
APW	average price of water for residential customers, real value, converted to 1995 USD, using PPP rates, by firm	2000 USD/m3	0.60	0.4339	2.67	0.0003
MPW	estimated marginal price of water, real value, converted to 1995 USD, using PPP rates, by firm	2000 USD/m3	0.58	0.4249	2.67	0.0003
APWW	average price of wastewater for residential customers, real value, converted to 1995 USD, using PPP rates, by firm	2000 USD/m3	0.59	0.9062	17.37	0.0005
GNI	real GNI per capita using purchasing power parity rates when converted to 1995 USD, by country	2000 USD/capita	9164.2	3565.88	17150.2	1321.118
RAINY	average daily rainfall per year, by country (total yearly amount of rain/ days in a year)	0.1 mm/day	15.43	7.0869	31.2	0.75
ATEMPY	average daily temperature in a year, by country	°C	9.38	3.0907	16.28	4.99
DRAIN13	number of rainy days in a year with precipitation above 1.3 mm, by country	#	74.48	31.1842	121	3
RWW	ratio of population with sewerage service to the population with water service	proportion	0.78	0.2593	2	0
RUR	dummy, being one if service area is defined as either 'rural' or 'rural and urban', by firm	binary	0.4	0.4906		
DWANDWW	dummy, being one if the ratio of population with sewerage service to the population with water service is positive and not greater than one	binary	0.88	0.3205		
DFPW	dummy, being one if a monthly fix charge for residential consumers of water is indicated, by firm	binary	0.23	0.4206		
Q	total volume of water produced	Million m3/year	30.62	59.6002	557	0.01
TCW	total operational expenses connected to water service, without depreciation, interest and capital repayments	Million 2000 USD/year	12.74	18.2002	172.83	0.04
LPRICE	all labor related costs (with social benefits) per full time equivalent staff per year	Thousand 2000 USD/capita/year	10.02	12.8439	351.95	0.01
EPRICE	estimated amount of electrical costs, related to water supply, compared to total volume of water produced	2000 USD/ 1000 m3	69.13	56.7884	561.18	0.02
SPRICE	estimated amount of contracted out service costs, related to water supply, compared to total volume of water produced	2000 USD/ 1000 m3	52.64	101.5503	1167.6	0
OPRICE	estimated amount of all other cost, related to water supply, compared to total volume of water produced	2000 USD/ 1000 m3	189.03	183.5694	1306.117	0
NETW	length of water distribution network	km	565.08	870.7435	9976	2
FA	gross book value of fixed assets, related to water supply, including work in progress	Million 2000 USD	34.26	71.6797	755.68	0.0005
CUST	population served with water	1000 capita	184.55	275.554	1785.9	1
LOSS	difference of water produced and sold, as a ratio of the amount of water produced	proportion	0.34	0.1974	0.97	0.0005

The omitted outliers are: CONSW>300 and APW>2.7 in the demand estimation and Q>600 in the cost estimation

The calculation of the marginal price was as follows: As only residential fixed prices and total number of connections (i.e. residential and nonresidential together) were given, I had to assume that the ratio of revenue from fixed price is the same in the two consumer groups.

Consequently the product of residential fixed tariff and total number of connections was deducted from the total water revenue. This measure divided by total water revenue gave the ratio of variable price in revenue, which allowed calculating the estimated residential water revenue coming from variable price. Dividing this by the volume of residential water consumption gave the estimated marginal price.

Appendix B – Calibration and further results of the numerical example

Table 7: Calibration used for the numerical example

Variable	Value	Unit of measurement
average price of water for residential customers	0.53	USD/m3
per capita water consumption of residential customers	71.98	m3/capita
water loss ratio	0.29	proportion
number of residential customers	561.51	1000 capita
volume of nonresidential consumption	27.63	million m3
average price of water for nonresidential customers	0.67	USD/m3
ratio of operating profit over total water revenue	0.04	proportion
ratio of operating profit net infrastructural investments over total water revenue	-0.04	proportion

Table 8: Results of the numerical example with the modified value of the discount factor: β=0.6

	UI_1	ΔP_1	ΔCS_1	$\Delta \pi_1$	ΔW_1	ΔCS_2	$\Delta \pi_2$	ΔW_2	ΔCS	$\Delta \pi$	ΔW
Scenario II/1	0,41	-38,0%	48,1%	-41,4%	6,7%	0%	0%	0%	48,1%	-41,4%	6,7%
Scenario III/1	0,41	4,3%	-4,3%	3,6%	-0,7%	0%	0%	0%	-4,3%	3,6%	-0,7%
Scenario I/2	-0,23	0%	0%	-8,6%	-8,6%	5,7%	3,2%	8,9%	109,0%	52,0%	4,8%
Scenario II/2	-0,23	-38,0%	48,1%	-50,0%	-1,9%	5,7%	3,2%	8,9%	157,1%	10,6%	11,5%
Scenario III/2	-0,23	14,9%	-14,6%	3,6%	-10,9%	5,7%	3,2%	8,9%	94,4%	64,3%	5,1%

