



4.1 Historical Review of the Water Pricing Debate

In the face of the deterioration of water availability in different regions of the world, the discussion on the introduction of economic policy instruments has become increasingly important (Hanemann 2004). The issue of adequate water supply for all was first addressed at the UN Water Conference in Mar del Plata (Argentina) in 1977. This convention resulted in the United Nations' commitment to a human right to drinking water in a quantity and quality appropriate to basic needs. The conference elaborated an action plan which clarified the link between water management measures and their socio-economic impact. This includes, among other things, the demand to reflect economic costs through the water price. Furthermore, economic incentives for an efficient and balanced use of water via the water price were declared to be useful. However, there were no explicit recommendations for the use of concrete instruments.

Another milestone in water policy was the 1992 International Conference on Water and Environment in Dublin, Ireland. The conference culminated in the formulation of the four important Dublin principles, which set out the conditions for sustainable water resources management. Principle 4 declares the economic value of water. Since water that was used for one process might be unsuitable to be used for other processes, a competition between different forms of use arises. Hence, water should be considered as an economic good, implying some form of cost coverage for water supply. The Dublin Conference was instrumental in a substantive and institutional reorientation of global water policy.

The outcome of the Dublin Conference formed the basis of Chap. 18 ("Water Management") of the Agenda 21, adopted at the 1992 UN Conference on Environment and Development in Rio de Janeiro. Representatives from 178 countries took part in the conference to discuss key environmental and development policy issues of the twenty-first century. The cost recovery principle was anchored in Agenda 21 as

a component of sustainable water resources management. In addition to production costs, external environmental costs must also be taken into account.¹

The Agenda therefore calls for tariff systems that take the actual costs of water as well as the consumer's assumed ability to pay into account. The inclusion of social concerns in the pricing of water is therefore explicitly required. While the Dublin Conference's outcome essentially considered water to be an economic good, Agenda 21 also considered water as a social good. These different views form the basis for the subsequent discourse on water pricing policy. However, it can be noted that the content of both the Dublin Principles and Agenda 21 strongly influenced the subsequent water policy (Dinar et al. 2015). The European Water Framework Directive provides an example of this.

The European Water Framework Directive, which was adopted in 2000, provides the legal basis for securing water resources and ensuring sustainable development within the European Union. The directive was adopted as a reaction to an increasing disparity between the available water supply and water demand. The principle of cost recovery referred to in Article 9 is an important part of the directive. It aims to cover the production costs, as well as the environmental and resource costs associated with the use of water resources. Furthermore, under this directive, pricing policy must be designed in a way that incentivizes an efficient water usage.

In the General Assembly of the United Nations on August 3, 2010, the members decided that the right to clean drinking water and sanitation should be a human right (United Nations 2010). This human right is in accordance with the content of Agenda 21 and the Dublin Principles, as it does not require free water and sanitation. It rather assures affordable access to adequate water and sanitation to satisfy the basic needs. The resulting challenge is to determine the extent to which poorer sections of the population can be involved in cost recovery to ensure the affordability of water supplies.

4.2 Criteria for Water Tariffs

Water pricing policy can pursue multiple objectives, which are revenue sufficiency, economic efficiency, environmental sustainability, and social concerns, including affordability and fairness considerations. Further important aspects of water pricing policy, which are not addressed in detail in this section, are the public and political acceptance as well as the simplicity and transparency of the water pricing policy (Boland and Whittington 2000a). In the following section, we give a brief overview of the four main goals that are also described as the sustainability dimensions of water pricing policy (Massarutto 2007b).

¹Chapter 18.16 of Agenda 21 reads: "A prerequisite for the sustainable management of water as a scarce vulnerable resource is the obligation to acknowledge in all planning and development its full costs".

4.2.1 Revenue Sufficiency

This goal is of special importance for the water suppliers because it relates to the claim that a tariff system should cover all the incurred costs. If costs are not fully covered, incoming cash flows are not sufficient to guarantee an effective and efficient operation and management of the water supply system. Furthermore, the absence of full cost recovery could result in a lack of financial resources, which would be necessary to make sufficient investments in the water supply infrastructure. This leads, in consequence, to a worse water supply service and hence an increasing dissatisfaction, which is accompanied by a decreasing willingness to pay by consumers. Hence, this goal is important for guaranteeing the long-term reproduction of the physical assets. Not only the pricing level, but also the stability is a matter of the tariff-setting process.

The full supply costs that have to be covered are those associated with providing water services to users. This contains the following types of cost (OECD 2010):

- Operation and maintenance costs, resulting from the day-to-day operations of the water supply system, such as electricity for pumping but also labor and repair costs.
- Capital costs, covering both, investments in existing infrastructure as well as capital for new investments, and servicing debt.

However, further cost components, such as opportunity cost and economic externalities, should also be addressed. In detail, these components include

- Opportunity cost that reflects the scarcity value of the water resource. They refer to the cost of not serving the next possible user.
- Economic externalities, which are benefits and costs associated with water management. It is possible to distinguish between positive external benefits (e.g., groundwater recharge benefits from irrigation or water reuse) and negative external costs (e.g., upstream diversion of water or the release of pollutants downstream within an irrigation or urban water system).

The cost components are illustrated in Fig. 4.1. The sum of the full supply cost, opportunity cost, and economic externalities is termed the full economic cost of the water supply service (OECD 2010). Furthermore, the operation and management of the water supply system could negatively impact the aquatic and non-aquatic environment, for instance, an increased water shortage in the ecosystem due to an over-exploitation of water resources. Those occurring negative externalities to the environment have to be addressed as another important cost component. The sum of the full economic cost and the environmental externalities forms the full cost of water supply service (OECD 2010).

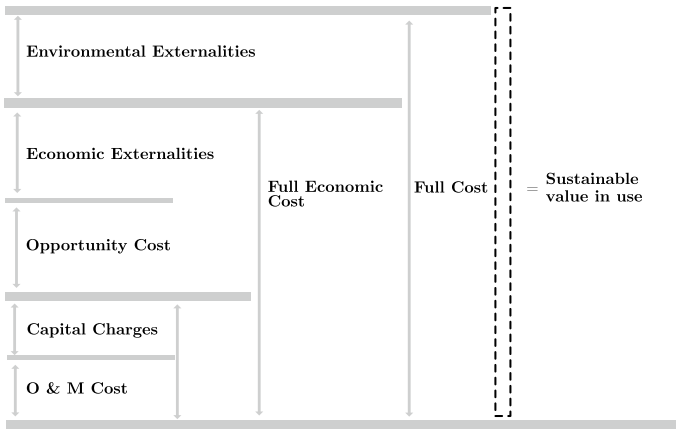


Fig. 4.1 General principles for the costs of water. *Source* Rogers et al. (1998)

4.2.2 Economic Efficiency

Water pricing policy should be conceived in a way that water is allocated to those users that benefit the most from receiving water resources. Hence, this goal implies the maximization of the aggregated economic rents of all water consumers. If water is allocated to users with low marginal benefits of water consumption, while other users with higher marginal benefits are not supplied, the principle of economic efficiency is violated. Furthermore, the pricing policy should disincentivize the wasteful usage of economic resources, because marginal benefits exceed the marginal cost of each unit consumed water.

4.2.3 Environmental Sustainability

The water resource in the environment is essential for the aquatic and non-aquatic nature and provides important ecosystem services (e.g., fishery) for the human society. As water resource conservation plays a crucial role in achieving environmental sustainability, the pricing policy should set incentives to protect the water in the nature. For example, over-exploitation from surface or subsurface water stocks has to be avoided.

4.2.4 Social Concerns

Acceptable levels of the water supply service should be accessible and affordable to all consumers, because water can be seen as a good of public interest. The focus of this goal is mainly the protection of vulnerable groups with low incomes. The reallocation of costs across different groups through the tariff structure is an important means to achieve this objective (OECD 2009).

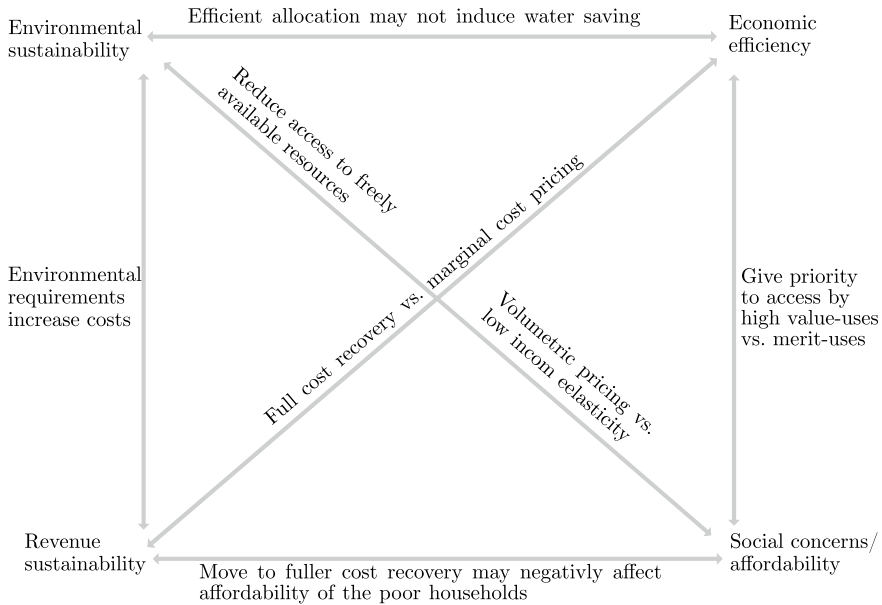


Fig. 4.2 Relation between goals of water pricing policy. *Source* Massarutto (2007a)

This means that, regardless of the budget, everybody should have access to the subsistence level. World Health Organization (2012) states the short-term subsistence level with 20 liter per capita and day. To evaluate the affordability of water supply, defining service indicators (e.g., the relation between expenditures for the water supply service and the overall budget) and determining threshold levels for these indicators are common procedures. Usually, the expenditures for the water service should not exceed three to five percent of the household income (OECD 2010; Walker 2009).

These four goals of water policy are related to one another in different ways. On the one hand, some goals can be achieved in accordance with another goal. For instance, an increase in the volumetric price may not only lead to higher revenues for the water supplier, which supports the revenue sufficiency goal, but it might also result in less exploited water resources due to a reduced water demand, which facilitates achieving the environmental sustainability goal.

On the other hand, one has to consider the trade-offs between the four objectives. Figure 4.2 illustrates these trade-offs. One important trade-off exists, for instance, between the goals of revenue sufficiency and social concerns. If water access has to be guaranteed to all consumers at low prices or maybe even for free in order to meet the affordability requirement, the revenues generated by the water supplier may not be sufficient to cover the full supply cost. Similarly, achieving the goal of economic efficiency counteracts revenue sufficiency. A pricing policy relying on the marginal cost usually constitutes the first best pricing policy because it maximizes economic efficiency. However, due to the crucial importance of fixed cost in water supply, average cost is higher than marginal cost, which is the main reason why water

supply is considered a natural monopoly in most cases. If the price is equal to the marginal cost but lower than the average cost, the break-even is not reached and the water supplier faces revenue deficits. There exist a variety of examples for further trade-offs between the four main goals of water pricing policy.

4.3 Water Tariff Design

4.3.1 Tariff Structures

4.3.1.1 Overview

Various forms of water tariff systems are conceivable and implemented in practice. These tariffs generate revenues for the water suppliers and can consist of various components. The most important components applied in practice are (OECD 2010)

- A one-time **connection fee**, to gain access to the service.
- A recurrent **fixed charge** (sometimes known as a standing charge or flat fee) that can be uniform across customers or linked to the customer's characteristic (e.g., size of supply pipe, meter flow capacity, property value, or number of water-using appliances).
- If a metering system is in place, a **volumetric rate**, which, when multiplied by the volume of water consumed in a charging period, gives rise to the volumetric charge for that period.
- In some circumstances, a **minimum charge** is paid for each period, regardless of consumption.

Based on the composition of these four tariff components, various tariff structures can be implemented, which yield different expenditure functions, $R(w)$, for each tariff structure. The expenditure function describes the payments, symbolized by R , of a water-consuming household to the water supplier depending on the consumption level of the household, w .² Based on the expenditure function, the average expenditure function, $AR(w)$, and the marginal expenditure function, $MR(w)$, can be derived. The average expenditures are the average payments of a household per unit of water consumed (usually measured in cubic meter), while the marginal expenditure represents the payments of the household for the consumption of one additional unit (cubic meter) of water. The average and marginal expenditure function can be calculated by the following algebraic relations: $AR(w) = \frac{R(w)}{w}$ and $MR(w) = \frac{\partial R(w)}{\partial w}$. The most common tariff structures and their expenditure, average expenditure, and marginal expenditure functions are listed in the following.

²The expenditure incurred by the household is equivalent to the revenue obtained by the water supplier; hence, the expenditures are denoted by R .

4.3.1.2 Flat Rates

If there is no water meter available to measure the water consumption, a flat rate is usually the only feasible tariff structure. The customers pay a rate regardless of their consumption. This rate can be uniform, or differentiated with respect to the customer's characteristics, e.g., the rateable value of the property being served. The expenditure, average expenditure, and marginal expenditure functions are

$$\begin{aligned}R(w) &= L \\AR(w) &= \frac{L}{w} \\MR(w) &= 0\end{aligned}$$

where L represents the flat rate (lump sum).

4.3.1.3 Single Volumetric Rates

A single volume rate per consumed amount (e.g. cubic meter) of water is charged. The level of the volume rate does not change with consumption as it is independent of the consumption level. In addition to this single volumetric rate, a recurrent fixed charge (base price) might exist, which represents a payment to the water supplier regardless of the customer's water consumption level. The expenditure, average expenditure, and marginal expenditure functions are

$$\begin{aligned}R(w) &= L + p \cdot Q \\AR(w) &= p + \frac{L}{w} \\MR(w) &= p\end{aligned}$$

where L represents the base price and p stands for the volume price. The flat rate is a special form of this kind of tariff with $p = 0$.

4.3.1.4 Block Rates

Zone Block Rates

The volumetric charge is adjusted step-wise with increasing volumes of water consumed. In the case of increasing block rates, the volume rate rises with successively higher consumption blocks; for decreasing block rates, volume rates decline with higher consumption blocks. In addition, a recurrent fixed charge (base price) may exist in this form of tariff structure. Furthermore, it is possible to differentiate between zone tariffs and relay tariffs, while the former is the most commonly applied form of block rates. Under a zone tariff, the consumers pay the volume price of the respective block for each unit (e.g., cubic meter) of the quantity consumed. Under a relay tariff, the volume price of the highest consumption block has to be paid for the whole quantity consumed.

Given a block rate with N blocks in total, every block is separated by threshold values, denoted by $q_1, q_2, \dots, q_i, \dots, q_N$, with $q_1 < q_2 < \dots < q_i < \dots < q_N$. The i th block is defined within the range $[q_{i-1}, q_i]$ and the volume price in this block is p_i . The relation $p_1 < p_2 < \dots < p_i < \dots < p_N$ is valid for increasing block rates, while the contrary situation with $p_1 > p_2 > \dots > p_i > \dots > p_N$ occurs for decreasing block rates.

Based on the presented general case with an arbitrary number of blocks, a tariff with three blocks is specified by the two threshold consumption levels q_1 and q_2 . The first block is defined for the interval $[0, q_1]$ and the relevant volume price in this first block is p_1 , while the second block which is defined within the range of both threshold levels $[q_1, q_2]$ is characterized by the volume price p_2 . Finally, the third block is relevant for a consumption which exceeds the second threshold level. The volume price of this third block is p_3 . Based on this specified tariff with three blocks and an observed consumption level of w , the following total, average, and marginal expenditure functions could be set up for a zone block tariff:

$$R(Q) = \begin{cases} L + p_1 \cdot Q & \text{for } w \leq q_1 \\ L + p_1 \cdot q_1 + p_2 \cdot (w - q_1) & \text{for } q_1 < w \leq q_2 \\ L + p_1 \cdot q_1 + p_2 \cdot (q_2 - q_1) + p_3 \cdot (w - q_2) & \text{for } w > q_2 \end{cases}$$

$$AR(w) = \begin{cases} p_1 + \frac{L}{w} & \text{for } w \leq q_1 \\ p_2 + \frac{(p_1 - p_2) \cdot q_1 + L}{w} & \text{for } q_1 < w \leq q_2 \\ p_3 + \frac{(p_1 \cdot q_1 + (p_2 - p_3) \cdot q_2 + L)}{w} & \text{for } w > q_2 \end{cases}$$

$$MR(Q) = \begin{cases} p_1 & \text{for } w \leq q_1 \\ p_2 & \text{for } q_1 < w \leq q_2 \\ p_3 & \text{for } w > q_2 \end{cases}$$

Relay Block Rates

The variable L represents the base price whose level is independent of the consumption level in this presented example. Regardless of the total observed consumption level w , the consumption within the first, second, and third blocks is priced with p_1 , p_2 , and p_3 , respectively, for a zone block tariff. However, in contrast to a zone block rate, a relay block rate is characterized by the fact that the entire consumption is priced with p_1 or p_2 or p_3 if the entire consumption level is within the first, $w \leq q_1$, or second, $w = [q_1, q_2]$, or third block, $w > q_2$, respectively. Therefore, the expenditure for a relay block rate differs from the expenditure for a zone block rate, if the total consumption level exceeds the first block. The total, average, and marginal expenditure functions of a relay block rate with three blocks have the following form:

$$R(w) = \begin{cases} L + p_1 \cdot w & \text{for } w \leq q_1 \\ L + p_2 \cdot w & \text{for } q_1 < w \leq q_2 \\ L + p_3 \cdot w & \text{for } w > q_2 \end{cases}$$

$$AR(w) = \begin{cases} p_1 + \frac{L}{w} & \text{for } w \leq q_1 \\ p_2 + \frac{L}{w} & \text{for } q_1 < w \leq q_2 \\ p_3 + \frac{L}{w} & \text{for } w > q_2 \end{cases}$$

$$MR(w) = \begin{cases} p_1 & \text{for } w \leq q_1 \\ p_2 & \text{for } q_1 < w \leq q_2 \\ p_3 & \text{for } w > q_2 \end{cases}$$

4.3.1.5 Adjusted Block Rates

This tariff structure is quite similar to block rates, but, in contrast to conventional block rates, adjusted block rates feature volumetric rates or block sizes that are adjusted depending on the consumer's characteristics (e.g., income or household size).

4.3.2 Price Discrimination

The distinction of consumer prices with respect to the individual consumer's characteristics is termed as price discrimination. Different forms of price discrimination can be distinguished (Varian and Repcheck 2010).

4.3.2.1 First-Degree Price Discrimination

Under first-degree price discrimination, prices for each unit of a good are set such that the price charged for each unit is equal to the consumer's willingness to pay for that unit. Therefore, the consumer's surplus will be skimmed off fully by the producer. This approach is also known as perfect price discrimination. Due to information asymmetries between the players in the market (e.g., the supplier does not know the willingness to pay of each consumer), perfect price discrimination is generally not applicable in practice.

4.3.2.2 Second-Degree Price Discrimination

For second-degree price discrimination, the price depends on the bought amounts. Block rates are a common example of this form of price discrimination, as the price schedule involves different prices for different amounts of water sold.

Optional tariffs are another example of second-degree price discrimination. Here, consumers can choose from different tariff options offered (e.g., tariff options that differ in their single volumetric rates and fixed charges). The consumer's decision on the choice of tariff depends on the consumer's consumption level. The lower the consumption level, the higher the preference for choosing a tariff with a higher volume rate and lower base price.

4.3.2.3 Third-Degree Price Discrimination

Different consumers are charged with different prices based on their individual characteristics. This is a form of price discrimination commonly practiced for a multitude of products and services, e.g., discounts for students or social welfare recipients. The adjusted block rate is an application of this form of price discrimination. The price adjustment should usually be determined in the way that poor households or especially large households receive a financial relief.

A further example of third-degree price discrimination is social tariffs, under which low-income households get a discount on the volumetric rate or fixed charge they have to pay. Therefore, they get a financial relief on their water bill, relative to other non-low-income households.

4.3.2.4 Spatial/Regional Price Discrimination

Spatial price discrimination occurs when prices depend on the location of the consumer. Under this form of price discrimination, the water supplier can establish various price zones within its water supply area. Regional price discrimination is applicable if the pricing regime is based on the cost-by-cause principle. A consumer that is further away from the waterworks may cause higher water distribution network cost than a consumer closer to the water treatment facility. Similarly, water delivery to consumers located on a hill or on top of a mountain is accompanied by higher pumping costs than the supply of consumers living in a valley. Therefore, consumers which are located further away from the waterworks or that are located on a mountain have to pay a higher volume price.

4.3.2.5 Temporal/Seasonal Price Discrimination

The price for water differs depending on the time point (or season) of consumption. Based on this approach, seasonal pricing schemes can be implemented. Temporal price discrimination can be useful due to a multitude of reasons, for instance, the costs for water supply in summer months can be higher than in the winter months, because of lower groundwater tables in summer. Another reason can be higher water requirements in summer months than in the residual year, especially in regions with a high share of agricultural water demand. Capital-intensive pumping equipment may be just exhausted in the summer periods, where irrigation is intensified due to the growth of the agricultural plants. Therefore, if the tariff system based on the cost-by-cause principle is implemented, the volume prices in the summer month would be higher than in the residual months of the year.

4.3.3 Two-Part Tariff Versus One-Part Tariff

Single two-part tariffs are characterized by a single volumetric rate, i.e., the price per volume of water consumed, and a recurrent fixed rate, which is independent of the consumption level usually paid monthly or yearly. The single one-part tariff is a special form of a single two-part tariff, because just one tariff component is relevant

in this kind of tariff. Either the one-part tariff is characterized by a single volumetric rate without a recurrent fixed rate (volumetric tariff) or by a recurrent fixed rate without a volumetric price (flat rate).

4.3.3.1 Two-Part Tariff

Assume that a representative consumer group has to be served with water by a supplier. The situation in the market should be cleared, which means that the volumes demanded by the consumers equal the water volumes offered by the supplier. The water supply to the consumer group causes costs, $C(w)$, where w stands for the level of consumption and supply. The average costs can be expressed as $AC(w) = \frac{C(w)}{w}$ and the marginal costs are $MC(w) = C'(w)$. We want to design a tariff that maximizes the total surplus. The total surplus is equal to the difference between the generated benefit due to consumption, $B(w)$, and the costs for supplying the consumer group, $C(w)$.

From household theory, we know that the consumers want to maximize their surplus, which is equivalent to the difference between the benefits from consumption $B(w)$ and the expenditures for water delivery which is $p \cdot w - L$. Hence, a consumer solves the following optimization problem:

$$\max_{\{w\}} [B(w) - p \cdot w - L]$$

From the consumer's perspective, water consumption, denoted by w , is the sole decision variable. By solving this optimization problem, we get $p = B'(w)$; hence, we can express water consumption in terms of volumetric prices, i.e., $w(p)$. The inverse form of $w(p)$ which is $p(w)$ is nothing else than the demand function which is determined by the marginal benefit $p(w) = B'(w)$.

If we furthermore assume that in a water supply system, total surplus should be maximized and water delivery is not restricted by any capacity or scarcity constraints, it is also possible to derive that the price is equal to the marginal cost level, $p = C'(w)$. This results from the fact that in a situation where total surplus is maximized, the demand function which is determined by the marginal benefit function should, according to household theory, be equal to the marginal cost.

The water supplier should be profitable and should generate enough revenue to cover its costs (financial sustainability goal of pricing policy). However, if we restrict the supplier (e.g., by regulation) in the way that the supplier cannot make any profits, the revenues must be equal to the total costs, which means $C(w) = L + p \cdot w$. Based on this assumption, it is possible to find the optimal recurrent fixed charge, L^* :

$$L^* = C(w^*) - p \cdot w^* \quad (4.1)$$

Based on the optimal fixed charge in Eq. (4.1), the average lump sum per consumed quantity, denoted by AL^* , can be calculated as the difference between the average costs and the price, as illustrated by Eq. (4.2).

$$AL^* = \frac{L^*}{w^*} = \frac{C(w^*) - p \cdot w^*}{w^*} = AC(w^*) - p \quad (4.2)$$

Water supply is usually a natural monopoly due to a high proportion of fixed costs to total costs. Hence, the average cost function is decreasing at the optimal consumption level w^* , i.e., $AC'(w^*) < 0$. In this case, the marginal costs are below the average cost level in the optimum, such that $C'(w^*) < AC(w^*)$ holds.³ Furthermore, we know that $p = C'(w^*)$, and hence $p < AC(w^*)$. Therefore, it follows that the average lump sum is positive, i.e., $AL^* > 0$, which also means that the fixed charge must be positive, i.e., $L^* > 0$.⁴

4.3.3.2 One-Part Tariff with Single Volumetric Rate

If, instead of a single two-part tariff, a single one-part tariff with just a volumetric rate is implemented, the costs can only be covered by revenues from the single volumetric price. As the water supplier is assumed to be regulated, the supplier should not make any profits; hence, revenues should be equal to cost, which yields the following condition:

$$C(w) = p \cdot w \quad (4.3)$$

From Eq. (4.3) follows that the volumetric price should cover the average cost:

$$p = \frac{C(w)}{w} = AC(w) \quad (4.4)$$

In the case of a single one-part tariff with just a volumetric rate, the consumer solves the following optimization problem:

$$\max_{\{w\}} [B(w) - p \cdot w] \quad (4.5)$$

The solution is $p = B'(w)$. Because of this optimality condition, the price determines the quantity level of consumption. The marginal benefit function $B'(w)$ is nothing else than the demand function. Therefore, similar to the two-part tariff, the quantity level w^V is determined by the price p^V , hence, $w^V(p^V)$.

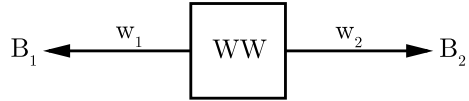
$$p^V = AC(w^V) = B'(w^V) \quad (4.6)$$

The optimal price and quantity, p^V and w^V , result from the intersection point of the average cost function and the demand function.

³We know that $AC(w) = \frac{C(w)}{w}$. Because of $AC'(w) < 0$, it is possible to write $AC'(w) = \frac{\partial \frac{C(w)}{w}}{\partial w} < 0$. Solving $\frac{\partial \frac{C(w)}{w}}{\partial w} < 0$, we get the following result: $\frac{C'(w) \cdot w - C(w)}{w^2} < 0$. This is $\frac{C'(w)}{w} - \frac{C(w)}{w^2} < 0$ and hence $C'(w) < \frac{C(w)}{w}$, which is $C'(w) < AC(w)$.

⁴ $L^* = \underbrace{AL^*}_{>0} \cdot \underbrace{w^*}_{\geq 0} > 0$.

Fig. 4.3 Universal service provider: The basic setup.
Source own illustration



4.3.3.3 Flat Rate

However, if a flat rate is implemented, only the recurrent base price, L , has to be paid for any amount of water consumption, as the volumetric price is zero, $p = 0$. For this case, the consumers solve the following optimization problem:

$$\max_{\{w\}} [B(w) - L] \tag{4.7}$$

the result of which is $B'(w) = 0$. Therefore, the consumption level w^L is determined by the maximum demand. Similar to the pricing regimes analyzed previously, revenues should cover the cost, while profits should not be generated. Hence, the revenues arising from the flat rate have to be set equal to the costs, which implies

$$L^L = C(w^L) \tag{4.8}$$

Based on the fixed rate, the average fixed rate per amount of water consumed is equal to the average cost level:

$$AL^L = \frac{C(w^L)}{w^L} = AC(w^L) \tag{4.9}$$

4.3.4 Universal Service Provider

4.3.4.1 Two Consumer Groups

The universal service provider is a service operator that offers infrastructure services such as water supply at uniform and affordable conditions. These principles are enforced by appropriate price regulation, either by the provider being a public enterprise or by a private operator being regulated by a price regulator. The concept of a universal service provider is the most common form in practice, because price discrimination based on the cost-by-cause principle is often not enforceable, whether due to political, social, economical, or fairness reasons.⁵

Figure 4.3 illustrates an exemplary situation of two consumer groups which are served with water through one water supply system. Due to delivery and consumption of water, w_1 and w_2 , both consumer groups obtain benefits, which are represented by $B_1(w_1)$ and $B_2(w_2)$. The service provision to the consumers causes costs: On the one hand, there are cost components that are caused by both consumer

⁵See the survey in Cremer et al. (2001).

groups, e.g., treatment costs in the waterworks, which are symbolized by the variable $C_{12}(w_{12}(w_1, w_2))$. The variable w_{12} represents the total amount of the consumed water level, hence:

$$w_{12}(w_1, w_2) = w_1 + w_2 \quad (4.10)$$

On the other hand, there are also cost components that are incurred by only one consumer group. The specific cost of consumer group 1, $C_1(w_1)$, depends solely on the amount of water consumed by group 1. Similarly, the specific cost of consumer group 2, $C_2(w_2)$, depends only on the amount of water consumed by group 2. These specific costs which are caused by just one consumer group are, for instance, pumping costs in the water networks.⁶ The optimization problem for maximizing total surplus in the water supply area is

$$\max_{\{w_1, w_2\}} [B_1(w_1) + B_2(w_2) - C_1(w_1) - C_2(w_2) - C_{12}(w_{12}(w_1, w_2))] \quad (4.11)$$

The KKT conditions resulting from the optimization problem are

$$B'_1(w_1) - C'_1(w_1) - \underbrace{C'_{12}(w_{12}) \cdot w'_{12}(w_1)}_{=1} \leq 0 \perp w_1 \geq 0 \quad (4.12)$$

$$B'_2(w_2) - C'_2(w_2) - \underbrace{C'_{12}(w_{12}) \cdot w'_{12}(w_2)}_{=1} \leq 0 \perp w_2 \geq 0 \quad (4.13)$$

It seems plausible to assume that both consumer groups consume positive amounts of water, i.e., $w_1 \geq 0$ and $w_2 \geq 0$. Under this assumption, the following optimality conditions can be formulated:

$$B'_1(w_1) = C'_1(w_1) + C'_{12}(w_{12}) \quad (4.14)$$

$$B'_2(w_2) = C'_2(w_2) + C'_{12}(w_{12}) \quad (4.15)$$

According to Eqs. (4.14) and (4.15), marginal benefit should equal marginal cost for each consumer group; hence, the optimal consumption level can be derived from the intersection point between the demand function and the marginal cost function for each addressed consumer group. Hence, for the optimal solution, the consumer groups have to pay different volumetric water prices if the marginal cost levels differ. A situation in which the marginal cost levels do not change with the output level (horizontal directed marginal costs functions) is depicted in Fig. 4.4. The component $C'_{12}(w_{12})$ is represented by the parameter c , which is, for instance, the cost rate for the treatment of 1 unit of water. However, the specific marginal (pumping) costs for serving consumer group 1 and 2 differ: $C'_1(w_1) < C'_2(w_2)$. Therefore, the volumetric

⁶We assume that the consumer group 2 has a higher geodetic level than consumer group 1. Hence, the pumping cost for serving consumer group 2 with one amount of water are higher than for the serving of consumer group 1.

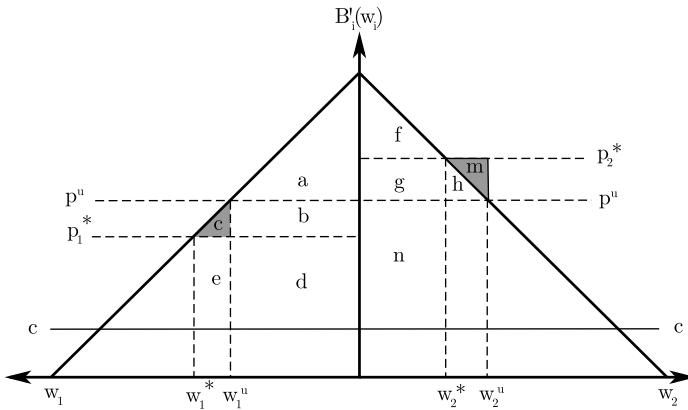


Fig. 4.4 Universal service provider with two consumer groups. *Source* own illustration

water price for consumer group 1 is lower than the one for consumer group 2, $p_1^* < p_2^*$, which means that a price discrimination has to be applied between the two exemplary consumer groups. This can be realized by, for instance, a regional price discrimination where two price zones are defined. The optimal consumption levels of both consumer groups in Fig. 4.4 are represented by the variables w_1^* and w_2^* .

4.3.4.2 Uniform Pricing

Suppose the water supplier is a universal service provider who offers water at a uniform price to all consumers, then the price p^u is set, which lies between the optimal price levels under price discrimination, i.e., $p_1^* < p^u < p_2^*$, as illustrated in Fig. 4.4. Compared to the case of price discrimination, setting a uniform price induces a decrease in the consumption level of group 1 from w_1^* to w_1^u , while consumer group 2 experiences an increase in its consumption level from w_2^* to w_2^u . Under a uniform price, the changes in prices and consumption levels relative to the optimal solution result in consumer group 1 losing some of its surplus, which is represented by the areas $b + c$ in Fig. 4.4, while consumer group 2 gains additional surplus, depicted by the areas $g + h$.

With respect to the supply side, the supplier may gain or lose by supplying groups 1 and 2. The price increase for group 1 impacts the producer surplus positively, whereas the consumption level decrease influences the producer surplus negatively. The area b , which is part of the consumer surplus under an optimal pricing regime, becomes producer surplus under a uniform pricing policy due to the price increase in group 1. In total, there is a loss of social welfare in supplying price zone 1 under a uniform price, which is symbolized by the area c .

A similar analysis can be done for consumer group 2. Due to the decreased price and increased consumption in group 2, the supplier’s loss in surplus amounts to the areas $g + h + m$. Therefore, there is a loss of social welfare induced by setting a uniform price in group 2, represented by area m . We conclude that a uniform pricing policy leads to economic losses in accordance with the areas $c + m$ in the whole water

Table 4.1 Distributional effects due to optimal and uniform pricing

		Group 1	Group 2
Consumer surplus (CS)	Optimum	$a + b + c$	f
	Uniform	a	$f + g + h$
	Δ CS	$-b - c$	$+g + h$
Producer surplus (PS)	Optimum	–	–
	Uniform	$+b$	$-g - h - m$
	Δ PS	$+b$	$-g - h - m$
Change social welfare		$-c$	$-m$

supply area compared to an optimal pricing policy. A more detailed overview of the distributional effects under the two addressed pricing regimes is given in Table 4.1.

Equity in pricing policy is often seen as fairer than (regional) price discrimination based on the cost-by-cause principle in a water supply area. However, the gain of fairness is associated with a loss of economic efficiency which is represented by the areas $c + m$. This is a matter of fairness preference, or of inequity aversion, which means how much loss of efficiency a society wants to accept in order to achieve a fair water allocation.

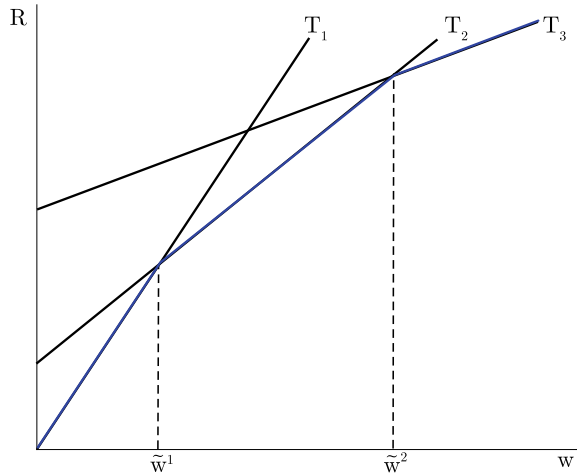
4.3.5 Optional Tariffs

4.3.5.1 The Concept

If the water supplier offers optional tariffs, consumers have the possibility to choose between various pricing options. The effectiveness of optional tariffs is explained with the help of an example: Fig. 4.5 depicts three pricing options whose expenditure functions depend on the level of consumption. Each of these pricing options contains a recurrent fixed charge and a single volumetric charge. The first pricing option, T_1 , is characterized by a relatively low fixed charge and a relatively high volumetric charge, while the third pricing option, T_3 , contains a relatively high fixed charge and a relatively low volumetric charge. Pricing option T_2 is characterized by a relatively moderate fixed charge and volumetric charge. If L_i and p_i stand for the fixed and volumetric charge of the i th pricing option, respectively, it is possible to characterize the optional tariff by $(L_1 < L_2 < L_3) \wedge (p_1 > p_2 > p_3)$.