Bibliography

Antonioli, Barbara, and Filippini, Massimo. 2001. "The use of a variable cost function in the regulation of the Italian water industry" *Utilities Policy,* 10(3-4): 181-187

Arbués, Fernando, García-Valiñas, María Ángeles, and Martínez-Espiñeira, Roberto. 2003. "Estimation of residential water demand: a state-of-the-art review" *Journal of Socio-Economics*, 32: 81-102

Arbués, Fernando, Barberán, Ramón, and Villanúa, Immaculada. 2004. "Price impact on urban residential water demand: A dynamic panel data approach" *Water Resources Research*, 40, W11402

Billings, R. Bruce, and Agthe, Donald E. 1980. "Price Elasticities for Water: A Case of Increasing Block Rates" *Land Economics*, 56(1): 73-84

Bisztray, Márta, Kis, András, Muraközy, Balázs, and Ungvári, Gábor. 2009. "Statistical Analysis of the Performance Data of Central and Eastern European Water Utilities" Regional Center for Energy Policy Research Working Paper

Bobadilla, Silvia Aguilar, and Camrova, Lenka. 2006. "Public water supply systems and sewerages in the Czech Republic: A Comparative Analysis" IREAS, Institute for Structural Policy

Buehler, Stefan, Schmutzler, Armin, and Benz, Men-Andri. 2004. "Infrastructure quality in deregulated industries: is there an underinvestment problem?" *International Journal of Industrial Organization*, 22: 253-267

Filippini, Massimo, Hrovatin, Nevenka, and Zorić, Jelena. 2008. "Cost efficiency of Slovenian water distribution utilities: an application of stochastic frontier methods" *Journal of Productivity Analysis*, 29(2): 169-182

Garcia, Serge, and Reynaud, Arnaud. 2004. "Estimating the benefits of efficient water pricing in France" *Resource and Energy Economics*, 26: 1-25

Garcia, Serge, and Thomas, Alban. 2001. "The Structure of Municipal Water Supply Costs: Application to a Panel of French Local Communities" *Journal of Productivity Analysis,* 16: 5-29

Garcia, Serge, and Thomas, Alban. 2003. "Regulation of Public Utilities under Asymmetric Information" *Environmental and Resource Economics,* 26: 145-162

García-Valiñas, Maria A. 2005. "Efficiency and Equity in Natural Resources Pricing: A Proposal for Urban Water Distribution Service" *Environmental and Resource Economics*, 32: 183-204

Kopańska, Agnieszka. 2009. "Public private partnership in water supply and sewage services in Poland. Analysis of possible application" Paper presented at the 17th NISPAcee Annual Conference, Budva, Montenegro

Maidment, DR, and Miaou, S-P. 1986. “Daily water use in nine cities” *Water Resources Research,* 22(6): 845-851

Martinez-Espiñeira, R. 2002. “Residential Water Demand in the Northwest of Spain” *Environmental and Resource Economics,* 21: 161-187

Montginoul, Marielle. 2007. “Analysing the Diversity of Water Pricing Structures: The case of France” *Water Resources Management,* 21(5): 861-871

Nauges, Céline, and van den Berg, Caroline. 2007. “How “natural” are natural monopolies in the water supply and sewerage sector? Case studies from developing and transition economies” World Bank Policy Research Working Paper 4137

Nauges, Céline, and Thomas, Alban. 2000. “Privately Operated Water Utilities, Municipal Price Negotiation, and Estimation of Residential Water Demand: The Case of France” *Land Economics*, 76(1): 68-85

Nieswiadomy, Michael L. and Molina, David J. 1989. “Comparing Residential Water Demand Estimates under Decreasing and Increasing Block Rates Using Household Data” *Land Economics,* 65(3): 280-289

Nordin, John A. 1976. “A Proposed Modification of Taylor's Demand Analysis: Comment” *The Bell Journal of Economics,* 7(2): 719-721

Renzetti, Steven. 1999. „Municipal water supply and sewage treatment: costs, prices, and distortions” *The Canadian Journal of Economics*, 32(3): 688-704

Rioja, Felix K. 2003. “The Penalties of Inefficient Infrastructure” *Review of Development Economics*, 7(1): 127-137

Shin, Jeong-Shik. 1985. „Perception of Price When Price Information Is Costly: Evidence from Residential Electricity Demand” *The Review of Economics and Statistics*, 67(4): 591-598

Taylor, Lester D. 1975. „The Demand for Electricity: A Survey” *The Bell Journal of Economics,* 6(1): 74-110

Taylor, R. G., McKean, John R., and Young, Robert A. 2004. „Alternate Price Specifications for Estimating Residential Water Demand with Fixed Fees” *Land Economics*, 80(3): 463-475

Timmins, Christopher. 2002. “Measuring the Dynamic Efficiency Costs of Regulators’ Preferences: Municipal Water Utilities in the Arid West” *Econometrica*, 70(2): 603-629

Viscusi, Kip, Vernon, James, and Harrington, James. 1995. *The Economics of Regulation and Antitrust*. Cambridge: MIT Press

Worthington, Andrew C., and Hoffman, Mark. 2006. “A state of the art review of residential water demand modelling” University of Wollongong Working Papers Series 06/27