The consumers have the option to choose between pricing options T_1 , T_2 , and T_3 , whereas the decision of the consumer depends on the consumption level. If the consumer consumes less than \tilde{w}^1 , which means $w \leq \tilde{w}^1$, the consumer minimizes its expenses by choosing option T_1 . Similarly, consumers who have a moderate consumption within the range $\tilde{w}^1 \leq w \leq \tilde{w}^2$ choose pricing option T_2 to minimize the expenditures. Consumers with a consumption above \tilde{w}^2 , i.e., $w \geq \tilde{w}^2$, minimize their expenditures by choosing price option T_3 . It becomes obvious that a higher consumption level is associated with a higher preference to choose a tariff with lower volumetric rate. This is the analogous logic to a volume discount, where the price

Fig. 4.5 Concept of optional tariffs. *Source* own illustration



per unit also decreases with an increase in units bought. This volume discount rule for an optional tariff can be formalized by the statement $\forall m \forall n : (w_m < w_n) \implies (p_m \geq p_n)$.⁷

This simple example shows that with increasing consumption the consumers choose pricing options with higher fixed fees. This is a typical characteristic of well-defined optional tariffs, which can be explained by the previously mentioned volume discount rule. Under the assumption that the volumetric price of the m th pricing option is greater than or equal to the one of the n th option, i.e., $p_m \geq p_n$, the following undesirable situations are possible:

- One or more pricing options have an absolute disadvantage compared to another/to other pricing option(s). Pricing options with an absolute disadvantage are never chosen and are, therefore, useless as an optional tariff.
- At least two pricing options are identical. If at least two options are identical, the consumer is indifferent between choosing the respective options. Hence, at least one option is useless in the optional tariff.

If one of these situations occurs, the optional tariff is not well defined. If the volumetric price of the m th pricing option is higher than the one of the n th pricing option, such that $p_m > p_n$, which means that the preference for choosing the m th pricing option instead of the n th pricing option will increase with a decreasing consumption level, the fixed charge of the m th pricing option has to be lower than the one of the n th pricing option in a well-defined optional tariff. This can be described in a formal way by the following statement: $\forall m \forall n : (p_m > p_n) \iff (L_m < L_n)$.

⁷If $w_n > w_m$, then $p_n \leq p_m$.

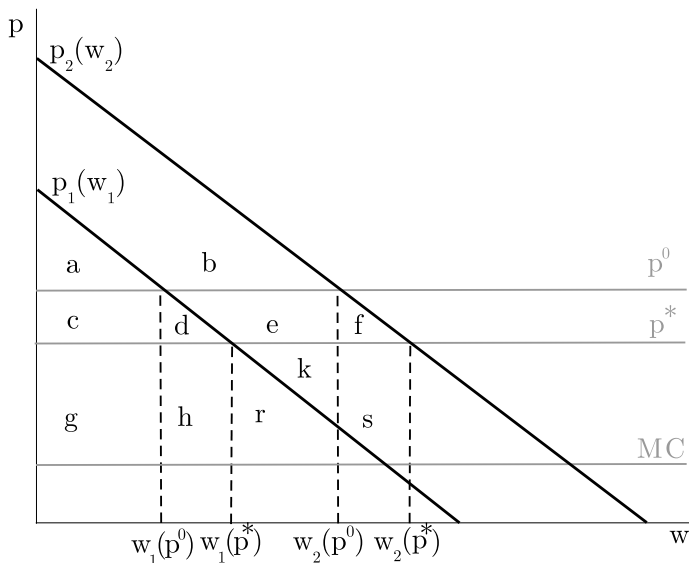


Fig. 4.6 Example with optional tariffs. *Source* own illustration

4.3.5.2 Effects on Consumers and Producers

Under the implementation of an optional tariff both, the consumer and the producer can gain social welfare. This hypothesis can be fostered with the help of a simple example illustrated in Fig. 4.6: Given there are two consumer groups $i = \{1, 2\}$ whose demand functions are known. Consumer group 1 has a lower demand than consumer group 2 at every price level. In the initial situation, only one tariff $T^0 = [L^0, p^0]$ is offered by the water supplier. L^0 and p^0 stand for the fixed and volumetric charges in the initial tariff, respectively. Based on this volumetric price, the consumer groups 1 and 2 consume the amounts of $w_1(p^0)$ and $w_2(p^0)$, respectively. The generated consumer surplus for consumer group 1 is represented by the area a , while the one for group 2 is given by areas $a + b$. The expenditures from the base price reduce the consumer surplus by L^0 for each consumer group. Therefore, the consumer surpluses for groups 1 and 2 are represented by $a - L^0$ and $a + b - L^0$, respectively. The water supplier generates a profit equal to the area $2 \cdot (c + g) + d + h$ from the volumetric charge plus $2 \cdot L^0$ from the fixed charge to cover its fixed costs.⁸

Now, we assume that the water supplier offers an optional tariff in which the initial pricing scheme $T^0 = [L^0, p^0]$ is supplemented by an alternative pricing option $T^* = [L^*, p^*]$, where L^* and p^* symbolize the fixed charge and the volumetric charge of pricing option T^* , respectively. We assume that $p^0 > p^*$, and hence $L^0 < L^*$, such that the optional tariffs are well defined. Every consumer has the choice between the pricing options T^0 and T^* . If consumer group $i = \{1, 2\}$ chooses option T^0 , this

⁸The profit generated from the volumetric price paid by groups 1 and 2 is symbolized by the areas $c + g$ and $c + d + g + h$, respectively.

group consumes the amount $w_i(p^0)$ and receives the consumer surplus CS_i^0 , while if it opts for option T^* , it consumes $w_i(p^*)$, which leads to a consumer surplus of CS_i^* . The consumer group $i = \{1, 2\}$ chooses the option that maximizes the consumer surplus. We assume that $CS_1^0 > CS_1^*$ and $CS_2^0 < CS_2^*$. Hence, consumer group 1 will choose option T^0 , while consumer group 2 decides for option T^* , which is the option with the lower volumetric and higher base price (volume discount rule).⁹ In the design of option T^* , the water supplier anticipates the reaction of group 2. Due to the reduction of the volumetric price by switching from T^0 to T^* , revenues from this charge will be lost, maximally amounting to $(p^0 - p^*) \cdot w_2(p^0)$, which is symbolized by the areas $c + d + e$. These revenue losses can be fully compensated by a higher fixed charge in option T^* , hence, $L^* = L^0 + (p^0 - p^*) \cdot w_2(p^0)$.

If the optional tariff is offered, the situation does not change for the consumer group 1, because they are still priced under T^0 and obtain a consumer surplus represented by the area $a - L^0$. However, group 2 will switch from tariff T^0 to T^* and increase consumption from $w_2(p^0)$ to $w_2(p^*)$. Therefore, the surplus increases by the areas $c + d + e + f$. Furthermore, due to the change of pricing, expenditures from the base price increase from L^0 to L^* . These additional expenditures are represented by the areas $c + d + e$.¹⁰ Hence, the consumer surplus increases by the area f to the level $a + b + f - L^0$ because of the introduction of the optional tariff. Therefore, consumer group 2 benefits from the introduction of the optional tariff, while consumer group 1 is not affected. Offering the optional tariff is also advantageous for the water supplier, as its profits increase by s to the level $2 \cdot (c + g) + d + e + h + r + k + s$ plus $2 \cdot L^0$.

Based on the example described above, we can conclude that the introduction of optional tariffs can lead to a situation where nobody is worsened and specific actors have an advantage compared to the initial situation without an optional tariff. Tables 4.2 and 4.3 give a detailed overview of the distributional effects for consumers and the water suppliers under both addressed pricing regimes, respectively.

4.3.6 Seasonal Pricing

In the case that a seasonal pricing scheme is implemented, the price for water changes with the time period of supply. For instance, the price in the summer month could be higher than in the winter month. We would like to term the period where the price has the highest level as peak period/peak season (e.g., summer), while the residual period is termed as off-peak period/off-peak season. If we implement this form of temporal price discrimination, the different price levels between the peak and off-peak seasons can be based on various reasons, for instance,

⁹ Assume that $CS_1^0 < CS_1^*$. Because of the volume discount rule, it is certain that $CS_2^0 < CS_2^*$.

¹⁰ $L^* = L^0 + (p^0 - p^*) \cdot w_2(p^0)$ and hence $L^* - L^0 = (p^0 - p^*) \cdot w_2(p^0)$. If there is a switch from pricing T^0 to T^* , the fixed charge increased from L^0 to L^* , which is nothing else than $L^* - L^0 = (p^0 - p^*) \cdot w_2(p^0)$. The additional revenues from the fixed charge are $(p^0 - p^*) \cdot w_2(p^0)$ which is equal to the areas $c + d + e$.

Table 4.2 Effects of surplus on consumer side due to the introduction of an optional tariff

		Group 1	Group 2
Initial tariff	Surplus from fixed charge	$-L^0$	$-L^0$
	Surplus from volumetric charge	a	$a + b$
	Consumer surplus	$a - L^0$ $2 \cdot a + b - 2 \cdot L^0$	$a + b - L^0$
Optional tariff	Surplus from fixed charge	$-L^0$	$-c - d - e - L^0$
	Surplus from volumetric charge	a	$a + b + c + d + e + f$
	Consumer surplus	$a - L^0$ $2 \cdot a + b + f - 2 \cdot L^0$	$a + b + f - L^0$
Change consumer surplus		/	$+f$ $+f$

Table 4.3 Effects of surplus for water supplier due to the introduction of an optional tariff

		Group 1	Group 2
Initial tariff	Surplus from fixed charge	L^0	L^0
	Surplus from volumetric charge	$c + g$	$c + d + e + g + h + r + k$
	Profit margin	$c + g + L^0$ $2 \cdot (c + g) + d + e + h + r + k + 2 \cdot L^0$	$c + d + e + g + h + r + k + L^0$
Optional tariff	Surplus from fixed charge	L^0	$c + d + e + L^0$
	Surplus from volumetric charge	$c + g$	$g + h + r + k + s$
	Profit margin	$c + g + L^0$ $2 \cdot (c + g) + d + e + h + r + k + s + 2 \cdot L^0$	$c + d + e + g + h + r + k + s + L^0$
Change profit margin		/	$+s$ $+s$

- Less water is available in the summer (peak season) than in the residual months (off-peak season). The different prices between these time intervals result from the different scarcity price levels.
- It is more expensive to deliver water in the summer (peak season) than in the residual months (off-peak season). Therefore, the marginal cost function in the peak season is higher than in the off-peak season. This results in a higher equilibrium price in the peak season compared to the off-peak season.
- The demand for water in the summer (peak season) is higher than in the residual months (off-peak season), because in summer people are more thirsty, more water is required for plants watering, more water is needed for filling pools, etc. Therefore, the demand function of the peak season is higher than the one of the off-peak season. This results in a higher equilibrium price in the peak season compared to the off-peak season.
- The provision of capacity for delivering water is related to costs. The higher the capacity, the higher the capacity costs. These capacity costs have to be covered by the revenues from the water price. It is thinkable that the capacity is just financed in the peak season. However, it is also possible that capacity is financed during the peak and off-peak seasons. A temporal price discrimination scheme which is based on the financing of capacity costs for water delivery infrastructure is presented in the following model.

4.3.6.1 Temporal Price Discrimination for Financing Capacity Costs

Suppose the two time seasons 1 and 2 where we have the water consumption w_1 and w_2 , respectively. In both seasons, benefits are generated from the water consumption related to the functions $B_1(w_1)$ and $B_2(w_2)$. We already know from household theory that the demand function is determined by the marginal benefit function. Hence, the demand function of season 1 is $p_1(w_1)$, while the demand function for season 2 is $p_2(w_2)$. The demand in season 2 is higher than in season 1 if

- for every consumption level w : $p_2 > p_1$, the price level in season 2 is higher than in season 1;
- for every price level p : $w_2 > w_1$, the consumption amount in season 2 is higher than in season 1.

Under this assumption, we term the season 2 as peak season and the season 1 as the off-peak season.

The supply of water is related to costs. On the one hand, we suppose the cost rate c for delivering water. Hence, the annual costs for water delivery are therefore represented by the term $c \cdot (w_1 + w_2)$. Furthermore, there also exist costs for the provision of capacity in the supply system. This capacity (e.g., pumping capacity) is needed for the delivery of water to the consumers. The capacity cost rate r represents the cost for the provision of one unit of capacity. The capacity level is represented by the variable k . Hence, the total annual capacity costs are therefore $r \cdot k$.

In the optimization, we want to calculate the optimal consumption and capacity levels in the way that we maximize the total surplus of 1 year, which includes the peak and off-peak seasons. The total surplus results from the difference of benefits and costs:

$$\begin{aligned} & \max_{\{w_1, w_2, k\}} [B_1(w_1) + B_2(w_2) - c \cdot (w_1 + w_2) - r \cdot k] \\ & s.t. \ w_1 \leq k \quad (\lambda_1) \\ & \quad \quad w_2 \leq k \quad (\lambda_2) \end{aligned}$$

Of course, the water delivery in both seasons is restricted by the chosen capacity level. Based on the optimization problem, the Lagrangian function can be set up:

$$L = B_1(w_1) + B_2(w_2) - c \cdot (w_1 + w_2) - r \cdot k + \lambda_1 \cdot [k - w_1] + \lambda_2 \cdot [k - w_2]$$

and finally the KKT conditions can be formulated:

$$B'_1(w_1) - c - \lambda_1 \leq 0 \perp w_1 \geq 0 \quad (4.16)$$

$$B'_2(w_2) - c - \lambda_2 \leq 0 \perp w_2 \geq 0 \quad (4.17)$$

$$\lambda_1 + \lambda_2 - r \leq 0 \perp k \geq 0 \quad (4.18)$$

$$k - w_1 \geq 0 \perp \lambda_1 \geq 0 \quad (4.19)$$

$$k - w_2 \geq 0 \perp \lambda_2 \geq 0 \quad (4.20)$$

We assume that we have a consumption in both seasons, hence $w_1 \geq 0$ and $w_2 \geq 0$. Because of Eqs.(4.19) and (4.20), it follows that we have to assume that $k \geq 0$. Therefore, we know from Eqs.(4.16)–(4.18) that

$$B'_1(w_1) = c + \lambda_1$$

$$B'_2(w_2) = c + \lambda_2$$

$$\lambda_1 + \lambda_2 = r$$

Regarding the conditions Eqs.(4.19) and (4.20), we have four possible options for assumption:

- $\lambda_1 = \lambda_2 = 0$: The capacity is exploited neither in the off-peak nor in the peak season;
- $\lambda_1 \geq 0, \lambda_2 = 0$: The capacity is exploited just in the off-peak season;
- $\lambda_1 = 0, \lambda_2 \geq 0$: The capacity is exploited just in the peak season;
- $\lambda_1 \geq 0, \lambda_2 \geq 0$: The capacity is exploited in the off-peak and peak seasons.

The assumption $\lambda_1 = \lambda_2 = 0$ can never lead to optimality, because of Eq.(4.18). We know that $\lambda_1 + \lambda_2 = r$ is not met, because $\lambda_1 + \lambda_2 = 0$, while capacity cost rate r is positive, i.e., $r > 0$.

The assumption $\lambda_1 \geq 0$, $\lambda_2 = 0$ can also never lead to optimality. This seems to be quite plausible, because under this assumption the capacity is only exploited in the off-peak season. This means that in the peak season the consumption is lower than in the off-peak season, even though the demand in the peak season is higher than in the off-peak season. Therefore, this assumption intuitively does not make much sense. However, there is also a mathematical way for finding a contradiction under this assumption. From conditions Eqs. (4.19) and (4.20), we know that $w_1 = k$ and $w_2 \leq k$, respectively. Hence, $w_2 \leq w_1$. From Eq. (4.18), it becomes obvious that $\lambda_1 = r$. Based on Eq. (4.17), we find that $B'_2(w_2) = c$, while from Eq. (4.16) it follows that $B'_1(w_1) = c + \lambda_1$, which is nothing else than $B'_1(k) = c + r$. Let us define the variable η which gives the difference between the consumption level in the peak season and the capacity, hence $\eta = k - w_2$. Therefore, we know that $k = \eta + w_2$. For sure, we are able to reformulate $B'_1(k)$ into the expression $B'_1(\eta + w_2)$. Furthermore, we also know that the benefit functions are concave, which means for the benefit function of user 1: $B''_1(w) < 0$. Due to this concavity condition we know that $B'_1(\eta + w_2) < B'_1(w_2)$. The marginal benefit in the off-peak season exceeds the one in the peak season, $c + r = B'_1(k) > B'_2(w_2) = c$. Combining the latter two relations, we can formulate that $B'_1(w_2) > B'_1(\eta + w_2) > B'_2(w_2)$, and hence $B'_1(w_2) > B'_2(w_2)$, which is a contradiction to a former assumption. We suppose that season 1 is the off-peak season with lower demand, while season 2 is the peak demand with higher demand. Because of this relation we know that for every consumption level w : $B'_1(w) < B'_2(w)$, the marginal benefit of the peak season 2 exceeds the one of the off-peak season 1, and therefore $B'_1(w_2) < B'_2(w_2)$.

Under the assumption $\lambda_1 = 0$ and $\lambda_2 \geq 0$, the capacity is just exploited in the peak season. Therefore, the consumption level in the peak season is higher than in the off-peak season which seems to be quite plausible. Based on Eqs. (4.16)–(4.20), we are able to set up the following optimality conditions for this case:

$$w_2 = k \quad (4.21)$$

$$w_1 \leq k \quad (4.22)$$

$$\lambda_2 = r \quad (4.23)$$

$$B'_1(w_1) = c \quad (4.24)$$

$$B'_2(w_2) = c + r \quad (4.25)$$

The situation is illustrated in Fig. 4.7. The left figure illustrates the situation where the made assumption $\lambda_1 = 0$ and $\lambda_2 \geq 0$ leads to optimality. The optimal price and consumption levels in the off-peak season result from the intersection between the marginal benefit in season 1 and the cost rate for water delivering c , while the optimal price and consumption levels in the peak season (summer) result from the marginal benefit in season 2 and the sum of the cost rate for water delivering and the capacity cost rate $c + r$.

The price in the peak season, which is equal to the marginal benefit $B'_2(w_2)$, is bigger than in the off-peak season which is $B'_1(w_1)$, because of the capacity cost rate r , which is only relevant in the peak season. Hence, the revenues for covering the

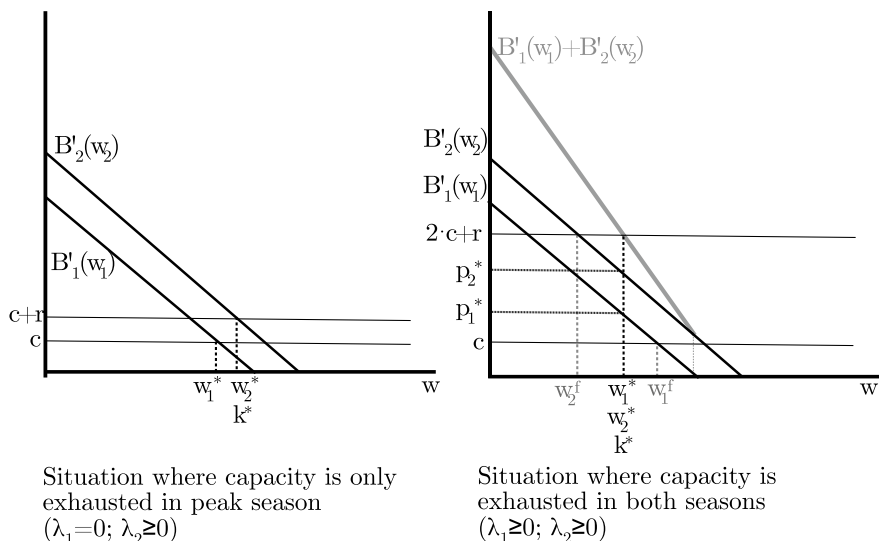


Fig. 4.7 Illustration of optimal seasonal pricing. *Source* own illustration

capacity costs are only earned in the peak season. The capacity cost rate does not impact the consumption level in the off-peak season, but the consumption level in the peak season. Of course, the higher the capacity cost rate, the lower the consumption in the off-peak season. If the capacity cost rate would be sufficiently high, the consumption in the peak season could fall below the consumption level in the off-peak season.

This situation is pictured in the right figure of Fig. 4.7. If we would set the price in the off-peak and peak season equal to c and $c + r$, respectively, it results in a consumption level w_1^f in the off-peak season which exceeds the consumption level w_2^f in the peak season. If the consumption level in the off-peak season exceeds the one of the peak season, we do not meet the condition Eq. (4.19), and hence, the other plausible case ($\lambda_1 \geq 0$ and $\lambda_2 \geq 0$) would lead to optimality.

The assumption $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ states that the capacity is exploited in both seasons, in the peak and the off-peak seasons. The consumption levels in both seasons are equal to the capacity. Based on Eqs. (4.16)–(4.18), we know that

$$\begin{aligned} \lambda_1 + \lambda_2 &= r \\ B'_1(w_1) &= c + \lambda_1 \\ B'_2(w_2) &= c + \lambda_2 \end{aligned}$$

which can be combined by summation of $B'_1(w_1)$ and $B'_2(w_2)$ to the following expression:

$$B'_1(w_1) + B'_2(w_2) = 2 \cdot c + r$$

Therefore, we can formulate the following optimality conditions:

$$w_1 = w_2 = k \quad (4.26)$$

$$B'_1(w_1) + B'_2(w_2) = 2 \cdot c + r \quad (4.27)$$

$$\lambda_1 = B'_1(w_1) - c \quad (4.28)$$

$$\lambda_2 = B'_2(w_2) - c \quad (4.29)$$

Here, we have the same consumption levels in both seasons. Because the demand in the peak season is higher than in the off-peak season, the price in the peak season is certainly higher than in the off-peak season. However, the revenues for financing the capacity costs are earned in both seasons.

The right figure of Fig. 4.7 pictures the situation for which the assumption $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ leads to optimality. The function $B'_1(w_1) + B'_2(w_2)$ results from the summation of the seasonal marginal benefit functions in vertical direction. The intersection point of the $B'_1(w_1) + B'_2(w_2)$ -curve with $2 \cdot c + r$ determines the optimal capacity k^* as well as the optimal consumption levels w_1^* and w_2^* of both seasons. The seasonal prices result from the consumption levels found. The respective marginal benefit functions are illustrated in Fig. 4.7 by p_1^* and p_2^* for the off-peak and peak seasons, respectively.

4.4 Increasing Block Tariffs

4.4.1 The Concept

The increasing block tariff (IBT) is an important tariff form which is quite often implemented especially in developing countries. This tariff form is characterized by a volumetric charge that increases with rising consumption level. In some cases, even the level of the fixed charge depends on the consumption level. The characterization of IBTs can be formalized by

$$w_j > w_i \rightarrow p_j \geq p_i \quad (4.30)$$

If the j th consumption level is higher than the i th consumption level, the volumetric price of the j th consumption level must not fall below the one of the i th consumption level.

The popularity of this tariff form is attributed to the combination of some alleged advantages. Regions in developing countries are often characterized by a high proportion of poor population class with low income and low water availability leading to water shortage. Therefore, addressing social concerns and environmental protection goals is important. A well-designed increasing block tariff can enforce a simultaneous achievement of both goals. The first water volumes per household can be provided at a low price, which makes water affordable even for the poorest. Hence, the implementation of a well-defined increasing block tariff may promote the access to the public water supply for even those, with the lowest income. A secured access

to water may lead to increased prosperity and well-being as well as the promotion of public health especially for the poorest people (Boland and Whittington 2000b).

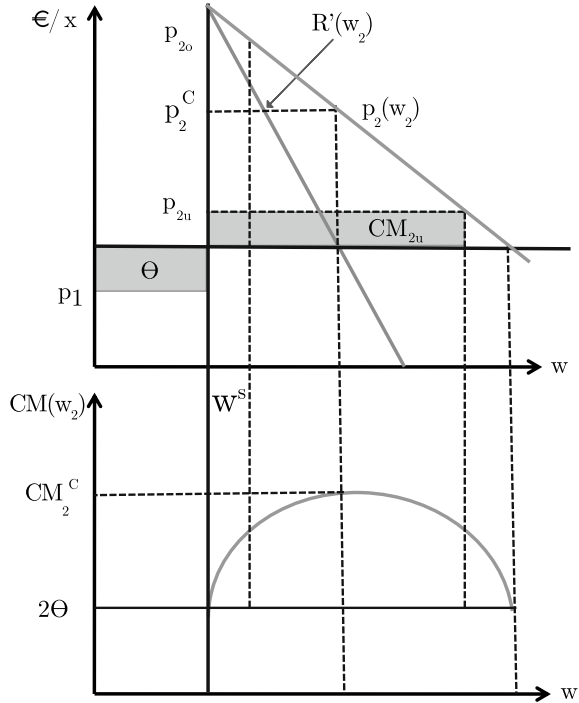
Starting from a low volume price for the first consumed water amounts, the increasing block tariff is characterized by a rising volume price with increasing water consumption. A sufficiently high volume price for a certain defined consumption level sets an incentive for avoiding wasteful use and fosters the implementation of water-saving technologies. Households equate their marginal willingness to pay to the price valid in the block in which consumption falls. With increasing tariffs, water consumption can be pushed back. Hence, the tariff system meets two objectives at the same time: It prevents water wastage, thus helping to conserve resources and guarantees access to water for the poor.

Figure 4.8 displays the case of a two-block tariff where the price increase from the first to the second block leads to an implicit cross-subsidization of the poor by the middle class. We assume first that there are only two households: a poor household and a household with a middle class income. The upper part of the figure shows two blocks. The width of the first block from 0 to w_s is equal to the subsistence level of an average household. The corresponding water price is p_1 , where p_1 is lower than the marginal costs of water supply (c). Water consumption beyond the lifeline falls into the second block. This demand is charged either by p_{2u} or p_{2o} . The demand for water in the second block follows the price-quantity function $p_2(w_2)$ for $w_2 \geq w_s$. While the poor household cannot afford to consume more water than the subsistence minimum, the middle-income household has sufficient income to consume more than the lifeline depending on the price in the second block.

Furthermore, we assume that the price in the second block is p_{2u} . The expense for the non-poor household is $p_1 w_s + p_{2u} \cdot (w_2 - w_s)$; the expenses for the poor households are simply $p_1 \cdot w_s$. Obviously, the first block generates a deficit of $2(c - p_1)w_s$ ($2 \times \text{Area } \Theta$) which must be covered by the contribution margin in the second block $p_{2u}(w_2 - w_s)$ (area CM_{2u}). The contribution margin is depicted on the vertical axis of the lower half of the picture as a function of the water consumption in the second block. As the price of block 2 is increased, the contribution margin (profits in the second block) rises until it reaches a maximum. If the price increases further, profits in the second block decrease until demand in the second block is choked off. The lines are drawn such that p_{2o} and p_{2u} generate the same amount of contribution margin. In addition, these two prices are chosen such that the deficit from the first block, 2Θ , is exactly covered.

It remains to choose one of the two price options in block 2. The municipality can choose either a flat increase of the price from block 1 to block 2 or decide to implement a strong uplift which seems more egalitarian. Figure 4.8 shows that the strive for a more egalitarian outcome does not improve the situation of the poor household. The increase in the price of the second block from p_{2u} to p_{2o} does not lead to a decrease of p_1 . Hence, if we follow the Pareto principle, we would choose the price in the second block such that the consumer rent of the middle class household is maximized subject to economic viability, i.e., that no deficit occurs. This is achieved if the water utility chooses p_{2u} . If a community adheres strictly to Egalitarianism, then p_{2o} is chosen at the expense of efficiency. In this case, water consumption is far

Fig. 4.8 Increasing two-block tariff. *Source* own illustration



less than the efficient level which is where the marginal willingness to pay is equal to marginal cost c . Water pricing then becomes a political issue and depends on how a society deals with justice and coherence issues.¹¹

4.4.2 Potential Adverse Effects on the Poor

If a poor household consumes less water than richer households—which is quite plausible, because water is a normal good and income elasticity for water is positive—the wealthier households are burdened with more costs than the poor households. This promotes equity and the reduction of income disparities between the households. Therefore, to conclude, the wealthier households would cross-subsidize the poorer households, which can be considered as fair. However, notice that both prices p_{2u} and p_{2o} lead to the same amount of cross-subsidization. Increasing the price of the second block does not necessarily contribute to more cross-subsidization and, hence,

¹¹There is a branch of cultural theories that explains the way of water management and, hence, of water allocation rules with the help of cultural configurations. In some cultural environments, the efficiency criterion in the sense economists use it (equalization of the marginal willingness to pay across all members of a society) plays only a minor role, see Hoekstra (1998).

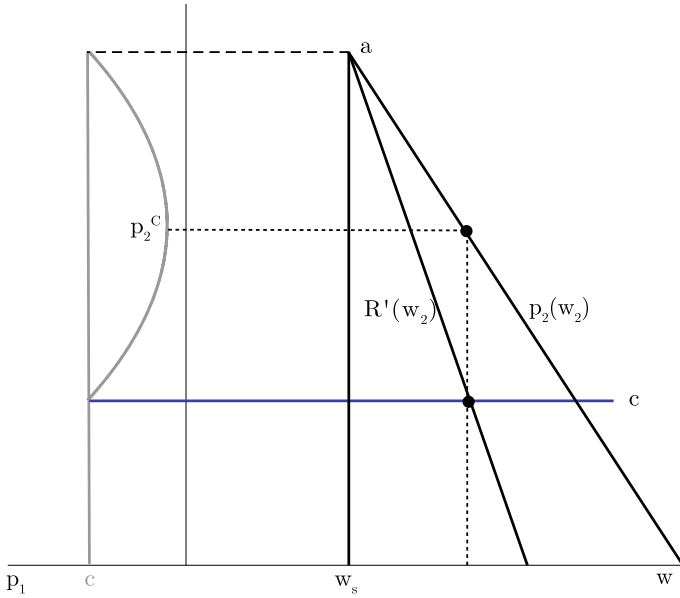


Fig. 4.9 The relation between block prices. *Source* own illustration

a lower price of the first block. Figure 4.9 shows that an exaggerated tariff progression can be detrimental to the poor.

In Fig. 4.9, it is assumed that the price-quantity function is linear

$$p_2(w_2) = a - b(w_2 - w_s) \quad \text{for } w_2 \geq w_s \tag{4.31}$$

The contribution margin of the second block is therefore

$$CM_2(w_2) = (p_2(w_2) - c)(w_2 - w_s) = (a - b(w_2 - w_s) - c)(w_2 - w_s) \quad \text{for } w_2 \geq w_s \tag{4.32}$$

Braking even requires the water utility to set

$$2(c - p_1)w_s = CM_2(w_2) \tag{4.33}$$

The contribution margin can also be expressed as a function of p_2 . Simply solve Eq. (4.31) for $w_2 - w_s$ which yields

$$w_2 - w_s = \frac{a - p_2}{b} \tag{4.34}$$

Inserting into Eq. (4.32) yields

$$CM_2(p_2) = (p_2 - c)\left(\frac{a - p_2}{b}\right) \quad \text{for } c \leq p_2 \leq a \tag{4.35}$$

Finally, inserting this expression into the break-even condition and solving for p_1 gives

$$p_1 = c - \frac{(p_2 - c)\left(\frac{a-p_2}{b}\right)}{2w_s} \quad (4.36)$$

This function is drawn in the left half of Fig. 4.9. We can observe that for the cases $p_2 = c$ and $p_2 = a$ the price in the first block is $p_1 = c$. For the open interval $c < p_2 < a$ we have $p_1 < c$. We observe also that the function defined in Eq. (4.36) is not monotonous. For all $p_2 > p_2^C$, an increase of p_2 leads to an increase of p_1 , i.e., strengthening the ascent of block prices goes to the detriment of the low-income customers. The reason for this is the elasticity of demand in the second block. A price increase leads to such a large reduction of demand in block 2 that the contribution margin is declining and, hence, the cross-subsidization goes back.

Instead of looking at the exact progression of the tariff, a policy dedicated to secure water access for the poor should make sure that the price of the first block is as low as possible. This can be achieved by maximizing the contribution margin in the second block. The lower half of Fig. 4.8 suggests how to choose this price optimally. It corresponds to the Cournot point p_2^C leading to the maximum cross-subsidy CM_2^C .¹²

Another potential pitfall of IBTs is that poorer households may have a larger size; thus, despite a lower per capita consumption, they could have a higher overall consumption level per household, compared to a wealthier household. In that case, the desired cross-subsidy mechanism for the support of poor households would not work anymore. Since households with larger size have a higher consumption, those families would therefore cross-subsidize households with a lower consumption level. The consequence is that poorer families have to carry a heavier burden than wealthier households which leads to the promotion of inequality.

4.4.3 Further Considerations

There are also some further issues regarding the conception of a well-defined increasing block tariff, see Boland and Whittington (2000b) and Meran and von Hirschhausen (2014):

- **Setting the initial block:** Because of political and other pressure, it is difficult for water companies to limit the initial block. A large initial block directly benefits not just the poor, but also the middle class and maybe even upper income households. If the majority of private connections to public water supply is held by middle and upper income households, these households receive the vast majority of water sold at subsidized prices. International standards for basic water needs are usually

¹²An exercise at the end of the chapter deals with the Cournot point and shows how the result depends on the demand elasticity.

in the range of 25–30 l per capita per day. It can be observed that most cities deliver households more water than the basic water need at the lowest price. Of course, also the household size is relevant for the desired sizing of the initial block. The higher the household size, the higher the required initial block size for purchasing the basic water needs at the lowest price.

- **Simplicity and transparency:** Increasing block tariffs are neither simple nor transparent. The more sophisticated the tariff structure, the harder to deduce the average or marginal price that is actually paid for a certain amount of water. The confusion about the marginal or average water price may lead to a restriction of the signal effect function of the price, whereby the consumers concerned no longer behave completely rational in accordance with expectations. Sophisticated tariffs may also create customer relation problems, making it more difficult for representatives of a water agency to explain bills. A tariff with a single volumetric rate, independent from the consumption level, is simple, transparent, robust, and easy to implement. This leads to consistent and understandable price signals.
- **Shared connections:** Increasing block rates are only implementable if customers have a metered water connection to measure the consumption quantities. In many cities of developing countries, the water meters are just available for the upper and middle-income households. However, the poor obtain water from vendors or shared connections. If several households share a metered water connection, water use by the consumers is quickly pushed to the higher priced blocks. The consequence may be that poor households pay a higher average price than the rich who have a private metered connection.
- **Reselling:** If poor households do not have a private metered water connection, so that an increasing block rate is not applicable for them, they can buy their water from households (neighbors) who have a metered connection. If a household sells water to other households, their water consumption is quickly pushed into the higher priced blocks. There is a similar situation like with shared connections: the more the water sold, the higher the average price. If this case occurs, the household which resells the water can capture the benefits from the first block and charge the resold water with a price that will recover the highest per unit charge plus some markup for inconvenience of water selling.

4.5 Pricing in Unconnected Water Markets

4.5.1 Stylized Facts

Despite some progress, there are many areas in the world, specifically in sub-Saharan Africa and South and South-East Asia, where the access to safe water and to adequate sanitation is not given. Not only in rural areas where water connections are very expensive due to the low settlement density, but also in urban regions with a growing population living in informal settlements, the rate of connected households is low. People in those areas have to rely on other sources for their water supply, such as public or private taps, and water vendors selling water door to door as well as own

wells. Furthermore, there may be other available sources like leakages in water pipes, harvested rainwater, and collected surface water. Hence, many people are dependent on water of unsafe origins. Due to the poor water quality, diseases can spread easily, worsening the living conditions of those affected. People are well aware of these risks but they simply do not have the money (or the time) to secure themselves access to clean freshwater. This seems to support the comprehensive expansion of a water distribution infrastructure. Since investments in water distribution are very expensive, financing in the context of a poor population becomes a major issue.

There are additional reasons why the settlements of the poor are often not connected to a pipe-based water distribution system:

1. The lack of water infrastructure is attributed to staff incompetence and a lack of motivation due to a salary system without incentive schemes. Also, water utilities may lack the required skills.
2. There is a lack of political interest in the poor. The water supply is geared to the needs of the middle class and upper income families.
3. But even if there is a political will to improve the water supply for the poor, there are obstacles to extend pipe-borne services to low-income urban areas. Often, the ownership of land and property is not well defined. In addition, utilities have problems to collect revenue from metered customers and they cannot prevent people from illegally tapping water.
4. There are also economic motives that prevent the extension of pipe-borne water supply. The water supply system consists of many actors with divergent objectives. A close-meshed water distribution system can be against their interest.

In the following, we will describe a water distribution system in more detail that is based on a decentralized supply mode.¹³ Figure 4.10 depicts the different actors.

- At the beginning of the production and distribution chain is the water utility. Utilities provide clean water through a pipe system to households. However, many empirical studies on the water conditions in urban and peri-urban areas in developing countries show that the connection rate is not very high. Households of the middle and high income (in urban areas) are connected to the water utility and the waste water treatment plant. Since the coverage of utility networks is often limited, low-income groups regularly have to rely on various other service supply systems.
- Taps and standpipes are connected to the pipes of the water utility. They are also called “water kiosks” and are often run by private or public managers. Either these managers are employed by the water utility or the kiosks are privately run. In this case, the owner/manager has to pay a license fee to the utility. The contract may include a lump sum fee as well as a volumetric part. These payments can also be

¹³We confine our analysis to the water supply system. A similar analysis can be done for the sanitation system.

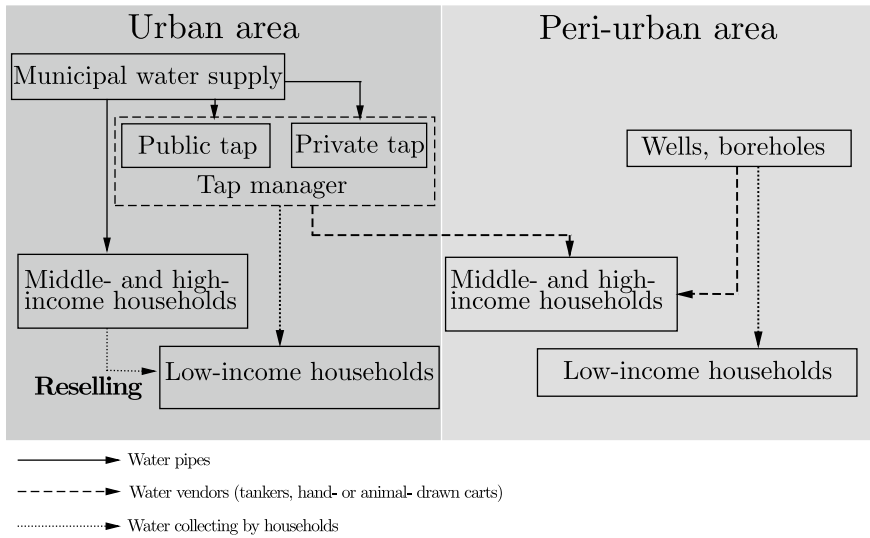


Fig. 4.10 Decentralized water sector in urban and peri-urban areas. *Source* own illustration

formal or, specifically if the prices are publicly regulated, informal, i.e., they are clandestine side payments. The manager, in turn, sells water to water vendors and to households in the vicinity of the tap as well.

- **Reselling from connected households:** This activity can often be observed and can be explained by the price spread between pipe-borne water and the water price in the informal market. We know from Sect. 4.4 that increasing block tariffs may be tailored in favor of the middle class offering water to a low price, sometimes even below costs. In addition, blocks are structured such that households can draw more water than needed without exhausting the block. Thus, there is scope for profitable water trading. Household water connections are therefore similar to private water kiosks.
- **Water vendors:** A main source of water for poor households is the services of water vendors. There exist mainly two types of vendors: wholesale vendors, often serving by truck, and distributing vendors that in turn sell water door to door. The technology of these street peddlers is rather simple. They carry water in plastic jerricans that are hauled by handcarts or bicycles. Capital costs of this mobile vending system seem to be lower than the piped water supply, at least for short and medium distances.
- **Private wells and boreholes:** In many urban and rural areas, households receive water from private wells. These wells can either be historic or recently built. This means that households resort to groundwater, which has decreasing quality standards due to population growth. Households are aware of this relationship. However, due to the high water price they cannot rely on any other water sources.

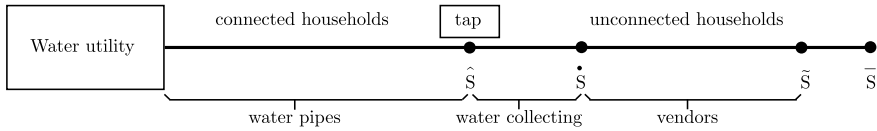


Fig. 4.11 Linear city. *Source* own illustration

4.5.2 Model

4.5.2.1 The Linear City

To analyze the characteristics of various modes of water supply, we utilize an economic model one can find in spatial economics: the linear city. It is assumed that all water customers are arranged along a line, the linear city, (Fig. 4.11). The customer density is constant along the line (identical distribution). One customer lives at each location, and each customer consumes exactly one unit of water, say 1 m^3 per month. The willingness to pay $V(s) = a - bs$ of each customer is decreasing along the linear city from the left to the right side. This property stems from the assumption that income of households decreases from left to right. On the left, the high-income households live followed by the middle class and finally the less fortunate, which settle on the right side. This model structure makes the analysis transparent and allows to identify the economic drivers that determine water prices along the linear city (Fig. 4.11).¹⁴

The water utility is located to the left. It conveys water to the connected households up to \hat{s} . At \hat{s} the pipe-borne distribution ends with a kiosk or water taps that can be accessed by customers without pipe connection to fetch water, or by mobile water vendors. Water collecting customers are located in the interval $[\hat{s}, \dot{s}]$. Customers to the right of \dot{s} do not fetch water at the water kiosk because collection costs are too high. Instead, they buy water from vendors operating between \dot{s} and \tilde{s} . The length of the city is \bar{s} . If $\tilde{s} < \bar{s}$, some customers remain without access to safe water and they have to rely on other sources like wells or water abstraction from surface water. This scenario contradicts the sustainable development goals, and the human right to water, and must be prevented.¹⁵

Let us proceed by defining the costs of the various actors of the linear city. Roughly, the water utility incurs two cost components. The water supply costs depend on the total amount of water provided. We assume that the water provision and the water distribution exhibit constant returns to scale, i.e., the cost function is linear.¹⁶ This applies also to the distribution costs, i.e., the costs of connecting households. Thus,

¹⁴This subsection is based on Meran et al. (2020).

¹⁵The simple model does not include other sources, e.g., water wells, boreholes, or the collection of surface water. Also, we do not consider reselling from connected households and illegal tapping. However, despite the simplicity we can derive some insightful results.

¹⁶See the literature at the end of this chapter. Our results also hold if water utilities exhibit a cost structure that cannot be approximated by a linear function.

the cost function of the water utility is

$$C_{WU} = \int_0^{\tilde{s}} m ds + \int_0^{\hat{s}} k ds = m\tilde{s} + k\hat{s} \quad (4.37)$$

where m are water treatment costs and k represents distribution costs per cubic meter. Collecting water is rather cumbersome. Often it is the women who fetch water with the help of canisters. The costs relate not only to the purchase price, but also to the lost time, which is missing for other productive activities. These opportunity costs are taken into account in the cost function. For a single household located at $s \in [\hat{s}, \tilde{s}]$ to fetch 1 m^3 of water per month costs δs , where δ indicates the opportunity costs per distance walked, taking into account that the distance s has to be taken twice (to the kiosk and back). Aggregating over all households yields total fetching costs

$$D_F = \delta \int_0^{\tilde{s}-\hat{s}} s ds = \frac{\delta}{2} (\tilde{s} - \hat{s})^2 \quad (4.38)$$

Vendors' costs are twofold. There is the time vendors lose when filling cans at the kiosk and decanting for each household served. If we weigh this amount of time with the income per hour attainable in other occupations (opportunity costs), we can derive the first cost component which is symbolized by c_2 . Besides the costs unrelated to distance, there are hauling costs which depend on the distance. It takes a certain time to haul 1 m^3 over, say, 100 meter and to return to the tap for the next delivery. Let us assume that there are two customers in distance s_i from the tap and to the left of \hat{s} , i.e., they do not fetch water from the tap. Total hauling costs for the vendor are $c_1(s_1 + s_2)$, where c_1 is lost income per distance unit. This calculation also takes into account that the vendor must undertake two trips per customer.¹⁷ Then, total vending costs for two customers are $C_{ven} = 2c_2 + c_1(s_1 + s_2)$.

In the model, we have to transpose the calculation into the continuous stretch.

$$C_{ven} = c_1 \int_{\hat{s}-\tilde{s}}^{\tilde{s}-\hat{s}} s ds + c_2 \int_{\hat{s}-\tilde{s}}^{\tilde{s}-\hat{s}} ds = \frac{c_1}{2} [(\tilde{s} - \hat{s})^2 - (\hat{s} - \tilde{s})^2] + c_2(\tilde{s} - \hat{s}) \quad (4.39)$$

The first item on the r.h.s. represents the aggregated hauling costs of all customers served in the interval $[\hat{s}, \tilde{s}]$. This term is quadratic since the aggregation takes place over distances that increase successively. The second item is the purchasing and selling costs which only depend on filling and decanting time and not on distance. Hence, this part is linear.

¹⁷We assume that the vendor has a limited tank capacity so that he has to cover the distance to the kiosk for each customer supplied. Alternatively, it is possible that two vendors supply one customer each.

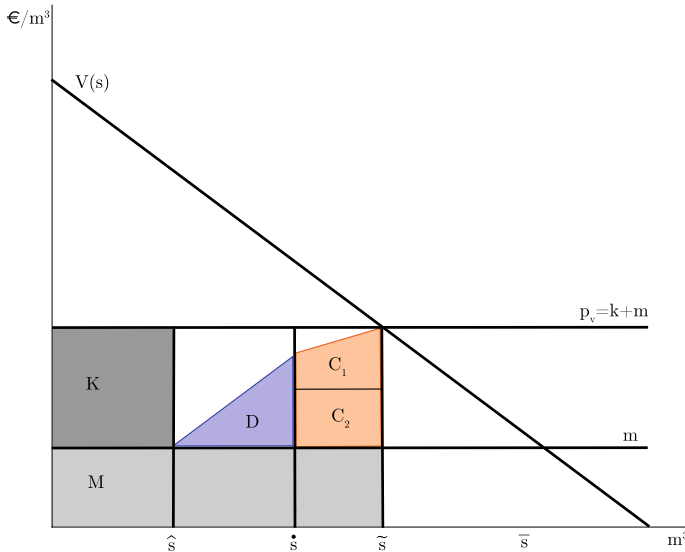


Fig. 4.12 Optimal modal split. *Source* own illustration

4.5.2.2 The Optimal Modal Split

The optimal modal split can be derived with the help of the integrated water resource management approach. How far should the pipe-borne water supply be extended, how many customers should ideally fetch water from the tap, and what distance should water vendors cover? The answer to these questions is the optimal modal split which can be derived from the following maximization program:

$$\max_{\{\hat{s}, \dot{s}, \tilde{s}\}} \left[\int_0^{\tilde{s}} V(s) ds - C_{WU} - D_F - C_{ven} \right], \quad \text{s.t.} \quad \hat{s} \leq \dot{s} \leq \tilde{s} \quad (4.40)$$

Assuming that all instrument variables are strictly positive and that all stretches are nonempty ($\hat{s} < \dot{s} < \tilde{s}$), the Kuhn–Tucker conditions¹⁸ are

$$a - b\tilde{s} - m - c_1(\tilde{s} - \hat{s}) - c_2 = 0 \quad (4.41)$$

$$c_1(\dot{s} - \hat{s}) + c_2 - \delta(\dot{s} - \hat{s}) = 0 \quad (4.42)$$

$$-k + c_1 [(\tilde{s} - \hat{s}) - (\dot{s} - \hat{s})] + \delta(\dot{s} - \hat{s}) = 0 \quad (4.43)$$

From these three equations, we can derive the optimal extension of the pipe-borne water supply, the optimal range of water fetchers, and the optimal operating area of vendors. This is called the optimal modal split of water supply consisting of $\{\hat{s}, \dot{s}, \tilde{s}\}$. Figure 4.12 depicts the optimal cost structure.

¹⁸The Kuhn–Tucker conditions are explained in Appendix A.

From 0 to \hat{s} , households are connected to the pipe-borne water supply. At \hat{s} , the water utility has installed a water kiosk where adjacent customers can fetch water. Fetching water in the range of the first yards is cheaper than having customers served by a vendor. This is due to the cost structure of both water supply modes. The vendor incurs two time-related costs, filling water into jerrycans or in canisters at the tap and at the selling point, whereas the collecting customers lose only one filling time. Of course, the household has distance-related costs δ per distance unit that are higher than the hauling costs c_1 per distance unit and m^3 of the vendor. Therefore, at \hat{s} the water supply mode switches to the vendor system. Hauling water from \hat{s} and selling it within the stretch $[\bar{s} - \hat{s}]$ is less costly than having customers in this interval collect the water at the position of the tap \hat{s} . This can be inferred from Eq. (4.42), where both marginal costs are set equal. In turn, the optimal range of connected households, i.e., the line $[0, \hat{s}]$, is determined by equalizing the respective marginal costs. Setting Eq. (4.43) into Eq. (4.42) yields $c_1(\bar{s} - \hat{s}) + c_2 + m = k + m$. The optimal modal split between the vendor's water supply range and the extension of pipe-borne water supply is determined by equalizing the respective marginal costs.

Once we have determined the optimal modal split, we can graphically represent the optimal cost structure. Areas K and M in Fig. 4.12 represent the distribution costs for connected households and costs of treating and providing the water to all households from $[0, \bar{s}]$, respectively. The triangle D represents total collecting costs of households in the line section $[\hat{s} - \bar{s}]$. Vendors' costs consist of time-related purchasing and selling costs C_2 and distance depending hauling costs C_1 .

The integrated water resource management usually applies a planning approach where economic rents are maximized taking into account technical constraints, e.g., hydrological laws. However, one must be careful when implementing this concept in practice. Two points are of particular importance.

- The pure maximization of the economic rent does not take into account the indispensable human right to water access. The result of Eq. (4.40) may lead to $\tilde{s} < \bar{s}$. If customers are excluded from the water supply system, we have to correct the optimization procedure by introducing the constraint $\tilde{s} \geq \bar{s}$. Then, we end up with a different optimal modal split that covers all customers in the linear city.¹⁹
- The planning approach sets water quantities and the line length of the various service modes in the linear city. In reality, however, consumers and also vendors are not quantity regulated, but only indirectly incentivized through prices. Therefore, it has to be clarified, which prices in the various sections of the route should be fixed or indirectly induced. The price determination in turn depends on whether the vendors are employees of the water company or whether they operate independently in a market. In the following subsection, we deal with this issue in more detail.

¹⁹An annotated exercise at the end of this chapter will lead the reader to the results (Problem 4.3).

4.5.2.3 Pricing and Regulation

Stylized Facts

In a decentralized water market consisting of collecting customers and vendors, it is very difficult to regulate all prices directly. Moreover, regulating prices may have repercussions on the market price at which vendors sell the water. Thus, to provide water to the customers of the linear city, the water utility has to follow a cautious regulation policy. In the context of our simple model, the water utility has the following instruments available: The extent of household water connections \hat{s} , the price of water at the tap point \hat{s} (we assume that the utility can differentiate between usual customers and water vendors), and the water price for connected households.

The water price in the area where vendors operate cannot directly be regulated. The water utility can try to influence the market outcome by setting a proper water tap price. For this, the regulator has to take into account the degree of competition in this market. There are many examples of highly cartelized water markets in urban, peri-urban, and rural areas. These cartels can be very effective in preventing market entry. Often, they operate beyond legal limits. In addition, both the kiosk manager and the water utility may be part of cartels. However, there are also examples where the water market around taps and in the vending area is competitive. Box 4.1 provides some empirical evidence.

Box 4.1 Small-scale water providers: Pioneers or predators?

In a study entitled “Small Scale Water Providers: Pioneers or Predators?”, Degol Hailu and colleagues (Hailu et al. 2011) empirically examine whether small-scale providers are an effective substitute for a missing pipe-borne water provision for the unconnected population, as proponents claim, or are simply predators, as the skeptics argue. The empirical investigation was conducted in Kenya, in a survey of nine communities within Nairobi city. The criteria for their choice are related to their settlement characteristics, i.e., urban or peri-urban locations, where a piped water supply is conceivable and their demographic characteristics. The survey sample comprised 576 households and about 159 small-scale water providers interviewed.

The supply side was structured as follows:

Types of small-scale water providers		
Type of provision	Sample size	Share (%)
Pushcart vendors	17	11
Tanker truck	15	09
Borehole	28	18
Tap water vendor	62	39
Water kiosk	37	23
Total	159	100

Source Hailu et al. (2011)

The above table shows that the fixed-point water suppliers (tap water vendors and water kiosks) make up 62 percent of total water supply. The mobile suppliers (pushcart vendors and tanker trucks) make up 20 percent of providers, receiving the water from Nairobi City Water and Sewerage Company (NCWCS). Pushcart vendors supply water by manual and donkey-pulled pushcarts and obtain water mostly from boreholes, water kiosks, or through an illegal connection to the piped network. Tanker trucks supply water in bulk to end users who possess storage tanks. Some of these households resell the water. Tanker trucks obtain water either from private boreholes or directly from the utility company. The remaining water supply comes from borehole water vendors. This water is often unsafe regarding quality. However, the advantage of vendors for households is that they lead to time savings. The opportunity costs of time are very high for households.

Small-scale water providers are not price takers who take the water price for given. Rather, they set the water price directly taking into account the price behavior of competitors. The study also showed that pricing followed a cost markup approach.

Mean water price across small-scale water providers

Type of provision	Mean price	Poverty premium
Pushcart vendor	12.15	30.28
Tanker truck	7.90	18.75
Borehole	6.11	14.28
Tap water vendor	3.18	6.95
Water kiosk	2.81	6.03

Source Hailu et al. (2011)

The above table shows that push car vendors charge on average the highest water prices. If you put this price in relation to the water price for connected households, you obtain the poverty premium, calculated as $(p_i/p_{NCWCS}) - 1$ where the index i refers to the types of small-scale water providers and NCWCS is the Nairobi water utility. The official rate NCWCS is Kshs 0.40 per 20l (1 € \approx 120 Kshs). The poverty premium of push car vendors is 30, which means that low-income households are paying 30 times as much as connected households are being charged.

Competitive Versus Cartel Solution

We start with the assumption that the water market for vendors is fully competitive. They offer water in the segment $[\underline{s}, \bar{s}]$. At each location along this segment, vendors have the same costs to supply one m^3 of water.²⁰ Take two points in close vicinity, s_1 and s_2 , where $s_1 < s_2$ (see Fig. 4.13, left half of the picture.). Costs per m^3 of water

²⁰Recall the assumption that at each point along the linear city customers want to buy one m^3 of water.

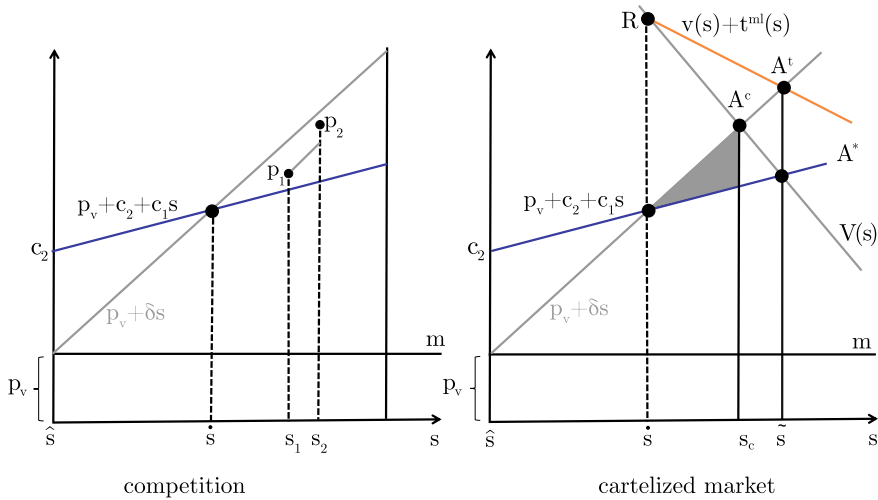


Fig. 4.13 Competition versus cartel. *Source* own illustration

sold are

$$p_v + c_2 + c_1 s_1 < p_v + c_2 + c_1 s_2 \tag{4.44}$$

where p_v is the price for the vendor at the tap. If a seller at s_2 wants to sell the water for a price higher than these costs, lets say p_2 , either another seller occurs at the same point undercutting this prices²¹ or she offers the water at point s_1 where the first seller does not operate. But if the price at s_1 is such that $p_1 + \delta(s_2 - s_1) < p_2$ the first seller would lose the demand at s_2 because customers at s_2 will move to s_1 to buy the water from the vendor there. Due to the mobility of customers, all selling points are in competition. It does not matter whether we have many or only two vendors in each segment. The price competition drives all prices down until

$$p(s) = p_v + c_2 + c_1 s \tag{4.45}$$

Hence, the price is depending on distance and follows the costs defined in Eq. (4.45). Notice that the highest price is reached at $p(\tilde{s}) = p_v + c_2 + c_1 \tilde{s}$. Here, a customers willingness to pay is equal to the price, i.e., $V(\tilde{s}) = p(\tilde{s})$. Beyond \tilde{s} , there is water supply by vendors because price would be higher than the willingness to pay.

That the water prices correspond to costs per m^3 follows from the assumed openness of the water market which allows newcomers to enter the market. Openness is not only due to the absence of legal constraints but also a matter of the very nature of costs. If entry and exit are costless, markets are contestable. This is the case when no costs are sunk for firms leaving the market, and no investment into specific capital is necessary for entering the market.²² This is not always the case in unconnected

²¹Competition in prices is called Bertrand competition.

²²The concept of contestability was introduced and elaborated by Baumol (1982).

water markets served by vendors. Vendor's investment exhibits a certain degree of specificity (jerry cans, hand trolley with a tank mounted on it, etc.).

However, if a sunk cost structure is more or less absent, the number of firms operating in the market is not of importance. Low prices are simply the result of the competition of potential newcomers. If an incumbent charges higher prices than marginal costs (see Eq. (4.45)), newcomers immediately invade the market at s driving down the vending price until the equilibrium is reached. Since incumbent vendors anticipate the potential threat by newcomers, they keep their prices equal or close to marginal costs. Thus, even if there are only few vendors serving customers, the pressure of potential competition drives the price down to marginal costs.

The water prices customers and vendors have to pay at the kiosk are still left to be fixed. After the implementation of the optimal water supply infrastructure, i.e., setting the optimal \hat{s} , the policy-maker has to assure that the division of collecting and vending zones follows the optimal pattern. In other words, she has to determine the optimal \hat{s} . This must be achieved indirectly by fixing the water price at the kiosk.

Unconnected households decide either to buy water from water taps or to purchase it from water vendors. The marginal customer is indifferent between both options, i.e.,

$$p_{col} + \delta(\dot{s} - \hat{s}) = p(\dot{s}) = p_v + c_2 + c_1(\dot{s} - \hat{s}) \quad (4.46)$$

where p_{col} (p_v) is the water price charged to the collecting customer (vendor). From Eq. (4.42), optimality of \dot{s} requires that $p_{col} = p_v$, i.e., a price discrimination policy between water collectors and vendors would be non-optimal. Finally, we can determine the level of the tap water price. From Eq. (4.41), it follows that the marginal willingness to pay must be equal to the marginal costs $c_1\tilde{s} + c_2 + m$. The water price vendors charge at \tilde{s} is $p(\tilde{s}) = c_1(\tilde{s} - \hat{s}) + c_2 + p_v$. Hence, the optimal tap water price is $p_{col} = p_v = m$.

Finally, we have to examine the economic viability of the price system. Total costs of the water utility including transport and distribution costs should be covered by the revenue raised. Cost coverage is defined as

$$p_{ch}\hat{s} - k\hat{s} - m\tilde{s} + p_v(\tilde{s} - \hat{s}) \quad (4.47)$$

where p_{ch} is the water price for connected households, which can be rewritten as

$$p_{ch}\hat{s} - k\hat{s} - m\hat{s} - m(\tilde{s} - \hat{s}) + p_v(\tilde{s} - \hat{s}) = p_{ch}\hat{s} - k\hat{s} - m\hat{s} \quad (4.48)$$

since $p_v = m$. Setting $p_{ch} = k + m$ leads to cost coverage of the water utility. Notice that $p_{ch} > p_v$, i.e., water for connected households is more expensive than for collectors at the tap. This might lead to trade between the kiosk manager and the adjacent connected households.

So far we have assumed that the market for water vending is competitive. There is a growing literature from scientists, development workers, and journalists about

the role of small-scale private water providers in non-competitive markets.²³ Proponents of the mobile and decentralized supply of water see vendors as pioneers and entrepreneurs that supply water to those, who otherwise would never have access to a reliable water source, even though they charge prices, well above costs. However, skeptics see the vendors solely as predators, who exploit the poor by charging high prices for water of poor quality.

Regulatory Options

Anyway, it is to be expected that water provision via a decentralized infrastructure of kiosks and mobile vendors will keep on persisting in the future. This leaves us with the following question: How can one force cartelized vendors to supply the poor at an affordable price? There are two approaches to reduce the negative economic and social effects of water supply cartels. The decision on which of the two should be adopted depends on the authority's compliance-monitoring capacity.

The first approach simply consists of introducing a zonal price cap for water vendors. The price²⁴ $\bar{p}(s)$ per, say, liter is optimally set, such that

$$\bar{p}(s) = p_v + c_2 + c_1(s - \hat{s}), \quad \text{for } s \geq \hat{s} > \hat{s} \quad (4.49)$$

If the vendors' cartel complies with the regulated price, the optimal modal split is replicated and customers are not exploited. Of course, this type of regulation is only enforceable if the authorities are capable to monitor the price of vendors. How can it be determined whether the supplier does not charge a too high price if customers complain? One way could be to issue water coupons that people can buy for $\bar{p}(s)$ at an issuing office of the water authority.²⁵ They can buy as many as they want. Vendors have to accept these coupons in exchange for water. The collected coupons can be redeemed at the same public authority office for the same price. Of course, this mechanism only works if the cartel is not able to force further receipts in addition to the coupons. This would be the case if the cartel could charge an effectively higher price than the coupon price for the water. If the public institutions are not able to prevent the abuse of this system, price regulation is undermined and the water customers pay a water price that an unregulated monopolist would demand.

Thus, to guarantee access to water for an affordable price, the regulation approach should rely much more on economic incentives instead of direct price regulation, which cannot be enforced if the necessary institutional capacity does not exist. The economic incentives must be set such that the desired regulatory effect on water prices is self-enforcing. Low prices for poor customers must be in the interest of cartelized vendors. This is a second more promising approach.

To develop an incentive scheme, we first have to ask what prices prevail in a cartelized vendor market. Cartelized vendors behave like a monopoly. They fix the

²³See Hailu et al. (2011) and the further readings at the end of this chapter.

²⁴As the linear city model is continuous with respect to the distance s , zonal pricing is expressed by a continuous price function.

²⁵The coupon can also be bought from authorized shops located along the stretch of the linear city.

prices along the segment beginning from \hat{s} to the end of their area of operation. We have assumed in our simple model that each household along the linear supply line demands 1m^3 of water. The maximum price they are willing to pay is

$$p_{max}(s) = p_{col} + \delta(s - \hat{s}), \quad \text{for } s \geq \hat{s} > \hat{s} \quad (4.50)$$

This case is depicted in the right section of Fig. 4.13. The profit depends on the supply distance:

$$\int_{\hat{s}-\hat{s}}^{s_c-\hat{s}} [p_{col} + \delta s - p_v - c_2 - c_1 s] ds \quad (4.51)$$

where s_c is the stretch the cartel is willing to supply. This distance can be derived from the maximizing behavior of the cartel. The cartel maximizes Eq. (4.51) with respect to $\{\hat{s}, s_c\}$ subject to the constraint

$$V(s_c) = a - bs_c \geq p_{col} + \delta(s_c - \hat{s}) \quad (4.52)$$

The charged price cannot be higher than the willingness to pay. The KKT conditions are

$$-\delta(\hat{s} - \hat{s}) + c_2 + c_1(\hat{s} - \hat{s}) \leq 0 \perp \hat{s} \quad \geq 0 \quad (4.53)$$

$$\delta(s_c - \hat{s}) - c_2 - c_1(s_c - \hat{s}) - \lambda(b + \delta) \leq 0 \perp s_c \quad \geq 0 \quad (4.54)$$

where λ is the Lagrangian to the constraint.²⁶

The picture in the right half of Fig. 4.13 shows the solutions to Eqs. (4.53) and (4.54). \hat{s} is chosen such that it equates the costs of water fetching to the costs of supplying customers by vending. The distance of supply s_c is chosen such that the marginal willingness to pay is equal to the price $p_{max}(s_c)$ charged by the cartel (See point A^c). In sum, the cartel will earn profits indicated as gray triangle in Fig. 4.13.

The water utility can try to reduce the cartel's prices $p_{max}(s)$ by reducing the price p_{col} , for which it sells water at the kiosk to households. However, this would lead to an inefficient cost structure. Decreasing the water price for collecting households (and at the same time leaving the purchasing price for vendors constant) leads to a reduction of consumer surplus exploitation, but at the same time results in an inefficient supply structure because the collecting segment is too large while the vending area is too narrow. Of course, it is an effective policy with regard to customer protection, but as we have argued an inefficient solution.

An alternative indirect mechanism is to subsidize customers in the vending area such that the effective marginal willingness to pay increases. This subsidy depends on the location of customers, i.e., it is a zonal subsidy²⁷:

$$t_{LM}(s) = (\delta - c_1)(s - \hat{s}) - c_2, \quad \text{for } s \geq \hat{s} > \hat{s} \quad (4.55)$$

²⁶Notice that $\lambda > 0$, otherwise Eqs. (4.53) and (4.54) are identical leading to $s_c = \hat{s}$. But this would imply that the monopolized vendor (cartel of vendors) does not sell water.

²⁷This subsidy system is based on a regulatory mechanism proposed by Loeb and Magat (1979) to regulate monopolies.

Hence, the effective marginal willingness to pay is $V(s) + t_{LM}(s)$ (see the orange line in Fig. 4.13). This line rotates around the point R until it intersects with $p_v + \delta(s - \hat{s})$ at A^t . This subsidy driven increase in the willingness to pay expands the vending segment to \tilde{s} , i.e., brings the vending area back to its efficient extent.

What changes in comparison to the water market without subsidies? First, the modal split is efficient, i.e., water is supplied up to \tilde{s} , and, second, consumers are protected by the subsidy introduced, i.e., they effectively pay a water price under full competition. This can be seen by the following calculation utilized by Eqs. (4.50) and (4.55):

$$p_{max}(s) - t_{LM}(s) = p_{col} + \delta(s - \hat{s}) - (\delta - c_1)(s - \hat{s}) + c_2 = p_{col} + c_1(s - \hat{s}) + c_2 \quad (4.56)$$

where $p_{col} = p_v = m$. This elegant mechanism comes at a price. The solution is rather expensive and, from a political standpoint, provocative. Customers must be subsidized and, at the same time, the cartel reaps a monopoly rent. Further, the issue of financing the subsidy remains.

4.6 Water Scarcity: Prices Versus Rationing

4.6.1 Options to Deal with Scarcity

The importance of water demand management increases in times of water scarcity. There are now many examples of how demand-side management can be designed. For example, California has developed numerous conservation strategies to reduce water demand, inter alia by utilizing pricing schemes, subsidies for water-efficient equipment, educational measures, water rationing, and water trading. Similarly, Australia has taken measures to cope with severe drought by developing a mix of water instruments to reduce effective demand for water. For instance, the Cairnes regional council has launched a campaign to use water wisely (information on water-saving behavior for households) in addition to mandatory restrictions (regulated sprinkling times).²⁸ In this section, we analyze the characteristics of water demand management based on prices vis a vis a non-price approach. The results of this comparison depend strongly on the evaluation criteria applied. In principle, the allocation of scarce water should comply with various criteria. The literature mentions efficiency, justice, technical feasibility, political enforceability, and environmental sustainability. The importance of each criterion depends on the situation under consideration. If, for example, water scarcity leads to a pronounced plight of the population, the criterion of just allocation of water is of greater importance than allocative efficiency, i.e., an allocation according to the marginal willingness to pay.

²⁸<http://www.cairnes.qld.gov.au/water-waste-roads/water/save-water>.

4.6.2 Rationing

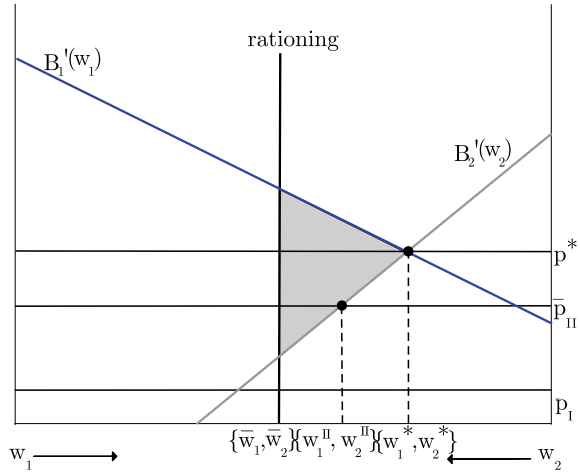
We know from war periods that the distribution of basic goods is often done through food stamps that cannot be transferred. This applies also to other emergency situations, e.g., water scarcity, where personal rationing is often viewed as a just allocation procedure. There are also other forms of rationing: In many poor urban areas in developing countries, water supply to households is rationed by interruption. Often, households receive water only for about 1–2 h per day. This can be understood as a non-price demand management approach sometimes deliberately chosen by local authorities to meet fairness criteria; sometimes, it is the result of aging network pipes and weak institutional management structures. In the following, we will introduce various water rationing methods in practice and compare them with a price-based water demand management approach.

Rationing can take up various forms²⁹:

- Rationing by fixed allotment: The scarce resource is distributed in fixed quantities to individuals or households depending on the household size. These allotments can vary seasonally depending on the scarcity situation. Also, the portions allotted can vary with respect to membership to specific economic sectors, e.g., industry, commerce, or the public area (school, etc.). It is important to differentiate the allotments with regard to their transferability. There exist various designs: allotments without transferability, allotments with transferability under regulated prices, and allotments completely tradeable in a free market. We will analyze these specifications below.
- Proportional allotments: Water use rights are allotted in proportion to water usage prior to the rationing. This method is easy to implement because it does not require lots of information on the characteristics of the user. The water utility simply needs a record of the historic water use profile to determine the allotment. Despite the relatively small amount of information required, this method suffers from some implementation problems. First, it is very difficult to allot water use rights to newcomers without a historic record; second, the reliance on historic water use may discourage water conservation. If water users know that the historic water use is utilized to build the distribution key for water use rights, users they will behave strategically by deliberately wasting water.
- Water rationing by increasing block tariffs: We have already analyzed increasing block tariffs. Block prices are only valid within the boundaries of a block. If a household wants to consume more water than allowed for the given block price, it exceeds the upper boundary and pays the higher prices of the following block. In this sense, increasing block tariffs also exhibit a rationing property.
- A more differentiated version of this approach is water budgets as applied in the USA. Each household gets a monthly water budget assigned, which is based on several characteristics, including the number of residents in the home, or the usage

²⁹We follow Lund and Reed (1995) and also Olmstead and Stavins (2009b).

Fig. 4.14 Pricing versus rationing. *Source* own illustration



type, e.g., indoor water use, garden sprinkling, etc. The effective price depends on the percentage utilization of the allocated water budget. If the household exceeds its water budget, the respective price increases.

- Water can also be rationed with respect to the type of water use, e.g., car washing, garden watering or luxurious applications like a fountain, etc. In times of severe drought, the municipality simply forbids certain types of water use which are in the higher part of the need hierarchy.
- Rationing by outage: This method is often applied in developing countries because it needs no institutional body that gathers and calculates the necessary information from the customers. Either water is provided only few hours a day or the water pressure is reduced. There is almost no differentiation among households except, perhaps, by rotating the outage geographically among districts allowing differentiated service times.

4.6.3 Comparison

In many cases, water scarcity can be managed by a deliberate pricing policy. If water gets scarce, simply increase the water price and the water allocation will take place in an efficient manner. Or one simply introduces a market where water can be traded without any regulation. Conversely, water rationing leads to welfare losses because the water is not allocated according to the marginal willingness to pay of customers.³⁰ But, as mentioned above, efficiency is only one of many criteria that have to be taken into account to find an allocation which is capable of approval by residents. Figure 4.14 shows the constellation for two users.

³⁰There are empirical studies estimating the exact amount of welfare losses. See, e.g., Grafton and Ward (2008).

Let us begin with the assumption that there are no water treatment costs. The sustainable water availability is given by \bar{W} which is the width of the diagram. From left to right, the water allocation of user 1 can be read up and for user 2 from right to left. If water is allocated within an unregulated market allocation, the result is $\{w_1^*, w_2^*\}$. The respective equilibrium price is p^* . We know from Chap. 3 that $\{w_1^*, w_2^*\}$ is optimal in the sense that this allocation maximizes the aggregated benefits of both users. Note that the positive water price mirrors solely the scarcity of water, whereas water treatment and distribution costs are not yet considered. However, this allocation could be regarded as unjust. In a situation of scarcity, the willingness to accept inequality is reduced. As an alternative, rationing could be implemented. In this figure, we assume that both users get the same amount of water, i.e., rationing by fixed allotment. If these allotments are not transferable, the final allocation is $\{\bar{w}_1, \bar{w}_2\}$. From a traditional welfare theoretical point of view, this allotment is connected to welfare losses (gray triangle).

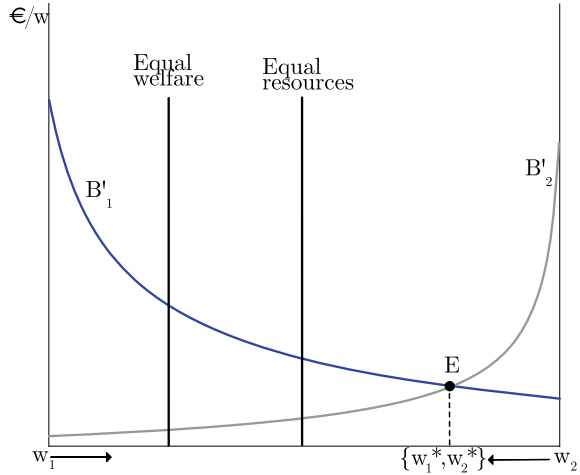
Some rationing schemes are combined with tradability either under regulated prices or with completely free pricing. The figure depicts two cases of price regulation. If water rations are transferable and the regulated price is \bar{p}_I , we can see immediately that no trade will occur. The price is too low for user 2 to sell some of his water allotments to user 1. If the regulated price is \bar{p}_{II} , restricted trade takes place leading to the allocation $\{w_1^{II}, w_2^{II}\}$. Only if the price is completely free, we reach the optimal allocation $\{w_1^*, w_2^*\}$. The key difference between the pure market solution and the rationing scheme with transferability is

1. that the scarcity rent accrues to the customers. This is also the case if water treatment costs are taken into account. Let the horizontal line \bar{p}_I represent marginal treatment costs and charge customers with a uniform tariff in the amount of marginal costs and distribute water coupons³¹ according to the water allotment. Customers pay the tariff and get credits (debits) depending on whether they use less (more) water than their allotments. The coupon price emerges on the trading floor for scarce water. In Fig. 4.14, the scarcity rent, i.e., the trading price for water coupons, is equal to $p^* - \bar{p}_I$. Since coupons have been distributed, the scarcity rents remain with the user, for user 1 $(p^* - \bar{p}_I)w_1^*$ and for user 2 $(p^* - \bar{p}_I)w_2^*$;
2. that efficiency and equity are no longer in a tight trade-off relationship. If water rights are equitably distributed, trading leads to higher total welfare without harming the principle of fairness,³² at least at the first glance. The final judgment depends on how fairness is defined. Do we refer to the distribution of resources, in our case water, or do we refer to the result of the allotment in terms of utility that accrues to customers? In the following, we take a closer look at this distinction.

³¹This proposal was made by Collinge (1994). Of course, these coupons need not be physically distributed to customers. They can be handled electronically on the individual account of customers.

³²See the discussion about the various fairness criteria in Chap. 3.

Fig. 4.15 Equality: welfare or resources. *Source* own illustration



4.6.4 Discussion

In the philosophical and in the economic literature, one can find a discussion of what is the right “equalisandum” of just distribution.³³ Do we want fairness through the equal distribution of resources, or does fairness refer to welfare equality? The issue can best be described with the help of Fig. 4.15.

Consider two farmers producing crop that needs water. From left to right, the marginal profit function of farmer 1 is depicted ($B'_1(w)$) and, vice versa, from right to left marginal profits of the second farmer ($B'_2(w)$). Obviously, farmer 1 is more productive than farmer 2. With the same amount of water, farmer 1 makes more profit than farmer 2. Since water is scarce, the question of how water should be allocated arises. If one follows the usual welfaristic approach, the allocation of point E is optimal maximizing aggregated profits leading to $B_1(w_1^*) > B_2(w_2^*)$. But if only fairness considerations matter, aggregate profits are irrelevant as a criterion of fairness. If one adheres to the principle of equality of resource distribution, each farmer gets half of it. But what if the equality of profits (welfare) counts? Then, the water allocation must be asymmetric in favor of farmer 2 who is less productive compared to farmer 1. The line named “Equal welfare” exactly depicts the resource allocation where both profits are equal. But why should profits be equal? Because it is just that both individuals bear the burden of water scarcity equally. Therefore, trade of water after allotment of water rights is often not allowed, because it can dilute fairness, even if we would observe a Pareto improvement in the case of allotment trade. This is the reason why water allotment is sometimes not transferable.³⁴

But why does the lower productivity of farmer 2 entitle him to receive more water? Much depends on the causes of the lower productivity. If farmer 2 is poor without

³³Roemer (1996) provides a thorough analysis.

³⁴In Exercise 4.5, the reader finds an analysis of the question if the efficiency criterion is always in conflict with the principle of fairness.

access to the capital market to finance a modern technology and farmer 1 has the opportunity to invest in water-saving production methods, then the uneven allocation of water might be justified. But what if farmer 2 is simply lazy or ignorant? In this case, the water allocation according to equal profits is not fair. Thus, it depends much on the water user's responsibility. If farmer 2 is not accountable for his low productivity, the fair water allocation might follow the principle of equalization of profits (or utility). If society believes in individual responsibilities, then the fair distribution will relate more to the resource side.

Box 4.2 The water-wise rules

Many municipalities in Australia have implemented so-called water-wise rules. These rules which are in fact regulations are aimed to save water in the everyday life of households. For instance, Sidney Water prescribes that households must use hoses fitted with a trigger nozzle, sprinklers, and irrigation systems when irrigating the garden. The irrigation time is restricted from 4 pm until 10 am. Breaching this rule can lead to a fine of 220 \$ for households. This regulation is a typical non-price approach and can be considered as a soft form of rationing. It prescribes a technological standard. Its very aim is to save water by increasing the efficiency of water use. Trigger nozzles slow down the water current through the hose. Per time unit less water gets distributed into the garden. As a result, there is no wasted water in the form of ponds evaporating into the air or runoffs. The water from the hose gets to the roots of the plants, with less unproductive water loss.

However, one has to be careful when implementing this kind of water-saving technology. Often the water-saving effects have been smaller than expected. The reason for that lies in behavioral changes that partially offset the efficiency effect of the water-saving device. The implicit water price decreases with improved water efficiency. This leads to the so-called rebound effect which is also very well known from energy consumption. To explain this effect, we apply a very simple microeconomic model.

Households derive utility from a blossoming garden. Let us assume that the extent and the intensity of the vegetation depend positively on the amount of irrigated water (of course, we exclude over-irrigation). In turn, the quality of the vegetation, its beauty, and its range creates benefit to households. If we take into account this interrelation, we express the benefit as a function of the effective water w_e used, i.e., the water that reaches the vegetation. The relation between water from the tap w and effective water use depends on the efficiency of the irrigation technology. We have

$$w_e = w\epsilon \quad (4.57)$$

where ϵ is the technological productivity. If $\epsilon = 1$, there is no water loss. In the simplest case, households maximize their benefit with regard to the amount

of water used for irrigation.

$$\max_w B(w\epsilon) - pw \quad \text{or} \quad \max_{w_e} B(w_e) - pw_e/\epsilon \quad (4.58)$$

where p is the water price and $p\epsilon$ is the effective price, i.e., the price per liter water reaching the plants. The optimality condition is

$$B'(w_e) = p/\epsilon \Rightarrow \hat{w}_e(\epsilon) \quad (4.59)$$

where $\hat{w}_e(\epsilon)$ is the effective water use. A rising ϵ leads to more consumption of water due to the decrease of the effective water price. This is the so-called rebound effect.

On the other side, the increased water efficiency lessens the water use which can be derived from Eq. (4.57). Both effects then determine the water consumption w .

$$\frac{dw}{d\epsilon} = \frac{d\hat{w}_e(\epsilon)/d\epsilon \cdot \epsilon - \hat{w}_e(\epsilon)}{\epsilon^2} \quad (4.60)$$

The first term on the right-hand side is the rebound effect, the second term is the counter-directed efficiency effect. That is the reason why the water-saving effect of increased water productivity is less than the calculated efficiency effect.

There are other examples where the rebound effect appears. Households with low-flow showerheads take longer showers. The “double flush” was observed when households installed low-flow toilets.

Source: Olmstead and Stavins (2009b), www.sydneywater.com.au

4.7 Exercises

Exercise 4.1 Designing an increasing two-block tariff

Increasing block tariffs are designed to allow poor income groups to access water. Often, the price of the first block is below marginal cost so as to render the access possible to even the poorest households. As a consequence, the upper income groups have to provide the necessary cross-subsidies to let the water utility break-even. There are four poor households and one household with sufficient income to consume water beyond the lifeline. The demand of the wealthy household can be captured by the price-quantity function $p_2(w_2) = 12 - (w_2 - w_s)$ where $w_s = 10$ is the lifeline. The task of water utility management is to secure the access to water for the poor by minimizing the water price in the first block and to assure at the same time that

no deficit occurs. Let us assume that fix costs are $F = 20$. Marginal costs are set at $c = 2$.

First, the utility maximizes the contribution margin CM_2 in the second block. The contribution margin is defined as

$$CM_2(w_2) = (p_2(w_2) - c)(w_2 - w_s) = (12 - (w_2 - w_s) - c)(w_2 - w_s) \quad (4.61)$$

To maximize $CM_2(w_2)$, we have to set the first derivation equal to zero and solve for w_2 .

$$12 - 2(w_2 - w_s) = c \Rightarrow w_2^* = 15 \quad (4.62)$$

Reinserting the result into the price function yields $p_2^* = 7$. The maximum surplus which can be extracted in block 2 is therefore $CM_2^* = CM_2(w_2^*) = (7 - 2)(15 - 10) = 25$.

To calculate the price of the first block, we have to ensure that the surplus of the second block covers the deficits of the first, that is to say

$$(4 + 1)(c - p_1) + F = CM_2^* \Rightarrow p_1^* = c - \frac{(F - CM_2^*)}{(4 + 1)} \quad (4.63)$$

Inserting the numerical values yields $p_1^* = 1$.

Exercise 4.2 Universal service provider

In many countries, the universal service obligation requires water utilities to provide water to spatially distinct customer groups at the same tariff. The tariff design does not reflect the differing connection costs of customers. Assume that marginal costs of supplying consumer group 1 are 2 € per m^3 , and the marginal costs of water provision to group 2 are 4 € per m^3 . To keep the calculations simple, assume that both groups have the same size $n = 1$ and that their marginal willingness to pay is identical, i.e., $p(w) = a - bw$, where $a = 10$ and $b = 0.5$. Furthermore, we assume that fixed costs can be covered by a uniform and constant access fee. Thus, it remains to fix the volumetric part of the tariff. If we follow the welfare-oriented approach of IWRM, we maximize the aggregated willingness to pay to determine the optimal allocation. This is a straightforward exercise requiring to set the marginal willingness to pay equal to marginal costs.

$$a - bw_1 = c_1 \Rightarrow w_1 = 16 \quad (4.64)$$

and

$$a - bw_2 = c_2 \Rightarrow w_2 = 12 \quad (4.65)$$

Hence, setting zonal prices such that $p_1 = c_1 = 2$ and $p_2 = c_2 = 4$ will maximize total economic rent. However, this is in contrast to the principle of universal service obligation that requires an equal treatment of both groups. Thus, we have to find the cost covering volumetric price, i.e., the price that covers all operating costs (fixed

costs are covered by the access fee). This price can be calculated by deriving the marginal willingness to pay from the demand functions $\hat{w}_i(p) = (a - p)/b$, $i = \{1, 2\}$. Cost coverage requires

$$(p - c_1)\hat{w}_1(p) + (p - c_2)\hat{w}_2(p) = (2p - c_1 - c_2)\hat{w}_2(p) = 0 \quad (4.66)$$

which yields $p = 3$. Each group is charged 3 € per m^3 whereby group 1 (2) pays more (less) than their marginal costs. Hence, group 2 is cross-subsidized by group 1. From a welfare theoretical viewpoint, this price is suboptimal. If one places emphasis on equal treatment of customers and if no other redistributing instruments are available, the implementation of the universal service obligation principle has its price in the form of welfare losses.

Exercise 4.3 Optimal modal split

Imagine a village with residents living along a straight road that runs from west where high- and middle-income people live to east where poor people live. In the west, there is a water utility, which processes water and distributes it to the inhabitants through an underground pipeline. However, the waterworks are still in an investment phase, and it is necessary to consider how many households are to be provided with a water connection and how many households are to be supplied by water kiosks and mobile water sellers. Residents are distributed evenly along the main road. The length of the road (linear city) where households dwell is $\bar{s} = 200$ length units, say, measured in 100 m. Water consumption per household is assumed to be $1 \text{ m}^3/\text{month}$. The willingness to pay for water depends in our example solely on income. We assume that water demand is completely price inelastic. The income distribution is reflected in the geography (west: upper incomes, east: lower incomes). The willingness to pay for water is geographically distributed according to the function $V(s) = a - bs$, where $a = 50$ and $b = 0.25$. Let us assume that treatment cost per m^3 water is $m = 0.5$ and distribution cost for connected households is $k = 16$ per length unit. Collecting costs refer to time costs and are $\delta = 1.5$ per length unit. Vendors' costs consist of purchasing/selling costs $c_2 = 4$ per m^3 and hauling costs $c_1 = 0.5$ per length unit.

The first task is to derive the optimal modal split as specified in Eq. (4.40). If you insert the numerical values of the parameters and solve the equation system Eqs. (4.41)–(4.43), you get the optimal values $\{\tilde{s} = 134, \dot{s} = 114, \hat{s} = 110\}$. From these values, we can derive the length of the collecting segment, i.e., $\dot{s} - \hat{s} = 4$ and the vending area $\tilde{s} - \dot{s} = 20$. The pure economic approach based on the marginal willingness to pay leads to an under-supply of the village. Households along the stretch of $200 - 134 = 66$ length units are not provided with water from the utility and have to take care of themselves.

If we follow the Social Development Goal 6, and its aim to give access to safe water sources for everybody, the water infrastructure has to be enlarged. From Eq. (4.42), it follows that $\dot{s} - \hat{s} = c_2/(\delta - c_1) = 4$, i.e., the collecting area is independent of the total length of the village. To calculate the length of the segment where households are connected, we take Eq. (4.43) and substitute for \tilde{s} the total stretch of the village, i.e., 200. Solving the equation yields $\hat{s} = 176$. Since $\dot{s} - \hat{s} = 4$ we have in this case

$\dot{s} = 180$. Thus, the vendor stretch $\bar{s} - \dot{s}$ is still 20. What we see is that the full coverage of water supply was solely achieved by enlarging the area of connected households.

What happens when vendors are cartelized and act like a monopolist? We have shown that in this case water prices in the vending area $s \geq \dot{s} > \hat{s}$ follow exactly the marginal collecting costs, i.e., $p_c = m + \delta(s - \hat{s})$. The cartel will equate the willingness to pay of the marginal customer at s_c with the collecting costs, i.e., $V(s)a - bs = m + \delta(s - \hat{s})$ which yields $s_c = 122.571$ which is less than the optimal value $\bar{s} = 134$ (see point A^c in Fig. 4.13).

The subsidy that makes $s_c = \bar{s}$ can be calculated from Eq. (4.55): Inserting the given numerical values of all parameters yields $t_{LM} = (s - \hat{s}) - 4$.

Exercise 4.4 Seasonal pricing

Suppose there are two time periods 1 and 2 where we have the water consumption w_1 and w_2 , respectively. Period 1 is the winter period (off-peak period), while period 2 is the summer period (peak period). The demand functions (which are equal to the marginal benefit functions) in both periods are

$$p_1(w_1) = 110 - w_1 \quad p_2(w_2) = 150 - w_2$$

The cost rate for the delivery of one amount of water is given with $c = 10$, while the cost rate for the provision of one unit of capacity is $r = 20$. In the optimization, we want to calculate the optimal consumption and capacity levels in the way that we maximize the total surplus of 1 year, which includes the peak and off-peak period:

$$\begin{aligned} & \max_{\{w_1, w_2, k\}} [B_1(w_1) + B_2(w_2) - c \cdot (w_1 + w_2) - r \cdot k] \\ & s.t. \quad w_1 \leq k \quad (\lambda_1) \\ & \quad \quad w_2 \leq k \quad (\lambda_2) \end{aligned}$$

Of course, the water delivery in both periods is restricted by the chosen capacity level which is addressed in the constraints of the optimization problem. Based on the optimization problem, the Lagrangian function can be set up:

$$L = B_1(w_1) + B_2(w_2) - c \cdot (w_1 + w_2) - r \cdot k + \lambda_1 \cdot [k - w_1] + \lambda_2 \cdot [k - w_2]$$

and finally the KKT conditions can be formulated:

$$B'_1(w_1) - c - \lambda_1 \leq 0 \perp w_1 \geq 0 \quad (4.16)$$

$$B'_2(w_2) - c - \lambda_2 \leq 0 \perp w_2 \geq 0 \quad (4.17)$$

$$\lambda_1 + \lambda_2 - r \leq 0 \perp k \geq 0 \quad (4.18)$$

$$k - w_1 \geq 0 \perp \lambda_1 \geq 0 \quad (4.19)$$

$$k - w_2 \geq 0 \perp \lambda_2 \geq 0 \quad (4.20)$$

We suppose that capacity is only exploited in the summer month; hence, we have to assume $w_1 \geq 0$, $w_2 \geq 0$, $k \geq 0$, $\lambda_1 = 0$ and $\lambda_2 \geq 0$. Therefore, we can calculate

$$\begin{aligned}(k) : \lambda_2 &= r = 20 \\(w_1) : B'_1(w_1) &= c \rightarrow 110 - w_1 = 10 \rightarrow w_1 = 100 \\(w_2) : B'_2(w_2) &= c + r \rightarrow 150 - w_2 = 30 \rightarrow w_2 = 120 \\(\lambda_2) : k &= w_2 = 120\end{aligned}$$

The solution does not violate the constraint Eq.(4.19), which states that $w_1 \leq k$ (hence: $100 \leq 120$). Therefore, we found no contradiction in the KKT conditions and this case leads to optimality. The prices in the winter/(off-peak) period (p_1) and summer/(peak) period (p_2) are

$$p_1 = B'_1(w_1) = c = 10 \quad p_2 = B'_2(w_2) = c + r = 30$$

For this case, the capacity is completely financed by the revenues from the summer month (peak period).

Suppose now that the capacity cost rate increases to the level of $r = 50$. Having the same assumption as before, $w_1 \geq 0$, $w_2 \geq 0$, $k \geq 0$, $\lambda_1 = 0$ and $\lambda_2 \geq 0$, where we suppose that capacity is only exploited in the peak period, we get the following results:

$$\begin{aligned}(k) : \lambda_2 &= r = 50 \\(w_1) : B'_1(w_1) &= c \rightarrow 110 - w_1 = 10 \rightarrow w_1 = 100 \\(w_2) : B'_2(w_2) &= c + r \rightarrow 150 - w_2 = 60 \rightarrow w_2 = 90 \\(\lambda_2) : k &= w_2 = 90\end{aligned}$$

The change of the capacity cost rate does not impact the consumption level in the winter period ($w_1 = 100$). However, due to the increase of the capacity cost rate, the consumption in the summer period decreases to the level of $w_2 = 90$. Therefore, we do not meet the constraint Eq.(4.19), which states that $w_1 \leq k$, because the consumption level in the winter period ($w_1 = 100$) is higher than the chosen capacity level ($k = 90$). Due to the contradiction, this case does not lead to optimality.

Therefore, we suppose that capacity is exploited during the entire year. Hence, we assume that $w_1 \geq 0$, $w_2 \geq 0$, $k \geq 0$, $\lambda_1 \geq 0$, and $\lambda_2 \geq 0$. Because of this assumption, we are able to set up the following system of equations based on the KKT conditions:

$$\begin{aligned}(\lambda_1) \wedge (\lambda_2) : w_1 &= w_2 = k \\(w_1) \wedge (w_2) : B'_1(k) + B'_2(k) &= 2 \cdot c + r \rightarrow 110 - k + 150 - k = 2 \cdot 10 + 50\end{aligned}$$

The solution is $w_1 = w_2 = k = 95$. The value of the dual variables λ_1 and λ_2 can be calculated from constraint Eqs. (4.16) and (4.17):

$$\begin{aligned}(w_1) : \lambda_1 &= B'_1(w_1) - c = 110 - 95 - 10 = 5 \\(w_2) : \lambda_2 &= B'_2(w_2) - c = 150 - 95 - 10 = 45\end{aligned}$$

The dual variables are non-negative, hence, we do not find a contradiction and this case leads to optimality. The price in the winter/(off-peak) period, which is p_1 , is lower than in the summer/(peak) period, being p_2 :

$$p_1 = B'_1(w_1) = c + \lambda_1 = 15 \quad p_2 = B'_2(w_2) = c + \lambda_2 = 55$$

Therefore, the capacity is financed by 90% during the summer period, because $\frac{\lambda_2}{r} = \frac{45}{50} = 0.9$, while the capacity is also financed by 10% during the winter period, because $\frac{\lambda_1}{r} = \frac{5}{50} = 0.1$.

Exercise 4.5 Proportional water right allotments

The allotment of tradable water rights may be a promising instrument to reconcile efficiency and fairness. Let us assume that the entitlements to water use are reduced in a proportional way along with increased water scarcity. In this exercise, we want to analyze how the proportional allotment is able to fulfill the criteria of efficiency and fairness. Let us assume that there are two water users, say, firms. Marginal profits are

$$B'_i = a_i - b_i w_i, \quad a_1 = 40, a_2 = 20, b_1 = b_2 = 1 \quad (4.67)$$

If water is abundant and there are no treatment costs of water, firms set marginal profits equal to zero which leads to

$$\hat{w}_1 = a_1/b_1 = 40 \quad \text{and} \quad \hat{w}_2 = a_2/b_2 = 20 \quad (4.68)$$

What are the profits? The calculation needs the profit function which can be achieved by integrating the marginal profit function with respect to w .

$$B_i = \int_0^{w_i} [a_i - b_i v] dv = w_i(a - (b_i/2)w_i) \quad (4.69)$$

Inserting the calculated water usage yields profits $B_1 = 800$ and $B_2 = 200$. The profit of firm 1 is four times as high as that of firm 2.

In the course of increasing water scarcity firms total water use of $40 + 20 = 60$ cannot be covered any more. Instead there is only water available of a total of $\bar{W} = 30$, i.e., half of the former total use. Water entitlements \bar{w}_i will be allotted proportionally³⁵:

$$\bar{w}_i = \frac{\hat{w}_i}{\hat{w}_1 + \hat{w}_2} \bar{W} = \hat{w}_i \frac{\bar{W}}{\hat{w}_1 + \hat{w}_2} = \frac{1}{2} \hat{w}_i \quad (4.70)$$

³⁵Notice that $\bar{W} = (1/2)(\hat{w}_1 + \hat{w}_2)$.

Utilizing Eq. (4.68) yields $\bar{w}_1 = 20$ and $\bar{w}_2 = 10$. If these allotments are not tradable, the resulting profits can be calculated by halving the former water use, and inserting the results into the profit function Eq. (4.69) which yields $\bar{B}_1 = 600$ and $\bar{B}_2 = 150$. The distributional effects of the proportional allotment can be captured by the profit ratio $\bar{B}_1/\bar{B}_2 = 4$. Hence, water rationing has not changed the profit distribution, according to the proportionality rule. However, the water allocation is not optimal if one takes the efficiency criterion into account. Hence, we maximize total profits under the constraint that total water use does not exceed the sustainability constraint $\bar{W} = 30$.

$$\max_{w_1, w_2} [B_1(w_1) + B_2(w_2)], \quad \text{s.t. } w_1 + w_2 \leq \bar{W} \quad (4.71)$$

The first-order conditions require w_i^* to be set in a way that $B'_1 = B'_2$ while meeting the sustainability constraint. Inserting all relevant values leads to the solution $w_1^* = 25$ and $w_2^* = 5$. We see immediately that these values differ from the proportional allotments $\bar{w}_1 = 20$ and $\bar{w}_2 = 10$.

The efficient solution can be implemented by introducing a market for water entitlements. The market equilibrium is characterized by an excess demand for water by firm 1 of $w_1^* - \bar{w}_1 = 25 - 20$. Firm 2 is net seller of entitlements $\bar{w}_2 - w_2^* = 10 - 5$. The equilibrium price can be calculated by inserting w_1^* into the marginal profit function of firm 1, i.e.,

$$p^* = B'_1(w_1^*) = a_1 - b_1 w_1^* = 15 \quad (4.72)$$

To calculate net profits, we insert the market solution into the profit functions and deduct the net demand for water

$$B_i^n = w_i^* (a - (b_i/2)w_i^*) - p^*(w_i^* - \bar{w}_i) \quad (4.73)$$

which yields $B_1^n = 612.5$ and $B_2^n = 162.5$. If we compare these values with profits under the rationing system without trading, we see that both profits have risen. This is not surprising, because trade is voluntary and therefore only takes place when both parties are better off after a trade. By close inspection, we also see that the profit of the small firm 1 has risen stronger than that of the bigger firm 2. The profit ratio is now $B_1^n/B_2^n \approx 3.8$. Hence, in our example, the enhancement of efficiency by allowing to trade water allotments leads to a profit distribution in favor of the smaller firm. Obviously, the principle of efficiency and the principle of justice need not always be in conflict. In addition, the introduction of trade fulfills Rawls' principle of difference: The worst placed user improves her situation under the observed regime.

4.8 Further Reading

Water is a multidimensional resource that not only serves as a private good. Hane-mann (2004) gives an instructive historical outline of the economic dimension of water, taking into account not only efficiency aspects but also the primary supply based on human rights. The design of tariffs for water services (drinking water, sanitation) should take into account various criteria. Boland and Whittington (2000a) and Massarutto (2007b) present the various criteria and evaluate different tariff structures on the basis of these criteria. In particular, IBTs are taken into account (see also OECD 2010; Walker 2009). Rogers et al. (1998) review the tariff building criteria underlying the Dublin Water Principles. Designed as practical guides, OECD (2009) and OECD (2010) provide a problem-oriented introduction and an overview of the financing of water infrastructures with special attention to tariff policy. Not only internal costs but also social costs of water supply (environmental costs, etc.) are taken into account.

There are a number of tariff structures that are used in the supply of infrastructure goods (water, energy, transportation). The basic analysis techniques of the effects of tariff variants are introduced in the microeconomic textbook of Varian and Repcheck (2010). But there are not only economic aspects to be considered. We often observe political and legal requirements, which have to be taken into account in the tariff structure, particularly in the case of network services. Cremer et al. (2001) analyze the universal service provider which is subject to the universal service obligation, i.e., the provision to serve all customers at affordable (and equal) rates.

IBTs are widespread in the water sector, especially in Asia. There is a large number of studies on how these tariffs work, some more practical, others more theoretical, among the latter Boland and Whittington (2000a, b), Whittington (2003), Dahan and Nisan (2007), and Monteiro and Rosetta-Palma (2011). Meran and von Hirschhausen (2017) develop a microeconomic model with social preferences where a strong inequity aversion leads to IBTs as tariff system.

Unconnected water markets play a major role in developing countries. There are many case studies investigating the precise institutional, cultural, and political characteristics. A comprehensive study is Kjellén and McGranahan (2006), which examines the water supply in Dar es Salaam, Tanzania, Hailu et al. (2011) for Nigeria, Kenya, and for Cochabamba, Bolivia, Wutich et al. (2016). An analytical economic analysis of rent extracting behavior is given by Lovei and Whittington (1991). Baumol (1982) analyzes cost structures that allow for competitive behavior of suppliers even if only few actors operate in the market.

Rationing is often used in the case of water scarcity. Lund and Reed (1995) and Olmstead and Stavins (2009a) provide an overview of the different types and an economic assessment. An analysis that takes into account aspects of equity and efficiency in times of severe scarcity (e.g., war times) is Tobin (1970).

4.9 Chapter-Annex: Overview of Water Tariff Structures

Table 4.4 Tariff structures for water supply and sanitation and policy objectives: a synthesis based on OECD (2010)

Tariff structure	Examples	Ecological sustainability	Economic efficiency	Financial sustainability	Equity/affordability
Uniform flat fee	Sub-areas of two water supply companies in the United Kingdom. Still used by many sampled non-OECD utilities	Very poor. No incentives to water saving nor to other aspects of sustainable water use	Poor for drinking water (no linkage between fee structure and behavior that may help minimize investment). OK for water-borne sanitation (costs do not depend on water consumption)	Potentially OK, but commitment to cost recovery is what really matters. Avoid political determination of fees	Very regressive (unless properly integrated with other elements of a social security system)
Non-uniform flat rate linked with specific aspects of households, e.g., (i) property value or other income proxy, (ii) dwelling characteristics linked with water use	Still used by 70% of UK households, common in the former Soviet Union	Poor if linked with income-related variable. Good if linked with dwelling characteristics linked with water use (e.g., use of water recycling devices) or with specific behavior that wants to be encouraged (e.g., rainwater harvesting)	As above	As above, provided that total revenues are guaranteed	Potentially good effects, provided that criteria used correspond to personal wealth. Regressive otherwise (unless properly integrated with other elements of a social security system)

(continued)

Table 4.4 (continued)

Tariff structure	Examples	Ecological sustainability	Economic efficiency	Financial sustainability	Equity/affordability
Uniform volumetric rate + 0 fixed charge	Still present in numerous OECD countries. Most recurrent in sample of non-OECD utilities	As above; higher, since 0 fixed charge means a larger marginal rate (for the same revenue levels)	Efficient if water is scarce or infrastructure nearing capacity (i.e., if there is rivalry in consumption) or if variable costs are high compared to fixed costs. Not very efficient if otherwise it would discourage users but this would reduce societal benefits. Inefficiency depends on demand elasticity (the lower the elasticity, the lower the inefficiency)	Good potential for financial recovery. Can have (temporary) negative impact on revenue in case of a sudden move from flat charges due to impact on demand (e.g., Berlin experience)	Depends on income elasticity. If this is low, it can hit large poor households hard
Uniform volumetric rate + fixed charge > 0	Classic, e.g., Germany (structure enshrined in law)	High, depending on the marginal rate (impact on demand only if it is high enough) + individual metering	Optimal provided the following applies: volumetric rate = SRMC (short-run marginal cost) and fixed charge = lump sum. Particularly suited in case SRMC is constant (e.g., electricity)	As above	Depends on size of fixed charge, but tends to be regressive (not so only if marginal cost is high and income elasticity is high which is rare). Size of fixed charge can be differentiated based on income

(continued)

Table 4.4 (continued)

Tariff structure	Examples	Ecological sustainability	Economic efficiency	Financial sustainability	Equity/affordability
Uniform volumetric rate + rebate (fixed charge < 0)	No known application. May have been applied in municipalities in the United States	As above. Highest if rebates take into account specific circumstances (e.g., use of water recycling devices, drip irrigation or water-saving sprinklers in gardens) or with specific behavior that wants to be encouraged (e.g., rainwater harvesting, use of less pollutant detergents)	As above. In turn, could be efficient in combination with a positive fixed fee (idea: $r = SRMC$; fixed cost redistributed including a rebate for the poor)	As above	Progressive and useful for reducing impact on poor. But only if rebate is targeted; otherwise, distributive effect depending on income elasticity, just like with IBTs
Traditional IBT (both block widths and prices fixed) + fixed charge	Italy. Increasing number of developing countries	Highest, provided that metering is individual and marginal rates in the upper blocks are high	Potentially the best solution provided $r = SRMC$ and fixed charge = lump sum. Particularly suited in case $SRMC$ is increasing (e.g., costly extra supply to be purchased)	As above	Can be very regressive if: (i) low demand elasticity to income; (ii) resulting average tariff is below cost recovery levels and this discourages extension of network; (iii) many households sharing the same tap

(continued)

Table 4.4 (continued)

Tariff structure	Examples	Ecological sustainability	Economic efficiency	Financial sustainability	Equity/affordability
IBT + fixed charge + exact occupancy amendment	Flanders, Brussels Malta, some communes in Luxembourg	As above, but reduced incentives for large families	Depends on how closely the resulting average volumetric charge reflects SRMC. Rest as above	As above	Reduces impact on large families (best if accompanied by reduction of leaks and improved efficiency of appliances). Depends on correlation of size and income of households. Problem (ii) above remains
IBT + fixed charge + low-income households may apply for extension	Proposed Social Tariff Plan in Portugal	As above, but reduced incentives for low-income households that apply for extension of blocks	Good for reducing demand in peak periods and optimizing capacity use	Uncertainty about number of households applying (may be reduced over time)	Successful, if all eligible claim and block width reflect consumption patterns of the poor. Problem (ii) above remains
IBT + fixed charge + larger households (e.g., N = 4) may apply for extension	Some Spanish cities. Greek DEYA, cities. Proposed option in Portugal	As above, but reduced incentives for large families that apply for extension of blocks	Depends on whether there is a fixed charge or not	As above	Depends on correlation of size and income of households. Problem (ii) above remains
IBT + fixed charge + targeted subsidies to low income	Chile	Highest, provided that metering is individual and marginal rates in the upper blocks are high	As above	As above	Depends on the capacity to target the poor. Problem (ii) above remains

References

- Baumol, W. (1982). Contestable markets: An uprising in the theory of industry structure. *American Economic Review*, 72(1), 1–15.
- Boland, J. J., & Whittington, D. (2000a). The political economy of water tariff design in developing countries: Increasing block tariffs versus uniform price with rebate. Pages 215–235 of: A. Dinar (Ed.). *The political economy of water pricing reforms*. Oxford: Oxford University Press.
- Boland, J. J., & Whittington, D. (2000b). *Water tariff design in developing countries: disadvantages of increasing block tariffs (IBTs) and advantages of uniform price with rebate (UPR) designs*. Ottawa: The International Development Research Centre (IDRC).
- Collinge, R. A. (1994). Transferable rate entitlements: The overlooked opportunity in municipal water pricing. *Public Finance Quarterly*, 22(12), 46–64.
- Cremer, H., Gasmı, F., Grimaud, A., & Laffont, J. J. (2001). Universal service: An economic perspective. *Annals of Public and Cooperative Economics*, 72(1), 5–43.
- Dahan, M., & Nisan, U. (2007). Unintended consequences of increasing block tariffs: Pricing policy in urban water. *Water Resources Research*, 43(3), 1–10.
- Dinar, A., Pochat, V., Albiac-Murillo, J., & others. (2015). *Water pricing experiences and innovations*. New York City, NY, U.S.A.: Springer.
- Grafton, R. Q., & Ward, M. B. (2008). Prices versus rationing: Marshallian surplus and mandatory water restrictions. *The Economic Record*, 84, 57–65.
- Hailu, D., Rendtorff-Smith, S., & Tsukada, R. (2011). *Small-scale water providers in Kenya: Pioneers or predators?*. New York: United Nations Development Programme.
- Hanemann, W. A. (2004). The economic conception of water. Pages 61–91 of: P. Rogers, M. Ramón Llamas, L. Martínez-Cortina, (Eds.), *Water crisis: Myth or reality?* London: Taylor and Francis plc.
- Hoekstra, A. Y. (1998). Appreciation of water: Four perspectives. *Water Policy*, 1, 605–622.
- Kjellén, M., & McGranahan, G. (2006). *Informal water vendors and the urban poor*. Human Settlements Discussion Paper Series, London: International Institute for Environment and Development.
- Loeb, M., & Magat, W. A. (1979). A decentralized method for utility regulation. *The Journal of Law and Economics*, 22(2), 399–404.
- Lovei, L., & Whittington, D. (1991). *Rent seeking in water supply*. Washington, DC: World Bank Infrastructure and Urban Development Department Report INU 85.
- Lund, J. R., & Reed, R. U. (1995). Drought water rationing and transferable rations. *Journal of Water Resources Planning and Management*, 121(6), 429–437.
- Massarutto, A. (2007a). Abstraction charges: How can the theory guide us? In: *Presentation made at the OECD expert meeting on sustainable financing for affordable water services: From theory to practice*.
- Massarutto, A. (2007b). Water pricing and full cost recovery of water services: Economic incentive or instrument of public finance? *Water Policy*, 9(6), 591–613.
- Meran, G., & von Hirschhausen, C. (2014). Increasing block tariffs in the water sector—an interpretation in terms of social preferences. *The BE Journal of Economic Analysis & Policy*, 17(3).
- Meran, G., & von Hirschhausen, C. (2017). Increasing block tariffs in the water sector – an interpretation in terms of social preferences. *The B.E. Journal of Economic Analysis and Policy*, 17.
- Meran, G., Siehlow, M., & von Hirschhausen, C. (2020). *Pipes, taps and vendors: An integrated water management approach* (p. 1856). DIW Berlin Discussion Paper.
- Monteiro, H., & Rosetta-Palma, C. (2011). Pricing for scarcity? An efficiency analysis of increasing block tariffs. *Water Resources Research*, 47, 1–11.
- OECD. (2009). *Managing water for all: An OECD perspective on pricing and financing*. Paris: OECD Publishing.
- OECD. (2010). *Pricing water resources and water and sanitation services*. Paris: OECD Publishing.
- Olmstead, S. M., & Stavins, R. N. (2009a). Comparing price and nonprice approaches to urban water conservation. *Water Resources Research*, 45, 1–10.

- Olmstead, S. M., & Stavins, R. N. (2009b). Comparing price and nonprice approaches to urban water conservation. *Water Resources Research*, 45(4).
- Roemer, J. E. (1996). *Theories of distributive justice*. Cambridge, MA: Harvard University Press.
- Rogers, P., Bhatia, R., & Huber, A. (1998). *Water as a social and economic good: How to put the principle into practice*. Stockholm: Global Water Partnership/Swedish International Development Cooperation Agency.
- Tobin, J. (1970). On limiting the domain of inequity. *Journal of Law and Economics*, 13.
- United Nations. (2010). *The human right to water and sanitation*. New York: G.A. Res. 64/292, UN DOC. A/RES/64/292 (August 3, 2010).
- Varian, H. R., & Repcheck, J. (2010). *Intermediate microeconomics: A modern approach* (Vol. 6). New York: WW Norton & Company.
- Walker, A. (2009). *The independent review of charging for household water and sewerage services: Final report*. London: Department for Environment, Food and Rural Affairs.
- Whittington, D. (2003). Municipal water pricing and tariff design: A reform agenda for South Asia. *Water Policy*, 5, 61–76.
- World Health Organization (WHO). (2012). How much water is needed in emergencies. *WHO (Geneva, Switzerland)*, 9, 5.
- Wutich, A., Beresford, M., & Carvajal, C. (2016). Can informal water vendors deliver on the promise of a human right to water? Results from Cochabamba, Bolivia. *World Development*, 79, 14–24.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

